BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Nine Number One February 2023 ISSN: 1350-7265

CONTENTS

(continued)

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

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Nine Number One February 2023 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 29 Number 1 February 2023 Pages 1–874 ISI/BS

Volume 29 Number 1 February 2023 ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope

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

Bernoulli Society for Mathematical Statistics and Probability

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

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

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

Executive Committee

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

---

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

In 2023 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **29**(1), 2023, 1–23
https://doi.org/10.3150/21-BEJ1447

Conformal prediction: A unified review of theory and new challenges

MATTEO FONTANA[1,2,3,c], GIANLUCA ZENI[1,a] and SIMONE VANTINI[1,b]

[1]*MOX-Department of Mathematics, Politecnico di Milano, Italy.* [a]*gianluca.zeni@mail.polimi.it,* [b]*simone.vantini@polimi.it*
[2]*Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Italy*
[3]*now at European Commission, Joint Research Centre (JRC), Ispra, Italy.* [c]*matteo.fontana@ec.europa.eu*

In this work we provide a review of basic ideas and novel developments about Conformal Prediction — an innovative distribution-free, non-parametric forecasting method, based on minimal assumptions — that is able to yield in a very straightforward way prediction sets that are valid in a statistical sense also in the finite sample case. The discussion provided in the paper covers the theoretical underpinnings of Conformal Prediction, and then proceeds to list the more advanced developments and adaptations of the original idea.

Keywords: Conformal prediction; nonparametric statistics; prediction intervals; review

References

Bahadur, R.R. and Savage, L.J. (1956). The nonexistence of certain statistical procedures in nonparametric problems. *Ann. Math. Stat.* **27** 1115–1122. MR0084241 https://doi.org/10.1214/aoms/1177728077

Balasubramanian, V., Ho, S.-S. and Vovk, V. (2014). *Conformal Prediction for Reliable Machine Learning: Theory, Adaptations and Applications*. London: Newnes.

Barber, R.F., Candes, E.J., Ramdas, A. and Tibshirani, R.J. (2019a). Predictive inference with the jackknife+. arXiv preprint. Available at arXiv:1904.06019.

Barber, R.F., Candes, E.J., Ramdas, A. and Tibshirani, R.J. (2019b). Conformal prediction under covariate shift. arXiv preprint. Available at arXiv:1904.06019.

Barber, R.F., Candès, E.J., Ramdas, A. and Tibshirani, R.J. (2021). The limits of distribution-free conditional predictive inference. *Inf. Inference* **10** 455–482. MR4270755 https://doi.org/10.1093/imaiai/iaaa017

Burnaev, E. and Vovk, V. (2014). Efficiency of conformalized ridge regression. In *Conference on Learning Theory* 605–622.

Capuccini, M., Carlsson, L., Norinder, U. and Spjuth, O. (2015). Conformal prediction in Spark: Large-scale machine learning with confidence. In *2015 IEEE/ACM 2nd International Symposium on Big Data Computing (BDC)* 61–67. IEEE.

Carlsson, L., Ahlberg, E., Boström, H., Johansson, U. and Linusson, H. (2015). Modifications to p-values of conformal predictors. In *International Symposium on Statistical Learning and Data Sciences* 251–259. Springer.

Chen, W., Chun, K.-J. and Barber, R.F. (2018). Discretized conformal prediction for efficient distribution-free inference. *Stat* **7** e173, 11. MR3769053 https://doi.org/10.1002/sta4.173

Chen, W., Wang, Z., Ha, W. and Barber, R.F. (2016a). Trimmed conformal prediction for high-dimensional models. arXiv preprint. Available at arXiv:1611.09933.

Chen, X., Monfort, M., Liu, A. and Ziebart, B.D. (2016b). Robust covariate shift regression. In *Artificial Intelligence and Statistics* 1270–1279.

Chernozhukov, V., Wuthrich, K. and Zhu, Y. (2018). Exact and robust conformal inference methods for predictive machine learning with dependent data. arXiv preprint. Available at arXiv:1802.06300.

Dashevskiy, M. and Luo, Z. (2011). Time series prediction with performance guarantee. *IET Commun.* **5** 1044–1051.

Devetyarov, D. and Nouretdinov, I. (2010). Prediction with confidence based on a random forest classifier. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* 37–44. Springer.

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.* **14** 1–67. With a discussion and a rejoinder by the authors. MR0829555 https://doi.org/10.1214/aos/1176349830

Diquigiovanni, J., Fontana, M. and Vantini, S. (2021a). Conformal prediction bands for multivariate functional data. *J. Multivariate Anal.* in press.

Diquigiovanni, J., Fontana, M. and Vantini, S. (2021b). The importance of being a band: finite-sample exact distribution-free prediction sets for functional data. arXiv preprint. Available at arXiv:2102.06746.

Diquigiovanni, J., Fontana, M. and Vantini, S. (2021c). Distribution-free prediction bands for multivariate functional time series: an application to the Italian gas market. arXiv preprint. Available at arXiv:2107.00527.

Donoho, D.L. (1988). One-sided inference about functionals of a density. *Ann. Statist.* **16** 1390–1420. MR0964930 https://doi.org/10.1214/aos/1176351045

Dunn, R. and Wasserman, L. (2018). Distribution-free prediction sets with random effects. arXiv preprint. Available at arXiv:1809.07441.

Fontana, M., Zeni, G. and Vantini, S. (2023). Supplement to "Conformal prediction: A unified review of theory and new challenges." https://doi.org/10.3150/21-BEJ1447SUPP

Gammerman, A., Vovk, V. and Vapnik, V. (1998). Learning by transduction. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* 148–155. Morgan Kaufmann Publishers Inc.

Hebiri, M. (2010). Sparse conformal predictors. *Stat. Comput.* **20** 253–266. MR2610776 https://doi.org/10.1007/s11222-009-9167-2

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

Ho, S.-S. and Wechsler, H. (2004). Learning from data streams via online transduction. Ma et al, 45–52.

Johansson, U., Boström, H., Löfström, T. and Linusson, H. (2014). Regression conformal prediction with random forests. *Mach. Learn.* **97** 155–176. MR3252831 https://doi.org/10.1007/s10994-014-5453-0

Johansson, U., Ahlberg, E., Boström, H., Carlsson, L., Linusson, H. and Sönströd, C. (2015). Handling small calibration sets in Mondrian inductive conformal regressors. In *International Symposium on Statistical Learning and Data Sciences* 271–280. Springer.

Lambrou, A., Nouretdinov, I. and Papadopoulos, H. (2015). Inductive Venn Prediction. *Ann. Math. Artif. Intell.* **74** 181–201. MR3353902 https://doi.org/10.1007/s10472-014-9420-z

Lambrou, A., Papadopoulos, H., Nouretdinov, I. and Gammerman, A. (2012). Reliable probability estimates based on support vector machines for large multiclass datasets. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* 182–191. Springer.

Lei, J. (2019). Fast exact conformalization of the lasso using piecewise linear homotopy. *Biometrika* **106** 749–764. MR4031197 https://doi.org/10.1093/biomet/asz046

Lei, J., Rinaldo, A. and Wasserman, L. (2015). A conformal prediction approach to explore functional data. *Ann. Math. Artif. Intell.* **74** 29–43. MR3353895 https://doi.org/10.1007/s10472-013-9366-6

Lei, J., Robins, J. and Wasserman, L. (2013). Distribution-free prediction sets. *J. Amer. Statist. Assoc.* **108** 278–287. MR3174619 https://doi.org/10.1080/01621459.2012.751873

Lei, J. and Wasserman, L. (2014). Distribution-free prediction bands for non-parametric regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 71–96. MR3153934 https://doi.org/10.1111/rssb.12021

Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J. and Wasserman, L. (2018). Distribution-free predictive inference for regression. *J. Amer. Statist. Assoc.* **113** 1094–1111. MR3862342 https://doi.org/10.1080/01621459.2017.1307116

Linusson, H., Norinder, U., Boström, H., Johansson, U. and Löfström, T. (2014a). On the calibration of aggregated conformal predictors. *Proc. Mach. Learn. Res.* **60** 261–270.

Linusson, H., Johansson, U., Boström, H. and Löfström, T. (2014b). Efficiency comparison of unstable transductive and inductive conformal classifiers. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* 261–270. Springer.

Medarametla, D. and Candès, E. (2021). Distribution-free conditional median inference. *Electron. J. Stat.* **15** 4625–4658. MR4316665 https://doi.org/10.1214/21-ejs1910

Melluish, T., Vovk, V. and Gammerman, A. (1999). Transduction for regression estimation with confidence. In *Neural Information Processing Systems, NIPS'99*.

Melluish, T., Saunders, C., Nouretdinov, I. and Vovk, V. (2001). Comparing the Bayes and typicalness frameworks. In *European Conference on Machine Learning* 360–371. Springer.

Nouretdinov, I., Melluish, T. and Vovk, V. (2001). Ridge regression confidence machine. In *ICML* 385–392.

Nouretdinov, I. and Vovk, V. (2006). Criterion of calibration for transductive confidence machine with limited feedback. *Theoret. Comput. Sci.* **364** 3–9. MR2268298 https://doi.org/10.1016/j.tcs.2006.07.037

Nouretdinov, I., Vovk, V., Vyugin, M. and Gammerman, A. (2001). Pattern recognition and density estimation under the general i.i.d. assumption. In *Computational Learning Theory (Amsterdam, 2001). Lecture Notes in Computer Science* **2111** 337–353. Berlin: Springer. MR2042045 https://doi.org/10.1007/3-540-44581-1_22

Nouretdinov, I., Volkhonskiy, D., Lim, P., Toccaceli, P. and Gammerman, A. (2018). Inductive Venn-Abers predictive distribution. *Proc. Mach. Learn. Res.* **91** 1–22.

Nouretdinov, I., Gammerman, J., Fontana, M. and Rehal, D. (2020). Multi-level conformal clustering: A distribution-free technique for clustering and anomaly detection. *Neurocomputing* **397** 279–291.

Papadopoulos, H. (2008). Inductive conformal prediction: Theory and application to neural networks. In *Tools in Artificial Intelligence*. InTech.

Papadopoulos, H., Gammerman, A. and Vovk, V. (2008). Normalized nonconformity measures for regression conformal prediction. In *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA 2008)* 64–69.

Papadopoulos, H. and Haralambous, H. (2010). Neural networks regression inductive conformal predictor and its application to total electron content prediction. In *International Conference on Artificial Neural Networks* 32–41. Springer.

Papadopoulos, H. and Haralambous, H. (2011). Reliable prediction intervals with regression neural networks. *Neural Netw.* **24** 842–851.

Papadopoulos, H., Vovk, V. and Gammerman, A. (2002). Qualified prediction for large data sets in the case of pattern recognition. In *ICMLA* 159–163. MR2805257

Papadopoulos, H., Vovk, V. and Gammerman, A. (2011). Regression conformal prediction with nearest neighbours. *J. Artificial Intelligence Res.* **40** 815–840. MR2805257 https://doi.org/10.1613/jair.3198

Papadopoulos, H., Proedrou, K., Vovk, V. and Gammerman, A. (2002). Inductive confidence machines for regression. In *Machine Learning: ECML 2002. Lecture Notes in Computer Science* **2430** 345–356. Berlin: Springer. MR2050303 https://doi.org/10.1007/3-540-36755-1_29

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

Riabko, D. (2005). On the flexibility of theoretical models for pattern recognition, Ph.D. thesis, Citeseer.

Romano, Y., Patterson, E. and Candès, E.J. (2019). Conformalized quantile regression. arXiv preprint. Available at arXiv:1905.03222.

Ryabko, D., Vovk, V. and Gammerman, A. (2003). Online region prediction with real teachers. Submitted for publication. Criterion of Calibration for Transductive Confidence Machine, 267.

Saunders, C., Gammerman, A. and Vovk, V. (1999). Transduction with confidence and credibility. In *Proceedings of the International Joint Conference on Artificial Intelligence* **2** 722–726.

Saunders, C., Gammerman, A. and Vovk, V. (2000). Computationally efficient transductive machines. In *International Conference on Algorithmic Learning Theory* 325–337. Springer.

Shafer, G. and Vovk, V. (2008). A tutorial on conformal prediction. *J. Mach. Learn. Res.* **9** 371–421. MR2417240

Shen, J., Liu, R.Y. and Xie, M. (2018). Prediction with confidence—a general framework for predictive inference. *J. Statist. Plann. Inference* **195** 126–140. MR3760843 https://doi.org/10.1016/j.jspi.2017.09.012

Shimodaira, H. (2000). Improving predictive inference under covariate shift by weighting the log-likelihood function. *J. Statist. Plann. Inference* **90** 227–244. MR1795598 https://doi.org/10.1016/S0378-3758(00)00115-4

Steinberger, L. and Leeb, H. (2016). Leave-one-out prediction intervals in linear regression models with many variables. arXiv preprint. Available at arXiv:1602.05801.

Steinberger, L. and Leeb, H. (2018). Conditional predictive inference for high-dimensional stable algorithms. arXiv preprint. Available at arXiv:1809.01412.

Valiant, L.G. (1984). A theory of the learnable. *Commun. ACM* **27** 1134–1142.

Vapnik, V.N. (1998). *Statistical Learning Theory. Adaptive and Learning Systems for Signal Processing, Communications, and Control*. New York: Wiley. MR1641250

Vovk, V. (2002). Online confidence machines are well-calibrated. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on* 187–196. IEEE.

Vovk, V. (2013a). Conditional validity of inductive conformal predictors. *Mach. Learn.* **92** 349–376. MR3080332 https://doi.org/10.1007/s10994-013-5355-6

Vovk, V. (2013b). Transductive conformal predictors. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* 348–360. Springer.

Vovk, V. (2015). Cross-conformal predictors. *Ann. Math. Artif. Intell.* **74** 9–28. MR3353894 https://doi.org/10.1007/s10472-013-9368-4

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

Vovk, V., Nouretdinov, I. and Gammerman, A. (2009). On-line predictive linear regression. *Ann. Statist.* **37** 1566–1590. MR2509084 https://doi.org/10.1214/08-AOS622

Vovk, V., Shafer, G. and Nouretdinov, I. (2004). Self-calibrating probability forecasting. In *Advances in Neural Information Processing Systems* 1133–1140.

Vovk, V., Lindsay, D., Nouretdinov, I. and Gammerman, A. (2003). Mondrian confidence machine. Technical report.

Vovk, V., Fedorova, V., Nouretdinov, I. and Gammerman, A. (2016). Criteria of efficiency for conformal prediction. In *Conformal and Probabilistic Prediction with Applications* 23–39. Springer International Publishing.

Vovk, V., Nouretdinov, I., Manokhin, V. and Gammerman, A. (2018). Cross-conformal predictive distributions. In *Conformal and Probabilistic Prediction and Applications* 37–51.

Vovk, V., Shen, J., Manokhin, V. and Xie, M. (2019). Nonparametric predictive distributions based on conformal prediction. *Mach. Learn.* **108** 445–474. MR3917055 https://doi.org/10.1007/s10994-018-5755-8

Wang, D., Wang, P. and Shi, J. (2018). A fast and efficient conformal regressor with regularized extreme learning machine. *Neurocomputing* **304** 1–11.

Bernoulli **29**(1), 2023, 24–51
https://doi.org/10.3150/21-BEJ1448

Statistics of the two star ERGM

SUMIT MUKHERJEE[a] and YUANZHE XU[b]

Columbia University, New York, NY, USA. [a]*sm3949@columbia.edu,* [b]*yuanzhe.xu@columbia.edu*

In this paper, we explore the two star Exponential Random Graph Model, which is a two parameter exponential family on the space of simple labeled graphs. We introduce auxiliary variables to express the two star model as a mixture of the β model on networks. Using this representation, we study asymptotic distribution of the number of edges, and the sampling variance of the degrees. In particular, the limiting distribution for the number of edges has similar phase transition behavior to that of the magnetization in the Curie-Weiss Ising model of Statistical Physics. Using this, we show existence of consistent estimates for both parameters. Finally, we prove that the centered partial sum of degrees converges as a process to a Brownian bridge in all parameter domains, irrespective of the phase transition.

Keywords: ERGM; auxiliary variables; phase transition; consistent estimation; two-star

References

[1] Andersen, H.C. and Diaconis, P. (2007). Hit and run as a unifying device. *J. Soc. Fr. Stat. & Rev. Stat. Appl.* **148** 5–28. MR2502361

[2] Anderson, C.J., Wasserman, S. and Crouch, B. (1999). A p* primer: Logit models for social networks. *Soc. Netw.* **21** 37–66.

[3] Bhamidi, S., Bresler, G. and Sly, A. (2008). Mixing time of exponential random graphs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science* 803–812. IEEE.

[4] Blitzstein, J. and Diaconis, P. (2010). A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet Math.* **6** 489–522. MR2809836 https://doi.org/10.1080/15427951.2010.557277

[5] Chatterjee, S. and Diaconis, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41** 2428–2461. MR3127871 https://doi.org/10.1214/13-AOS1155

[6] Chatterjee, S., Diaconis, P. and Sly, A. (2011). Random graphs with a given degree sequence. *Ann. Appl. Probab.* **21** 1400–1435. MR2857452 https://doi.org/10.1214/10-AAP728

[7] Chatterjee, S. and Mukherjee, S. (2019). Estimation in tournaments and graphs under monotonicity constraints. *IEEE Trans. Inf. Theory* **65** 3525–3539. MR3959003 https://doi.org/10.1109/TIT.2019.2893911

[8] Comets, F. and Gidas, B. (1991). Asymptotics of maximum likelihood estimators for the Curie-Weiss model. *Ann. Statist.* **19** 557–578. MR1105836 https://doi.org/10.1214/aos/1176348111

[9] Deb, N. and Mukherjee, S. (2020). Fluctuations in Mean-Field Ising models. ArXiv Preprint. Available at arXiv:2005.00710.

[10] DeMuse, R., Larcomb, D. and Yin, M. (2018). Phase transitions in edge-weighted exponential random graphs: Near-degeneracy and universality. *J. Stat. Phys.* **171** 127–144. MR3773854 https://doi.org/10.1007/s10955-018-1991-3

[11] Edwards, R.G. and Sokal, A.D. (1988). Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Phys. Rev. D* **38** 2009–2012. MR0965465 https://doi.org/10.1103/PhysRevD.38.2009

[12] Ellis, R.S. and Newman, C.M. (1978). The statistics of Curie-Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332 https://doi.org/10.1007/BF01012508

[13] Frank, O. and Strauss, D. (1986). Markov graphs. *J. Amer. Statist. Assoc.* **81** 832–842. MR0860518

[14] Ganguly, S. and Nam, K. (2019). Sub-critical Exponential random graphs: Concentration of measure and some applications. ArXiv Preprint. Available at arXiv:1909.11080.

[15] Ghosal, P. and Mukherjee, S. (2020). Joint estimation of parameters in Ising model. *Ann. Statist.* **48** 785–810. MR4102676 https://doi.org/10.1214/19-AOS1822

[16] Götze, F., Sambale, H. and Sinulis, A. (2021). Concentration inequalities for polynomials in α-sub-exponential random variables. *Electron. J. Probab.* **26** Paper No. 48. MR4247973 https://doi.org/10.1214/21-ejp606

[17] Handcock, M.S., Robins, G., Snijders, T., Moody, J. and Besag, J. (2003). Assessing degeneracy in statistical models of social networks. Technical Report, Working paper.

[18] Holland, P.W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *J. Amer. Statist. Assoc.* **76** 33–65. MR0608176

[19] Mukherjee, R., Mukherjee, S. and Yuan, M. (2018). Global testing against sparse alternatives under Ising models. *Ann. Statist.* **46** 2062–2093. MR3845011 https://doi.org/10.1214/17-AOS1612

[20] Mukherjee, S. (2020). Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics. *Bernoulli* **26** 1016–1043. MR4058359 https://doi.org/10.3150/19-BEJ1135

[21] Mukherjee, S. Xu, Y. (2023). Supplement to "Statistics of the two star ERGM." https://doi.org/10.3150/21-BEJ1448SUPP

[22] Papangelou, F. (1989). On the Gaussian fluctuations of the critical Curie-Weiss model in statistical mechanics. *Probab. Theory Related Fields* **83** 265–278. MR1012501 https://doi.org/10.1007/BF00333150

[23] Park, J. and Newman, M.E.J. (2004). Solution of the two-star model of a network. *Phys. Rev. E (3)* **70** 066146. MR2133810 https://doi.org/10.1103/PhysRevE.70.066146

[24] Park, J. and Newman, M.E.J. (2004). Statistical mechanics of networks. *Phys. Rev. E (3)* **70** 066117. MR2133807 https://doi.org/10.1103/PhysRevE.70.066117

[25] Radin, C. and Yin, M. (2013). Phase transitions in exponential random graphs. *Ann. Appl. Probab.* **23** 2458–2471. MR3127941 https://doi.org/10.1214/12-AAP907

[26] Rinaldo, A., Petrović, S. and Fienberg, S.E. (2013). Maximum likelihood estimation in the β-model. *Ann. Statist.* **41** 1085–1110. MR3113804 https://doi.org/10.1214/12-AOS1078

[27] Robins, G., Pattison, P., Kalish, Y. and Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Soc. Netw.* **29** 173–191.

[28] Sambale, H. and Sinulis, A. (2020). Logarithmic Sobolev inequalities for finite spin systems and applications. *Bernoulli* **26** 1863–1890. MR4091094 https://doi.org/10.3150/19-BEJ1172

[29] Schweinberger, M. (2011). Instability, sensitivity, and degeneracy of discrete exponential families. *J. Amer. Statist. Assoc.* **106** 1361–1370. MR2896841 https://doi.org/10.1198/jasa.2011.tm10747

[30] Schweinberger, M. and Stewart, J. (2020). Concentration and consistency results for canonical and curved exponential-family models of random graphs. *Ann. Statist.* **48** 374–396. MR4065166 https://doi.org/10.1214/19-AOS1810

[31] Shalizi, C.R. and Rinaldo, A. (2013). Consistency under sampling of exponential random graph models. *Ann. Statist.* **41** 508–535. MR3099112 https://doi.org/10.1214/12-AOS1044

[32] Snijders, T.A., Pattison, P.E., Robins, G.L. and Handcock, M.S. (2006). New specifications for exponential random graph models. *Sociol. Method.* **36** 99–153.

[33] Swendsen, R.H. and Wang, J.-S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58** 86.

[34] Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications* **8**. Cambridge: Cambridge University Press.

[35] Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks. I. An introduction to Markov graphs and p. *Psychometrika* **61** 401–425. MR1424909 https://doi.org/10.1007/BF02294547

Bernoulli **29**(1), 2023, 52–74
https://doi.org/10.3150/21-BEJ1449

Template matching with ranks

ERY ARIAS-CASTRO[a] and LIN ZHENG[b]

Department of Mathematics, University of California, San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0112, USA. [a]*eariascastro@ucsd.edu*, [b]*liz176@ucsd.edu*

We consider the problem of matching a template to a noisy signal. Motivated by some recent proposals in the signal processing literature, we suggest a rank-based method and study its asymptotic properties using some well-established techniques in empirical process theory combined with Hájek's projection method. The resulting estimator of the shift is shown to achieve a parametric rate of convergence and to be asymptotically normal. Some numerical simulations corroborate these findings.

Keywords: Matched filter; template matching; scan statistics; Spearman rank correlation; empirical processes; minimax optimality

References

[1] Arias-Castro, E., Castro, R.M., Tánczos, E. and Wang, M. (2018). Distribution-free detection of structured anomalies: Permutation and rank-based scans. *J. Amer. Statist. Assoc.* **113** 789–801. MR3832227 https://doi.org/10.1080/01621459.2017.1286240

[2] Arias-Castro, E. and Zheng, L. (2020). Template matching and change point detection by M-estimation. ArXiv preprint. Available at arXiv:2009.04072.

[3] Ayinde, O. and Yang, Y.-H. (2002). Face recognition approach based on rank correlation of Gabor-filtered images. *Pattern Recognit.* **35** 1275–1289.

[4] Banks, J. and Bennamoun, M. (2001). Reliability analysis of the rank transform for stereo matching. *IEEE Trans. Syst. Man Cybern. B* **31** 870–880.

[5] Banks, J., Bennamoun, M., Kubik, K. and Corke, P. (1999). A constraint to improve the reliability of stereo matching using the rank transform. In *IEEE International Conference on Acoustics, Speech, and Signal Processing* **6** 3321–3324. Phoenix, AZ, USA: IEEE.

[6] Chen, J., Cai, C. and Li, C. (2012). A multi-window stereo matching algorithm in rank tranform domain. In *IEEE International Conference on Signal Processing* **2** 997–1000. Beijing, China: IEEE.

[7] Darkhovskh, B.S. (1976). A nonparametric method for the a posteriori detection of the "disorder" time of a sequence of independent random variables. *Theory Probab. Appl.* **21** 178–183.

[8] Draper, D. (1988). Rank based robust analysis of linear models. I. Exposition and review. *Statist. Sci.* **3** 239–271. MR0968391

[9] Galea, C. and Farrugia, R.A. (2016). Face photo-sketch recognition using local and global texture descriptors. In *European Signal Processing Conference* 2240–2244. Budapest, Hungary: IEEE.

[10] Geng, N. and Gou, Q. (2012). Adaptive color stereo matching based on rank transform. In *International Conference on Industrial Control and Electronics Engineering* 1701–1704. Xi'an, China: IEEE.

[11] Gerstenberger, C. (2018). Robust Wilcoxon-type estimation of change-point location under short-range dependence. *J. Time Series Anal.* **39** 90–104. MR3743095 https://doi.org/10.1111/jtsa.12268

[12] Gibbons, J.D. and Chakraborti, S. (2011). *Nonparametric Statistical Inference*, 5th ed. *Statistics: Textbooks and Monographs*. Boca Raton, FL: CRC Press. MR2681063

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

[14] Giraitis, L., Koul, H.L. and Surgailis, D. (1996). Asymptotic normality of regression estimators with long memory errors. *Statist. Probab. Lett.* **29** 317–335. MR1409327 https://doi.org/10.1016/0167-7152(95)00188-3

[15] Gombay, E. and Hušková, M. (1998). Rank based estimators of the change-point. *J. Statist. Plann. Inference* **67** 137–154. MR1624682 https://doi.org/10.1016/S0378-3758(97)00099-2

[16] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Stat.* **39** 325–346. MR0222988 https://doi.org/10.1214/aoms/1177698394

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

[18] Heiler, S. and Willers, R. (1988). Asymptotic normality of *R*-estimates in the linear model. *Statistics* **19** 173–184. MR0945375 https://doi.org/10.1080/02331888808802084

[19] Hettmansperger, T.P. and McKean, J.W. (2010). *Robust Nonparametric Statistical Methods*. Boca Raton, FL, USA: CRC Press.

[20] Hinkley, D.V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* **57** 1–17. MR0273727 https://doi.org/10.1093/biomet/57.1.1

[21] Hušková, M. (1997). Limit theorems for rank statistics. *Statist. Probab. Lett.* **32** 45–55. MR1439496 https://doi.org/10.1016/S0167-7152(96)00055-7

[22] Jurečková, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Stat.* **42** 1328–1338. MR0295487 https://doi.org/10.1214/aoms/1177693245

[23] Kong, C.H., Soeller, C. and Cannell, M.B. (2008). Increasing sensitivity of Ca^{2+} spark detection in noisy images by application of a matched-filter object detection algorithm. *Biophys. J.* **95** 6016–6024.

[24] Kordelas, G. and Daras, P. (2009). Robust SIFT-based feature matching using Kendall's rank correlation measure. In *IEEE International Conference on Image Processing* 325–328. Cairo, Egypt: IEEE.

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

[26] Koul, H.L. and Mukherjee, K. (1993). Asymptotics of *R*-, M*D*- and L*AD*-estimators in linear regression models with long range dependent errors. *Probab. Theory Related Fields* **95** 535–553. MR1217450 https://doi.org/10.1007/BF01196733

[27] Lehmann, E.L. (2006). *Nonparametrics: Statistical Methods Based on Ranks*. New York: Springer. MR2279708

[28] Lung-Yut-Fong, A., Lévy-Leduc, C. and Cappé, O. (2015). Homogeneity and change-point detection tests for multivariate data using rank statistics. *J. SFdS* **156** 133–162. MR3436651

[29] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

[30] Wang, Y., Wang, Z. and Zi, X. (2020). Rank-based multiple change-point detection. *Comm. Statist. Theory Methods* **49** 3438–3454. MR4107611 https://doi.org/10.1080/03610926.2019.1589515

[31] Xiong, X., Xu, Q., Jin, G., Zhang, H. and Gao, X. (2020). Rank-based local self-similarity descriptor for optical-to-SAR image matching. *IEEE Geosci. Remote Sens. Lett.* **17** 1742–1746.

[32] Yao, Y.-C. and Au, S.T. (1989). Least-squares estimation of a step function. *Sankhyā Ser. A* **51** 370–381. MR1175613

Bernoulli **29**(1), 2023, 75–104
https://doi.org/10.3150/21-BEJ1450

Mixing time guarantees for unadjusted Hamiltonian Monte Carlo

NAWAF BOU-RABEE[1,a] and ANDREAS EBERLE[2,b]

[1] *Department of Mathematical Sciences, Rutgers University Camden, 311 N 5th Street, Camden, NJ 08102 USA.*
[a] *nawaf.bourabee@rutgers.edu*
[2] *Institut für Angewandte Mathematik, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany.*
[b] *eberle@uni-bonn.de*

We provide quantitative upper bounds on the total variation mixing time of the Markov chain corresponding to the unadjusted Hamiltonian Monte Carlo (uHMC) algorithm. For two general classes of models and fixed time discretization step size h, the mixing time is shown to depend only logarithmically on the dimension. Moreover, we provide quantitative upper bounds on the total variation distance between the invariant measure of the uHMC chain and the true target measure. As a consequence, we show that an ε-accurate approximation of the target distribution μ in total variation distance can be achieved by uHMC: (i) for a broad class of models with $O\left(d^{3/4}\varepsilon^{-1/2}\log(d/\varepsilon)\right)$ gradient evaluations; and (ii) for mean field models with weak interactions with $O\left(d^{1/2}\varepsilon^{-1/2}\log(d/\varepsilon)\right)$ gradient evaluations. The proofs are based on the construction of successful couplings for uHMC that realize the upper bounds.

Keywords: MCMC; Hamiltonian Monte Carlo; mixing time; couplings; variational integrators

References

[1] Armstrong, S. and Mourrat, J.-C. (2019). Variational methods for the kinetic Fokker-Planck equation. Available at arXiv:1902.04037v2.

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

[3] Bou-Rabee, N. and Eberle, A. (2023). Supplement to "Mixing time guarantees for unadjusted Hamiltonian Monte Carlo." https://doi.org/10.3150/21-BEJ1450SUPP

[4] Bou-Rabee, N. and Eberle, A. (2021). Two-scale coupling for preconditioned Hamiltonian Monte Carlo in infinite dimensions. *Stoch. Partial Differ. Equ. Anal. Comput.* **9** 207–242. MR4218791 https://doi.org/10.1007/s40072-020-00175-6

[5] Bou-Rabee, N., Eberle, A. and Zimmer, R. (2020). Coupling and convergence for Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **30** 1209–1250. MR4133372 https://doi.org/10.1214/19-AAP1528

[6] Bou-Rabee, N. and Sanz-Serna, J.M. (2018). Geometric integrators and the Hamiltonian Monte Carlo method. *Acta Numer.* **27** 113–206. MR3826507 https://doi.org/10.1017/s0962492917000101

[7] Bou-Rabee, N. and Schuh, K. (2020). Convergence of Unadjusted Hamiltonian Monte Carlo for Mean-Field Models. ArXiv preprint. Available at arXiv:2009.08735.

[8] Chen, Y., Dwivedi, R., Wainwright, M.J. and Yu, B. (2020). Fast mixing of metropolized Hamiltonian Monte Carlo: Benefits of multi-step gradients. *J. Mach. Learn. Res.* **21** Paper No. 92. MR4119160

[9] Chen, Z. and Vempala, S.S. (2019). Optimal convergence rate of Hamiltonian Monte Carlo for strongly logconcave distributions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. LIPIcs. Leibniz Int. Proc. Inform.* **145** Art. No. 64. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4012714

[10] Chewi, S., Lu, C., Ahn, K., Cheng, X., Gouic, T.L. and Rigollet, P. (2020). Optimal dimension dependence of the Metropolis-Adjusted Langevin Algorithm. ArXiv preprint. Available at arXiv:2012.12810.

[11] Dolbeault, J., Mouhot, C. and Schmeiser, C. (2015). Hypocoercivity for linear kinetic equations conserving mass. *Trans. Amer. Math. Soc.* **367** 3807–3828. MR3324910 https://doi.org/10.1090/S0002-9947-2015-06012-7

[12] Duane, S., Kennedy, A.D., Pendleton, B.J. and Roweth, D. (1987). Hybrid Monte Carlo. *Phys. Lett. B* **195** 216–222. MR3960671 https://doi.org/10.1016/0370-2693(87)91197-x

[13] Durmus, A. and Eberle, A. (2021). Asymptotic bias of inexact Markov Chain Monte Carlo Methods in High Dimension. Available at arXiv:2108.00682 [math.PR].

[14] Durmus, A. and Moulines, É. (2019). High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli* **25** 2854–2882. MR4003567 https://doi.org/10.3150/18-BEJ1073

[15] Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041 https://doi.org/10.1007/s00440-015-0673-1

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

[17] Eberle, A. and Majka, M.B. (2019). Quantitative contraction rates for Markov chains on general state spaces. *Electron. J. Probab.* **24** Paper No. 26. MR3933205 https://doi.org/10.1214/19-EJP287

[18] Friedland, S. and Lim, L.-H. (2018). Nuclear norm of higher-order tensors. *Math. Comp.* **87** 1255–1281. MR3766387 https://doi.org/10.1090/mcom/3239

[19] Hairer, E., Lubich, C. and Wanner, G. (2010). *Geometric Numerical Integration*. *Springer Series in Computational Mathematics* **31**. Heidelberg: Springer. MR2840298

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

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

[22] Heng, J. and Jacob, P.E. (2019). Unbiased Hamiltonian Monte Carlo with couplings. *Biometrika* **106** 287–302. MR3949304 https://doi.org/10.1093/biomet/asy074

[23] Kac, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. III* 171–197. Berkeley and Los Angeles, CA: Univ. California Press. MR0084985

[24] Lee, Y.T. and Vempala, S.S. (2017). Geodesic walks in polytopes. In *STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing* 927–940. New York: ACM. MR3678240

[25] Lee, Y.T. and Vempala, S.S. (2018). Convergence rate of Riemannian Hamiltonian Monte Carlo and faster polytope volume computation. In *STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* 1115–1121. New York: ACM. MR3826321 https://doi.org/10.1145/3188745.3188774

[26] Leimkuhler, B. and Reich, S. (2004). *Simulating Hamiltonian Dynamics*. *Cambridge Monographs on Applied and Computational Mathematics* **14**. Cambridge: Cambridge Univ. Press. MR2132573

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

[28] Lindvall, T. (2002). *Lectures on the Coupling Method*. Mineola, NY: Dover. Corrected reprint of the 1992 original. MR1924231

[29] Lindvall, T. and Rogers, L.C.G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872. MR0841588

[30] Lu, J. and Wang, L. (2021). On explicit L^2-convergence rate estimate for piecewise deterministic Markov processes in MCMC algorithms. Available at arXiv:2007.14927.

[31] Madras, N. and Sezer, D. (2010). Quantitative bounds for Markov chain convergence: Wasserstein and total variation distances. *Bernoulli* **16** 882–908. MR2730652 https://doi.org/10.3150/09-BEJ238

[32] Mangoubi, O. and Smith, A. (2021). Mixing of Hamiltonian Monte Carlo on strongly log-concave distributions: Continuous dynamics. *Ann. Appl. Probab.* **31** 2019–2045. MR4332690 https://doi.org/10.1214/20-aap1640

[33] Marsden, J.E. and West, M. (2001). Discrete mechanics and variational integrators. *Acta Numer.* **10** 357–514. MR2009697 https://doi.org/10.1017/S096249290100006X

[34] Mattingly, J.C., Stuart, A.M. and Tretyakov, M.V. (2010). Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.* **48** 552–577. MR2669996 https://doi.org/10.1137/090770527

[35] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2509253 https://doi.org/10.1017/CBO9780511626630

[36] Mischler, S. and Mouhot, C. (2013). Kac's program in kinetic theory. *Invent. Math.* **193** 1–147. MR3069113 https://doi.org/10.1007/s00222-012-0422-3

[37] Monmarché, P. (2021). High-dimensional MCMC with a standard splitting scheme for the underdamped Langevin diffusion. *Electron. J. Stat.* **15** 4117–4166. MR4309974 https://doi.org/10.1214/21-ejs1888

[38] Montenegro, R. and Tetali, P. (2006). Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.* **1** x+121. MR2341319 https://doi.org/10.1561/0400000003

[39] Müller, S. and Ortiz, M. (2004). On the Γ-convergence of discrete dynamics and variational integrators. *J. Nonlinear Sci.* **14** 279–296. MR2061537 https://doi.org/10.1007/s00332-004-0585-1

[40] Neal, R.M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo*. *Chapman & Hall/CRC Handb. Mod. Stat. Methods* 113–162. Boca Raton, FL: CRC Press. MR2858447

[41] Roberts, G.O. and Rosenthal, J.S. (2002). One-shot coupling for certain stochastic recursive sequences. *Stochastic Process. Appl.* **99** 195–208. MR1901153 https://doi.org/10.1016/S0304-4149(02)00096-0

[42] Rump, S.M. (2018). Estimates of the determinant of a perturbed identity matrix. *Linear Algebra Appl.* **558** 101–107. MR3854189 https://doi.org/10.1016/j.laa.2018.08.009

[43] Saloff-Coste, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1996)*. *Lecture Notes in Math.* **1665** 301–413. Berlin: Springer. MR1490046 https://doi.org/10.1007/BFb0092621

[44] Vempala, S. (2005). Geometric random walks: A survey. In *Combinatorial and Computational Geometry*. *Math. Sci. Res. Inst. Publ.* **52** 577–616. Cambridge: Cambridge Univ. Press. MR2178341

[45] Villani, C. (2009). Hypocoercivity. *Mem. Amer. Math. Soc.* **202** iv+141. MR2562709 https://doi.org/10.1090/S0065-9266-09-00567-5

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

Bernoulli **29**(1), 2023, 105–131
https://doi.org/10.3150/21-BEJ1451

High dimensional generalized linear models for temporal dependent data

YUEFENG HAN[1,a], RUEY S. TSAY[2,b] and WEI BIAO WU[3,c]

[1]*Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA.* [a]*yuefeng.han@rutgers.edu*
[2]*Booth School of Business, University of Chicago, Chicago, IL 60637, USA.* [b]*ruey.tsay@chicagobooth.edu*
[3]*Department of Statistics, University of Chicago, Chicago, IL 60637, USA.* [c]*wbwu@galton.uchicago.edu*

High dimensional generalized linear models are widely applicable in many scientific fields with data having heavy tails. However, little is known about statistical guarantees on the estimates of such models in a time series setting. In this article, we establish statistical error bounds and support recovery guarantees of the classical ℓ_1 regularized procedure for generalized linear model with temporal dependent data. We also propose a new robust M-estimator for high dimensional time series. Properties of the proposed robust procedure are investigated both theoretically and numerically. As an extension, we introduce a robust estimator for linear regression and show that the proposed robust estimator achieves nearly the optimal rate as that for i.i.d sub-Gaussian data. Simulation results show that the proposed method performs well numerically in the presence of heavy-tailed and serially dependent covariates and/or errors, and it significantly outperforms the classical Lasso method. For applications, we demonstrate, in the supplementary material, the regularized robust procedure via analyzing high-frequency trading data in finance.

Keywords: High dimensional analysis; time series analysis; generalized linear model; robust estimation; support recovery

References

[1] Avella-Medina, M. and Ronchetti, E. (2018). Robust and consistent variable selection in high-dimensional generalized linear models. *Biometrika* **105** 31–44. MR3768863 https://doi.org/10.1093/biomet/asx070
[2] Basu, S. and Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567. MR3357870 https://doi.org/10.1214/15-AOS1315
[3] Bianco, A.M. and Yohai, V.J. (1996). Robust estimation in the logistic regression model. In *Robust Statistics, Data Analysis, and Computer Intensive Methods (Schloss Thurnau, 1994)*. *Lect. Notes Stat.* **109** 17–34. New York: Springer. MR1491394 https://doi.org/10.1007/978-1-4612-2380-1_2
[4] Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. MR2533469 https://doi.org/10.1214/08-AOS620
[5] Blundell, C., Beck, J. and Heller, K.A. (2012). Modelling reciprocating relationships with Hawkes processes. In *Advances in Neural Information Processing Systems* 2600–2608.
[6] Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334** 495–500. MR1890640 https://doi.org/10.1016/S1631-073X(02)02292-6
[7] Brown, E.N., Kass, R.E. and Mitra, P.P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nat. Neurosci.* **7** 456.
[8] Brownlees, C., Joly, E. and Lugosi, G. (2015). Empirical risk minimization for heavy-tailed losses. *Ann. Statist.* **43** 2507–2536. MR3405602 https://doi.org/10.1214/15-AOS1350
[9] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. *Springer Series in Statistics*. Heidelberg: Springer. MR2807761 https://doi.org/10.1007/978-3-642-20192-9
[10] Cantoni, E. and Ronchetti, E. (2001). Robust inference for generalized linear models. *J. Amer. Statist. Assoc.* **96** 1022–1030. MR1947250 https://doi.org/10.1198/016214501753209004

[11] Catoni, O. (2012). Challenging the empirical mean and empirical variance: A deviation study. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 1148–1185. MR3052407 https://doi.org/10.1214/11-AIHP454
[12] Chen, L. and Wu, W.B. (2017). Concentration inequalities for empirical processes of linear time series. *J. Mach. Learn. Res.* **18** Paper No. 231, 46 pp. MR3845530
[13] 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 https://doi.org/10.1214/13-AOS1182
[14] Chen, X., Xu, M. and Wu, W.B. (2016). Regularized estimation of linear functionals of precision matrices for high-dimensional time series. *IEEE Trans. Signal Process.* **64** 6459–6470. MR3566612 https://doi.org/10.1109/TSP.2016.2605079
[15] Ding, M., Mo, J., Schroeder, C.E. and Wen, X. (2011). Analyzing coherent brain networks with granger causality. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 5916–5918. IEEE.
[16] Fan, J., Gong, W. and Zhu, Z. (2019). Generalized high-dimensional trace regression via nuclear norm regularization. *J. Econometrics* **212** 177–202. MR3994013 https://doi.org/10.1016/j.jeconom.2019.04.026
[17] Fan, J., Li, Q. and Wang, Y. (2017). Estimation of high dimensional mean regression in the absence of symmetry and light tail assumptions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 247–265. MR3597972 https://doi.org/10.1111/rssb.12166
[18] Fan, J., Liu, H., Sun, Q. and Zhang, T. (2018). I-LAMM for sparse learning: Simultaneous control of algorithmic complexity and statistical error. *Ann. Statist.* **46** 814–841. MR3782385 https://doi.org/10.1214/17-AOS1568
[19] Fan, J., Wang, W. and Zhu, Z. (2021). A shrinkage principle for heavy-tailed data: High-dimensional robust low-rank matrix recovery. *Ann. Statist.* **49** 1239–1266. MR4298863 https://doi.org/10.1214/20-aos1980
[20] Fan, Y. and Lv, J. (2013). Asymptotic equivalence of regularization methods in thresholded parameter space. *J. Amer. Statist. Assoc.* **108** 1044–1061. MR3174683 https://doi.org/10.1080/01621459.2013.803972
[21] Genkin, A., Lewis, D.D. and Madigan, D. (2007). Large-scale Bayesian logistic regression for text categorization. *Technometrics* **49** 291–304. MR2408634 https://doi.org/10.1198/004017007000000245
[22] Guo, S., Wang, Y. and Yao, Q. (2016). High-dimensional and banded vector autoregressions. *Biometrika* **103** 889–903. MR3620446 https://doi.org/10.1093/biomet/asw046
[23] Hall, E.C., Raskutti, G. and Willett, R.M. (2019). Learning high-dimensional generalized linear autoregressive models. *IEEE Trans. Inf. Theory* **65** 2401–2422. MR3928208 https://doi.org/10.1109/TIT.2018.2884673
[24] Hampel, F.R. (1971). A general qualitative definition of robustness. *Ann. Math. Stat.* **42** 1887–1896. MR0301858 https://doi.org/10.1214/aoms/1177693054
[25] Hampel, F.R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* **69** 383–393. MR0362657
[26] Han, F., Lu, H. and Liu, H. (2015). A direct estimation of high dimensional stationary vector autoregressions. *J. Mach. Learn. Res.* **16** 3115–3150. MR3450535
[27] Han, Y. and Tsay, R.S. (2020). High-dimensional linear regression for dependent data with applications to nowcasting. *Statist. Sinica* **30** 1797–1827. MR4260745 https://doi.org/10.5705/ss.202018.0044
[28] Han, Y., Tsay, R.S. and Wu, W.B. (2023). Supplement to "High dimensional generalized linear models for temporal dependent data." https://doi.org/10.3150/21-BEJ1451SUPP
[29] Hill, R.W. (1977). *Robust Regression when There Are Outliers in the Carriers*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)—Harvard University. MR2940734
[30] Hodges, J.L. Jr. (1967). Efficiency in normal samples and tolerance of extreme values for some estimates of location. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I* 163–186. Berkeley, CA: Univ. California Press. MR0214251
[31] Hsu, D. and Sabato, S. (2016). Loss minimization and parameter estimation with heavy tails. *J. Mach. Learn. Res.* **17** Paper No. 18, 40 pp. MR3491112
[32] Huang, J., Sun, T., Ying, Z., Yu, Y. and Zhang, C.-H. (2013). Oracle inequalities for the LASSO in the Cox model. *Ann. Statist.* **41** 1142–1165. MR3113806 https://doi.org/10.1214/13-AOS1098
[33] Huang, J. and Zhang, C.-H. (2012). Estimation and selection via absolute penalized convex minimization and its multistage adaptive applications. *J. Mach. Learn. Res.* **13** 1839–1864. MR2956344
[34] Huang, S.-J. and Shih, K.-R. (2003). Short-term load forecasting via arma model identification including non-Gaussian process considerations. *IEEE Trans. Power Syst.* **18** 673–679.

[35] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Stat.* **35** 73–101. MR0161415 https://doi.org/10.1214/aoms/1177703732

[36] Huber, P.J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), Vol. I: Statistics* 221–233. Berkeley, CA: Univ. California Press. MR0216620

[37] Ivanoff, S., Picard, F. and Rivoirard, V. (2016). Adaptive Lasso and group-Lasso for functional Poisson regression. *J. Mach. Learn. Res.* **17** Paper No. 55, 46 pp. MR3504615

[38] Jiang, X., Raskutti, G. and Willett, R. (2015). Minimax optimal rates for Poisson inverse problems with physical constraints. *IEEE Trans. Inf. Theory* **61** 4458–4474. MR3372365 https://doi.org/10.1109/TIT.2015.2441072

[39] Krishnapuram, B., Carin, L., Figueiredo, M.A.T. and Hartemink, A.J. (2005). Sparse multinomial logistic regression: Fast algorithms and generalization bounds. *IEEE Trans. Pattern Anal. Mach. Intell.* **27** 957–968.

[40] Linderman, S. and Adams, R. (2014). Discovering latent network structure in point process data. In *International Conference on Machine Learning* 1413–1421.

[41] Loh, P.-L. (2017). Statistical consistency and asymptotic normality for high-dimensional robust M-estimators. *Ann. Statist.* **45** 866–896. MR3650403 https://doi.org/10.1214/16-AOS1471

[42] Loh, P.-L. and Wainwright, M.J. (2017). Support recovery without incoherence: A case for nonconvex regularization. *Ann. Statist.* **45** 2455–2482. MR3737898 https://doi.org/10.1214/16-AOS1530

[43] Lokhorst, J. (1999). *The Lasso and Generalised Linear Models*. *Honors Project*. Australia: The Univ. Adelaide.

[44] Mallows, C.L. (1975). On some topics in robustness. Unpublished memorandum, Bell Telephone Laboratories, Murray Hill, NJ.

[45] Mark, B., Raskutti, G. and Willett, R. (2019). Network estimation from point process data. *IEEE Trans. Inf. Theory* **65** 2953–2975. MR3951378 https://doi.org/10.1109/TIT.2018.2875766

[46] Massart, P. (2000). About the constants in Talagrand's concentration inequalities for empirical processes. *Ann. Probab.* **28** 863–884. MR1782276 https://doi.org/10.1214/aop/1019160263

[47] McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR3223057 https://doi.org/10.1007/978-1-4899-3242-6

[48] Meier, L., van de Geer, S. and Bühlmann, P. (2008). The group Lasso for logistic regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 53–71. MR2412631 https://doi.org/10.1111/j.1467-9868.2007.00627.x

[49] Merlevède, F., Peligrad, M. and Rio, E. (2011). A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probab. Theory Related Fields* **151** 435–474. MR2851689 https://doi.org/10.1007/s00440-010-0304-9

[50] Merrill, H.M. and Schweppe, F.C. (1971). Bad data suppression in power system static state estimation. *IEEE Trans. Power Appar. Syst.* **6** 2718–2725.

[51] Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** 2308–2335. MR3378468 https://doi.org/10.3150/14-BEJ645

[52] Negahban, S.N., Ravikumar, P., Wainwright, M.J. and Yu, B. (2012). A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers. *Statist. Sci.* **27** 538–557. MR3025133 https://doi.org/10.1214/12-STS400

[53] Ogata, Y. (1999). Seismicity analysis through point-process modeling: A review. In *Seismicity Patterns, Their Statistical Significance and Physical Meaning* 471–507. Springer.

[54] Pillow, J.W., Shlens, J., Paninski, L., Sher, A., Litke, A.M., Chichilnisky, E. and Simoncelli, E.P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454** 995.

[55] Priestley, M.B. (1988). *Nonlinear and Nonstationary Time Series Analysis*. London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0991969

[56] Raginsky, M., Jafarpour, S., Harmany, Z.T., Marcia, R.F., Willett, R.M. and Calderbank, R. (2011). Performance bounds for expander-based compressed sensing in Poisson noise. *IEEE Trans. Signal Process.* **59** 4139–4153. MR2865974 https://doi.org/10.1109/TSP.2011.2157913

[57] Raginsky, M., Willett, R.M., Harmany, Z.T. and Marcia, R.F. (2010). Compressed sensing performance bounds under Poisson noise. *IEEE Trans. Signal Process.* **58** 3990–4002. MR2780163 https://doi.org/10.1109/TSP.2010.2049997

[58] Raginsky, M., Willett, R.M., Horn, C., Silva, J. and Marcia, R.F. (2012). Sequential anomaly detection in the presence of noise and limited feedback. *IEEE Trans. Inf. Theory* **58** 5544–5562. MR2957451 https://doi.org/10.1109/TIT.2012.2201375

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

[60] Rosenblatt, M. (1971). *Markov Processes. Structure and Asymptotic Behavior*. *Die Grundlehren der Mathematischen Wissenschaften* **184**. New York: Springer. MR0329037

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

[62] Roth, V. (2004). The generalized LASSO. *IEEE Trans. Neural Netw.* **15** 16–28. https://doi.org/10.1109/TNN.2003.809398

[63] Shao, X. and Wu, W.B. (2007). Asymptotic spectral theory for nonlinear time series. *Ann. Statist.* **35** 1773–1801. MR2351105 https://doi.org/10.1214/009053606000001479

[64] Shevade, S.K. and Keerthi, S.S. (2003). A simple and efficient algorithm for gene selection using sparse logistic regression. *Bioinformatics* **19** 2246–2253.

[65] Silva, J. and Willett, R. (2008). Hypergraph-based anomaly detection of high-dimensional co-occurrences. *IEEE Trans. Pattern Anal. Mach. Intell.* **31** 563–569.

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

[67] Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* **16** 385–395.

[68] Tong, H. (1990). *Nonlinear Time Series: A Dynamical System Approach*. *Oxford Statistical Science Series* **6**. New York: Oxford Univ. Press, The Clarendon Press. MR1079320

[69] Tsay, R.S. (2005). *Analysis of Financial Time Series*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR2162112 https://doi.org/10.1002/0471746193

[70] Tukey, J.W. (1960). A survey of sampling from contaminated distributions. In *Contributions to Probability and Statistics* 448–485. Stanford, CA: Stanford Univ. Press. MR0120720

[71] Tukey, J.W. (1962). The future of data analysis. *Ann. Math. Stat.* **33** 1–67. MR0133937 https://doi.org/10.1214/aoms/1177704711

[72] van de Geer, S. and Müller, P. (2012). Quasi-likelihood and/or robust estimation in high dimensions. *Statist. Sci.* **27** 469–480. MR3025129 https://doi.org/10.1214/12-STS397

[73] van de Geer, S.A. (2008). High-dimensional generalized linear models and the lasso. *Ann. Statist.* **36** 614–645. MR2396809 https://doi.org/10.1214/009053607000000929

[74] Vere-Jones, D. and Ozaki, T. (1982). Some examples of statistical estimation applied to earthquake data. *Ann. Inst. Statist. Math.* **34** 189–207.

[75] Wiener, N. (1958). *Nonlinear Problems in Random Theory*. *Technology Press Research Monographs*. New York: Wiley. MR0100912

[76] Wu, W.-B. and Wu, Y.N. (2016). Performance bounds for parameter estimates of high-dimensional linear models with correlated errors. *Electron. J. Stat.* **10** 352–379. MR3466186 https://doi.org/10.1214/16-EJS1108

[77] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. MR2172215 https://doi.org/10.1073/pnas.0506715102

[78] Wu, W.B. (2007). M-Estimation of linear models with dependent errors. *Ann. Statist.* **35** 495–521. MR2336857 https://doi.org/10.1214/009053606000001406

[79] Wu, W.B. and Min, W. (2005). On linear processes with dependent innovations. *Stochastic Process. Appl.* **115** 939–958. MR2138809 https://doi.org/10.1016/j.spa.2005.01.001

[80] Wu, W.B. and Shao, X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41** 425–436. MR2052582 https://doi.org/10.1239/jap/1082999076

[81] Zhang, C., Guo, X., Cheng, C. and Zhang, Z. (2014). Robust-BD estimation and inference for varying-dimensional general linear models. *Statist. Sinica* **24** 653–673. MR3235393

[82] Zhang, D. (2021). Robust estimation of the mean and covariance matrix for high dimensional time series. *Statist. Sinica* **31** 797–820. MR4286195 https://doi.org/10.5705/ss.20

[83] Zhang, D. and Wu, W.B. (2017). Gaussian approximation for high dimensional time series. *Ann. Statist.* **45** 1895–1919. MR3718156 https://doi.org/10.1214/16-AOS1512

[84] Zhou, H.H. and Raskutti, G. (2019). Non-parametric sparse additive auto-regressive network models. *IEEE Trans. Inf. Theory* **65** 1473–1492. MR3923181 https://doi.org/10.1109/TIT.2018.2849988

[85] Zhou, K., Zha, H. and Song, L. (2013). Learning social infectivity in sparse low-rank networks using multi-dimensional Hawkes processes. In *Artificial Intelligence and Statistics* 641–649.

Bernoulli **29**(1), 2023, 132–152
https://doi.org/10.3150/21-BEJ1452

Infinite-color randomly reinforced urns with dominant colors

HRISTO SARIEV[1,a], SANDRA FORTINI[2,b] and SONIA PETRONE[2,c]

[1] *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, "Akad. Georgi Bonchev" bl. 8, Sofia 1113, Bulgaria.* [a] *h.sariev@math.bas.bg*
[2] *Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milano, Italy.*
[b] *sandra.fortini@unibocconi.it*, [c] *sonia.petrone@unibocconi.it*

We define and prove limit results for a class of *dominant Pólya sequences*, which are randomly reinforced urn processes with *color-specific* random weights and *unbounded* number of possible colors. Under fairly mild assumptions on the expected reinforcement, we show that the predictive and the empirical distributions converge almost surely (a.s.) in total variation to the same random probability measure $\tilde{P}$; moreover, $\tilde{P}(\mathcal{D}) = 1$ a.s., where $\mathcal{D}$ denotes the set of *dominant* colors for which the expected reinforcement is maximum. In the general case, the predictive probabilities and the empirical frequencies of any δ-neighborhood of $\mathcal{D}$ converge a.s. to one. That is, although *non*-dominant colors continue to be regularly observed, their distance to $\mathcal{D}$ converges in probability to zero. We refine the above results with rates of convergence. We further hint potential applications of dominant Pólya sequences in randomized clinical trials and species sampling, and use our central limit results for Bayesian inference.

Keywords: Reinforced processes; Pólya sequences; random probability measures; species sampling; Bayesian nonparametrics

References

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

[2] Aletti, G., May, C. and Secchi, P. (2009). A central limit theorem, and related results, for a two-color randomly reinforced urn. *Adv. in Appl. Probab.* **41** 829–844. MR2571318 https://doi.org/10.1239/aap/1253281065

[3] Alexander, J.M., Skyrms, B. and Zabell, S.L. (2012). Inventing new signals. *Dyn. Games Appl.* **2** 129–145. MR2891628 https://doi.org/10.1007/s13235-011-0027-2

[4] Bandyopadhyay, A. and Thacker, D. (2017). Pólya urn schemes with infinitely many colors. *Bernoulli* **23** 3243–3267. MR3654806 https://doi.org/10.3150/16-BEJ844

[5] Bassetti, F., Crimaldi, I. and Leisen, F. (2010). Conditionally identically distributed species sampling sequences. *Adv. in Appl. Probab.* **42** 433–459. MR2675111 https://doi.org/10.1239/aap/1275055237

[6] Beggs, A.W. (2005). On the convergence of reinforcement learning. *J. Econom. Theory* **122** 1–36. MR2131871 https://doi.org/10.1016/j.jet.2004.03.008

[7] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P. (2010). Central limit theorems for multicolor urns with dominated colors. *Stochastic Process. Appl.* **120** 1473–1491. MR2653262 https://doi.org/10.1016/j.spa.2010.04.005

[8] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P. (2011). A central limit theorem and its applications to multicolor randomly reinforced urns. *J. Appl. Probab.* **48** 527–546. MR2840314 https://doi.org/10.1239/jap/1308662642

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

[10] Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. *Ann. Math. Stat.* **33** 882–886. MR0149577 https://doi.org/10.1214/aoms/1177704456

[11] Blackwell, D. and MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

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

[13] Chen, L.H.Y. (1978). A short note on the conditional Borel-Cantelli lemma. *Ann. Probab.* **6** 699–700. MR0496420

[14] Crimaldi, I. (2009). An almost sure conditional convergence result and an application to a generalized Pólya urn. *Int. Math. Forum* **4** 1139–1156. MR2524635

[15] Dubins, L.E. and Freedman, D.A. (1965). A sharper form of the Borel-Cantelli lemma and the strong law. *Ann. Math. Stat.* **36** 800–807. MR0182041 https://doi.org/10.1214/aoms/1177700054

[16] Fortini, S. and Petrone, S. (2012). Predictive construction of priors in Bayesian nonparametrics. *Braz. J. Probab. Stat.* **26** 423–449. MR2949087 https://doi.org/10.1214/11-BJPS176

[17] Fortini, S. and Petrone, S. (2020). Quasi-Bayes properties of a procedure for sequential learning in mixture models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 1087–1114. MR4136504

[18] Fortini, S., Petrone, S. and Sariev, H. (2021). Predictive constructions based on measure-valued Pólya urn processes. *Mathematics* **9** 2845.

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

[20] Hansen, B. and Pitman, J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.* **46** 251–256. MR1745692 https://doi.org/10.1016/S0167-7152(99)00109-1

[21] Häusler, E. and Luschgy, H. (2015). *Stable Convergence and Stable Limit Theorems*. *Probability Theory and Stochastic Modelling* **74**. Cham: Springer. MR3362567 https://doi.org/10.1007/978-3-319-18329-9

[22] Hoppe, F.M. (1984). Pólya-like urns and the Ewens' sampling formula. *J. Math. Biol.* **20** 91–94. MR0758915 https://doi.org/10.1007/BF00275863

[23] Janson, S. (2019). Random replacements in Pólya urns with infinitely many colours. *Electron. Commun. Probab.* **24** Paper No. 23, 11. MR3953824 https://doi.org/10.1214/19-ECP226

[24] Kallenberg, O. (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16** 508–534. MR0929061

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

[26] Mailler, C. and Marckert, J.-F. (2017). Measure-valued Pólya urn processes. *Electron. J. Probab.* **22** Paper No. 26, 33. MR3629870 https://doi.org/10.1214/17-EJP47

[27] May, C. and Flournoy, N. (2009). Asymptotics in response-adaptive designs generated by a two-color, randomly reinforced urn. *Ann. Statist.* **37** 1058–1078. MR2502661 https://doi.org/10.1214/08-AOS596

[28] Muliere, P., Paganoni, A.M. and Secchi, P. (2006). A randomly reinforced urn. *J. Statist. Plann. Inference* **136** 1853–1874. MR2255601 https://doi.org/10.1016/j.jspi.2005.08.009

[29] Pemantle, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79. MR2282181 https://doi.org/10.1214/07-PS094

[30] Pemantle, R. and Volkov, S. (1999). Vertex-reinforced random walk on **Z** has finite range. *Ann. Probab.* **27** 1368–1388. MR1733153 https://doi.org/10.1214/aop/1022677452

[31] Pitman, J. (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 245–267. Hayward, CA: IMS. MR1481784 https://doi.org/10.1214/lnms/1215453576

[32] Rosenberger, W.F. and Lachin, J.M. (2016). *Randomization in Clinical Trials: Theory and Practice*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR3443072 https://doi.org/10.1002/9781118742112

[33] Zhang, L.-X., Hu, F., Cheung, S.H. and Chan, W.S. (2014). Asymptotic properties of multicolor randomly reinforced Pólya urns. *Adv. in Appl. Probab.* **46** 585–602. MR3215547 https://doi.org/10.1239/aap/1401369708

Bernoulli **29**(1), 2023, 153–180
https://doi.org/10.3150/21-BEJ1453

On admissible estimation of a mean vector when the scale is unknown

YUZO MARUYAMA[1,a] and WILLIAM E. STRAWDERMAN[2,b]

[1]*Graduate School of Business Administration, Kobe University, Kobe, Japan.* [a]*maruyama@port.kobe-u.ac.jp*
[2]*Department of Statistics and Biostatistics, Rutgers University, New Brunswick, NJ, USA.*
[b]*straw@stat.rutgers.edu*

We consider admissibility of generalized Bayes estimators of the mean of a p-variate normal distribution when the scale is unknown, and the loss is quadratic. The priors considered put the improper invariant prior on the scale while the prior on the mean has a hierarchical normal structure conditional on the scale. This conditional hierarchical prior is indexed by a hyperparameter, a. In earlier studies, the authors established admissibility/inadmissibility of the generalized Bayes estimator under the proper/improper conditional prior ($a > -1$ / $a < -2$), respectively. In this paper we complete the admissibility/inadmissibility characterization for this class of priors by establishing admissibility for the improper conditional prior ($-2 \le a \le -1$). This boundary, $a = -2$, with admissibility for $a \ge -2$ and inadmissibility for $a < -2$ corresponds exactly to that in the known scale case for this class of conditional priors, and which follows from Brown's 1971 paper. As a notable benefit of this enlargement of the class of admissible generalized Bayes estimators, we give admissible and minimax estimators for $p \ge 3$ as opposed to an earlier study which required $p \ge 5$. In one particularly interesting special case, we establish that the joint Stein prior for the unknown scale case leads to a minimax admissible estimator for $p \ge 3$.

Keywords: Admissibility; Bayes estimators; Minimaxity

References

[1] Berger, J.O. (1976). Admissible minimax estimation of a multivariate normal mean with arbitrary quadratic loss. *Ann. Statist.* **4** 223–226. MR0397940

[2] Blyth, C.R. (1951). On minimax statistical decision procedures and their admissibility. *Ann. Math. Stat.* **22** 22–42. MR0039966 https://doi.org/10.1214/aoms/1177729690

[3] Brewster, J.F. and Zidek, J.V. (1974). Improving on equivariant estimators. *Ann. Statist.* **2** 21–38. MR0381098

[4] Brown, L.D. (1971). Admissible estimators, recurrent diffusions, and insoluble boundary value problems. *Ann. Math. Stat.* **42** 855–903. MR0286209 https://doi.org/10.1214/aoms/1177693318

[5] Brown, L.D. and Hwang, J.T. (1982). A unified admissibility proof. In *Statistical Decision Theory and Related Topics, III, Vol. 1 (West Lafayette, Ind., 1981)* 205–230. New York: Academic Press. MR0705290

[6] Faith, R.E. (1978). Minimax Bayes estimators of a multivariate normal mean. *J. Multivariate Anal.* **8** 372–379. MR0512607 https://doi.org/10.1016/0047-259X(78)90060-X

[7] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 361–379. Berkeley, CA: Univ. California Press. MR0133191

[8] Johnstone, I.M. (2019). Larry Brown's work on admissibility. *Statist. Sci.* **34** 657–668. MR4048596 https://doi.org/10.1214/19-STS744

[9] Kubokawa, T. (1991). An approach to improving the James-Stein estimator. *J. Multivariate Anal.* **36** 121–126. MR1094273 https://doi.org/10.1016/0047-259X(91)90096-K

[10] Lin, P.E. and Tsai, H.L. (1973). Generalized Bayes minimax estimators of the multivariate normal mean with unknown covariance matrix. *Ann. Statist.* **1** 142–145. MR0331573

[11] Maruyama, Y. and Strawderman, W.E. (2005). A new class of generalized Bayes minimax ridge regression estimators. *Ann. Statist.* **33** 1753–1770. MR2166561 https://doi.org/10.1214/009053605000000327

[12] Maruyama, Y. and Strawderman, W.E. (2017). A sharp boundary for SURE-based admissibility for the normal means problem under unknown scale. *J. Multivariate Anal.* **162** 134–151. MR3719340 https://doi.org/10.1016/j.jmva.2017.09.002

[13] Maruyama, Y. and Strawderman, W.E. (2020). Admissible Bayes equivariant estimation of location vectors for spherically symmetric distributions with unknown scale. *Ann. Statist.* **48** 1052–1071. MR4102687 https://doi.org/10.1214/19-AOS1837

[14] Maruyama, Y. and Strawderman, W.E. (2021). Admissible estimators of a multivariate normal mean vector when the scale is unknown. *Biometrika* **108** 997–1003. MR4341365 https://doi.org/10.1093/biomet/asaa102

[15] Maruyama, Y., Strawderman, W.E. (2023). Supplement to "On admissible estimation of a mean vector when the scale is unknown." https://doi.org/10.3150/21-BEJ1453SUPP

[16] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. I* 197–206. Berkeley and Los Angeles, CA: Univ. California Press. MR0084922

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

[18] Strawderman, W.E. (1971). Proper Bayes minimax estimators of the multivariate normal mean. *Ann. Math. Stat.* **42** 385–388. MR0397939 https://doi.org/10.1214/aoms/1177693528

Bernoulli **29**(1), 2023, 181–204
https://doi.org/10.3150/21-BEJ1454

On a projection-based class of uniformity tests on the hypersphere

EDUARDO GARCÍA-PORTUGUÉS[1,a], PAULA NAVARRO-ESTEBAN[2,b] and JUAN A. CUESTA-ALBERTOS[2,c]

[1]*Departamento de Estadística, Universidad Carlos III de Madrid, Leganés, Spain.* [a]*edgarcia@est-econ.uc3m.es*
[2]*Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Santander, Spain.* [b]*paula.navarro@unican.es,* [c]*juan.cuesta@unican.es*

We introduce a projection-based class of uniformity tests on the hypersphere. The class employs statistics that integrate, along all possible directions, a weighted quadratic discrepancy between the empirical cumulative distribution of the projections and the projected uniform distribution. Simple expressions for several test statistics are obtained for the circle and the sphere, as well as relatively tractable forms for higher dimensions. Despite their different origins, variants of the proposed class are shown to contain and be contained in variants of the Sobolev class of uniformity tests. Our new class proves itself advantageous by allowing the derivation of new tests that neatly extend the circular-only tests by Watson, Ajne, and Rothman, and by introducing the first instance of an Anderson–Darling-like test for directional data. We obtain usable asymptotic distributions and the local asymptotic optimality against certain alternatives of the new tests. A simulation study evaluates the theoretical findings and provides evidence that the new testing proposals are competitive. An application to the study of the crater distribution on Rhea illustrates the usage of the new tests.

Keywords: Circular data; directional data; projection; uniformity

References

[1] Ajne, B. (1968). A simple test for uniformity of a circular distribution. *Biometrika* **55** 343–354. MR0235662 https://doi.org/10.1093/biomet/55.2.343

[2] Alvarellos, J.L., Zahnle, K.J., Dobrovolskis, A.R. and Hamill, P. (2005). Fates of satellite ejecta in the Saturn system. *Icarus* **178** 104–123. https://doi.org/10.1016/j.icarus.2005.04.017

[3] Anderson, T.W. and Darling, D.A. (1954). A test of goodness of fit. *J. Amer. Statist. Assoc.* **49** 765–769. MR0069459

[4] Bakshaev, A. (2010). N-distance tests of uniformity on the hypersphere. *Nonlinear Anal. Model. Control* **15** 15–28. MR2677458

[5] Barlow, N.G. (2015). Constraining geologic properties and processes through the use of impact craters. *Geomorphology* **240** 18–33. https://doi.org/10.1016/j.geomorph.2014.08.027

[6] Beran, R.J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Probab.* **5** 177–195. MR0228098 https://doi.org/10.1017/s002190020003237x

[7] Bingham, C. (1974). An antipodally symmetric distribution on the sphere. *Ann. Statist.* **2** 1201–1225. MR0397988

[8] Cai, T., Fan, J. and Jiang, T. (2013). Distributions of angles in random packing on spheres. *J. Mach. Learn. Res.* **14** 1837–1864. MR3104497

[9] Cuesta-Albertos, J.A., Cuevas, A. and Fraiman, R. (2009). On projection-based tests for directional and compositional data. *Stat. Comput.* **19** 367–380. MR2565311 https://doi.org/10.1007/s11222-008-9098-3

[10] Cuesta-Albertos, J.A., Fraiman, R. and Ransford, T. (2007). A sharp form of the Cramér-Wold theorem. *J. Theoret. Probab.* **20** 201–209. MR2324526 https://doi.org/10.1007/s10959-007-0060-7

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

[12] D'Agostino, R.B. and Stephens, M.A., eds. (1986). *Goodness-of-Fit Techniques*. *Statistics: Textbooks and Monographs* **68**. Marcel Dekker, New York.

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

[14] Fassett, C.I. (2016). Analysis of impact crater populations and the geochronology of planetary surfaces in the inner solar system. *J. Geophys. Res. Planets* **121** 1900–1926. https://doi.org/10.1002/2016JE005094

[15] Fay, M.P. (2010). Two-sided exact tests and matching confidence intervals for discrete data. *R J.* **2** 53–58. https://doi.org/10.32614/RJ-2010-008

[16] Figueiredo, A. (2007). Comparison of tests of uniformity defined on the hypersphere. *Statist. Probab. Lett.* **77** 329–334. MR2339036 https://doi.org/10.1016/j.spl.2006.07.012

[17] Figueiredo, A. and Gomes, P. (2003). Power of tests of uniformity defined on the hypersphere. *Comm. Statist. Simulation Comput.* **32** 87–94. MR1963295 https://doi.org/10.1081/SAC-120013113

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

[19] García-Portugués, E., Navarro-Esteban, P. and Cuesta-Albertos, J.A. (2021). A Cramér–von Mises test of uniformity on the hypersphere. In *Statistical Learning and Modeling in Data Analysis* (S. Balzano, G.C. Porzio, R. Salvatore, D. Vistocco and M. Vichi, eds.). *Studies in Classification, Data Analysis and Knowledge Organization* 107–116. Cham: Springer. https://doi.org/10.1007/978-3-030-69944-4_{1}2

[20] García-Portugués, E., Navarro-Esteban, P. Cuesta-Albertos, J.A. (2023). Supplement to "On a projection-based class of uniformity tests on the hypersphere." https://doi.org/10.3150/21-BEJ1454SUPP

[21] García-Portugués, E., Paindaveine, D. and Verdebout, T. (2020). On optimal tests for rotational symmetry against new classes of hyperspherical distributions. *J. Amer. Statist. Assoc.* **115** 1873–1887. MR4189764 https://doi.org/10.1080/01621459.2019.1665527

[22] García-Portugués, E., Paindaveine, D. and Verdebout, T. (2021). rotasym: Tests for Rotational Symmetry on the Hypersphere. R package version 1.1.3.

[23] García-Portugués, E. and Verdebout, T. (2018). An overview of uniformity tests on the hypersphere. Available at arXiv:1804.00286.

[24] García-Portugués, E. and Verdebout, T. (2021). sphunif: Uniformity Tests on the Circle, Sphere, and Hypersphere. R package version 1.0.1.

[25] Giné M., E. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. *Ann. Statist.* **3** 1243–1266. MR0388663

[26] Hirata, N. (2016). Differential impact cratering of Saturn's satellites by heliocentric impactors. *J. Geophys. Res. Planets* **121** 111–117. https://doi.org/10.1002/2015JE004940

[27] Jammalamadaka, S.R., Meintanis, S. and Verdebout, T. (2020). On Sobolev tests of uniformity on the circle with an extension to the sphere. *Bernoulli* **26** 2226–2252. MR4091107 https://doi.org/10.3150/19-BEJ1191

[28] Jupp, P.E. (2008). Data-driven Sobolev tests of uniformity on compact Riemannian manifolds. *Ann. Statist.* **36** 1246–1260. MR2418656 https://doi.org/10.1214/009053607000000541

[29] Jupp, P.E. and Kume, A. (2020). Measures of goodness of fit obtained by almost-canonical transformations on Riemannian manifolds. *J. Multivariate Anal.* **176** 104579. MR4045270 https://doi.org/10.1016/j.jmva.2019.104579

[30] Ley, C. and Verdebout, T. (2017). *Modern Directional Statistics*. *Chapman & Hall/CRC Interdisciplinary Statistics Series*. Boca Raton, FL: CRC Press. MR3752655

[31] Ley, C. and Verdebout, T., eds. (2017). *Modern Directional Statistics*. *Chapman & Hall/CRC Interdisciplinary Statistics Series*. Boca Raton, FL: CRC Press. MR3752655

[32] Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR1828667

[33] Pewsey, A. and García-Portugués, E. (2021). Recent advances in directional statistics. *TEST* **30** 1–58. MR4242171 https://doi.org/10.1007/s11749-021-00759-x

[34] Prentice, M.J. (1978). On invariant tests of uniformity for directions and orientations. *Ann. Statist.* **6** 169–176. MR0458721

[35] Rayleigh, Lord (1919). On the problem of random vibrations, and of random flights in one, two, or three dimensions. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **37** 321–347. https://doi.org/10.1080/14786440408635894

[36] Rothman, E.D. (1972). Tests for uniformity of a circular distribution. *Sankhyā Ser. A* **34** 23–32. MR0339396

[37] Squyres, S.W., Howell, C., Liu, M.C. and Lissauer, J.J. (1997). Investigation of crater "saturation" using spatial statistics. *Icarus* **125** 67–82. https://doi.org/10.1006/icar.1996.5560

[38] Stephens, M.A. (1969). A goodness-of-fit statistic for the circle, with some comparisons. *Biometrika* **56** 161–168. https://doi.org/10.1093/biomet/56.1.161

[39] Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. *J. Amer. Statist. Assoc.* **69** 730–737. https://doi.org/10.2307/2286009

[40] Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48** 109–114. MR0131930 https://doi.org/10.1093/biomet/48.1-2.109

[41] Zahnle, K., Schenk, P., Levison, H. and Dones, L. (2003). Cratering rates in the outer Solar System. *Icarus* **163** 263–289. https://doi.org/10.1016/S0019-1035(03)00048-4

Bernoulli **29**(1), 2023, 205–228
https://doi.org/10.3150/21-BEJ1455

Donsker results for the empirical process indexed by functions of locally bounded variation and applications to the smoothed empirical process

ERIC BEUTNER[1,a] and HENRYK ZÄHLE[2,b]

[1]*Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands.* [a]*e.a.beutner@vu.nl*
[2]*Department of Mathematics, Saarland University, Germany.* [b]*zaehle@math.uni-sb.de*

Recently, Radulovic and Wegkamp introduced a new technique to show convergence in distribution of the empirical process indexed by functions of bounded variation. This method of proof allows to directly extend convergence results known for the canonical empirical process to convergence in distribution of the empirical process indexed by functions of bounded variation. The purpose of this article is twofold. First, we extend the mentioned technique to index functions of *locally* bounded variation. Second, and more importantly, we demonstrate that this technique provides a new approach to show convergence in distribution of the smoothed empirical process based on kernel density estimators. Using this approach we can prove to the best of our knowledge the first results on convergence in distribution of the smoothed empirical process of *dependent data*. Our results cover both weak and strong dependence as well as index sets of functions of locally bounded variation. Moreover our results cover an MISE optimal choice of the bandwidth for the kernel density estimator which to some extent is the plug-in property in the Bickel–Ritov sense. In the case of i.i.d. data our results extend a seminal result of Giné and Nickl.

Keywords: Weak convergence; empirical process; function of locally bounded variation; smoothed empirical process; kernel smoothing; optimal bandwidth; plug-in property

References

[1] Bartlett, M.S. (1963). Statistical estimation of density functions. *Sankhyā Ser. A* **25** 245–254. MR0170445
[2] Berghaus, B., Bücher, A. and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli* **23** 743–772. MR3556791 https://doi.org/10.3150/15-BEJ751
[3] Beutner, E., Wu, W.B. and Zähle, H. (2012). Asymptotics for statistical functionals of long-memory sequences. *Stochastic Process. Appl.* **122** 910–929. MR2891441 https://doi.org/10.1016/j.spa.2011.10.006
[4] Beutner, E. and Zähle, H. (2012). Deriving the asymptotic distribution of U- and V-statistics of dependent data using weighted empirical processes. *Bernoulli* **18** 803–822. MR2948902 https://doi.org/10.3150/11-BEJ358
[5] Beutner, E. and Zähle, H. (2014). Continuous mapping approach to the asymptotics of U- and V-statistics. *Bernoulli* **20** 846–877. MR3178520 https://doi.org/10.3150/13-BEJ508
[6] Beutner, E. and Zähle, H. (2016). Functional delta-method for the bootstrap of quasi-Hadamard differentiable functionals. *Electron. J. Stat.* **10** 1181–1222. MR3499525 https://doi.org/10.1214/16-EJS1140
[7] Beutner, E., Zähle, H. (2023). Supplement to "Donsker results for the empirical process indexed by functions of locally bounded variation and applications to the smoothed empirical process." https://doi.org/10.3150/21-BEJ1455SUPP
[8] Bickel, P.J. and Ritov, Y. (2003). Nonparametric estimators which can be "plugged-in". *Ann. Statist.* **31** 1033–1053. MR2001641 https://doi.org/10.1214/aos/1059655904

[9] Billingsley, P. (1999). *Convergence of Probability Measures*. New York: Wiley.
[10] Bretagnolle, J. and Huber, C. (1979). Estimation des densités: Risque minimax. *Z. Wahrsch. Verw. Gebiete* **47** 119–137. MR0523165 https://doi.org/10.1007/BF00535278
[11] Chigansky, P. (2012). Solutions to the exercises from 'Introduction to nonparametric estimation' by A. B. Tsybakov. Online material to a Lecture Course at Hebrew University.
[12] Dede, S. (2009). An empirical central limit theorem in $\mathbf{L}^1$ for stationary sequences. *Stochastic Process. Appl.* **119** 3494–3515. MR2568284 https://doi.org/10.1016/j.spa.2009.06.006
[13] Dehling, H. and Taqqu, M.S. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17** 1767–1783. MR1026312 https://doi.org/10.1214/aos/1176347394
[14] Dehling, H. and Taqqu, M.S. (1991). Bivariate symmetric statistics of long-range dependent observations. *J. Statist. Plann. Inference* **28** 153–165. MR1115815 https://doi.org/10.1016/0378-3758(91)90031-9
[15] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999 https://doi.org/10.1214/aop/1022677394
[16] Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52. MR0550122 https://doi.org/10.1007/BF00535673
[17] Dovgoshey, O., Martio, O., Ryazanov, V. and Vuorinen, M. (2006). The Cantor function. *Expo. Math.* **24** 1–37. MR2195181 https://doi.org/10.1016/j.exmath.2005.05.002
[18] Dudley, R.M. (1992). Fréchet differentiability, p-variation and uniform Donsker classes. *Ann. Probab.* **20** 1968–1982. MR1188050
[19] Dudley, R.M. (1999). *Uniform Central Limit Theorems*. *Cambridge Studies in Advanced Mathematics* **63**. Cambridge: Cambridge Univ. Press. MR1720712 https://doi.org/10.1017/CBO9780511665622
[20] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358 https://doi.org/10.1017/CBO9780511755347
[21] Estévez, G. and Vieu, P. (2003). Nonparametric estimation under long memory dependence. *J. Nonparametr. Stat.* **15** 535–551. MR2017486 https://doi.org/10.1080/10485250310001604668
[22] Gill, R.D., van der Laan, M.J. and Wellner, J.A. (1995). Inefficient estimators of the bivariate survival function for three models. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 545–597. MR1338452
[23] Giné, E. and Nickl, R. (2008). Uniform central limit theorems for kernel density estimators. *Probab. Theory Related Fields* **141** 333–387. MR2391158 https://doi.org/10.1007/s00440-007-0087-9
[24] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics, [40]*. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862
[25] Hairer, M. and Labbé, C. (2015). A simple construction of the continuum parabolic Anderson model on $\mathbf{R}^2$. *Electron. Commun. Probab.* **20** no. 43. MR3358965 https://doi.org/10.1214/ECP.v20-4038
[26] Hall, P. and Hart, J.D. (1990). Convergence rates in density estimation for data from infinite-order moving average processes. *Probab. Theory Related Fields* **87** 253–274. MR1080492 https://doi.org/10.1007/BF01198432
[27] Hallin, M. and Tran, L.T. (1996). Kernel density estimation for linear processes: Asymptotic normality and optimal bandwidth derivation. *Ann. Inst. Statist. Math.* **48** 429–449. MR1424774 https://doi.org/10.1007/BF00050847
[28] Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24** 992–1024. MR1401834 https://doi.org/10.1214/aos/1032526953
[29] Hoffmann-Jørgensen, J. (1984). *Stochastic Processes in Polish Spaces*. Unpublished manuscript.
[30] Johannes, J. and Rao, S.S. (2011). Nonparametric estimation for dependent data. *J. Nonparametr. Stat.* **23** 661–681. MR2836283 https://doi.org/10.1080/10485252.2010.484491
[31] Kerkyacharian, G., Lepski, O. and Picard, D. (2001). Nonlinear estimation in anisotropic multi-index denoising. *Probab. Theory Related Fields* **121** 137–170. MR1863916 https://doi.org/10.1007/PL00008800
[32] Krätschmer, V., Schied, A. and Zähle, H. (2015). Quasi-Hadamard differentiability of general risk functionals and its application. *Stat. Risk Model.* **32** 25–47. MR3331780 https://doi.org/10.1515/strm-2014-1174
[33] Major, P. (1981). Limit theorems for nonlinear functionals of Gaussian sequences. *Z. Wahrsch. Verw. Gebiete* **57** 129–158. MR0623458 https://doi.org/10.1007/BF00533717

[34] Mayer, U., Zähle, H. and Zhou, Z. (2020). Functional weak limit theorem for a local empirical process of non-stationary time series and its application. *Bernoulli* **26** 1891–1911. MR4091095 https://doi.org/10.3150/19-BEJ1174

[35] Mielniczuk, J. (1997). On the asymptotic mean integrated squared error of a kernel density estimator for dependent data. *Statist. Probab. Lett.* **34** 53–58. MR1457496 https://doi.org/10.1016/S0167-7152(96)00165-4

[36] Mikulevicius, R. and Pragarauskas, H. (2003). On Cauchy-Dirichlet problem in half-space for parabolic SPDEs in weighted Hölder spaces. *Stochastic Process. Appl.* **106** 185–222. MR1989627 https://doi.org/10.1016/S0304-4149(03)00042-5

[37] Mikulevicius, R. and Pragarauskas, H. (2005). On Cauchy-Dirichlet problem in half-space for linear integro-differential equations in weighted Hölder spaces. *Electron. J. Probab.* **10** 1398–1416. MR2191633 https://doi.org/10.1214/EJP.v10-292

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

[39] Nadaraja, È.A. (1965). On non-parametric estimates of density functions and regression. *Theory Probab. Appl.* **10** 186–190.

[40] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Stat.* **33** 1065–1076. MR0143282 https://doi.org/10.1214/aoms/1177704472

[41] Pham, T.D. and Tran, L.T. (1985). Some mixing properties of time series models. *Stochastic Process. Appl.* **19** 297–303. MR0787587 https://doi.org/10.1016/0304-4149(85)90031-6

[42] Pollard, D. (1984). *Convergence of Stochastic Processes*. *Springer Series in Statistics*. New York: Springer. MR0762984 https://doi.org/10.1007/978-1-4612-5254-2

[43] Radulović, D. and Wegkamp, M. (2000). Weak convergence of smoothed empirical processes: Beyond Donsker classes. In *High Dimensional Probability, II (Seattle, WA, 1999)* (E. Giné, D.M. Mason and J.A. Wellner, eds.). *Progress in Probability* **47** 89–105. Boston, MA: Birkhäuser. MR1857317

[44] Radulović, D. and Wegkamp, M. (2009). Uniform central limit theorems for pregaussian classes of functions. In *High Dimensional Probability V: The Luminy Volume*. *Inst. Math. Stat. (IMS) Collect.* **5** 84–102. Beachwood, OH: IMS. MR2797942 https://doi.org/10.1214/09-IMSCOLL507

[45] Radulović, D. and Wegkamp, M. (2018). Weak convergence of stationary empirical processes. *J. Statist. Plann. Inference* **194** 75–84. MR3725790 https://doi.org/10.1016/j.jspi.2017.09.006

[46] Radulović, D., Wegkamp, M. and Zhao, Y. (2017). Weak convergence of empirical copula processes indexed by functions. *Bernoulli* **23** 3346–3384. MR3654809 https://doi.org/10.3150/16-BEJ849

[47] Rio, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. *Mathématiques & Applications (Berlin) [Mathematics & Applications]* **31**. Berlin: Springer. MR2117923

[48] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* **27** 832–837. MR0079873 https://doi.org/10.1214/aoms/1177728190

[49] Rost, D. (2000). Limit theorems for smoothed empirical processes. In *High Dimensional Probability, II (Seattle, WA, 1999)* (E. Giné, D.M. Mason and J.A. Wellner, eds.). *Progress in Probability* **47** 107–113. Boston, MA: Birkhäuser. MR1857318

[50] Schilling, R.L. (2005). *Measures, Integrals and Martingales*. New York: Cambridge Univ. Press. MR2200059 https://doi.org/10.1017/CBO9780511810886

[51] Scott, D.W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1191168 https://doi.org/10.1002/9780470316849

[52] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838963

[53] Stupfler, G. (2016). On the weak convergence of the kernel density estimator in the uniform topology. *Electron. Commun. Probab.* **21** Paper No. 17. MR3485386 https://doi.org/10.1214/16-ECP4638

[54] Taqqu, M.S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302. MR0400329 https://doi.org/10.1007/BF00532868

[55] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50** 53–83. MR0550123 https://doi.org/10.1007/BF00535674

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

[57] van der Vaart, A. (1994). Weak convergence of smoothed empirical processes. *Scand. J. Stat.* **21** 501–504. MR1310093

[58] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

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

[60] Watson, G.S. and Leadbetter, M.R. (1963). On the estimation of the probability density. I. *Ann. Math. Stat.* **34** 480–491. MR0148149 https://doi.org/10.1214/aoms/1177704159

[61] Woodroofe, M. (1970). On choosing a delta-sequence. *Ann. Math. Stat.* **41** 1665–1671. MR0270515 https://doi.org/10.1214/aoms/1177696810

[62] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. MR2172215 https://doi.org/10.1073/pnas.0506715102

[63] Wu, W.B. (2008). Empirical processes of stationary sequences. *Statist. Sinica* **18** 313–333. MR2384990

[64] Yukich, J.E. (1992). Weak convergence of smoothed empirical processes. *Scand. J. Stat.* **19** 271–279. MR1183201

[65] Zähle, H. (2015). Qualitative robustness of statistical functionals under strong mixing. *Bernoulli* **21** 1412–1434. MR3352049 https://doi.org/10.3150/14-BEJ608

[66] Zolotarev, V.M. (1983). Probability metrics. *Theory Probab. Appl.* **28** 278–302.

Bernoulli **29**(1), 2023, 229–273
https://doi.org/10.3150/21-BEJ1456

Rank-based testing for semiparametric VAR models: A measure transportation approach

MARC HALLIN[1,a], DAVIDE LA VECCHIA[2,b] and HANG LIU[3,c]

[1]*ECARES and Département de Mathématique, Université libre de Bruxelles CP 114/4, Avenue F.D. Roosevelt 50, B-1050 Bruxelles, Belgium.* [a]*mhallin@ulb.ac.be*
[2]*Research Center for Statistics, University of Geneva, Bd du Pont d'Arve 40 - CH-1211 Geneva, Switzerland.* [b]*davide.lavecchia@unige.ch*
[3]*International Institute of Finance, School of Management, University of Science and Technology of China, Hefei, China.* [c]*hliu01@ustc.edu.cn*

We develop a class of tests for semiparametric vector autoregressive (VAR) models with unspecified innovation densities based on the recent measure-transportation-based concepts of multivariate *center-outward ranks* and *signs*. We show that these concepts, combined with Le Cam's asymptotic theory of statistical experiments, yield novel testing procedures, which (a) are valid under a broad class of innovation densities (possibly non-elliptical, skewed, and/or with infinite moments), (b) are optimal (locally asymptotically maximin or most stringent) at selected ones, and (c) are robust against additive outliers. In order to show this, we establish, for a general class of center-outward rank-based serial statistics, a Hájek asymptotic representation result, of independent interest, which allows for a rank-based reconstruction of central sequences. As an illustration, we consider the problems of testing the absence of serial correlation in multiple-output and possibly non-linear regression (an extension of the classical Durbin-Watson problem) and the sequential identification of the order p of a VAR(p) model. A Monte Carlo comparative study of our tests and their routinely-applied Gaussian competitors demonstrates the benefits (in terms of size, power, and robustness) of our methodology; these benefits are particularly significant in the presence of asymmetric and leptokurtic innovation densities. A real-data application concludes the paper.

Keywords: Multivariate ranks; distribution-freeness; Hájek representation; local asymptotic normality; Durbin-Watson test; VAR order selection

References

[1] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory (Tsahkadsor, 1971)* (B.N. Petrov and F. Csaki, eds.) 267–281. MR0483125

[2] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **65** 367–389. MR1983753 https://doi.org/10.1111/1467-9868.00391

[3] Benghabrit, Y. and Hallin, M. (1992). Optimal rank-based tests against first-order superdiagonal bilinear dependence. *J. Statist. Plann. Inference* **32** 45–61. MR1177358 https://doi.org/10.1016/0378-3758(92)90151-H

[4] Benghabrit, Y. and Hallin, M. (1996). Rank-based tests for autoregressive against bilinear serial dependence. *J. Nonparametr. Stat.* **6** 253–272. MR1383054 https://doi.org/10.1080/10485259608832674

[5] Bertsekas, D.P. (1991). *Linear Network Optimization: Algorithms and Codes*. Cambridge, MA: MIT Press. MR1201582

[6] Boldin, M.V. (2012). Robustness of sign tests in autoregression. *Moscow Math. Bull.* **67** 112–115.

[7] Boldin, M.V. (2013). Robustness of sign tests for testing hypotheses about order of autoregression. *Theory Probab. Appl.* **57** 663–670. MR3201671 https://doi.org/10.1137/S0040585X97986242

[8] Burkard, R., Dell'Amico, M. and Martello, S. (2009). *Assignment Problems*. Philadelphia, PA: SIAM. MR2488749 https://doi.org/10.1137/1.9780898717754

[9] Cassart, D., Hallin, M. and Paindaveine, D. (2010). On the estimation of cross-information quantities in rank-based inference. In *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková* (J. Antoch, M. Hušková and P.K. Sen, eds.). *Inst. Math. Stat. (IMS) Collect.* **7** 35–45. Beachwood, OH: IMS. MR2808364

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

[11] Croux, C. and Joossens, K. (2008). Robust estimation of the vector autoregressive model by a least trimmed squares procedure. *COMPSTAT* **2008** 489–501.

[12] Deb, N., Bhattacharya, B. and Sen, B. (2019). Efficiency lower bounds for distribution-free Hotelling-type two-sample tests based on optimal transport. Available at arXiv:2104.01986.

[13] Deb, N. and Sen, B. (2022). Multivariate rank-based distribution-free nonparametric testing using measure transportation. *J. Amer. Statist. Assoc.* To appear.

[14] del Barrio, E., González-Sanz, A. and Hallin, M. (2020). A note on the regularity of optimal-transport-based center-outward distribution and quantile functions. *J. Multivariate Anal.* **180** 104671. MR4147635 https://doi.org/10.1016/j.jmva.2020.104671

[15] Dick, J. and Pillichshammer, F. (2014). Discrepancy theory and quasi-Monte Carlo integration. In *A Panorama of Discrepancy Theory* (W. Chen, A. Sirvastava and G. Travaglini, eds.). *Lecture Notes in Math.* **2107** 539–619. Cham: Springer. MR3330331 https://doi.org/10.1007/978-3-319-04696-9_9

[16] Durbin, J. and Watson, G.S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika* **37** 409–428. MR0039210 https://doi.org/10.1093/biomet/37.3-4.409

[17] Durbin, J. and Watson, G.S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika* **38** 159–178. MR0042662 https://doi.org/10.1093/biomet/38.1-2.159

[18] Figalli, A. (2018). On the continuity of center-outward distribution and quantile functions. *Nonlinear Anal.* **177** 413–421. MR3886582 https://doi.org/10.1016/j.na.2018.05.008

[19] Garel, B. and Hallin, M. (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend. *Ann. Inst. Statist. Math.* **47** 551–579. MR1364260

[20] Garel, B. and Hallin, M. (1999). Rank-based autoregressive order identification. *J. Amer. Statist. Assoc.* **94** 1357–1371. MR1731496 https://doi.org/10.2307/2669948

[21] Ghosal, P. and Sen, B. (2022). Multivariate ranks and quantiles using optimal transportation and applications to goodness-of-fit testing. *Ann. Statist.* To appear.

[22] Hájek, J. (1961). Some extensions of the Wald-Wolfowitz-Noether theorem. *Ann. Math. Stat.* **32** 506–523. MR0130707 https://doi.org/10.1214/aoms/1177705057

[23] Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. New York: Academic Press. MR0229351

[24] Hallin, M. (1986). Nonstationary q-dependent processes and time-varying moving-average models: Invertibility properties and the forecasting problem. *Adv. in Appl. Probab.* **18** 170–210. MR0827335 https://doi.org/10.2307/1427242

[25] Hallin, M. (1994). On the Pitman non-admissibility of correlogram-based methods. *J. Time Series Anal.* **15** 607–611. MR1312324 https://doi.org/10.1111/j.1467-9892.1994.tb00215.x

[26] Hallin, M. (2017). On distribution and quantile functions, ranks and signs in $\mathbb{R}^d$. ECARES. Working paper.

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

[28] Hallin, M., Hlubinka, D. and Hudecová, Š. (2022). Fully distribution-free center-outward rank tests for multiple-output regression and MANOVA. *J. Amer. Statist. Assoc.* To appear.

[29] Hallin, M., Ingenbleek, J.-F. and Puri, M.L. (1985). Linear serial rank tests for randomness against ARMA alternatives. *Ann. Statist.* **13** 1156–1181. MR0803764 https://doi.org/10.1214/aos/1176349662

[30] Hallin, M. and La Vecchia, D. (2017). R-estimation in semiparametric dynamic location-scale models. *J. Econometrics* **196** 233–247. MR3589479 https://doi.org/10.1016/j.jeconom.2016.08.002

[31] Hallin, M. and La Vecchia, D. (2020). A simple R-estimation method for semiparametric duration models. *J. Econometrics* **218** 736–749. MR4149246 https://doi.org/10.1016/j.jeconom.2020.04.036

[32] Hallin, M., La Vecchia, D. and Liu, H. (2022). Center-outward R-estimation for semiparametric VARMA models. *J. Amer. Statist. Assoc.* To appear.

[33] Hallin, M. and Mélard, G. (1988). Rank-based tests for randomness against first-order serial dependence. *J. Amer. Statist. Assoc.* **83** 1117–1128. MR0997590

[34] Hallin, M., Oja, H. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. II. Optimal *R*-estimation of shape. *Ann. Statist.* **34** 2757–2789. MR2329466 https://doi.org/10.1214/009053606000000948

[35] Hallin, M. and Paindaveine, D. (2002). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks. *Ann. Statist.* **30** 1103–1133. MR1926170 https://doi.org/10.1214/aos/1031689019

[36] Hallin, M. and Paindaveine, D. (2004). Rank-based optimal tests of the adequacy of an elliptic VARMA model. *Ann. Statist.* **32** 2642–2678. MR2153998 https://doi.org/10.1214/009053604000000724

[37] Hallin, M. and Paindaveine, D. (2004). Multivariate signed-rank tests in vector autoregressive order identification. *Statist. Sci.* **19** 697–711. MR2185591 https://doi.org/10.1214/088342304000000602

[38] Hallin, M. and Paindaveine, D. (2005). Affine-invariant aligned rank tests for the multivariate general linear model with VARMA errors. *J. Multivariate Anal.* **93** 122–163. MR2119768 https://doi.org/10.1016/j.jmva.2004.01.005

[39] Hallin, M. and Paindaveine, D. (2006). Asymptotic linearity of serial and nonserial multivariate signed rank statistics. *J. Statist. Plann. Inference* **136** 1–32. MR2205969 https://doi.org/10.1016/j.jspi.2004.05.013

[40] Hallin, M. and Paindaveine, D. (2006). Semiparametrically efficient rank-based inference for shape. I. Optimal rank-based tests for sphericity. *Ann. Statist.* **34** 2707–2756. MR2329465 https://doi.org/10.1214/009053606000000731

[41] Hallin, M. and Paindaveine, D. (2008). Optimal rank-based tests for homogeneity of scatter. *Ann. Statist.* **36** 1261–1298. MR2418657 https://doi.org/10.1214/07-AOS508

[42] Hallin, M., Paindaveine, D. and Verdebout, T. (2010). Optimal rank-based testing for principal components. *Ann. Statist.* **38** 3245–3299. MR2766852 https://doi.org/10.1214/10-AOS810

[43] Hallin, M., Paindaveine, D. and Verdebout, T. (2013). Optimal rank-based tests for common principal components. *Bernoulli* **19** 2524–2556. MR3160563 https://doi.org/10.3150/12-BEJ461

[44] Hallin, M., Paindaveine, D. and Verdebout, T. (2014). Efficient R-estimation of principal and common principal components. *J. Amer. Statist. Assoc.* **109** 1071–1083. MR3265681 https://doi.org/10.1080/01621459.2014.880057

[45] Hallin, M. and Puri, M.L. (1988). Optimal rank-based procedures for time series analysis: Testing an ARMA model against other ARMA models. *Ann. Statist.* **16** 402–432. MR0924878 https://doi.org/10.1214/aos/1176350712

[46] Hallin, M. and Puri, M.L. (1991). Time series analysis via rank order theory: Signed-rank tests for ARMA models. *J. Multivariate Anal.* **39** 1–29. MR1128669 https://doi.org/10.1016/0047-259X(91)90002-J

[47] Hallin, M. and Puri, M.L. (1994). Aligned rank tests for linear models with autocorrelated error terms. *J. Multivariate Anal.* **50** 175–237. MR1293044 https://doi.org/10.1006/jmva.1994.1040

[48] Hallin, M. and Vermandele, C. (1996). An elementary proof of asymptotic normality for serial rank statistics. In *Research Developments in Probability and Statistics* (E. Brunner and M. Denker, eds.). 163–191. Utrecht: VSP. MR1462415

[49] Hallin, M. and Werker, B.J.M. (2003). Semi-parametric efficiency, distribution-freeness and invariance. *Bernoulli* **9** 137–165. MR1963675 https://doi.org/10.3150/bj/1068129013

[50] Judd, K.L. (1998). *Numerical Methods in Economics*. Cambridge, MA: MIT Press. MR1646331

[51] Kreiss, J.-P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133. MR0885727 https://doi.org/10.1214/aos/1176350256

[52] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. *Springer Series in Statistics*. New York: Springer. MR0856411 https://doi.org/10.1007/978-1-4612-4946-7

[53] Le Cam, L. and Yang, G.L. (2000). *Asymptotics in Statistics: Some Basic Concepts*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1784901 https://doi.org/10.1007/978-1-4612-1166-2

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

[55] Machado, J.A.F. (1993). Robust model selection and *M*-estimation. *Econometric Theory* **9** 478–493. MR1241985 https://doi.org/10.1017/S0266466600007775

[56] Mancini, L., Ronchetti, E. and Trojani, F. (2005). Optimal conditionally unbiased bounded-influence inference in dynamic location and scale models. *J. Amer. Statist. Assoc.* **100** 628–641. MR2160565 https://doi.org/10.1198/016214504000001402

[57] Mélard, G. (2022). An indirect proof for the asymptotic properties of VARMA model estimators. *Econom. Stat.* **21** 96–111. MR4366848 https://doi.org/10.1016/j.ecosta.2020.12.004

[58] Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **63**. Philadelphia, PA: SIAM. MR1172997 https://doi.org/10.1137/1.9781611970081

[59] Peyré, G. and Cuturi, M. (2019). Computational optimal transport with applications to data science. *Found. Trends Mach. Learn.* **11** 355–607.

[60] Pötscher, B.M. (1983). Order estimation in ARMA-models by Lagrangian multiplier tests. *Ann. Statist.* **11** 872–885. MR0707937

[61] Ronchetti, E. (1997). Robustness aspects of model choice. *Statist. Sinica* **7** 327–338. MR1466686

[62] Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6** 461–464. MR0468014

[63] Shi, H., Hallin, M., Drton, M. and Han, F. (2021). Center-outward sign- and rank-based quadrant, Spearman, and Kendall tests for multivariate independence. Available at arXiv:2111.15567.

[64] Shi, H., Hallin, M., Drton, M. and Han, F. (2022). On universally consistent and fully distribution-free rank tests of vector independence. *Ann. Statist.* To appear.

[65] Shorack, G.R. (2000). *Probability for Statisticians*. *Springer Texts in Statistics*. New York: Springer. MR1762415

[66] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

Bernoulli **29**(1), 2023, 274–299
https://doi.org/10.3150/21-BEJ1457

Cramér-type moderate deviation of normal approximation for unbounded exchangeable pairs

ZHUO-SONG ZHANG[a]

Department of Statistics and Applied Probability, National University of Singapore, Singapore 117546.
[a]*zszhang.stat@gmail.com*

In Stein's method, the exchangeable pair approach is commonly used to estimate the approximation errors in normal approximation. In this paper, we establish a Cramér-type moderate deviation theorem of normal approximation for unbounded exchangeable pairs. As applications, Cramér-type moderate deviation theorems for the sums of local statistics and general Curie–Weiss model are obtained.

Keywords: Stein's method; exchangeable pair approach; Cramér-type moderate deviation; sums of local statistics; general Curie–Weiss model

References

[1] Barbour, A.D., Chen, L.H.Y. and Choi, K.P. (1995). Poisson approximation for unbounded functions. I. Independent summands. *Statist. Sinica* **5** 749–766. MR1347617
[2] Barbour, A.D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. New York: The Clarendon Press, Oxford Univ. Press. Oxford Science Publications. MR1163825
[3] Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072 https://doi.org/10.1007/s00440-006-0029-y
[4] Chatterjee, S. (2014). A short survey of Stein's method. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV* 1–24. Seoul: Kyung Moon Sa. MR3727600
[5] Chatterjee, S., Diaconis, P. and Meckes, E. (2005). Exchangeable pairs and Poisson approximation. *Probab. Surv.* **2** 64–106. MR2121796 https://doi.org/10.1214/154957805100000096
[6] Chatterjee, S. and Meckes, E. (2008). Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 257–283. MR2453473
[7] Chatterjee, S. and Shao, Q.-M. (2011). Nonnormal approximation by Stein's method of exchangeable pairs with application to the Curie–Weiss model. *Ann. Appl. Probab.* **21** 464–483. MR2807964 https://doi.org/10.1214/10-AAP712
[8] Chen, L.H.Y. and Choi, K.P. (1992). Some asymptotic and large deviation results in Poisson approximation. *Ann. Probab.* **20** 1867–1876. MR1188045
[9] Chen, L.H.Y., Fang, X. and Shao, Q.-M. (2013). Moderate deviations in Poisson approximation: A first attempt. *Statist. Sinica* **23** 1523–1540. MR3222808
[10] Chen, L.H.Y., Fang, X. and Shao, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293. MR3059199 https://doi.org/10.1214/12-AOP746
[11] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. *Probability and Its Applications (New York)*. Heidelberg: Springer. MR2732624 https://doi.org/10.1007/978-3-642-15007-4
[12] Eichelsbacher, P. and Löwe, M. (2010). Stein's method for dependent random variables occurring in statistical mechanics. *Electron. J. Probab.* **15** 962–988. MR2659754 https://doi.org/10.1214/EJP.v15-777
[13] Eichelsbacher, P. and Martschink, B. (2014). Rates of convergence in the Blume–Emery–Griffiths model. *J. Stat. Phys.* **154** 1483–1507. MR3176416 https://doi.org/10.1007/s10955-014-0925-y

[14] Eichelsbacher, P. and Martschink, B. (2015). On rates of convergence in the Curie–Weiss–Potts model with an external field. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 252–282. MR3300970 https://doi.org/10.1214/14-AIHP599

[15] 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.

[16] Ellis, R.S. and Newman, C.M. (1978). The statistics of Curie–Weiss models. *J. Stat. Phys.* **19** 149–161. MR0503332 https://doi.org/10.1007/BF01012508

[17] Fang, X., Luo, L. and Shao, Q.-M. (2020). A refined Cramér-type moderate deviation for sums of local statistics. *Bernoulli* **26** 2319–2352. MR4091111 https://doi.org/10.3150/20-BEJ1195

[18] Holmes, S. and Reinert, G. (2004). Stein's method for the bootstrap. In *Stein's Method: Expository Lectures and Applications*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **46** 95–136. Beachwood, OH: IMS. MR2118605

[19] Kirkpatrick, K. and Meckes, E. (2013). Asymptotics of the mean-field Heisenberg model. *J. Stat. Phys.* **152** 54–92. MR3067076 https://doi.org/10.1007/s10955-013-0753-5

[20] Kirkpatrick, K. and Nawaz, T. (2016). Asymptotics of mean-field $O(N)$ models. *J. Stat. Phys.* **165** 1114–1140. MR3575640 https://doi.org/10.1007/s10955-016-1667-9

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

[22] Petrov, V.V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* **82**. New York: Springer. MR0388499

[23] Raič, M. (2007). CLT-related large deviation bounds based on Stein's method. *Adv. in Appl. Probab.* **39** 731–752. MR2357379 https://doi.org/10.1239/aap/1189518636

[24] Rinott, Y. and Rotar, V. (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted U-statistics. *Ann. Appl. Probab.* **7** 1080–1105. MR1484798 https://doi.org/10.1214/aoap/1043862425

[25] Shao, Q.-M., Zhang, M. and Zhang, Z.-S. (2021). Cramér-type moderate deviation theorems for nonnormal approximation. *Ann. Appl. Probab.* **31** 247–283. MR4254480 https://doi.org/10.1214/20-aap1589

[26] Shao, Q.-M. and Zhang, Z.-S. (2016). Identifying the limiting distribution by a general approach of Stein's method. *Sci. China Math.* **59** 2379–2392. MR3578962 https://doi.org/10.1007/s11425-016-0322-3

[27] Shao, Q.-M. and Zhang, Z.-S. (2019). Berry–Esseen bounds of normal and nonnormal approximation for unbounded exchangeable pairs. *Ann. Probab.* **47** 61–108. MR3909966 https://doi.org/10.1214/18-AOP1255

[28] 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: Univ. California Press. MR0402873

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

[30] Zhang, Z.-S. (2021). Berry–Esseen bounds for generalized U-statistics. Preprint. Available at arXiv:2104.03419.

[31] Zhang, Z.-S. (2023). Supplement to "Cramér-type moderate deviation of normal approximation for unbounded exchangeable pairs." https://doi.org/10.3150/21-BEJ1457SUPP

Bernoulli **29**(1), 2023, 300–322
https://doi.org/10.3150/21-BEJ1458

Variational formulas for asymptotic variance of general discrete-time Markov chains

LU-JING HUANG[1,a] and YONG-HUA MAO[2,b]

[1] *School of Mathematics and Statistics, Fujian Normal University, 350007 Fuzhou, P.R. China.* [a] *huanglj@fjnu.edu.cn*
[2] *School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, 100875 Beijing, P.R. China.* [b] *maoyh@bnu.edu.cn*

The asymptotic variance is an important criterion to evaluate the performance of Markov chains, especially for the central limit theorems. We give the variational formulas for the asymptotic variance of discrete-time (non-reversible) Markov chains on general state space. The variational formulas provide many applications, extending the classical Peskun's comparison theorem to non-reversible Markov chains, and obtaining several comparison theorems between Markov chains with various perturbations.

Keywords: Markov chain; asymptotic variance; variational formula; non-reversible; Peskun's theorem; comparison theorem

References

[1] Aldous, D.J. and Fill, J.A. (2002). *Reversible Markov Chains and Random Walks on Graphs*. Available at www.berkeley.edu/users/aldous/book.html.
[2] Andrieu, C. and Livingstone, S. (2021). Peskun–Tierney ordering for Markovian Monte Carlo: Beyond the reversible scenario. *Ann. Statist.* **49** 1958–1981. MR4319237 https://doi.org/10.1214/20-aos2008
[3] Bierkens, J. (2016). Non-reversible Metropolis–Hastings. *Stat. Comput.* **26** 1213–1228. MR3538633 https://doi.org/10.1007/s11222-015-9598-x
[4] Caracciolo, S., Pelissetto, A. and Sokal, A.D. (1990). Nonlocal Monte Carlo algorithm for self-avoiding walks with fixed endpoints. *J. Stat. Phys.* **60** 1–53. MR1063212 https://doi.org/10.1007/BF01013668
[5] Chen, T.-L. and Hwang, C.-R. (2013). Accelerating reversible Markov chains. *Statist. Probab. Lett.* **83** 1956–1962. MR3079029 https://doi.org/10.1016/j.spl.2013.05.002
[6] Cogburn, R. (1972). The central limit theorem for Markov processes. In *Proceedings of the Sixth Annual Berkeley Symposium on Mathematical Statistics and Probability* **2** 485–512.
[7] Derriennic, Y. and Lin, M. (2001). The central limit theorem for Markov chains with normal transition operators, started at a point. *Probab. Theory Related Fields* **119** 508–528. MR1826405 https://doi.org/10.1007/PL00008769
[8] Doyle, P.G. (1994). Energy for Markov chains. Available at https://math.dartmouth.edu/~doyle/docs/energy/energy.pdf.
[9] Duncan, A.B., Lelièvre, T. and Pavliotis, G.A. (2016). Variance reduction using nonreversible Langevin samplers. *J. Stat. Phys.* **163** 457–491. MR3483241 https://doi.org/10.1007/s10955-016-1491-2
[10] Fill, J.A. and Kahn, J. (2013). Comparison inequalities and fastest-mixing Markov chains. *Ann. Appl. Probab.* **23** 1778–1816. MR3114917 https://doi.org/10.1214/12-AAP886
[11] Frigessi, A., Hwang, C.-R. and Younes, L. (1992). Optimal spectral structure of reversible stochastic matrices, Monte Carlo methods and the simulation of Markov random fields. *Ann. Appl. Probab.* **2** 610–628. MR1177902
[12] Gaudillière, A. and Landim, C. (2014). A Dirichlet principle for non reversible Markov chains and some recurrence theorems. *Probab. Theory Related Fields* **158** 55–89. MR3152780 https://doi.org/10.1007/s00440-012-0477-5

[13] Gordin, M.I. (1969). The central limit theorem for stationary processes. *Dokl. Akad. Nauk SSSR* **188** 739–741. MR0251785

[14] Gordin, M.I. and Lifšic, B.A. (1978). Central limit theorem for stationary Markov processes. *Dokl. Akad. Nauk SSSR* **239** 766–767. MR0501277

[15] Griffeath, D. and Liggett, T.M. (1982). Critical phenomena for Spitzer's reversible nearest particle systems. *Ann. Probab.* **10** 881–895. MR0672290

[16] Häggström, O. (2005). On the central limit theorem for geometrically ergodic Markov chains. *Probab. Theory Related Fields* **132** 74–82. MR2136867 https://doi.org/10.1007/s00440-004-0390-7

[17] Huang, L.-J. and Mao, Y.-H. (2018). Variational principles of hitting times for non-reversible Markov chains. *J. Math. Anal. Appl.* **468** 959–975. MR3852560 https://doi.org/10.1016/j.jmaa.2018.08.036

[18] Huang, L.-J., Kim, K.-Y. and Mao, Y.-H. (2020). Variational principles for the exit time of non-symmetric diffusions. Preprint. Available at arXiv:2006.16446.

[19] Hwang, C.-R., Normand, R. and Wu, S.-J. (2015). Variance reduction for diffusions. *Stochastic Process. Appl.* **125** 3522–3540. MR3357618 https://doi.org/10.1016/j.spa.2015.03.006

[20] Hwang, C.-R. (2006). Nonreversible perturbations accelerate convergence. In *RIMS Kokyuroku 1462, "The 7th Workshop on Stochastic Numerics"* 26–34.

[21] Ibragimov, I.A. and Linnik, Y.V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff Publishing (English translation). MR0322926

[22] Isaacson, D. and Luecke, G.R. (1978). Strongly ergodic Markov chains and rates of convergence using spectral conditions. *Stochastic Process. Appl.* **7** 113–121. MR0483021 https://doi.org/10.1016/0304-4149(78)90042-x

[23] Kinateder, K.K.J. and McDonald, P. (1999). Variational principles for average exit time moments for diffusions in Euclidean space. *Proc. Amer. Math. Soc.* **127** 2767–2772. MR1600101 https://doi.org/10.1090/S0002-9939-99-04843-1

[24] Kipnis, C. and Varadhan, S.R.S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** 1–19. MR0834478

[25] Komorowski, T., Landim, C. and Olla, S. (2012). *Fluctuations in Markov Processes: Time Symmetry and Martingale Approximation*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **345**. Heidelberg: Springer. MR2952852 https://doi.org/10.1007/978-3-642-29880-6

[26] Leisen, F. and Mira, A. (2008). An extension of Peskun and Tierney orderings to continuous time Markov chains. *Statist. Sinica* **18** 1641–1651. MR2469328

[27] Leskelä, L. and Vihola, M. (2017). Conditional convex orders and measurable martingale couplings. *Bernoulli* **23** 2784–2807. MR3648045 https://doi.org/10.3150/16-BEJ827

[28] Maire, F., Douc, R. and Olsson, J. (2014). Comparison of asymptotic variances of inhomogeneous Markov chains with application to Markov chain Monte Carlo methods. *Ann. Statist.* **42** 1483–1510. MR3262458 https://doi.org/10.1214/14-AOS1209

[29] Marshall, A.W., Olkin, I. and Arnold, B.C. (1979). *Inequalities: Theory of Majorization and Its Applications*. *Mathematics in Science and Engineering* **143**. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0552278

[30] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2509253 https://doi.org/10.1017/CBO9780511626630

[31] Pai, H.-M. and Hwang, C.-R. (2013). Accelerating Brownian motion on N-torus. *Statist. Probab. Lett.* **83** 1443–1447. MR3041295 https://doi.org/10.1016/j.spl.2013.02.009

[32] Peskun, P.H. (1973). Optimum Monte-Carlo sampling using Markov chains. *Biometrika* **60** 607–612. MR0362823 https://doi.org/10.1093/biomet/60.3.607

[33] Pinsky, R.G. (1988). A generalized Dirichlet principle for second order nonselfadjoint elliptic operators. *SIAM J. Math. Anal.* **19** 204–213. MR0924555 https://doi.org/10.1137/0519015

[34] Roberts, G.O. and Rosenthal, J.S. (1997). Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.* **2** 13–25. MR1448322 https://doi.org/10.1214/ECP.v2-981

[35] Roberts, G.O. and Rosenthal, J.S. (2008). Variance bounding Markov chains. *Ann. Appl. Probab.* **18** 1201–1214. MR2418242 https://doi.org/10.1214/07-AAP486

[36] Sahu, S.K. and Zhigljavsky, A.A. (2003). Self-regenerative Markov chain Monte Carlo with adaptation. *Bernoulli* **9** 395–422. MR1997490 https://doi.org/10.3150/bj/1065444811

[37] Sun, Y., Gomez, F. and Schmidhuber, J. (2010). Improving the asymptotic performance of Markov chain Monte-Carlo by inserting vortices. In *Neural Information Processing Systems* 2235–2243.

[38] Tierney, L. (1998). A note on Metropolis–Hastings kernels for general state spaces. *Ann. Appl. Probab.* **8** 1–9. MR1620401 https://doi.org/10.1214/aoap/1027961031

Bernoulli **29**(1), 2023, 323–349
https://doi.org/10.3150/22-BEJ1459

Stratified incomplete local simplex tests for curvature of nonparametric multiple regression

YANGLEI SONG[1,a], XIAOHUI CHEN[2,b] and KENGO KATO[3,c]

[1]*Department of Mathematics and Statistics, Queen's University, Jeffery Hall, Kingston, ON, Canada, K7L 3N6.* [a]*yanglei.song@queensu.ca*
[2]*Department of Statistics, University of Illinois at Urbana-Champaign, 725 S. Wright Street, Champaign, IL 61820, USA.* [b]*xhchen@illinois.edu*
[3]*Department of Statistics and Data Science, Cornell University, 1194 Comstock Hall, Ithaca, NY 14853, USA.* [c]*kk976@cornell.edu*

Principled nonparametric tests for regression curvature in $\mathbb{R}^d$ are often statistically and computationally challenging. This paper introduces the stratified incomplete local simplex (SILS) tests for joint concavity of nonparametric multiple regression. The SILS tests with suitable bootstrap calibration are shown to achieve simultaneous guarantees on dimension-free computational complexity, polynomial decay of the uniform error-in-size, and power consistency for general (global and local) alternatives. To establish these results, we develop a general theory for incomplete U-processes with stratified random sparse weights. Novel technical ingredients include maximal inequalities for the supremum of multiple incomplete U-processes.

Keywords: Nonparametric regression; curvature testing; incomplete U-processes; stratification

References

[1] Abrevaya, J. and Jiang, W. (2005). A nonparametric approach to measuring and testing curvature. *J. Bus. Econom. Statist.* **23** 1–19. MR2108688 https://doi.org/10.1198/073500104000000316
[2] Arcones, M.A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426
[3] Arcones, M.A. and Giné, E. (1994). U-processes indexed by Vapnik–Červonenkis classes of functions with applications to asymptotics and bootstrap of U-statistics with estimated parameters. *Stochastic Process. Appl.* **52** 17–38. MR1289166 https://doi.org/10.1016/0304-4149(94)90098-1
[4] Blom, G. (1976). Some properties of incomplete U-statistics. *Biometrika* **63** 573–580. MR0474582 https://doi.org/10.1093/biomet/63.3.573
[5] Borenstein, S. and Farrell, J. (2007). Do investors forecast fat firms? Evidence from the gold-mining industry. *Rand J. Econ.* **38** 626–647.
[6] Borovskikh, Y.V. (1996). *U-Statistics in Banach Spaces*. Utrecht: VSP. MR1419498
[7] Bowman, A., Jones, M. and Gijbels, I. (1998). Testing monotonicity of regression. *J. Comput. Graph. Statist.* **7** 489–500.
[8] Brown, B.M. and Kildea, D.G. (1978). Reduced U-statistics and the Hodges-Lehmann estimator. *Ann. Statist.* **6** 828–835. MR0491556
[9] Cai, T.T. and Low, M.G. (2015). A framework for estimation of convex functions. *Statist. Sinica* **25** 423–456. MR3379081
[10] Chatterjee, S. (2014). A new perspective on least squares under convex constraint. *Ann. Statist.* **42** 2340–2381. MR3269982 https://doi.org/10.1214/14-AOS1254
[11] Chatterjee, S. (2016). An improved global risk bound in concave regression. *Electron. J. Stat.* **10** 1608–1629. MR3522655 https://doi.org/10.1214/16-EJS1151
[12] Chen, X. and Kato, K. (2019). Randomized incomplete U-statistics in high dimensions. *Ann. Statist.* **47** 3127–3156. MR4025737 https://doi.org/10.1214/18-AOS1773

[13] Chen, X. and Kato, K. (2020). Jackknife multiplier bootstrap: Finite sample approximations to the U-process supremum with applications. *Probab. Theory Related Fields* **176** 1097–1163. MR4087490 https://doi.org/10.1007/s00440-019-00936-y

[14] Chen, Y. and Wellner, J.A. (2016). On convex least squares estimation when the truth is linear. *Electron. J. Stat.* **10** 171–209. MR3466180 https://doi.org/10.1214/15-EJS1098

[15] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42** 1564–1597. MR3262461 https://doi.org/10.1214/14-AOS1230

[16] Chetverikov, D., Santos, A. and Shaikh, A.M. (2018). The econometrics of shape restrictions. *Ann. Rev. Econ.* **10** 31–63.

[17] De la Pena, V. and Giné, E. (2012). *Decoupling: From Dependence to Independence*. Berlin: Springer.

[18] Diack, C.A.T. and Thomas-Agnan, C. (1998). A nonparametric test of the non-convexity of regression. *J. Nonparametr. Stat.* **9** 335–362. MR1646904 https://doi.org/10.1080/10485259808832749

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

[20] Einmahl, U. and Mason, D.M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.* **33** 1380–1403. MR2195639 https://doi.org/10.1214/009053605000000129

[21] Fama, E.F. and French, K.R. (1993). Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* **33** 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

[22] Fang, Z. and Seo, J. (2021). A projection framework for testing shape restrictions that form convex cones. *Econometrica* **89** 2439–2458. MR4325255 https://doi.org/10.3982/ecta17764

[23] Frees, E.W. (1994). Estimating densities of functions of observations. *J. Amer. Statist. Assoc.* **89** 517–525. MR1294078

[24] Ghosal, S., Sen, A. and van der Vaart, A.W. (2000). Testing monotonicity of regression. *Ann. Statist.* **28** 1054–1082. MR1810919 https://doi.org/10.1214/aos/1015956707

[25] Giné, E. and Mason, D.M. (2007). On local U-statistic processes and the estimation of densities of functions of several sample variables. *Ann. Statist.* **35** 1105–1145. MR2341700 https://doi.org/10.1214/009053607000000154

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

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

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

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

[30] Gupta, P. and Bhattacharjee, G.P. (1984). An efficient algorithm for random sampling without replacement. In *Foundations of Software Technology and Theoretical Computer Science (Bangalore, 1984)*. *Lecture Notes in Computer Science* **181** 435–442. Berlin: Springer. MR0786338 https://doi.org/10.1007/3-540-13883-8_89

[31] Hall, P. (1991). On convergence rates of suprema. *Probab. Theory Related Fields* **89** 447–455. MR1118558 https://doi.org/10.1007/BF01199788

[32] Han, Q. and Wellner, J.A. (2016). Multivariate convex regression: Global risk bounds and adaptation. ArXiv preprint. Available at arXiv:1601.06844.

[33] Hannah, L.A. and Dunson, D.B. (2013). Multivariate convex regression with adaptive partitioning. *J. Mach. Learn. Res.* **14** 3261–3294. MR3144462

[34] Hanson, D.L. and Pledger, G. (1976). Consistency in concave regression. *Ann. Statist.* **4** 1038–1050. MR0426273

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

[36] Janson, S. (1984). The asymptotic distributions of incomplete U-statistics. *Z. Wahrsch. Verw. Gebiete* **66** 495–505. MR0753810 https://doi.org/10.1007/BF00531887

[37] Juditsky, A. and Nemirovski, A. (2002). On nonparametric tests of positivity/monotonicity/convexity. *Ann. Statist.* **30** 498–527. MR1902897 https://doi.org/10.1214/aos/1021379863

[38] Komarova, T. and Hidalgo, J. (2019). Testing nonparametric shape restrictions. ArXiv preprint. Available at arXiv:1909.01675.

[39] Kuosmanen, T. (2008). Representation theorem for convex nonparametric least squares. *Econom. J.* **11** 308–325.

[40] Kur, G., Dagan, Y. and Rakhlin, A. (2019). Optimality of maximum likelihood for log-concave density estimation and bounded convex regression. ArXiv preprint. Available at arXiv:1903.05315.

[41] Lee, S., Linton, O. and Whang, Y.-J. (2009). Testing for stochastic monotonicity. *Econometrica* **77** 585–602. MR2503039 https://doi.org/10.3982/ECTA7145

[42] Lehmann, E.L. and Romano, J.P. (2006). *Testing Statistical Hypotheses*. New York: Springer.

[43] Lim, E. and Glynn, P.W. (2012). Consistency of multidimensional convex regression. *Oper. Res.* **60** 196–208. MR2911667 https://doi.org/10.1287/opre.1110.1007

[44] Mammen, E. (1991). Nonparametric regression under qualitative smoothness assumptions. *Ann. Statist.* **19** 741–759. MR1105842 https://doi.org/10.1214/aos/1176348118

[45] Matzkin, R.L. (1991). Semiparametric estimation of monotone and concave utility functions for polychotomous choice models. *Econometrica* **59** 1315–1327. MR1133036 https://doi.org/10.2307/2938369

[46] Mazumder, R., Choudhury, A., Iyengar, G. and Sen, B. (2019). A computational framework for multivariate convex regression and its variants. *J. Amer. Statist. Assoc.* **114** 318–331. MR3941257 https://doi.org/10.1080/01621459.2017.1407771

[47] Murphy, K.M. and Welch, F. (1990). Empirical age-earnings profiles. *J. Labor Econ.* **8** 202–229.

[48] Nolan, D. and Pollard, D. (1987). U-processes: Rates of convergence. *Ann. Statist.* **15** 780–799. MR0888439 https://doi.org/10.1214/aos/1176350374

[49] Nolan, D. and Pollard, D. (1988). Functional limit theorems for U-processes. *Ann. Probab.* **16** 1291–1298. MR0942769

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

[51] Sherman, R.P. (1994). Maximal inequalities for degenerate U-processes with applications to optimization estimators. *Ann. Statist.* **22** 439–459. MR1272092 https://doi.org/10.1214/aos/1176325377

[52] Song, Y., Chen, X. and Kato, K. (2019). Approximating high-dimensional infinite-order U-statistics: Statistical and computational guarantees. *Electron. J. Stat.* **13** 4794–4848. MR4038726 https://doi.org/10.1214/19-EJS1643

[53] Song, Y., Chen, X. Kato, K. (2023). Supplement to "Stratified incomplete local simplex tests for curvature of nonparametric multiple regression." https://doi.org/10.3150/22-BEJ1459SUPP

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

[55] Wang, J.C. and Meyer, M.C. (2011). Testing the monotonicity or convexity of a function using regression splines. *Canad. J. Statist.* **39** 89–107. MR2815340 https://doi.org/10.1002/cjs.10094

[56] Zhang, D. (2001). Bayesian bootstraps for U-processes, hypothesis tests and convergence of Dirichlet U-processes. *Statist. Sinica* **11** 463–478. MR1844535

Bernoulli **29**(1), 2023, 350–376
https://doi.org/10.3150/22-BEJ1460

On algebraic Stein operators for Gaussian polynomials

EHSAN AZMOODEH[1,a], DARIO GASBARRA[2,b] and ROBERT E. GAUNT[3,c]

[1] *Department of Mathematical Sciences, Mathematical Sciences Bldg, The University of Liverpool, Liverpool L69 7ZL, UK.* [a] *ehsan.azmoodeh@liverpool.ac.uk*
[2] *Department of Mathematics and Statistics, University of Helsinki, P.O.Box 68 00014, Finland.*
[b] *dario.gasbarra@helsinki.fi*
[3] *Department of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, UK.*
[c] *robert.gaunt@manchester.ac.uk*

The first essential ingredient to build up Stein's method for a continuous target distribution is to identify a so-called *Stein operator*, namely a linear differential operator with polynomial coefficients. In this paper, we introduce the notion of *algebraic* Stein operators (see Definition 3.4), and provide a novel algebraic method to find *all* the algebraic Stein operators up to a given order and polynomial degree for a target random variable of the form $Y = h(X)$, where $X = (X_1, \dots, X_d)$ has i.i.d. standard Gaussian components and $h \in \mathbb{K}[X]$ is a polynomial with coefficients in the ring $\mathbb{K}$. Our approach links the existence of an algebraic Stein operator with *null controllability* of a certain linear discrete system. A `MATLAB` code checks the null controllability up to a given finite time T (the order of the differential operator), and provides all *null control* sequences (polynomial coefficients of the differential operator) up to a given maximum degree m. This is the first paper that connects Stein's method with computational algebra to find Stein operators for highly complex probability distributions, such as $H_{20}(X_1)$, where H_p is the p-th Hermite polynomial. Some examples of Stein operators for $H_p(X_1)$, $p = 3, 4, 5, 6$, are gathered in the Appendix and many other examples are given in the Supplementary Information.

Keywords: Stein's method; Stein operator; Gaussian integration by parts; Malliavin calculus; linear system theory; null controllability; symbolic computation; Hermite polynomials

References

[1] Antsaklis, P.J. and Michel, A.N. (2007). *A Linear Systems Primer*. Boston, MA: Birkhäuser, Inc. MR2375027
[2] Arras, A., Mijoule, G., Poly, G. and Swan, Y. (2017). A new approach to the Stein-Tikhomirov method: With applications to the second Wiener chaos and Dickman convergence. Available at https://arxiv.org/abs/1605.06819.
[3] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2019). A bound on the Wasserstein-2 distance between linear combinations of independent random variables. *Stochastic Process. Appl.* **129** 2341–2375. MR3958435 https://doi.org/10.1016/j.spa.2018.07.009
[4] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2020). Stein characterizations for linear combinations of gamma random variables. *Braz. J. Probab. Stat.* **34** 394–413. MR4093265 https://doi.org/10.1214/18-BJPS420
[5] Arras, B. and Houdré, C. (2019). *On Stein's Method for Infinitely Divisible Laws with Finite First Moment. SpringerBriefs in Probability and Mathematical Statistics*. Cham: Springer. MR3931309
[6] Azmoodeh, E. and Gasbarra, D. (2018). New moments criteria for convergence towards normal product/tetilla laws. Available at https://arxiv.org/abs/1708.07681.
[7] Azmoodeh, E., Gasbarra, D. and Gaunt, R.E. (2021). An asymptotic approach to proving sufficiency of Stein characterisations. Available at https://arxiv.org/abs/2109.08579.
[8] Azmoodeh, E., Gasbarra, D. Gaunt, R.E. (2023). Supplement to "On algebraic Stein operators for Gaussian polynomials." https://doi.org/10.3150/22-BEJ1460SUPP

[9] Azmoodeh, E., Peccati, G. and Poly, G. (2015). Convergence towards linear combinations of chi-squared random variables: A Malliavin-based approach. In *In Memoriam Marc Yor—Séminaire de Probabilités XLVII*. *Lecture Notes in Math.* **2137** 339–367. Cham: Springer. MR3444306 https://doi.org/10.1007/978-3-319-18585-9_16

[10] Azmoodeh, E., Peccati, G. and Yang, X. (2021). Malliavin-Stein method: A survey of some recent developments. *Mod. Stoch. Theory Appl.* **8** 141–177. MR4279874 https://doi.org/10.15559/21-vmsta184

[11] Baker, A. (1975). *Transcendental Number Theory*. London: Cambridge Univ. Press. MR0422171

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

[13] Bourlès, H. (2010). *Linear Systems*. *Control Systems, Robotics and Manufacturing Series*. London: ISTE. MR2798200

[14] Bourlès, H. and Marinescu, B. (2011). *Linear Time-Varying Systems: Algebraic-Analytic Approach*. *Lecture Notes in Control and Information Sciences* **410**. Berlin: Springer. MR2790708 https://doi.org/10.1007/978-3-642-19727-7

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

[16] Döbler, C. (2015). Stein's method of exchangeable pairs for the beta distribution and generalizations. *Electron. J. Probab.* **20** no. 109. MR3418541 https://doi.org/10.1214/EJP.v20-3933

[17] Döbler, C., Gaunt, R.E. and Vollmer, S.J. (2017). An iterative technique for bounding derivatives of solutions of Stein equations. *Electron. J. Probab.* **22** Paper No. 96. MR3724564 https://doi.org/10.1214/17-EJP118

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

[19] Eden, R. and Viens, F. (2013). General upper and lower tail estimates using Malliavin calculus and Stein's equations. In *Seminar on Stochastic Analysis, Random Fields and Applications VII* (R.C. Dalang, M. Dozzi and F. Russo, eds.). *Progress in Probability* **67** 55–84. Basel: Birkhäuser/Springer. MR3380095

[20] Eden, R. and Víquez, J. (2015). Nourdin-Peccati analysis on Wiener and Wiener-Poisson space for general distributions. *Stochastic Process. Appl.* **125** 182–216. MR3274696 https://doi.org/10.1016/j.spa.2014.09.001

[21] Eichelsbacher, P. and Thäle, C. (2015). Malliavin-Stein method for variance-gamma approximation on Wiener space. *Electron. J. Probab.* **20** Paper No. 123. MR3425543 https://doi.org/10.1214/EJP.v20-4136

[22] Fuhrmann, P.A. (1981). *Linear Systems and Operators in Hilbert Space*. New York: McGraw-Hill International Book Co. MR0629828

[23] Fuhrmann, P.A. and Helmke, U. (2015). *The Mathematics of Networks of Linear Systems*. *Universitext*. Cham: Springer. MR3362408 https://doi.org/10.1007/978-3-319-16646-9

[24] Gaunt, R.E. (2014). Variance-gamma approximation via Stein's method. *Electron. J. Probab.* **19** no. 38. MR3194737 https://doi.org/10.1214/EJP.v19-3020

[25] Gaunt, R.E. (2017). On Stein's method for products of normal random variables and zero bias couplings. *Bernoulli* **23** 3311–3345. MR3654808 https://doi.org/10.3150/16-BEJ848

[26] Gaunt, R.E. (2018). Products of normal, beta and gamma random variables: Stein operators and distributional theory. *Braz. J. Probab. Stat.* **32** 437–466. MR3787761 https://doi.org/10.1214/16-BJPS349

[27] Gaunt, R.E. (2019). Stein operators for variables form the third and fourth Wiener chaoses. *Statist. Probab. Lett.* **145** 118–126. MR3873897 https://doi.org/10.1016/j.spl.2018.09.001

[28] Gaunt, R.E., Mijoule, G. and Swan, Y. (2019). An algebra of Stein operators. *J. Math. Anal. Appl.* **469** 260–279. MR3857522 https://doi.org/10.1016/j.jmaa.2018.09.015

[29] Gaunt, R.E., Mijoule, G. and Swan, Y. (2020). Some new Stein operators for product distributions. *Braz. J. Probab. Stat.* **34** 795–808. MR4153642 https://doi.org/10.1214/19-BJPS460

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

[31] Indritz, J. (1961). An inequality for Hermite polynomials. *Proc. Amer. Math. Soc.* **12** 981–983. MR0132852 https://doi.org/10.2307/2034406

[32] Kailath, T. (1980). *Linear Systems*. *Prentice-Hall Information and System Sciences Series*. Englewood Cliffs, NJ: Prentice-Hall, Inc. MR0569473

[33] Kalman, R.E. (1960). On the general theory of control systems. In *Proceedings of the First World Congress of the International Federation of Automatic Control* 481–493.

[34] Kalman, R.E. (1963). Mathematical description of linear dynamical systems. *J. SIAM Control Ser. A* **1** 152–192. MR0152167

[35] Kawano, Y. and Ohtsuka, T. (2016). Commutative algebraic methods for controllability of discrete-time polynomial systems. *Internat. J. Control* **89** 343–351. MR3435204 https://doi.org/10.1080/00207179.2015.1076939

[36] Klamka, J. (1993). Controllability of dynamical systems—a survey. *Arch. Contol. Sci.* **2** 283–310. MR1325771

[37] Koroljuk, V.S. and Borovskich, Y.V. (1994). *Theory of U-Statistics*. *Mathematics and Its Applications* **273**. Dordrecht: Kluwer Academic. MR1472486 https://doi.org/10.1007/978-94-017-3515-5

[38] Kusuoka, S. and Tudor, C.A. (2012). Stein's method for invariant measures of diffusions via Malliavin calculus. *Stochastic Process. Appl.* **122** 1627–1651. MR2914766 https://doi.org/10.1016/j.spa.2012.02.005

[39] Lee, A.J. (1990). *U-Statistics: Theory and Practice*. *Statistics: Textbooks and Monographs* **110**. New York: Dekker. MR1075417

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

[41] Ley, C. and Swan, Y. (2013). Stein's density approach and information inequalities. *Electron. Commun. Probab.* **18** no. 7. MR3019670 https://doi.org/10.1214/ECP.v18-2578

[42] Nourdin, I., Nualart, D. and Peccati, G. (2016). Quantitative stable limit theorems on the Wiener space. *Ann. Probab.* **44** 1–41. MR3456331 https://doi.org/10.1214/14-AOP965

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

[44] Nourdin, I. and Peccati, G. (2010). Cumulants on the Wiener space. *J. Funct. Anal.* **258** 3775–3791. MR2606872 https://doi.org/10.1016/j.jfa.2009.10.024

[45] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. MR2962301 https://doi.org/10.1017/CBO9781139084659

[46] Nourdin, I. and Peccati, G. (2015). The optimal fourth moment theorem. *Proc. Amer. Math. Soc.* **143** 3123–3133. MR3336636 https://doi.org/10.1090/S0002-9939-2015-12417-3

[47] Nourdin, I., Peccati, G. and Réveillac, A. (2010). Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 45–58. MR2641769 https://doi.org/10.1214/08-AIHP308

[48] Nourdin, I. and Poly, G. (2012). Convergence in law in the second Wiener/Wigner chaos. *Electron. Commun. Probab.* **17** no. 36. MR2970700 https://doi.org/10.1214/ecp.v17-2023

[49] Nualart, D. and Nualart, E. (2018). *Introduction to Malliavin Calculus*. *Institute of Mathematical Statistics Textbooks* **9**. Cambridge: Cambridge Univ. Press. MR3838464 https://doi.org/10.1017/9781139856485

[50] Peccati, G. (2014). Quantitative CLTs on a Gaussian space: A survey of recent developments. In *Journées MAS 2012*. *ESAIM Proc.* **44** 61–78. Les Ulis: EDP Sci. MR3178608 https://doi.org/10.1051/proc/201444003

[51] Peköz, E.A., Röllin, A. and Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** 1188–1218. MR3076682 https://doi.org/10.1214/12-AAP868

[52] Pike, J. and Ren, H. (2014). Stein's method and the Laplace distribution. *ALEA Lat. Am. J. Probab. Math. Stat.* **11** 571–587. MR3283586 https://doi.org/10.1007/s11633-014-0767-8

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

[54] Rugh, W.J. (1993). *Linear System Theory*. *Prentice Hall Information and System Sciences Series*. Englewood Cliffs, NJ: Prentice Hall, Inc. MR1211190

[55] Schoutens, W. (2001). Orthogonal polynomials in Stein's method. *J. Math. Anal. Appl.* **253** 515–531. MR1808151 https://doi.org/10.1006/jmaa.2000.7159

[56] Sontag, E.D. (1998). *Mathematical Control Theory: Deterministic Finite-Dimensional Systems*, 2nd ed. *Texts in Applied Mathematics* **6**. New York: Springer. MR1640001 https://doi.org/10.1007/978-1-4612-0577-7

[57] 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. MR0402873

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

[59] Stein, C., Diaconis, P., Holmes, S. and Reinert, G. (2004). Use of exchangeable pairs in the analysis of simulations. In *Stein's Method: Expository Lectures and Applications* (P. Diaconis and S. Holmes, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **46** 1–26. Beachwood, OH: IMS. MR2118600 https://doi.org/10.1214/lnms/1196283797

[60] Triggiani, R. (1975). Controllability and observability in Banach space with bounded operators. *SIAM J. Control* **13** 462–491. MR0394386

[61] Triggiani, R. (1976). Extensions of rank conditions for controllability and observability to Banach spaces and unbounded operators. *SIAM J. Control Optim.* **14** 313–338. MR0479499 https://doi.org/10.1137/0314022

[62] Xu, L. (2019). Approximation of stable law in Wasserstein-1 distance by Stein's method. *Ann. Appl. Probab.* **29** 458–504. MR3910009 https://doi.org/10.1214/18-AAP1424

Bernoulli **29**(1), 2023, 377–402
https://doi.org/10.3150/22-BEJ1461

Targeted cross-validation

JIAWEI ZHANG[a], JIE DING[b] and YUHONG YANG[c]

School of Statistics, University of Minnesota, MN 55455, USA. [a]*zhan4362@umn.edu,* [b]*dingj@umn.edu,* [c]*yangx374@umn.edu*

In many applications, we have access to the complete dataset but are only interested in the prediction of a particular region of predictor variables. A standard approach is to find the globally best modeling method from a set of candidate methods. However, it is perhaps rare in reality that one candidate method is uniformly better than the others. A natural approach for this scenario is to apply a weighted L_2 loss in performance assessment to reflect the region-specific interest. We propose a targeted cross-validation (TCV) to select models or procedures based on a general weighted L_2 loss. We show that the TCV is consistent in selecting the best performing candidate under the weighted L_2 loss. Experimental studies are used to demonstrate the use of TCV and its potential advantage over the global CV or the approach of using only local data for modeling a local region.

Previous investigations on CV have relied on the condition that when the sample size is large enough, the ranking of two candidates stays the same. However, in many applications with the setup of changing data-generating processes or highly adaptive modeling methods, the relative performance of the methods is not static as the sample size varies. Even with a fixed data-generating process, it is possible that the ranking of two methods switches infinitely many times. In this work, we broaden the concept of the selection consistency by allowing the best candidate to switch as the sample size varies, and then establish the consistency of the TCV. This flexible framework can be applied to high-dimensional and complex machine learning scenarios where the relative performances of modeling procedures are dynamic.

Keywords: Consistency; cross-validation; model selection; regression

References

[1] Allen, D.M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics* **16** 125–127. MR0343481 https://doi.org/10.2307/1267500

[2] Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Stat. Surv.* **4** 40–79. MR2602303 https://doi.org/10.1214/09-SS054

[3] Arlot, S. and Celisse, A. (2011). Segmentation of the mean of heteroscedastic data via cross-validation. *Stat. Comput.* **21** 613–632. MR2826696 https://doi.org/10.1007/s11222-010-9196-x

[4] Arlot, S. and Lerasle, M. (2016). Choice of V for V-fold cross-validation in least-squares density estimation. *J. Mach. Learn. Res.* **17** 208. MR3595142

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

[6] Baraud, Y., Giraud, C. and Huet, S. (2014). Estimator selection in the Gaussian setting. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 1092–1119. MR3224300 https://doi.org/10.1214/13-AIHP539

[7] Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. *Wadsworth Statistics/Probability Series*. Belmont, CA: Wadsworth Advanced Books and Software. MR0726392

[8] Burman, P. (1989). A comparative study of ordinary cross-validation, v-fold cross-validation and the repeated learning-testing methods. *Biometrika* **76** 503–514. MR1040644 https://doi.org/10.1093/biomet/76.3.503

[9] Burman, P. (1990). Estimation of optimal transformations using v-fold cross validation and repeated learning-testing methods. *Sankhyā Ser. A* **52** 314–345. MR1178041

[10] Celisse, A. (2014). Optimal cross-validation in density estimation with the L^2-loss. *Ann. Statist.* **42** 1879–1910. MR3262471 https://doi.org/10.1214/14-AOS1240

[11] Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions. *Numer. Math.* **31** 377–403.
[12] Ding, J., Tarokh, V. and Yang, Y. (2018). Model selection techniques: An overview. *IEEE Signal Process. Mag.* **35** 16–34.
[13] Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* **32** 928–961. MR2065194 https://doi.org/10.1214/009053604000000256
[14] Feng, Y. and Yu, Y. (2019). The restricted consistency property of leave-n_v-out cross-validation for high-dimensional variable selection. *Statist. Sinica* **29** 1607–1630. MR3932532
[15] Geisser, S. (1975). The predictive sample reuse method with applications. *J. Amer. Statist. Assoc.* **70** 320–328.
[16] Harrison, D. and Rubinfeld, D.L. (1978). Hedonic housing prices and the demand for clean air. *J. Environ. Econ. Manage.* **5** 81–102.
[17] Lei, J. (2020). Cross-validation with confidence. *J. Amer. Statist. Assoc.* **115** 1978–1997. MR4189771 https://doi.org/10.1080/01621459.2019.1672556
[18] Li, K.-C. (1984). Consistency for cross-validated nearest neighbor estimates in nonparametric regression. *Ann. Statist.* **12** 230–240. MR0733510 https://doi.org/10.1214/aos/1176346403
[19] Li, K.-C. (1987). Asymptotic optimality for C_p, C_L, cross-validation and generalized cross-validation: Discrete index set. *Ann. Statist.* **15** 958–975. MR0902239 https://doi.org/10.1214/aos/1176350486
[20] Maillard, G., Arlot, S. and Lerasle, M. (2021). Aggregated hold-out. *J. Mach. Learn. Res.* **22** 20. MR4253713 https://doi.org/10.1080/15502287.2020.1834012
[21] Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.* **9** 141–142.
[22] Nemirovski, A. (2000). Topics in non-parametric statistics. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1998). Lecture Notes in Math.* **1738** 85–277. Berlin: Springer. MR1775640
[23] Picard, R.R. and Cook, R.D. (1984). Cross-validation of regression models. *J. Amer. Statist. Assoc.* **79** 575–583. MR0763576
[24] Racine, J. (2000). Consistent cross-validatory model-selection for dependent data: hv-block cross-validation. *J. Econometrics* **99** 39–61.
[25] Shao, J. (1993). Linear model selection by cross-validation. *J. Amer. Statist. Assoc.* **88** 486–494. MR1224373
[26] Shao, J. (1997). An asymptotic theory for linear model selection (with discussion). *Statist. Sinica* **7** 221–264. MR1466682
[27] Speckman, P. (1985). Spline smoothing and optimal rates of convergence in nonparametric regression models. *Ann. Statist.* **13** 970–983. MR0803752 https://doi.org/10.1214/aos/1176349650
[28] Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053. MR0673642
[29] Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *J. Roy. Statist. Soc. Ser. B* **36** 111–147. MR0356377
[30] Wong, W.H. (1983). On the consistency of cross-validation in kernel nonparametric regression. *Ann. Statist.* **11** 1136–1141. MR0720259 https://doi.org/10.1214/aos/1176346327
[31] Yang, Y. (2006). Comparing learning methods for classification. *Statist. Sinica* **16** 635–657. MR2267253
[32] Yang, Y. (2007). Consistency of cross validation for comparing regression procedures. *Ann. Statist.* **35** 2450–2473. MR2382654 https://doi.org/10.1214/009053607000000514
[33] Yang, Y. (2008). Localized model selection for regression. *Econometric Theory* **24** 472–492. MR2391618 https://doi.org/10.1017/S0266466608080195
[34] Zhan, Z. and Yang, Y. (2022). Profile electoral college cross-validation. *Inform. Sci.* **586** 24–40.
[35] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. MR2604701 https://doi.org/10.1214/09-AOS729
[36] Zhang, P. (1993). Model selection via multifold cross validation. *Ann. Statist.* **21** 299–313. MR1212178 https://doi.org/10.1214/aos/1176349027
[37] Zhang, Y. and Yang, Y. (2015). Cross-validation for selecting a model selection procedure. *J. Econometrics* **187** 95–112. MR3347297 https://doi.org/10.1016/j.jeconom.2015.02.006

Bernoulli **29**(1), 2023, 403–427
https://doi.org/10.3150/22-BEJ1462

Adaptive Bayesian inference for current status data on a grid

MINWOO CHAE[a]

Department of Industrial and Management Engineering, Pohang University of Science and Technology, 77 Cheongam-Ro, Nam-Gu, Pohang, Gyeongbuk 37673, South Korea. [a]*mchae@postech.ac.kr*

We study a Bayesian approach to the inference of an event time distribution in the current status model where observation times are supported on a grid of potentially unknown sparsity and multiple subjects share the same observation time. The model leads to a very simple likelihood, but statistical inferences are non-trivial due to the unknown sparsity of the grid. In particular, for an inference based on the maximum likelihood estimator, one needs to estimate the density of the event time distribution which is challenging because the event time is not directly observed. We consider Bayes procedures with a Dirichlet prior on the event time distribution. With this prior, the Bayes estimator and credible sets can be easily computed via a Gibbs sampler algorithm. Our main contribution is to provide thorough investigation of frequentist's properties of the posterior distribution. Specifically, it is shown that the posterior convergence rate is adaptive to the unknown sparsity of the grid. If the grid is sufficiently sparse, we further prove the Bernstein–von Mises theorem which guarantees frequentist's validity of Bayesian credible sets. A numerical study is also conducted for illustration.

Keywords: Adaptive procedure; Bayesian survival analysis; Bernstein–von Mises theorem; posterior convergence rate; current status model; interval-censored data

References

Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. *Springer Series in Statistics*. New York: Springer. MR1198884 https://doi.org/10.1007/978-1-4612-4348-9

Banerjee, M. (2007). Likelihood based inference for monotone response models. *Ann. Statist.* **35** 931–956. MR2341693 https://doi.org/10.1214/009053606000001578

Bennett, S. (1983). Analysis of survival data by the proportional odds model. *Stat. Med.* **2** 273–277.

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

Bochkina, N.A. and Green, P.J. (2014). The Bernstein-von Mises theorem and nonregular models. *Ann. Statist.* **42** 1850–1878. MR3262470 https://doi.org/10.1214/14-AOS1239

Cai, B., Lin, X. and Wang, L. (2011). Bayesian proportional hazards model for current status data with monotone splines. *Comput. Statist. Data Anal.* **55** 2644–2651. MR2802342 https://doi.org/10.1016/j.csda.2011.03.013

Castillo, I. (2012). A semiparametric Bernstein–von Mises theorem for Gaussian process priors. *Probab. Theory Related Fields* **152** 53–99. MR2875753 https://doi.org/10.1007/s00440-010-0316-5

Castillo, I. (2014). On Bayesian supremum norm contraction rates. *Ann. Statist.* **42** 2058–2091. MR3262477 https://doi.org/10.1214/14-AOS1253

Castillo, I. and Nickl, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. MR3262473 https://doi.org/10.1214/14-AOS1246

Castillo, I. and van der Pas, S. (2021). Multiscale Bayesian survival analysis. *Ann. Statist.* **49** 3559–3582. MR4352541 https://doi.org/10.1214/21-aos2097

Chae, M., De Blasi, P. and Walker, S.G. (2021). Posterior asymptotics in Wasserstein metrics on the real line. *Electron. J. Stat.* **15** 3635–3677. MR4298980 https://doi.org/10.1214/21-ejs1869

Chakraborty, M. and Ghosal, S. (2021). Coverage of credible intervals in nonparametric monotone regression. *Ann. Statist.* **49** 1011–1028. MR4255117 https://doi.org/10.1214/20-aos1989

Cox, D.R. (1972). Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B* **34** 187–220. MR0341758

De Blasi, P. and Hjort, N.L. (2009). The Bernstein-von Mises theorem in semiparametric competing risks models. *J. Statist. Plann. Inference* **139** 2316–2328. MR2507993 https://doi.org/10.1016/j.jspi.2008.10.018

Dunson, D.B. and Dinse, G.E. (2002). Bayesian models for multivariate current status data with informative censoring. *Biometrics* **58** 79–88. MR1891046 https://doi.org/10.1111/j.0006-341X.2002.00079.x

Finkelstein, D.M. and Wolfe, R.A. (1985). A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics* **41** 933–945. MR0833140 https://doi.org/10.2307/2530965

Gentleman, R. and Geyer, C.J. (1994). Maximum likelihood for interval censored data: Consistency and computation. *Biometrika* **81** 618–623. MR1311106 https://doi.org/10.1093/biomet/81.3.618

Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228

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

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

Groeneboom, P. (1991). Nonparametric maximum likelihood estimators for interval censoring and deconvolution. Technical Report.

Groeneboom, P. and Hendrickx, K. (2018). Current status linear regression. *Ann. Statist.* **46** 1415–1444. MR3819105 https://doi.org/10.1214/17-AOS1589

Groeneboom, P., Maathuis, M.H. and Wellner, J.A. (2008). Current status data with competing risks: Limiting distribution of the MLE. *Ann. Statist.* **36** 1064–1089. MR2418649 https://doi.org/10.1214/009053607000000983

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

Groeneboom, P. and Wellner, J.A. (2001). Computing Chernoff's distribution. *J. Comput. Graph. Statist.* **10** 388–400. MR1939706 https://doi.org/10.1198/10618600152627997

Janson, S. (2016). Large deviation inequalities for sums of indicator variables. Available at arXiv:1609.00533.

Jewell, N.P. and van der Laan, M. (2004). Current status data: Review, recent developments and open problems. In *Advances in Survival Analysis*. *Handbook of Statist.* **23** 625–642. Amsterdam: Elsevier. MR2065792 https://doi.org/10.1016/S0169-7161(03)23035-2

Kim, Y. (2006). The Bernstein-von Mises theorem for the proportional hazard model. *Ann. Statist.* **34** 1678–1700. MR2283713 https://doi.org/10.1214/009053606000000533

Kim, Y. (2009). A Bernstein-von Mises theorem for doubly censored data. *Statist. Sinica* **19** 581–595. MR2514176

Kim, Y. and Lee, J. (2001). On posterior consistency of survival models. *Ann. Statist.* **29** 666–686. MR1865336 https://doi.org/10.1214/aos/1009210685

Kim, Y. and Lee, J. (2003a). Bayesian analysis of proportional hazard models. *Ann. Statist.* **31** 493–511. MR1983539 https://doi.org/10.1214/aos/1051027878

Kim, Y. and Lee, J. (2003b). Bayesian bootstrap for proportional hazards models. *Ann. Statist.* **31** 1905–1922. MR2036394 https://doi.org/10.1214/aos/1074290331

Kim, Y. and Lee, J. (2004). A Bernstein-von Mises theorem in the nonparametric right-censoring model. *Ann. Statist.* **32** 1492–1512. MR2089131 https://doi.org/10.1214/009053604000000526

Kleijn, B. and Knapik, B. (2012). Semiparametric posterior limits under local asymptotic exponentiality. Available at arXiv:1210.6204.

Lindsey, J.C. and Ryan, L.M. (1998). Methods for interval-censored data. *Stat. Med.* **17** 219–238.

Panov, M. and Spokoiny, V. (2015). Finite sample Bernstein–von Mises theorem for semiparametric problems. *Bayesian Anal.* **10** 665–710. MR3420819 https://doi.org/10.1214/14-BA926

Rabinowitz, D., Tsiatis, A. and Aragon, J. (1995). Regression with interval-censored data. *Biometrika* **82** 501–513. MR1366277 https://doi.org/10.1093/biomet/82.3.501

Shen, X. (2000). Linear regression with current status data. *J. Amer. Statist. Assoc.* **95** 842–852. MR1804443 https://doi.org/10.2307/2669468

Sinha, D., Chen, M.-H. and Ghosh, S.K. (1999). Bayesian analysis and model selection for interval-censored survival data. *Biometrics* **55** 585–590. MR1705161 https://doi.org/10.1111/j.0006-341X.1999.00585.x

Tang, R., Banerjee, M. and Kosorok, M.R. (2012). Likelihood based inference for current status data on a grid: A boundary phenomenon and an adaptive inference procedure. *Ann. Statist.* **40** 45–72. MR3013179 https://doi.org/10.1214/11-AOS942

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

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

Wainwright, M.J. (2019). *High-Dimensional Statistics: A Non-asymptotic Viewpoint*. *Cambridge Series in Statistical and Probabilistic Mathematics* **48**. Cambridge: Cambridge Univ. Press. MR3967104 https://doi.org/10.1017/9781108627771

Wong, M., Lam, K. and Lo, E. (2005). Bayesian analysis of clustered interval-censored data. *J. Dent. Res.* **84** 817–821.

Bernoulli **29**(1), 2023, 428–450
https://doi.org/10.3150/22-BEJ1463

Bootstrapping the operator norm in high dimensions: Error estimation for covariance matrices and sketching

MILES E. LOPES[1,a], N. BENJAMIN ERICHSON[2,b] and MICHAEL W. MAHONEY[2,c]

[1] *Department of Statistics, University of California, Davis, Mathematical Sciences Building, Davis, CA, 95616, USA,* [a] *melopes@ucdavis.edu*
[2] *Department of Statistics, University of California, Berkeley, Evans Hall, Berkeley, CA, 94720, USA and International Computer Science Institute, Berkeley, CA, 94704, USA,* [b] *erichson@berkeley.edu,* [c] *mmahoney@stat.berkeley.edu*

Although the operator (spectral) norm is one of the most widely used metrics for covariance estimation, comparatively little is known about the fluctuations of error in this norm. To be specific, let $\widehat{\Sigma}$ denote the sample covariance matrix of n i.i.d. observations in $\mathbb{R}^p$ that arise from a population matrix Σ, and let $T_n = \sqrt{n}\|\widehat{\Sigma} - \Sigma\|_{\mathrm{op}}$. In the setting where the eigenvalues of Σ have a decay profile of the form $\lambda_j(\Sigma) \asymp j^{-2\beta}$, we analyze how well the bootstrap can approximate the distribution of T_n. Our main result shows that up to factors of $\log(n)$, the bootstrap can approximate the distribution of T_n with respect to the Kolmogorov metric at the rate of $n^{-\frac{\beta-1/2}{6\beta+4}}$, which does not depend on the ambient dimension p. In addition, we offer a supporting result of independent interest that establishes a high-probability upper bound for T_n based on flexible moment assumptions. More generally, we discuss the consequences of our work beyond covariance matrices, and show how the bootstrap can be used to estimate the errors of sketching algorithms in randomized numerical linear algebra (RandNLA). An illustration of these ideas is also provided with a climate data example.

Keywords: Bootstrap; error estimation; high-dimensional statistics; covariance estimation; randomized numerical linear algebra; sketching

References

Adamczak, R., Litvak, A.E., Pajor, A. and Tomczak-Jaegermann, N. (2011). Sharp bounds on the rate of convergence of the empirical covariance matrix. *C. R. Math. Acad. Sci. Paris* **349** 195–200. MR2769907 https://doi.org/10.1016/j.crma.2010.12.014

Ainsworth, M. and Oden, J.T. (2000). *A Posteriori Error Estimation in Finite Element Analysis*. *Pure and Applied Mathematics (New York)*. New York: Wiley Interscience. MR1885308 https://doi.org/10.1002/9781118032824

Ar, S., Blum, M., Codenotti, B. and Gemmell, P. (1993). Checking approximate computations over the reals. In *Annual ACM Symposium on Theory of Computing (STOC)*.

Babuška, I. and Rheinboldt, W.C. (1978). Error estimates for adaptive finite element computations. *SIAM J. Numer. Anal.* **15** 736–754. MR0483395 https://doi.org/10.1137/0715049

Bai, Z., Erichson, N.B., Meena, M.G., Taira, K. and Brunton, S.L. (2019). Randomized methods to characterize large-scale vortical flow networks. *PLoS ONE* **14** e0225265. https://doi.org/10.1371/journal.pone.0225265

Bertail, P. (1997). Second-order properties of an extrapolated bootstrap without replacement under weak assumptions. *Bernoulli* **3** 149–179. MR1466305 https://doi.org/10.2307/3318585

Bertail, P. and Politis, D.N. (2001). Extrapolation of subsampling distribution estimators: The i.i.d. and strong mixing cases. *Canad. J. Statist.* **29** 667–680. MR1888512 https://doi.org/10.2307/3316014

Bickel, P.J. and Levina, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227. MR2387969 https://doi.org/10.1214/009053607000000758

Bickel, P.J. and Sakov, A. (2002). Extrapolation and the bootstrap. *Sankhyā: The Indian Journal of Statistics, Series A* 640–652.

Bickel, P.J. and Yahav, J.A. (1988). Richardson extrapolation and the bootstrap. *J. Amer. Statist. Assoc.* **83** 387–393. MR0971364

Brunton, S.L., Proctor, J.L., Tu, J.H. and Kutz, J.N. (2015). Compressed sensing and dynamic mode decomposition. *J. Comput. Dyn.* **2** 165–191. MR3507892 https://doi.org/10.3934/jcd.2015002

Bunea, F. and Xiao, L. (2015). On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA. *Bernoulli* **21** 1200–1230. MR3338661 https://doi.org/10.3150/14-BEJ602

Cai, T.T., Zhang, C.-H. and Zhou, H.H. (2010). Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.* **38** 2118–2144. MR2676885 https://doi.org/10.1214/09-AOS752

Cangiani, A., Georgoulis, E.H., Pryer, T. and Sutton, O.J. (2017). A posteriori error estimates for the virtual element method. *Numer. Math.* **137** 857–893. MR3719046 https://doi.org/10.1007/s00211-017-0891-9

Chen, R.Y., Gittens, A. and Tropp, J.A. (2012). The masked sample covariance estimator: An analysis using matrix concentration inequalities. *Inf. Inference* **1** 2–20. MR3311439 https://doi.org/10.1093/imaiai/ias001

Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. MR3161448 https://doi.org/10.1214/13-AOS1161

Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42** 1564–1597. MR3262461 https://doi.org/10.1214/14-AOS1230

Chernozhukov, V., Chetverikov, D. and Kato, K. (2016). Empirical and multiplier bootstraps for suprema of empirical processes of increasing complexity, and related Gaussian couplings. *Stochastic Process. Appl.* **126** 3632–3651. MR3565470 https://doi.org/10.1016/j.spa.2016.04.009

Chernozhukov, V., Chetverikov, D. and Kato, K. (2017). Central limit theorems and bootstrap in high dimensions. *Ann. Probab.* **45** 2309–2352. MR3693963 https://doi.org/10.1214/16-AOP1113

Cohen, M.B., Nelson, J. and Woodruff, D.P. (2016). Optimal approximate matrix product in terms of stable rank. In *43rd International Colloquium on Automata, Languages, and Programming*. *LIPIcs. Leibniz Int. Proc. Inform.* **55** Art. No. 11. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR3577072

Dixon, J.D. (1983). Estimating extremal eigenvalues and condition numbers of matrices. *SIAM J. Numer. Anal.* **20** 812–814. MR0708458 https://doi.org/10.1137/0720053

Drineas, P. and Kannan, R. (2001). Fast Monte-Carlo algorithms for approximate matrix multiplication. In *42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001)* 452–459. Los Alamitos, CA: IEEE Computer Soc. MR1948734

Drineas, P., Kannan, R. and Mahoney, M.W. (2006). Fast Monte Carlo algorithms for matrices. I. Approximating matrix multiplication. *SIAM J. Comput.* **36** 132–157. MR2231643 https://doi.org/10.1137/S0097539704442684

Drineas, P. and Mahoney, M.W. (2018). Lectures on randomized numerical linear algebra. In *The Mathematics of Data. IAS/Park City Math. Ser.* **25** 1–48. Providence, RI: Amer. Math. Soc. MR3839166

El Karoui, N. and Purdom, E. (2019). The non-parametric bootstrap and spectral analysis in moderate and high-dimension. In *AISTATS 2019. Proceedings of Machine Learning Research* **89** 2115–2124.

Erichson, N.B., Mathelin, L., Kutz, J.N. and Brunton, S.L. (2019). Randomized dynamic mode decomposition. *SIAM J. Appl. Dyn. Syst.* **18** 1867–1891. MR4021266 https://doi.org/10.1137/18M1215013

Erichson, N.B., Zheng, P., Manohar, K., Brunton, S.L., Kutz, J.N. and Aravkin, A.Y. (2020). Sparse principal component analysis via variable projection. *SIAM J. Appl. Math.* **80** 977–1002. MR4091178 https://doi.org/10.1137/18M1211350

Freivalds, R. (1979). Fast probabilistic algorithms. In *Mathematical Foundations of Computer Science, 1979 (Proc. Eighth Sympos., Olomouc, 1979). Lecture Notes in Computer Science* **74** 57–69. Berlin: Springer. MR0570975

Gupta, V., Wang, S., Courtade, T. and Ramchandran, K. (2018). Oversketch: Approximate matrix multiplication for the cloud. In *2018 IEEE International Conference on Big Data (Big Data)* 298–304. IEEE.

Halko, N., Martinsson, P.G. and Tropp, J.A. (2011). Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.* **53** 217–288. MR2806637 https://doi.org/10.1137/090771806

Hall, P., Lee, Y.K., Park, B.U. and Paul, D. (2009). Tie-respecting bootstrap methods for estimating distributions of sets and functions of eigenvalues. *Bernoulli* **15** 380–401. MR2543867 https://doi.org/10.3150/08-BEJ154

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 https://doi.org/10.3150/16-BEJ912

Holodnak, J.T. and Ipsen, I.C.F. (2015). Randomized approximation of the Gram matrix: Exact computation and probabilistic bounds. *SIAM J. Matrix Anal. Appl.* **36** 110–137. MR3306014 https://doi.org/10.1137/130940116

Hsu, D., Kakade, S.M. and Zhang, T. (2012). Tail inequalities for sums of random matrices that depend on the intrinsic dimension. *Electron. Commun. Probab.* **17** no. 14. MR2900355 https://doi.org/10.1214/ECP.v17-1869

Johnstone, I.M. and Paul, D. (2018). PCA in high dimensions: An orientation. *Proc. IEEE Inst. Electr. Electron. Eng.* **106** 1277–1292. https://doi.org/10.1109/JPROC.2018.2846730

Kannan, R. and Vempala, S. (2017). Randomized algorithms in numerical linear algebra. *Acta Numer.* **26** 95–135. MR3653851 https://doi.org/10.1017/S0962492917000058

Koltchinskii, V., Löffler, M. and Nickl, R. (2020). Efficient estimation of linear functionals of principal components. *Ann. Statist.* **48** 464–490. MR4065170 https://doi.org/10.1214/19-AOS1816

Koltchinskii, V. and Lounici, K. (2017a). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23** 110–133. MR3556768 https://doi.org/10.3150/15-BEJ730

Koltchinskii, V. and Lounici, K. (2017b). Normal approximation and concentration of spectral projectors of sample covariance. *Ann. Statist.* **45** 121–157. MR3611488 https://doi.org/10.1214/16-AOS1437

Koltchinskii, V. and Lounici, K. (2017c). New asymptotic results in principal component analysis. *Sankhya A* **79** 254–297. MR3707422 https://doi.org/10.1007/s13171-017-0106-6

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 https://doi.org/10.1073/pnas.0709640104

Lopes, M.E. (2019). Estimating the algorithmic variance of randomized ensembles via the bootstrap. *Ann. Statist.* **47** 1088–1112. MR3909961 https://doi.org/10.1214/18-AOS1707

Lopes, M.E., Blandino, A. and Aue, A. (2019). Bootstrapping spectral statistics in high dimensions. *Biometrika* **106** 781–801. MR4031199 https://doi.org/10.1093/biomet/asz040

Lopes, M.E., Erichson, N.B. and Mahoney, M.W. (2019). Bootstrapping the operator norm in high dimensions: Error estimation for covariance matrices and sketching. Available at arXiv:1909.06120.

Lopes, M.E., Erichson, N.B. and Mahoney, M.W. (2023). Supplement to "Bootstrapping the Operator Norm in High Dimensions: Error Estimation for Covariance Matrices and Sketching." https://doi.org/10.3150/22-BEJ1463SUPP

Lopes, M.E., Lin, Z. and Müller, H.-G. (2020). Bootstrapping max statistics in high dimensions: Near-parametric rates under weak variance decay and application to functional and multinomial data. *Ann. Statist.* **48** 1214–1229. MR4102694 https://doi.org/10.1214/19-AOS1844

Lopes, M.E., Wang, S. and Mahoney, M.W. (2018). Error estimation for randomized least-squares algorithms via the bootstrap. In *International Conference on Machine Learning (ICML)* 3223–3232.

Lopes, M.E., Wang, S. and Mahoney, M.W. (2019). A bootstrap method for error estimation in randomized matrix multiplication. *J. Mach. Learn. Res.* **20** Paper No. 39. MR3948079

Lounici, K. (2014). High-dimensional covariance matrix estimation with missing observations. *Bernoulli* **20** 1029–1058. MR3217437 https://doi.org/10.3150/12-BEJ487

Lust-Piquard, F. (1986). Inégalités de Khintchine dans C_p $(1 < p < \infty)$. *C. R. Acad. Sci. Paris Sér. I Math.* **303** 289–292. MR0859804

Mahoney, M.W. (2011). Randomized algorithms for matrices and data. *Found. Trends Mach. Learn.* **3** 123–224.

Minsker, S. (2017). On some extensions of Bernstein's inequality for self-adjoint operators. *Statist. Probab. Lett.* **127** 111–119. MR3648301 https://doi.org/10.1016/j.spl.2017.03.020

Naumov, A., Spokoiny, V. and Ulyanov, V. (2019). Bootstrap confidence sets for spectral projectors of sample covariance. *Probab. Theory Related Fields* **174** 1091–1132. MR3980312 https://doi.org/10.1007/s00440-018-0877-2

NOAA (2019). National Oceanic and Atmospheric Administration.

Oliveira, R.I. (2010). Sums of random Hermitian matrices and an inequality by Rudelson. *Electron. Commun. Probab.* **15** 203–212. MR2653725 https://doi.org/10.1214/ECP.v15-1544

Pagh, R. (2013). Compressed matrix multiplication. *ACM Trans. Comput. Theory* **5** Art. 9. MR3100651 https://doi.org/10.1145/2493252.2493254

Pisier, G. (2016). *Martingales in Banach Spaces*. *Cambridge Studies in Advanced Mathematics* **155**. Cambridge: Cambridge Univ. Press. MR3617459

Reiss, M. and Wahl, M. (2020). Nonasymptotic upper bounds for the reconstruction error of PCA. *Ann. Statist.* **48** 1098–1123. MR4102689 https://doi.org/10.1214/19-AOS1839

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C. and Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *J. Climate* **15** 1609–1625.

Rudelson, M. (1999). Random vectors in the isotropic position. *J. Funct. Anal.* **164** 60–72. MR1694526 https://doi.org/10.1006/jfan.1998.3384

Rudelson, M. and Vershynin, R. (2007). Sampling from large matrices: An approach through geometric functional analysis. *J. ACM* **54** Art. 21. MR2351844 https://doi.org/10.1145/1255443.1255449

Saibaba, A.K. (2019). Randomized subspace iteration: Analysis of canonical angles and unitarily invariant norms. *SIAM J. Matrix Anal. Appl.* **40** 23–48. MR3896637 https://doi.org/10.1137/18M1179432

Sarlós, T. (2006). Improved approximation algorithms for large matrices via random projections. In *Annual IEEE Symposium on Foundations of Computer Science (FOCS)*.

Silin, I. and Fan, J. (2020). Hypothesis testing for eigenspaces of covariance matrix. Available at arXiv:2002.09810.

Talagrand, M. (2014). *Upper and Lower Bounds for Stochastic Processes: Modern Methods and Classical Problems*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. a Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. a Series of Modern Surveys in Mathematics]* **60**. Heidelberg: Springer. MR3184689 https://doi.org/10.1007/978-3-642-54075-2

Tropp, J.A., Yurtsever, A., Udell, M. and Cevher, V. (2019). Streaming low-rank matrix approximation with an application to scientific simulation. *SIAM J. Sci. Comput.* **41** A2430–A2463. MR3986561 https://doi.org/10.1137/18M1201068

van der Vaart, A.W. and Wellner, J.A. (2000). *Weak Convergence and Empirical Processes*. Berlin: Springer.

Verfürth, R. (1994). A posteriori error estimation and adaptive mesh-refinement techniques. *J. Comput. Appl. Math.* **50** 67–83.

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

Woodruff, D.P. (2014). Sketching as a tool for numerical linear algebra. *Found. Trends Theor. Comput. Sci.* **10** iv+157. MR3285427 https://doi.org/10.1561/0400000060

Woolfe, F., Liberty, E., Rokhlin, V. and Tygert, M. (2008). A fast randomized algorithm for the approximation of matrices. *Appl. Comput. Harmon. Anal.* **25** 335–366. MR2455599 https://doi.org/10.1016/j.acha.2007.12.002

Yao, J. and Lopes, M.E. (2022). Rates of bootstrap approximation for eigenvalues in high-dimensional PCA. *Statist. Sinica*. To appear. Available at arXiv:2104.07328.

Bernoulli **29**(1), 2023, 451–472
https://doi.org/10.3150/22-BEJ1464

On consistency and sparsity for high-dimensional functional time series with application to autoregressions

SHAOJUN GUO[1,a] and XINGHAO QIAO[2,b]

[1]*Institute of Statistics and Big Data, Renmin University of China, Beijing, 100872, P.R. China.* [a]*sjguo@ruc.edu.cn*
[2]*Department of Statistics, London School of Economics, London, WC2A 2AE, U.K..* [b]*x.qiao@lse.ac.uk*

Modelling a large collection of functional time series arises in a broad spectral of real applications. Under such a scenario, not only the number of functional variables can be diverging with, or even larger than the number of temporally dependent functional observations, but each function itself is an infinite-dimensional object, posing a challenging task. In this paper, we propose a three-step procedure to estimate high-dimensional functional time series models. To provide theoretical guarantees for the three-step procedure, we focus on multivariate stationary processes and propose a novel functional stability measure based on their spectral properties. Such stability measure facilitates the development of some useful concentration bounds on sample (auto)covariance functions, which serve as a fundamental tool for further convergence analysis in high-dimensional settings. As functional principal component analysis (FPCA) is one of the key dimension reduction techniques in the first step, we also investigate the non-asymptotic properties of the relevant estimated terms under a FPCA framework. To illustrate with an important application, we consider vector functional autoregressive models and develop a regularization approach to estimate autoregressive coefficient functions under the sparsity constraint. Using our derived non-asymptotic results, we investigate convergence properties of the regularized estimate under high-dimensional scaling. Finally, the finite-sample performance of the proposed method is examined through both simulations and a public financial dataset.

Keywords: Functional principal component analysis; functional stability measure; high-dimensional functional time series; non-asymptotics; sparsity; vector functional autoregression

References

Araya Valdivia, E. (2020). Relative concentration bounds for the spectrum of kernel matrices. Preprint. Available at arXiv:1812.02108v8.

Aue, A., Norinho, D.D. and Hörmann, S. (2015). On the prediction of stationary functional time series. *J. Amer. Statist. Assoc.* **110** 378–392. MR3338510 https://doi.org/10.1080/01621459.2014.909317

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

Basu, S., Shojaie, A. and Michailidis, G. (2015). Network Granger causality with inherent grouping structure. *J. Mach. Learn. Res.* **16** 417–453. MR3335801

Bathia, N., Yao, Q. and Ziegelmann, F. (2010). Identifying the finite dimensionality of curve time series. *Ann. Statist.* **38** 3352–3386. MR2766855 https://doi.org/10.1214/10-AOS819

Beck, A. and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2** 183–202. MR2486527 https://doi.org/10.1137/080716542

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

Billio, M., Casarin, R. and Rossini, L. (2019). Bayesian nonparametric sparse VAR models. *J. Econometrics* **212** 97–115. MR3994009 https://doi.org/10.1016/j.jeconom.2019.04.022

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

Cho, H., Goude, Y., Brossat, X. and Yao, Q. (2013). Modeling and forecasting daily electricity load curves: A hybrid approach. *J. Amer. Statist. Assoc.* **108** 7–21. MR3174599 https://doi.org/10.1080/01621459.2012.722900

Fan, Y., James, G.M. and Radchenko, P. (2015). Functional additive regression. *Ann. Statist.* **43** 2296–2325. MR3396986 https://doi.org/10.1214/15-AOS1346

Ghosh, S., Khare, K. and Michailidis, G. (2019). High-dimensional posterior consistency in Bayesian vector autoregressive models. *J. Amer. Statist. Assoc.* **114** 735–748. MR3963176 https://doi.org/10.1080/01621459.2018.1437043

Guo, S., Qiao, X. (2023). Supplement to "On consistency and sparsity for high-dimensional functional time series with application to autoregressions." https://doi.org/10.3150/22-BEJ1464SUPP

Guo, S., Wang, Y. and Yao, Q. (2016). High-dimensional and banded vector autoregressions. *Biometrika* **103** 889–903. MR3620446 https://doi.org/10.1093/biomet/asw046

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

Han, F., Lu, H. and Liu, H. (2015). A direct estimation of high dimensional stationary vector autoregressions. *J. Mach. Learn. Res.* **16** 3115–3150. MR3450535

Hörmann, S., Kidziński, Ł. and Hallin, M. (2015). Dynamic functional principal components. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 319–348. MR3310529 https://doi.org/10.1111/rssb.12076

Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. MR2662361 https://doi.org/10.1214/09-AOS768

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

Jirak, M. (2016). Optimal eigen expansions and uniform bounds. *Probab. Theory Related Fields* **166** 753–799. MR3568039 https://doi.org/10.1007/s00440-015-0671-3

Kock, A.B. and Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *J. Econometrics* **186** 325–344. MR3343790 https://doi.org/10.1016/j.jeconom.2015.02.013

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

Koltchinskii, V. and Lounici, K. (2017). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23** 110–133. MR3556768 https://doi.org/10.3150/15-BEJ730

Kong, D., Xue, K., Yao, F. and Zhang, H.H. (2016). Partially functional linear regression in high dimensions. *Biometrika* **103** 147–159. MR3465827 https://doi.org/10.1093/biomet/asv062

Li, D., Robinson, P.M. and Shang, H.L. (2020). Long-range dependent curve time series. *J. Amer. Statist. Assoc.* **115** 957–971. MR4107692 https://doi.org/10.1080/01621459.2019.1604362

Li, B. and Solea, E. (2018). A nonparametric graphical model for functional data with application to brain networks based on fMRI. *J. Amer. Statist. Assoc.* **113** 1637–1655. MR3902235 https://doi.org/10.1080/01621459.2017.1356726

Li, R., Cui, L., Li, J., Zhao, A., Fu, H., Wu, Y., Zhang, L., Kong, L. and Chen, J. (2017). Spatial and temporal variation of particulate matter and gaseous pollutants in China during 2014-2016. *Atmos. Environ.* **161** 235–246.

Loh, P.-L. and Wainwright, M.J. (2012). High-dimensional regression with noisy and missing data: Provable guarantees with nonconvexity. *Ann. Statist.* **40** 1637–1664. MR3015038 https://doi.org/10.1214/12-AOS1018

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

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

O'Donoghue, B. and Candès, E. (2015). Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.* **15** 715–732. MR3348171 https://doi.org/10.1007/s10208-013-9150-3

Panaretos, V.M. and Tavakoli, S. (2013). Fourier analysis of stationary time series in function space. *Ann. Statist.* **41** 568–603. MR3099114 https://doi.org/10.1214/13-AOS1086

Qiao, X., Guo, S. and James, G.M. (2019). Functional graphical models. *J. Amer. Statist. Assoc.* **114** 211–222. MR3941249 https://doi.org/10.1080/01621459.2017.1390466

Qiao, X., Qian, C., James, G.M. and Guo, S. (2020). Doubly functional graphical models in high dimensions. *Biometrika* **107** 415–431. MR4108937 https://doi.org/10.1093/biomet/asz072

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

Simon, N. and Tibshirani, R. (2012). Standardization and the group Lasso penalty. *Statist. Sinica* **22** 983–1001. MR2987480 https://doi.org/10.5705/ss.2011.075

Sun, Y., Li, Y., Kuceyeski, A. and Basu, S. (2018). Large spectral density matrix estimation by thresholding. Available at arXiv:1812.00532.

Voorman, A., Shojaie, A. and Witten, D. (2014). Graph estimation with joint additive models. *Biometrika* **101** 85–101. MR3180659 https://doi.org/10.1093/biomet/ast053

Wong, K.C., Li, Z. and Tewari, A. (2020). Lasso guarantees for β-mixing heavy-tailed time series. *Ann. Statist.* **48** 1124–1142. MR4102690 https://doi.org/10.1214/19-AOS1840

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 49–67. MR2212574 https://doi.org/10.1111/j.1467-9868.2005.00532.x

Zhang, L., Mykland, P.A. and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100** 1394–1411. MR2236450 https://doi.org/10.1198/016214505000000169

Bernoulli **29**(1), 2023, 473–495
https://doi.org/10.3150/22-BEJ1465

Suboptimality of constrained least squares and improvements via non-linear predictors

TOMAS VAŠKEVIČIUS[1,a] and NIKITA ZHIVOTOVSKIY[2,b]

[1]*Department of Statistics, University of Oxford, Oxford, OX1 3LB, United Kingdom,*
[a]*tomas.vaskevicius@stats.ox.ac.uk*
[2]*Department of Mathematics, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland,*
[b]*nikita.zhivotovskii@math.ethz.ch*

We study the problem of predicting as well as the best linear predictor in a bounded Euclidean ball with respect to the squared loss. When only boundedness of the data generating distribution is assumed, we establish that the least squares estimator constrained to a bounded Euclidean ball does not attain the classical $O(d/n)$ excess risk rate, where d is the dimension of the covariates and n is the number of samples. In particular, we construct a bounded distribution such that the constrained least squares estimator incurs an excess risk of order $\Omega(d^{3/2}/n)$ hence refuting a recent conjecture of Ohad Shamir [JMLR 2015]. In contrast, we observe that non-linear predictors can achieve the optimal rate $O(d/n)$ with no assumptions on the distribution of the covariates. We discuss additional distributional assumptions sufficient to guarantee an $O(d/n)$ excess risk rate for the least squares estimator. Among them are certain moment equivalence assumptions often used in the robust statistics literature. While such assumptions are central in the analysis of unbounded and heavy-tailed settings, our work indicates that in some cases, they also rule out unfavorable bounded distributions.

Keywords: Constrained least squares; ridge regression; empirical processes; average stability; Vovk-Azoury-Warmuth forecaster

References

Audibert, J.-Y. (2009). Fast learning rates in statistical inference through aggregation. *Ann. Statist.* **37** 1591–1646. MR2533466 https://doi.org/10.1214/08-AOS623

Audibert, J.-Y. and Catoni, O. (2010). Linear regression through PAC-Bayesian truncation. Preprint. Available at arXiv:1010.0072.

Audibert, J.-Y. and Catoni, O. (2011). Robust linear least squares regression. *Ann. Statist.* **39** 2766–2794. MR2906886 https://doi.org/10.1214/11-AOS918

Azoury, K.S. and Warmuth, M.K. (2001). Relative loss bounds for on-line density estimation with the exponential family of distributions. *Mach. Learn.* **43** 211–246.

Bach, F.R. and Moulines, E. (2013). Non-strongly-convex smooth stochastic approximation with convergence rate $O(1/n)$. In *Advances in Neural Information Processing Systems* **26**.

Balázs, G., György, A. and Szepesvári, C. (2016). Chaining bounds for empirical risk minimization. Preprint. Available at arXiv:1609.01872.

Bartlett, P.L., Bousquet, O. and Mendelson, S. (2005). Local Rademacher complexities. *Ann. Statist.* **33** 1497–1537. MR2166554 https://doi.org/10.1214/009053605000000282

Bartlett, P.L. and Mendelson, S. (2006). Empirical minimization. *Probab. Theory Related Fields* **135** 311–334. MR2240689 https://doi.org/10.1007/s00440-005-0462-3

Bellec, P.C. (2018). Sharp oracle inequalities for least squares estimators in shape restricted regression. *Ann. Statist.* **46** 745–780. MR3782383 https://doi.org/10.1214/17-AOS1566

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

Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probab. Theory Related Fields* **97** 113–150. MR1240719 https://doi.org/10.1007/BF01199316

Bousquet, O., Hanneke, S., Moran, S. and Zhivotovskiy, N. (2020). Proper learning, Helly number, and an optimal SVM bound. *Proc. Mach. Learn. Res.* **125** 582–609.

Breiman, L. and Freedman, D. (1983). How many variables should be entered in a regression equation? *J. Amer. Statist. Assoc.* **78** 131–136. MR0696857

Catoni, O. (2016). PAC-Bayesian bounds for the Gram matrix and least squares regression with a random design. Preprint. Available at arXiv:1603.05229.

Cesa-Bianchi, N. and Lugosi, G. (2006). *Prediction, Learning, and Games*. Cambridge: Cambridge Univ. Press. MR2409394 https://doi.org/10.1017/CBO9780511546921

Chatterjee, S. (2014). A new perspective on least squares under convex constraint. *Ann. Statist.* **42** 2340–2381. MR3269982 https://doi.org/10.1214/14-AOS1254

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

Chinot, G., Lecué, G. and Lerasle, M. (2020). Robust statistical learning with Lipschitz and convex loss functions. *Probab. Theory Related Fields* **176** 897–940. MR4087486 https://doi.org/10.1007/s00440-019-00931-3

Donoho, D.L. and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. MR1635414 https://doi.org/10.1214/aos/1024691081

Forster, J. and Warmuth, M.K. (2002). Relative expected instantaneous loss bounds *J. Comput. System Sci.* **64** 76–102. MR1896143 https://doi.org/10.1006/jcss.2001.1798

Foster, D.J., Kale, S., Luo, H., Mohri, M. and Sridharan, K. (2018). Logistic regression: The importance of being improper. *Proc. Mach. Learn. Res.* **75** 167–208.

Gonen, A. and Shalev-Shwartz, S. (2017). Average stability is invariant to data preconditioning. Implications to exp-concave empirical risk minimization. *J. Mach. Learn. Res.* **18** Paper No. 222, 13. MR3827110

Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. *Springer Series in Statistics*. New York: Springer. MR1920390 https://doi.org/10.1007/b97848

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

Hazan, E., Agarwal, A. and Kale, S. (2007). Logarithmic regret algorithms for online convex optimization. *Mach. Learn.* **69** 169–192.

Hazan, E., Koren, T. and Levy, K.Y. (2014). Logistic regression: Tight bounds for stochastic and online optimization. In *Conference on Learning Theory* 197–209.

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

Jézéquel, R., Gaillard, P. and Rudi, A. (2020). Efficient improper learning for online logistic regression. *Proc. Mach. Learn. Res.* **125** 2085–2108.

Juditsky, A., Rigollet, P. and Tsybakov, A.B. (2008). Learning by mirror averaging. *Ann. Statist.* **36** 2183–2206. MR2458184 https://doi.org/10.1214/07-AOS546

Klochkov, Y. and Zhivotovskiy, N. (2020). Uniform Hanson-Wright type concentration inequalities for unbounded entries via the entropy method. *Electron. J. Probab.* **25** Paper No. 22, 30. MR4073683 https://doi.org/10.1214/20-ejp422

Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *Ann. Statist.* **34** 2593–2656. MR2329442 https://doi.org/10.1214/009053606000001019

Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems: École d'Été de Probabilités de Saint-Flour*. *Lecture Notes in Math.* **2033**. Heidelberg: Springer. MR2829871 https://doi.org/10.1007/978-3-642-22147-7

Koren, T. and Levy, K. (2015). Fast rates for exp-concave empirical risk minimization. In *Advances in Neural Information Processing Systems* 1477–1485.

Kur, G., Rakhlin, A. and Guntuboyina, A. (2020). On suboptimality of least squares with application to estimation of convex bodies. In *Conference on Learning Theory* 2406–2424.

Lecué, G. (2013). Empirical risk minimization is optimal for the convex aggregation problem. *Bernoulli* **19** 2153–2166. MR3160549 https://doi.org/10.3150/12-BEJ447

Lecué, G. and Mendelson, S. (2013). Learning subgaussian classes: Upper and minimax bounds. Preprint. Available at arXiv:1305.4825.

Lecué, G. and Mendelson, S. (2016). Performance of empirical risk minimization in linear aggregation. *Bernoulli* **22** 1520–1534. MR3474824 https://doi.org/10.3150/15-BEJ701

Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces: Isoperimetry and Processes*. *Classics in Mathematics*. Berlin: Springer. MR2814399

Liang, T., Rakhlin, A. and Sridharan, K. (2015). Learning with square loss: Localization through offset Rademacher complexity. In *Conference on Learning Theory* 1260–1285.

Lugosi, G. and Mendelson, S. (2020). Risk minimization by median-of-means tournaments. *J. Eur. Math. Soc. (JEMS)* **22** 925–965. MR4055993 https://doi.org/10.4171/jems/937

McMahan, B. and Streeter, M. (2012). Open problem: Better bounds for online logistic regression. In *Conference on Learning Theory*.

Mendelson, S. (2015). Learning without concentration. *J. ACM* **62** Art. 21, 25. MR3367000 https://doi.org/10.1145/2699439

Mendelson, S. (2021). Extending the scope of the small-ball method. *Studia Math.* **256** 147–167. MR4165508 https://doi.org/10.4064/sm190420-21-11

Mendelson, S. and Zhivotovskiy, N. (2020). Robust covariance estimation under L_4-L_2 norm equivalence. *Ann. Statist.* **48** 1648–1664. MR4124338 https://doi.org/10.1214/19-AOS1862

Mourtada, J. (2019). Exact minimax risk for linear least squares, and the lower tail of sample covariance matrices. Preprint. Available at arXiv:1912.10754.

Mourtada, J. and Gaïffas, S. (2019). An improper estimator with optimal excess risk in misspecified density estimation and logistic regression. Preprint. Available at arXiv:1912.10784.

Mourtada, J., Vaškevičius, T. and Zhivotovskiy, N. (2021). Distribution-free robust linear regression. *Math. Stat. Learn.* **4** 253–292. MR4383735 https://doi.org/10.4171/msl/27

Nemirovski, A. (2000). Topics in non-parametric statistics. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1998)*. *Lecture Notes in Math.* **1738** 85–277. Berlin: Springer. MR1775640

Oliveira, R.I. (2010). Sums of random Hermitian matrices and an inequality by Rudelson. *Electron. Commun. Probab.* **15** 203–212. MR2653725 https://doi.org/10.1214/ECP.v15-1544

Oliveira, R.I. (2016). The lower tail of random quadratic forms with applications to ordinary least squares. *Probab. Theory Related Fields* **166** 1175–1194. MR3568047 https://doi.org/10.1007/s00440-016-0738-9

Orabona, F. (2019). A Modern Introduction to Online Learning. Preprint. Available at arXiv:1912.13213v2.

Rakhlin, A., Sridharan, K. and Tsybakov, A.B. (2017). Empirical entropy, minimax regret and minimax risk. *Bernoulli* **23** 789–824. MR3606751 https://doi.org/10.3150/14-BEJ679

Rudelson, M. (1999). Random vectors in the isotropic position. *J. Funct. Anal.* **164** 60–72. MR1694526 https://doi.org/10.1006/jfan.1998.3384

Saumard, A. (2018). On optimality of empirical risk minimization in linear aggregation. *Bernoulli* **24** 2176–2203. MR3757527 https://doi.org/10.3150/17-BEJ925

Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge Univ. Press.

Shamir, O. (2015). The sample complexity of learning linear predictors with the squared loss. *J. Mach. Learn. Res.* **16** 3475–3486. MR3450546

Tropp, J. A. (2015). An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.* **8** 1-230.

Tsybakov, A.B. (2003). Optimal rates of aggregation. In *Conference on Learning Theory* 303–313.

Vaškevičius, T. and Zhivotovskiy, N. (2023). Supplement to "Suboptimality of constrained least squares and improvements via non-linear predictors." https://doi.org/10.3150/22-BEJ1465SUPP

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

Vovk, V. (1998). Competitive on-line linear regression. In *Advances in Neural Information Processing Systems* 364–370.

Vovk, V. (2001). Competitive on-line statistics. *Int. Stat. Rev.* **69** 213–248.

Wainwright, M.J. (2019). *High-Dimensional Statistics: A Non-asymptotic Viewpoint*. *Cambridge Series in Statistical and Probabilistic Mathematics* **48**. Cambridge: Cambridge Univ. Press. MR3967104 https://doi.org/10.1017/9781108627771

Wang, Y.-X. (2018). Revisiting differentially private linear regression: Optimal and adaptive prediction and estimation in unbounded domain. In *Uncertainty in Artificial Intelligence*.

Bernoulli **29**(1), 2023, 496–522
https://doi.org/10.3150/22-BEJ1466

Quadratic variation and quadratic roughness

RAMA CONT[a] and PURBA DAS[b]

Mathematical Institute, University of Oxford, Oxford, England, [a]*Rama.Cont@maths.ox.ac.uk,* [b]*Purba.Das@maths.ox.ac.uk*

We study the concept of quadratic variation of a continuous path along a sequence of partitions and its dependence with respect to the choice of the partition sequence. We introduce the concept of *quadratic roughness* of a path along a partition sequence and show that for Hölder-continuous paths satisfying this roughness condition, the quadratic variation along balanced partitions is invariant with respect to the choice of the partition sequence. Typical paths of Brownian motion are shown to satisfy this quadratic roughness property almost-surely along any partition with a required step size condition. Using these results we derive a formulation of the pathwise Föllmer-Itô calculus which is invariant with respect to the partition sequence. We also derive an invarience of local time under quadratic roughness.

Keywords: Quadratic variation; pathwise integration; Brownian motion; Itô calculus; local time; roughness

References

[1] Ananova, A. (2018). Pathwise integration and functional calculus for paths with finite quadratic variation. PhD thesis, Imperial College London.

[2] Ananova, A. and Cont, R. (2017). Pathwise integration with respect to paths of finite quadratic variation. *J. Math. Pures Appl. (9)* **107** 737–757. MR3650323 https://doi.org/10.1016/j.matpur.2016.10.004

[3] Bertoin, J. (1987). Temps locaux et intégration stochastique pour les processus de Dirichlet. In *Séminaire de Probabilités, XXI*. *Lecture Notes in Math.* **1247** 191–205. Berlin: Springer. MR0941983 https://doi.org/10.1007/BFb0077634

[4] Catellier, R. and Gubinelli, M. (2016). Averaging along irregular curves and regularisation of ODEs. *Stochastic Process. Appl.* **126** 2323–2366. MR3505229 https://doi.org/10.1016/j.spa.2016.02.002

[5] Chiu, H. and Cont, R. (2018). On pathwise quadratic variation for càdlàg functions. *Electron. Commun. Probab.* **23** Paper No. 85. MR3882226 https://doi.org/10.1214/18-ECP186

[6] Chiu, H. and Cont, R. (2021). Causal functional calculus. Available at arXiv:1912.07951.

[7] Cont, R. and Fournié, D.-A. (2010). Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.* **259** 1043–1072. MR2652181 https://doi.org/10.1016/j.jfa.2010.04.017

[8] Cont, R. and Fournié, D.A. (2016). Functional Itô calculus and functional Kolmogorov equations. In *Stochastic Integration by Parts and Functional Itô Calculus*. *Adv. Courses Math. CRM Barcelona* 115–207. Cham: Birkhäuser/Springer. MR3497715

[9] Cont, R. and Perkowski, N. (2019). Pathwise integration and change of variable formulas for continuous paths with arbitrary regularity. *Trans. Amer. Math. Soc.* **6** 161–186. MR3937343 https://doi.org/10.1090/btran/34

[10] Davis, M., Obłój, J. and Raval, V. (2014). Arbitrage bounds for prices of weighted variance swaps. *Math. Finance* **24** 821–854. MR3274933 https://doi.org/10.1111/mafi.12021

[11] Davis, M., Obłój, J. and Siorpaes, P. (2018). Pathwise stochastic calculus with local times. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 1–21. MR3765878 https://doi.org/10.1214/16-AIHP792

[12] Dellacherie, C. and Meyer, P.-A. (1978). *Probabilities and Potential*. *North-Holland Mathematics Studies* **29**. Amsterdam: North-Holland. MR0521810

[13] Dudley, R.M. (1973). Sample functions of the Gaussian process. *Ann. Probab.* **1** 66–103. MR0346884 https://doi.org/10.1214/aop/1176997026

[14] Dudley, R.M. and Norvaiša, R. (2011). *Concrete Functional Calculus*. *Springer Monographs in Mathematics*. New York: Springer. MR2732563 https://doi.org/10.1007/978-1-4419-6950-7

[15] Fernandez de la Vega, W. (1974). On almost sure convergence of quadratic Brownian variation. *Ann. Probab.* **2** 551–552. MR0359029 https://doi.org/10.1214/aop/1176996675

[16] Föllmer, H. (1981). Calcul d'Itô sans probabilités. In *Seminar on Probability, XV (Univ. Strasbourg, Strasbourg, 1979/1980) (French)*. *Lecture Notes in Math.* **850** 143–150. Berlin: Springer. MR0622559

[17] Freedman, D. (1983). *Brownian Motion and Diffusion*, 2nd ed. New York-Berlin: Springer. MR0686607

[18] Friz, P.K. and Hairer, M. (2014). *A Course on Rough Paths*. *Universitext*. Cham: Springer. MR3289027 https://doi.org/10.1007/978-3-319-08332-2

[19] Geman, D. and Horowitz, J. (1976). Local times for real and random functions. *Duke Math. J.* **43** 809–828. MR0418210

[20] Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* **8** 1–67. MR0556414

[21] Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864 https://doi.org/10.1214/aoms/1177693335

[22] Imkeller, P. and Prömel, D.J. (2015). Existence of Lévy's area and pathwise integration. *Commun. Stoch. Anal.* **9** 93–111. MR3601245 https://doi.org/10.31390/cosa.9.1.06

[23] Karandikar, R.L. (1983). On the quadratic variation process of a continuous martingale. *Illinois J. Math.* **27** 178–181. MR0694639

[24] Karandikar, R.L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* **57** 11–18. MR1327950 https://doi.org/10.1016/0304-4149(95)00002-O

[25] Karandikar, R.L. and Rao, B.V. (2014). On quadratic variation of martingales. *Proc. Indian Acad. Sci. Math. Sci.* **124** 457–469. MR3258633 https://doi.org/10.1007/s12044-014-0179-2

[26] Kim, D. (2022). Local time for continuous paths with arbitrary regularity. *J. Theor. Probab.* **35**. https://doi.org/10.1007/s10959-022-01159-z

[27] Lévy, P. (1940). Le mouvement brownien plan. *Amer. J. Math.* **62** 487–550. MR0002734 https://doi.org/10.2307/2371467

[28] Lévy, P. (1965). *Processus Stochastiques et Mouvement Brownien*. Paris: Gauthier-Villars & Cie. MR0190953

[29] Perkowski, N. and Prömel, D.J. (2015). Local times for typical price paths and pathwise Tanaka formulas. *Electron. J. Probab.* **20** no. 46. MR3339866 https://doi.org/10.1214/EJP.v20-3534

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

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

[32] Taylor, S.J. (1972). Exact asymptotic estimates of Brownian path variation. *Duke Math. J.* **39** 219–241. MR0295434

[33] Wuermli, M. (1980). Lokalzeiten für martingale, diploma thesis, Universität Bonn.

Bernoulli **29**(1), 2023, 523–550
https://doi.org/10.3150/22-BEJ1467

Chung-type law of the iterated logarithm and exact moduli of continuity for a class of anisotropic Gaussian random fields

CHEUK YIN LEE[1,a] and YIMIN XIAO[2,b]

[1]*Institut de mathématiques, École Polytechnique Fédérale de Lausanne, Station 8, CH-1015 Lausanne, Switzerland.* [a]*cheuk.lee@epfl.ch*
[2]*Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, United States.* [b]*xiaoy@msu.edu*

We establish a Chung-type law of the iterated logarithm and the exact local and uniform moduli of continuity for a large class of anisotropic Gaussian random fields with a harmonizable-type integral representation and the property of strong local nondeterminism. Compared with the existing results in the literature, our results do not require the assumption of stationary increments and provide more precise upper and lower bounds for the limiting constants. The results are applicable to the solutions of a class of linear stochastic partial differential equations driven by a fractional-colored Gaussian noise, including the stochastic heat equation.

Keywords: Gaussian random fields; harmonizable representation; strong local nondeterminism; law of the iterated logarithm; modulus of continuity; stochastic heat equation

References

[1] Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. *Springer Monographs in Mathematics*. New York: Springer. MR2319516
[2] Anderson, T.W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176. MR0069229 https://doi.org/10.2307/2032333
[3] Balan, R.M. and Tudor, C.A. (2008). The stochastic heat equation with fractional-colored noise: Existence of the solution. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 57–87. MR2413088
[4] Dalang, R.C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** 6. MR1684157 https://doi.org/10.1214/EJP.v4-43
[5] Dalang, R.C., Lee, C.Y., Mueller, C. and Xiao, Y. (2021). Multiple points of Gaussian random fields. *Electron. J. Probab.* **26** 17. MR4235468 https://doi.org/10.1214/21-EJP589
[6] Dalang, R.C., Mueller, C. and Xiao, Y. (2017). Polarity of points for Gaussian random fields. *Ann. Probab.* **45** 4700–4751. MR3737922 https://doi.org/10.1214/17-AOP1176
[7] Herrell, R., Song, R., Wu, D. and Xiao, Y. (2020). Sharp space-time regularity of the solution to stochastic heat equation driven by fractional-colored noise. *Stoch. Anal. Appl.* **38** 747–768. MR4112745 https://doi.org/10.1080/07362994.2020.1721301
[8] Hinojosa-Calleja, A. and Sanz-Solé, M. (2021). Anisotropic Gaussian random fields: Criteria for hitting probabilities and applications. *Stoch. Partial Differ. Equ. Anal. Comput.* **9** 984–1030. MR4333508 https://doi.org/10.1007/s40072-021-00190-1
[9] Ledoux, M. (1996). Isoperimetry and Gaussian analysis. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1994)*. *Lecture Notes in Math.* **1648** 165–294. Berlin: Springer. MR1600888 https://doi.org/10.1007/BFb0095676
[10] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **23**. Berlin: Springer. MR1102015 https://doi.org/10.1007/978-3-642-20212-4

[11] Lee, C.Y. (2021). Local nondeterminism and local times of the stochastic wave equation driven by fractional-colored noise. Available at arXiv:2105.05084.

[12] Lee, C.Y. (2022). The Hausdorff measure of the range and level sets of Gaussian random fields with sectorial local nondeterminism. *Bernoulli* **28** 277–306. MR4337706 https://doi.org/10.3150/21-bej1342

[13] Lee, C.Y. and Xiao, Y. (2019). Local nondeterminism and the exact modulus of continuity for stochastic wave equation. *Electron. Commun. Probab.* **24** 52. MR4003126 https://doi.org/10.1214/19-ecp264

[14] Lee, C.Y. and Xiao, Y. (2023). Supplement to "Chung-type law of the iterated logarithm and exact moduli of continuity for a class of anisotropic Gaussian random fields." https://doi.org/10.3150/22-BEJ1467SUPP

[15] Li, W.V. and Shao, Q.-M. (2001). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes: Theory and Methods*. *Handbook of Statist.* **19** 533–597. Amsterdam: North-Holland. MR1861734 https://doi.org/10.1016/S0169-7161(01)19019-X

[16] Luan, N. and Xiao, Y. (2010). Chung's law of the iterated logarithm for anisotropic Gaussian random fields. *Statist. Probab. Lett.* **80** 1886–1895. MR2734256 https://doi.org/10.1016/j.spl.2010.08.016

[17] Luan, N. and Xiao, Y. (2012). Spectral conditions for strong local nondeterminism and exact Hausdorff measure of ranges of Gaussian random fields. *J. Fourier Anal. Appl.* **18** 118–145. MR2885561 https://doi.org/10.1007/s00041-011-9193-2

[18] Marcus, M.B. and Rosen, J. (2006). *Markov Processes, Gaussian Processes, and Local Times*. *Cambridge Studies in Advanced Mathematics* **100**. Cambridge: Cambridge Univ. Press. MR2250510 https://doi.org/10.1017/CBO9780511617997

[19] Meerschaert, M.M., Wang, W. and Xiao, Y. (2013). Fernique-type inequalities and moduli of continuity for anisotropic Gaussian random fields. *Trans. Amer. Math. Soc.* **365** 1081–1107. MR2995384 https://doi.org/10.1090/S0002-9947-2012-05678-9

[20] Monrad, D. and Rootzén, H. (1995). Small values of Gaussian processes and functional laws of the iterated logarithm. *Probab. Theory Related Fields* **101** 173–192. MR1318191 https://doi.org/10.1007/BF01375823

[21] Mueller, C. and Tribe, R. (2002). Hitting properties of a random string. *Electron. J. Probab.* **7** 10. MR1902843 https://doi.org/10.1214/EJP.v7-109

[22] Stein, E.M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series, No. 30*. Princeton, NJ: Princeton Univ. Press. MR0290095

[23] Talagrand, M. (1994). Sharper bounds for Gaussian and empirical processes. *Ann. Probab.* **22** 28–76. MR1258865

[24] Talagrand, M. (1995). Hausdorff measure of trajectories of multiparameter fractional Brownian motion. *Ann. Probab.* **23** 767–775. MR1334170

[25] Tudor, C.A. and Xiao, Y. (2017). Sample paths of the solution to the fractional-colored stochastic heat equation. *Stoch. Dyn.* **17** 1750004. MR3583991 https://doi.org/10.1142/S0219493717500046

[26] Xiao, Y. (1997). Hölder conditions for the local times and the Hausdorff measure of the level sets of Gaussian random fields. *Probab. Theory Related Fields* **109** 129–157. MR1469923 https://doi.org/10.1007/s004400050128

[27] Xiao, Y. (2008). Strong local nondeterminism and sample path properties of Gaussian random fields. In *Asymptotic Theory in Probability and Statistics with Applications*. *Adv. Lect. Math. (ALM)* **2** 136–176. Somerville, MA: Int. Press. MR2466984

[28] 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 https://doi.org/10.1007/978-3-540-85994-9_5

Bernoulli **29**(1), 2023, 551–579
https://doi.org/10.3150/22-BEJ1468

Multilevel bootstrap particle filter

KARI HEINE[a] and DANIEL BURROWS[b]

Department of Mathematical Sciences, University of Bath, Bath, UK, [a]*k.m.p.heine@bath.ac.uk,* [b]*dwb26@bath.ac.uk*

We consider situations where the applicability of sequential Monte Carlo particle filters is compromised due to the expensive evaluation of the particle weights. To alleviate this problem, we propose a new particle filter algorithm based on the multilevel approach. We show that the resulting multilevel bootstrap particle filter (MLBPF) retains the strong law of large numbers as well as the central limit theorem of classical particle filters under mild conditions. Our numerical experiments demonstrate up to 85% reduction in computation time compared to the classical bootstrap particle filter, in certain settings. While it should be acknowledged that this reduction is highly application dependent, and a similar gain should not be expected for all applications across the board, we believe that this substantial improvement in certain settings makes MLBPF an important addition to the family of sequential Monte Carlo methods.

Keywords: Sequential Monte Carlo; particle filter; multilevel; hidden Markov model

References

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

[2] Beaumont, M.A., Zhang, W. and Balding, D.J. (2002). Approximate Bayesian computation in population genetics. *Genetics* **162** 2025–2035.

[3] Beskos, A., Jasra, A., Law, K., Tempone, R. and Zhou, Y. (2017). Multilevel sequential Monte Carlo samplers. *Stochastic Process. Appl.* **127** 1417–1440. MR3630230 https://doi.org/10.1016/j.spa.2016.08.004

[4] Billingsley, P. (1995). *Probability and Measure*. New York: Wiley.

[5] Bogachev, V.I. (2007). *Measure Theory. Vol. I, II*. Berlin: Springer. MR2267655 https://doi.org/10.1007/978-3-540-34514-5

[6] Burkholder, D.L., Davis, B.J. and Gundy, R.F. (1972). Integral inequalities for convex functions of operators on martingales. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability Theory* 223–240. MR0400380

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

[8] Crisan, D. and Doucet, A. (2002). A survey of convergence results on particle filtering methods for practitioners. *IEEE Trans. Signal Process.* **50** 736–746. MR1895071 https://doi.org/10.1109/78.984773

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

[10] Del Moral, P., Doucet, A. and Jasra, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 411–436. MR2278333 https://doi.org/10.1111/j.1467-9868.2006.00553.x

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

[12] Douc, R. and Moulines, E. (2008). Limit theorems for weighted samples with applications to sequential Monte Carlo methods. *Ann. Statist.* **36** 2344–2376. MR2458190 https://doi.org/10.1214/07-AOS514

[13] Doucet, A., de Freitas, N. and Gordon, N., eds. (2001). *Sequential Monte Carlo Methods in Practice*. *Statistics for Engineering and Information Science*. New York: Springer. MR1847783 https://doi.org/10.1007/978-1-4757-3437-9

[14] Gelman, A., Bois, F. and Jian, J. (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Amer. Statist. Assoc.* **91** 1400–1412.

[15] Giles, M.B. (2008). Multi-level Monte Carlo path simulation. *Oper. Res.* **56** 607–617. MR2436856 https://doi.org/10.1287/opre.1070.0496

[16] Gordon, N.J., Salmond, D.J. and Smith, A.F.M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F* **140**.

[17] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*. New York: Academic Press. MR0624435

[18] Heine, K., Whiteley, N. and Cemgil, A.T. (2020). Parallelizing particle filters with butterfly interactions. *Scand. J. Stat.* **47** 361–396. MR4157144 https://doi.org/10.1111/sjos.12408

[19] Heinrich, S. (2001). Multilevel Monte Carlo methods. In *Large-Scale Scientific Computing* (S. Margenov, J. Waśniewski and P. Yalamov, eds.) Springer.

[20] Hong, Z. and Lian, H. (2012). Time-varying coefficient estimation in differential equation models with noisy time-varying covariates. *J. Multivariate Anal.* **103** 58–67. MR2823708 https://doi.org/10.1016/j.jmva.2011.06.007

[21] Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME Ser. D. J. Basic Eng.* **82** 35–45. MR3931993

[22] Karr, A.F. (1993). *Probability*. *Springer Texts in Statistics*. New York: Springer. MR1231974 https://doi.org/10.1007/978-1-4612-0891-4

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

[24] Liu, J.S. and Chen, R. (1998). Sequential Monte Carlo methods for dynamic systems. *J. Amer. Statist. Assoc.* **93** 1032–1044. MR1649198 https://doi.org/10.2307/2669847

[25] Prescott, T.P. and Baker, R.E. (2021). Multifidelity approximate Bayesian computation with sequential Monte Carlo parameter sampling. *SIAM/ASA J. Uncertain. Quantificat.* **9** 788–817. MR4273700 https://doi.org/10.1137/20M1316160

[26] Rebeschini, P. and van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *Ann. Appl. Probab.* **25** 2809–2866. MR3375889 https://doi.org/10.1214/14-AAP1061

[27] Toni, T., Welch, D., Strelkowa, N., Ipsen, A. and Stumpf, M.P.H. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J. R. Soc. Interface* **6** 187–202.

[28] Vinson, J.R. (1989). *The Behavior of Thin Walled Structures: Beams, Plates, and Shells*. *Mechanics of Surface Structures* **8**. Dordrecht: Kluwer Academic. MR0984128

[29] Apple Accelerate Framework. https://developer.apple.com/documentation/accelerate. Accessed: 2021-04-07.

Bernoulli **29**(1), 2023, 580–596
https://doi.org/10.3150/22-BEJ1469

Exponential and strong ergodicity for one-dimensional time-changed symmetric stable processes

TAO WANG[a]

School of Mathematical Sciences, Beijing Normal University, Laboratory of Mathematics and Complex Systems, Ministry of Education, Beijing, China. [a]*wang_tao@mail.bnu.edu.cn*

We obtain explicit criteria for both exponential ergodicity and strong ergodicity for one-dimensional time-changed symmetric stable processes with $\alpha \in (1,2)$. Explicit lower bounds for ergodic convergence rates are given.

Keywords: Stable process; time change; Dirichlet eigenvalue; strong ergodicity; exponential ergodicity; Green operator

References

[1] Bogdan, K. and Żak, T. (2006). On Kelvin transformation. *J. Theoret. Probab.* **19** 89–120. MR2256481 https://doi.org/10.1007/s10959-006-0003-8

[2] Byczkowska, H. and Byczkowski, T. (2001). One-dimensional symmetric stable Feynman-Kac semigroups. *Probab. Math. Statist.* **21** 381–404. MR1911445

[3] Chen, M. (1998). Estimate of exponential convergence rate in total variation by spectral gap. *Acta Math. Sin. New Ser.* **14** 9–16. MR1694048 https://doi.org/10.1007/BF02563878

[4] Chen, M. (2000). The principal eigenvalue for jump processes. *Acta Math. Sin. (Engl. Ser.)* **16** 361–368. MR1787092 https://doi.org/10.1007/s101140000067

[5] Chen, M. (2000). Explicit bounds of the first eigenvalue. *Sci. China Ser. A* **43** 1051–1059. MR1802148 https://doi.org/10.1007/BF02898239

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

[7] Chen, M.-F. and Wang, F.-Y. (2000). Cheeger's inequalities for general symmetric forms and existence criteria for spectral gap. *Ann. Probab.* **28** 235–257. MR1756004 https://doi.org/10.1214/aop/1019160118

[8] Chen, Z.-Q. and Fukushima, M. (2012). *Symmetric Markov Processes, Time Change, and Boundary Theory*. *London Mathematical Society Monographs Series* **35**. Princeton, NJ: Princeton Univ. Press. MR2849840

[9] Chen, Z.-Q. and Wang, J. (2014). Ergodicity for time-changed symmetric stable processes. *Stochastic Process. Appl.* **124** 2799–2823. MR3217425 https://doi.org/10.1016/j.spa.2014.04.003

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

[11] Döring, L. and Kyprianou, A.E. (2020). Entrance and exit at infinity for stable jump diffusions. *Ann. Probab.* **48** 1220–1265. MR4112713 https://doi.org/10.1214/19-AOP1389

[12] Döring, L., Kyprianou, A.E. and Weissmann, P. (2020). Stable processes conditioned to avoid an interval. *Stochastic Process. Appl.* **130** 471–487. MR4046505 https://doi.org/10.1016/j.spa.2019.01.004

[13] Grigor'yan, A. and Telcs, A. (2012). Two-sided estimates of heat kernels on metric measure spaces. *Ann. Probab.* **40** 1212–1284. MR2962091 https://doi.org/10.1214/11-AOP645

[14] Kyprianou, A.E. (2018). Stable Lévy processes, self-similarity and the unit ball. *ALEA Lat. Am. J. Probab. Math. Stat.* **15** 617–690. MR3808900 https://doi.org/10.30757/alea.v15-25

[15] Mao, Y.-H. (2002). Strong ergodicity for Markov processes by coupling methods. *J. Appl. Probab.* **39** 839–852. MR1938175 https://doi.org/10.1239/jap/1037816023

[16] Mao, Y.-H. (2006). Convergence rates in strong ergodicity for Markov processes. *Stochastic Process. Appl.* **116** 1964–1976. MR2307067 https://doi.org/10.1016/j.spa.2006.05.008

[17] Mao, Y.H. and Wang, T. (2022). Convergence rates in uniform ergodicity by hitting times and L^2-exponential convergence rates. *J. Theoret. Probab.* (to appear).

[18] Oshima, Y. (2013). *Semi-Dirichlet Forms and Markov Processes*. *De Gruyter Studies in Mathematics* **48**. Berlin: de Gruyter. MR3060116 https://doi.org/10.1515/9783110302066

[19] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[20] Shiozawa, Y. and Takeda, M. (2005). Variational formula for Dirichlet forms and estimates of principal eigenvalues for symmetric α-stable processes. *Potential Anal.* **23** 135–151. MR2139213 https://doi.org/10.1007/s11118-004-5392-7

[21] Wang, F.-Y. (1999). Existence of the spectral gap for elliptic operators. *Ark. Mat.* **37** 395–407. MR1714760 https://doi.org/10.1007/BF02412223

[22] Yano, K. (2010). Excursions away from a regular point for one-dimensional symmetric Lévy processes without Gaussian part. *Potential Anal.* **32** 305–341. MR2603019 https://doi.org/10.1007/s11118-009-9152-6

Bernoulli **29**(1), 2023, 597–624
https://doi.org/10.3150/22-BEJ1470

A necessary and sufficient condition for the convergence of the derivative martingale in a branching Lévy process

BASTIEN MALLEIN[1,a] and QUAN SHI[2,b]

[1] *Université Sorbonne Paris Nord, LAGA, UMR 7539, F-93430, Villetaneuse, France.*
[a] *mallein@math.univ-paris13.fr*
[2] *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, 100190, Beijing, China.*
[b] *quan.shi@amss.ac.cn*

A continuous-time particle system on the real line satisfying the branching property and an exponential integrability condition is called a branching Lévy process, and its law is characterized by a triplet (σ^2, a, Λ). We obtain a necessary and sufficient condition for the convergence of the derivative martingale of such a process to a non-trivial limit in terms of (σ^2, a, Λ). This extends previously known results on branching Brownian motions and branching random walks. To obtain this result, we rely on the spinal decomposition and establish a novel zero-one law on the perpetual integrals of centred Lévy processes conditioned to stay positive.

Keywords: Branching Lévy process; derivative martingale; spinal decomposition; Lévy process; perpetual integral

References

[1] Aïdékon, E. (2013). Convergence in law of the minimum of a branching random walk. *Ann. Probab.* **41** 1362–1426. MR3098680 https://doi.org/10.1214/12-AOP750
[2] Aïdékon, E. and Shi, Z. (2014). The Seneta–Heyde scaling for the branching random walk. *Ann. Probab.* **42** 959–993. MR3189063 https://doi.org/10.1214/12-AOP809
[3] Alsmeyer, G. and Iksanov, A. (2009). A log-type moment result for perpetuities and its application to martingales in supercritical branching random walks. *Electron. J. Probab.* **14** 289–312. MR2471666 https://doi.org/10.1214/EJP.v14-596
[4] Baguley, S., Döring, L. and Kyprianou, A. (2020). General path integrals and stable SDEs. Preprint. Available at arXiv:2012.07406.
[5] Bérard, J. and Gouéré, J.-B. (2011). Survival probability of the branching random walk killed below a linear boundary. *Electron. J. Probab.* **16** 396–418. MR2774095 https://doi.org/10.1214/EJP.v16-861
[6] Bertoin, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press. MR1406564
[7] Bertoin, J. (2001). Homogeneous fragmentation processes. *Probab. Theory Related Fields* **121** 301–318. MR1867425 https://doi.org/10.1007/s004400100152
[8] Bertoin, J. (2016). Compensated fragmentation processes and limits of dilated fragmentations. *Ann. Probab.* **44** 1254–1284. MR3474471 https://doi.org/10.1214/14-AOP1000
[9] Bertoin, J. and Mallein, B. (2018). Biggins' martingale convergence for branching Lévy processes. *Electron. Commun. Probab.* **23** Paper No. 83, 12 pp. MR3873790 https://doi.org/10.1214/18-ECP185
[10] Bertoin, J. and Mallein, B. (2019). Infinitely ramified point measures and branching Lévy processes. *Ann. Probab.* **47** 1619–1652. MR3945755 https://doi.org/10.1214/18-AOP1292
[11] Bertoin, J. and Rouault, A. (2005). Discretization methods for homogeneous fragmentations. *J. Lond. Math. Soc. (2)* **72** 91–109. MR2145730 https://doi.org/10.1112/S0024610705006423

[12] Bertoin, J. and Savov, M. (2011). Some applications of duality for Lévy processes in a half-line. *Bull. Lond. Math. Soc.* **43** 97–110. MR2765554 https://doi.org/10.1112/blms/bdq084

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

[14] Biggins, J.D. and Kyprianou, A.E. (2004). Measure change in multitype branching. *Adv. in Appl. Probab.* **36** 544–581. MR2058149 https://doi.org/10.1239/aap/1086957585

[15] Biggins, J.D. and Kyprianou, A.E. (2005). Fixed points of the smoothing transform: The boundary case. *Electron. J. Probab.* **10** 609–631. MR2147319 https://doi.org/10.1214/EJP.v10-255

[16] Bogus, K., Buraczewski, D. and Marynych, A. (2020). Self-similar solutions of kinetic-type equations: The boundary case. *Stochastic Process. Appl.* **130** 677–693. MR4046515 https://doi.org/10.1016/j.spa.2019.03.005

[17] Chaumont, L. and Doney, R.A. (2005). On Lévy processes conditioned to stay positive. *Electron. J. Probab.* **10** 948–961. MR2164035 https://doi.org/10.1214/EJP.v10-261

[18] Cheek, D. and Shneer, S. (2020). The empirical mean position of a branching Lévy process. *J. Appl. Probab.* **57** 1252–1259. MR4179608 https://doi.org/10.1017/jpr.2020.60

[19] Chen, L.H.Y. (1978). A short note on the conditional Borel–Cantelli lemma. *Ann. Probab.* **6** 699–700. MR0496420

[20] Chen, X. (2015). A necessary and sufficient condition for the nontrivial limit of the derivative martingale in a branching random walk. *Adv. in Appl. Probab.* **47** 741–760. MR3406606 https://doi.org/10.1239/aap/1444308880

[21] Dadoun, B. (2017). Asymptotics of self-similar growth-fragmentation processes. *Electron. J. Probab.* **22** Paper No. 27, 30 pp. MR3629871 https://doi.org/10.1214/17-EJP45

[22] Doney, R.A. and Maller, R.A. (2002). Stability of the overshoot for Lévy processes. *Ann. Probab.* **30** 188–212. MR1894105 https://doi.org/10.1214/aop/1020107765

[23] Döring, L. and Kyprianou, A.E. (2016). Perpetual integrals for Lévy processes. *J. Theoret. Probab.* **29** 1192–1198. MR3540494 https://doi.org/10.1007/s10959-015-0607-y

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

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

[26] Iksanov, A. and Mallein, B. (2019). A result on power moments of Lévy-type perpetuities and its application to the L_p-convergence of Biggins' martingales in branching Lévy processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **16** 315–331. MR3914353 https://doi.org/10.30757/alea.v16-11

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

[28] Jaffuel, B. (2012). The critical barrier for the survival of branching random walk with absorption. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 989–1009. MR3052402 https://doi.org/10.1214/11-AIHP453

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

[30] Kolb, M. and Savov, M. (2020). A characterization of the finiteness of perpetual integrals of Lévy processes. *Bernoulli* **26** 1453–1472. MR4058374 https://doi.org/10.3150/19-BEJ1167

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

[32] Lalley, S.P. and Sellke, T. (1987). A conditional limit theorem for the frontier of a branching Brownian motion. *Ann. Probab.* **15** 1052–1061. MR0893913

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

[34] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164

[35] Mallein, B. (2018). Genealogy of the extremal process of the branching random walk. *ALEA Lat. Am. J. Probab. Math. Stat.* **15** 1065–1087. MR3852245 https://doi.org/10.30757/alea.v15-39

[36] McKean, H.P. (1975). Application of Brownian motion to the equation of Kolmogorov–Petrovskii–Piskunov. *Comm. Pure Appl. Math.* **28** 323–331. MR0400428 https://doi.org/10.1002/cpa.3160280302

[37] Neveu, J. (1988). Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987 (Princeton, NJ, 1987). Progr. Probab. Statist.* **15** 223–242. Boston, MA: Birkhäuser. MR1046418 https://doi.org/10.1007/978-1-4684-0550-7_10

[38] Peyrière, J. (1974). Turbulence et dimension de Hausdorff. *C. R. Acad. Sci. Paris Sér. A* **278** 567–569. MR0431354

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

[40] Shi, Q. and Watson, A.R. (2019). Probability tilting of compensated fragmentations. *Electron. J. Probab.* **24** Paper No. 78, 39 pp. MR3991115 https://doi.org/10.1214/19-ejp316

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

[42] Uchiyama, K. (1982). Spatial growth of a branching process of particles living in $\mathbf{R}^d$. *Ann. Probab.* **10** 896–918. MR0672291

[43] Yang, T. and Ren, Y.-X. (2011). Limit theorem for derivative martingale at criticality w.r.t. branching Brownian motion. *Statist. Probab. Lett.* **81** 195–200. MR2748182 https://doi.org/10.1016/j.spl.2010.11.007

Bernoulli **29**(1), 2023, 625–651
https://doi.org/10.3150/22-BEJ1471

Almost-sure asymptotics for Riemannian random waves

LOUIS GASS[a]

Department, Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France. [a]*louis.gass@ens-rennes.fr*

We consider the Riemannian random wave model of Gaussian linear combinations of Laplace eigenfunctions on a general compact Riemannian manifold. With probability one with respect to the Gaussian coefficients, we establish that, both for large and small band models, the process properly rescaled and evaluated at an independently and uniformly chosen point X on the manifold, converges in distribution under the sole randomness of X towards an universal Gaussian field as the frequency tends to infinity. This result extends the celebrated central limit Theorem of Salem–Zygmund for trigonometric polynomials series to the more general framework of compact Riemannian manifolds. We then deduce from the above convergence the almost-sure asymptotics of the nodal volume associated with the random wave. To the best of our knowledge, in the real Riemannian case, these asymptotics were only known in expectation and not in the almost sure sense due to the lack of sufficiently accurate variance estimates. This in particular addresses a question of S. Zelditch regarding the almost sure equidistribution of nodal volume.

Keywords: Geometric probability; mathematical physics

References

[1] Álvarez Paiva, J.C. and Fernandes, E. (2007). Gelfand transforms and Crofton formulas. *Selecta Math. (N.S.)* **13** 369–390. MR2383600 https://doi.org/10.1007/s00029-007-0045-5

[2] Angst, J., Pham, V.-H. and Poly, G. (2018). Universality of the nodal length of bivariate random trigonometric polynomials. *Trans. Amer. Math. Soc.* **370** 8331–8357. MR3864378 https://doi.org/10.1090/tran/7255

[3] Angst, J. and Poly, G. (2021). Variations on Salem-Zygmund results for random trigonometric polynomials: Application to almost sure nodal asymptotics. *Electron. J. Probab.* **26** Paper No. 156. MR4350983 https://doi.org/10.1214/21-ejp716

[4] Armentano, D., Azaïs, J., Dalmao, F. and Léon, J.R. (2019). On the finiteness of the moments of the measure of level sets of random fields.

[5] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2017). A new approach to the Stein-Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence.

[6] Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: Wiley. MR2478201 https://doi.org/10.1002/9780470434642

[7] Beliaev, D. and Wigman, I. (2018). Volume distribution of nodal domains of random band-limited functions. *Probab. Theory Related Fields* **172** 453–492. MR3851836 https://doi.org/10.1007/s00440-017-0813-x

[8] Berry, M.V. (1977). Regular and irregular semiclassical wavefunctions. *J. Phys. A* **10** 2083–2091. MR0489542

[9] Bogomolny, E. and Schmit, C. (2002). Percolation model for nodal domains of chaotic wave functions. *Phys. Rev. Lett.* **88** 114102.

[10] Bourgain, J. (2014). On toral eigenfunctions and the random wave model. *Israel J. Math.* **201** 611–630. MR3265298 https://doi.org/10.1007/s11856-014-1037-z

[11] Buckley, J. and Wigman, I. (2016). On the number of nodal domains of toral eigenfunctions. *Ann. Henri Poincaré* **17** 3027–3062. MR3556515 https://doi.org/10.1007/s00023-016-0476-7

[12] Campese, S., Marinucci, D. and Rossi, M. (2018). Approximate normality of high-energy hyperspherical eigenfunctions. *J. Math. Anal. Appl.* **461** 500–522. MR3759554 https://doi.org/10.1016/j.jmaa.2017.11.051

[13] Canzani, Y. and Hanin, B. (2020). Local universality for zeros and critical points of monochromatic random waves. *Comm. Math. Phys.* **378** 1677–1712. MR4150887 https://doi.org/10.1007/s00220-020-03826-w

[14] Dierickx, G., Nourdin, I., Peccati, G. and Rossi, M. (2020). Small scale CLTs for the nodal length of monochromatic waves. ArXiv preprint. Available at arXiv:2005.06577.

[15] Donnelly, H. and Fefferman, C. (1988). Nodal sets of eigenfunctions on Riemannian manifolds. *Invent. Math.* **93** 161–183. MR0943927 https://doi.org/10.1007/BF01393691

[16] Gass, L. (2023). Supplement to "Almost-sure asymptotics for Riemannian random waves." https://doi.org/10.3150/22-BEJ1471SUPP

[17] Hörmander, L. (1968). The spectral function of an elliptic operator. *Acta Math.* **121** 193–218. MR0609014 https://doi.org/10.1007/BF02391913

[18] Jubin, B. (2019). Intrinsic volumes of sublevel sets.

[19] Keeler, B.J. (2019). A logarithmic improvement in the two-point Weyl law for manifolds without conjugate points. ArXiv preprint. Available at arXiv:1905.05136.

[20] Kunita, H. (1997). *Stochastic Flows and Stochastic Differential Equations*. *Cambridge Studies in Advanced Mathematics* **24**. Cambridge: Cambridge Univ. Press. Reprint of the 1990 original. MR1472487

[21] Letendre, T. (2016). Expected volume and Euler characteristic of random submanifolds. *J. Funct. Anal.* **270** 3047–3110. MR3470435 https://doi.org/10.1016/j.jfa.2016.01.007

[22] Logunov, A. (2018). Nodal sets of Laplace eigenfunctions: Polynomial upper estimates of the Hausdorff measure. *Ann. of Math. (2)* **187** 221–239. MR3739231 https://doi.org/10.4007/annals.2018.187.1.4

[23] Logunov, A. (2018). Nodal sets of Laplace eigenfunctions: Proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture. *Ann. of Math. (2)* **187** 241–262. MR3739232 https://doi.org/10.4007/annals.2018.187.1.5

[24] Logunov, A. and Malinnikova, E. (2016). Ratios of harmonic functions with the same zero set. *Geom. Funct. Anal.* **26** 909–925. MR3540456 https://doi.org/10.1007/s00039-016-0369-4

[25] Rudnick, Z. and Wigman, I. (2008). On the volume of nodal sets for eigenfunctions of the Laplacian on the torus. *Ann. Henri Poincaré* **9** 109–130. MR2389892 https://doi.org/10.1007/s00023-007-0352-6

[26] Salem, R. and Zygmund, A. (1954). Some properties of trigonometric series whose terms have random signs. *Acta Math.* **91** 245–301. MR0065679 https://doi.org/10.1007/BF02393433

[27] Shiffman, B. and Zelditch, S. (1999). Distribution of zeros of random and quantum chaotic sections of positive line bundles. *Comm. Math. Phys.* **200** 661–683. MR1675133 https://doi.org/10.1007/s002200050544

[28] Wigman, I. (2009). On the distribution of the nodal sets of random spherical harmonics. *J. Math. Phys.* **50** 013521. MR2492631 https://doi.org/10.1063/1.3056589

[29] Yau, S.T., ed. (1982). *Seminar on Differential Geometry*. *Annals of Mathematics Studies, No. 102*. Princeton, NJ: Princeton Univ. Press. MR0645728

[30] Zelditch, S. (2009). Real and complex zeros of Riemannian random waves. In *Spectral Analysis in Geometry and Number Theory*. *Contemp. Math.* **484** 321–342. Providence, RI: Amer. Math. Soc. MR1500155 https://doi.org/10.1090/conm/484/09482

[31] Zelditch, S. (2010). Recent developments in mathematical quantum chaos. In *Current Developments in Mathematics, 2009* 115–204. Somerville, MA: Int. Press. MR2757360

Bernoulli **29**(1), 2023, 652–668
https://doi.org/10.3150/22-BEJ1472

Design of c-optimal experiments for high-dimensional linear models

HAMID EFTEKHARI[a], MOULINATH BANERJEE[b] and YA'ACOV RITOV[c]

Department of Statistics, University of Michigan, Ann Arbor, Michigan, USA. [a]*hamidef@umich.edu,* [b]*moulib@umich.edu,* [c]*yritov@umich.edu*

We study randomized designs that minimize the asymptotic variance of a debiased lasso estimator when a large pool of unlabeled data is available but measuring the corresponding responses is costly. The optimal sampling distribution arises as the solution of a semidefinite program. The improvements in efficiency that result from these optimal designs are demonstrated via simulation experiments.

Keywords: Optimal design; inference; sparsity; compressed sensing

References

[1] Ahmed, N., Natarajan, T. and Rao, K.R. (1974). Discrete cosine transform. *IEEE Trans. Comput.* **C-23** 90–93. MR0356555 https://doi.org/10.1109/t-c.1974.223784

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

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

[4] Cai, T., Cai, T. and Guo, Z. (2019). Individualized treatment selection: An optimal hypothesis testing approach in high-dimensional models. arXiv preprint. Available at arXiv:1904.12891.

[5] Cai, T.T. and Guo, Z. (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *Ann. Statist.* **45** 615–646. MR3650395 https://doi.org/10.1214/16-AOS1461

[6] Černý, M. and Hladík, M. (2012). Two complexity results on c-optimality in experimental design. *Comput. Optim. Appl.* **51** 1397–1408. MR2891943 https://doi.org/10.1007/s10589-010-9377-8

[7] Chen, S.S., Donoho, D.L. and Saunders, M.A. (2001). Atomic decomposition by basis pursuit. *SIAM Rev.* **43** 129–159. MR1854649 https://doi.org/10.1137/S003614450037906X

[8] Chernoff, H. (1953). Locally optimal designs for estimating parameters. *Ann. Math. Stat.* **24** 586–602. MR0058932 https://doi.org/10.1214/aoms/1177728915

[9] Dette, H. et al. (1993). Elfving's theorem for D-optimality. *Ann. Statist.* **21** 753–766. MR1232517 https://doi.org/10.1214/aos/1176349149

[10] Dette, H. and Holland-Letz, T. et al. (2009). A geometric characterization of c-optimal designs for heteroscedastic regression. *Ann. Statist.* **37** 4088–4103. MR2572453 https://doi.org/10.1214/09-AOS708

[11] Dobriban, E. and Fan, J. (2016). Regularity properties for sparse regression. *Commun. Math. Stat.* **4** 1–19. MR3475839 https://doi.org/10.1007/s40304-015-0078-6

[12] Elfving, G. (1952). Optimum allocation in linear regression theory. *Ann. Math. Stat.* **23** 255–262. MR0047998 https://doi.org/10.1214/aoms/1177729442

[13] Fan, S.K. and Chaloner, K. (2003). A geometric method for singular c-optimal designs. *J. Statist. Plann. Inference* **113** 249–257. MR1963044 https://doi.org/10.1016/S0378-3758(01)00289-0

[14] Gu, L. and Yang, J.-F. (2013). Construction of nearly orthogonal Latin hypercube designs. *Metrika* **76** 819–830. MR3085834 https://doi.org/10.1007/s00184-012-0417-5

[15] Harman, R. and Jurík, T. (2008). Computing c-optimal experimental designs using the simplex method of linear programming. *Comput. Statist. Data Anal.* **53** 247–254. MR2649082 https://doi.org/10.1016/j.csda.2008.06.023

[16] Huang, Y., Kong, X. and Ai, M. (2020). Optimal designs in sparse linear models. *Metrika* **83** 255–273. MR4057422 https://doi.org/10.1007/s00184-019-00722-9

[17] Javanmard, A. and Lee, J.D. (2020). A flexible framework for hypothesis testing in high dimensions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 685–718. MR4112781

[18] Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909. MR3277152

[19] Kohavi, R., Longbotham, R., Sommerfield, D. and Henne, R.M. (2009). Controlled experiments on the web: Survey and practical guide. *Data Min. Knowl. Discov.* **18** 140–181. MR2469594 https://doi.org/10.1007/s10618-008-0114-1

[20] Eftekhari, H., Banerjee, M. and Ritov, Y. (2023). Supplement to "Design of c-optimal experiments for high-dimensional linear models." https://doi.org/10.3150/22-BEJ1472SUPP

[21] Lustig, M., Donoho, D. and Pauly, J.M. (2007). Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine* **58** 1182–1195.

[22] Misra, S. and Wu, Y. (2020). *Machine Learning Assisted Segmentation of Scanning Electron Microscopy Images of Organic-Rich Shales with Feature Extraction and Feature Ranking*. Gulf Professional Publishing. https://doi.org/10.1016/B978-0-12-817736-5.00010-7

[23] Pázman, A. and Pronzato, L. (2006). On the irregular behavior of LS estimators for asymptotically singular designs. *Statist. Probab. Lett.* **76** 1089–1096. MR2269278 https://doi.org/10.1016/j.spl.2005.12.010

[24] Pronzato, L. (2009). On the regularization of singular c-optimal designs. *Math. Slovaca* **59** 611–626. MR2557313 https://doi.org/10.2478/s12175-009-0151-2

[25] Pukelsheim, F. (2006). *Optimal Design of Experiments*. *Classics in Applied Mathematics* **50**. Philadelphia, PA: SIAM. MR2224698 https://doi.org/10.1137/1.9780898719109

[26] Pukelsheim, F. and Rieder, S. (1992). Efficient rounding of approximate designs. *Biometrika* **79** 763–770. MR1209476 https://doi.org/10.1093/biomet/79.4.763

[27] Ravi, S.N., Ithapu, V., Johnson, S. and Singh, V. (2016). Experimental design on a budget for sparse linear models and applications. In *International Conference on Machine Learning* 583–592.

[28] Rodríguez-Díaz, J.M. (2017). Computation of c-optimal designs for models with correlated observations. *Comput. Statist. Data Anal.* **113** 287–296. MR3662408 https://doi.org/10.1016/j.csda.2016.10.019

[29] Rudelson, M. and Zhou, S. (2012). Reconstruction from anisotropic random measurements. In *Conference on Learning Theory* 10–1.

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

[31] Sagnol, G. (2011). Computing optimal designs of multiresponse experiments reduces to second-order cone programming. *J. Statist. Plann. Inference* **141** 1684–1708. MR2763200 https://doi.org/10.1016/j.jspi.2010.11.031

[32] Seeger, M.W. (2008). Bayesian inference and optimal design for the sparse linear model. *J. Mach. Learn. Res.* **9** 759–813. MR2417254 https://doi.org/10.1017/s0143385700005320

[33] Stoica, P. and Babu, P. (2010). Algebraic derivation of elfving theorem on optimal experiment design and some connections with sparse estimation. *IEEE Signal Process. Lett.* **17** 743–745. https://doi.org/10.1109/LSP.2010.2053533

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

[35] van de Geer, S., Bühlmann, P., Ritov, Y. and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42** 1166–1202. MR3224285 https://doi.org/10.1214/14-AOS1221

[36] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. MR3837109 https://doi.org/10.1017/9781108231596

[37] Zhang, C.-H. and Zhang, S.S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 217–242. MR3153940 https://doi.org/10.1111/rssb.12026

Bernoulli **29**(1), 2023, 669–695
https://doi.org/10.3150/22-BEJ1474

Multivariate time series models for mixed data

ZINSOU-MAX DEBALY[a] and LIONEL TRUQUET[b]

ENSAI-CREST UMR 9194, Campus de Ker-Lann, Rue Blaise Pascal – BP 37203, 35712 BRUZ cedex, France,
[a]*zinsou-max.debaly@ensai.fr,* [b]*lionel.truquet@ensai.fr*

We introduce a general approach which unifies some previous attempts for modeling the dynamic of multivariate time series or for regression analysis when the data are of mixed type (binary/count/continuous). Our approach is quite flexible since conditionally on past values, each coordinate at time t can have a distribution compatible with a standard univariate time series model such as GARCH, ARMA, INGARCH or logistic models whereas past values of the other coordinates play the role of exogenous covariates in the dynamic. The simultaneous dependence in the multivariate time series can be modeled with a copula. Additional exogenous covariates are also allowed in the dynamic. We first study some usual stability properties of these models and then show that autoregressive parameters can be consistently estimated equation-by-equation using a pseudo-maximum likelihood method, leading to a fast implementation even when the number of time series is large. Moreover, we prove consistency results when a parametric copula model is fitted to the time series and in the case of Gaussian copulas, we show that the likelihood estimator of the correlation matrix is strongly consistent. We carefully check all our assumptions for two prototypical examples: a GARCH/INGARCH model and logistic/log-linear INGARCH model. Our results are illustrated with numerical experiments as well as two real data sets.

Keywords: Time series; mixed data; observation-driven models

References

[1] Alexander, R., Leon, D. and Carrière Chough, K. (2013). *Analysis of Mixed Data: Methods & Applications*. CRC Press.
[2] Bardet, J.-M., Doukhan, P. and Wintenberger, O. (2020). Contrast estimation of general locally stationary processes using coupling. arXiv preprint. Available at arXiv:2005.07397.
[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1
[4] Bouyé, E., Durrleman, V. and Nikeghbali, A. (2000). Gaël Riboulet, and Thierry Roncalli. Copulas for finance-a reading guide and some applications. Available at SSRN 1032533.
[5] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4
[6] Cox, D.R. and Snell, E.J. (1970). *The Analysis of Binary Data*. London: CRC Press.
[7] Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli* **25** 1013–1044. MR3920364 https://doi.org/10.3150/17-bej1011
[8] Davis, R.A. and Liu, H. (2016). Theory and inference for a class of nonlinear models with application to time series of counts. *Statist. Sinica* **26** 1673–1707. MR3586234
[9] Debaly, Z.-M. Truquet, L. (2023). Supplement to "Multivariate time series models for mixed data." https://doi.org/10.3150/22-BEJ1474SUPP
[10] Debaly, Z.M. and Truquet, L. (2021). Iterations of dependent random maps and exogeneity in nonlinear dynamics. *Econometric Theory* **37** 1135–1172. MR4348399 https://doi.org/10.1017/S0266466620000559
[11] Debaly, Z.M. and Truquet, L. (2021). A note on the stability of multivariate non-linear time series with an application to time series of counts. *Statist. Probab. Lett.* **179** 109196. MR4293765 https://doi.org/10.1016/j.spl.2021.109196
[12] Douc, R., Doukhan, P. and Moulines, E. (2013). Ergodicity of observation-driven time series models and consistency of the maximum likelihood estimator. *Stochastic Process. Appl.* **123** 2620–2647. MR3054539 https://doi.org/10.1016/j.spa.2013.04.010

[13] Doukhan, P., Neumann, M.H. and Truquet, L. (2020). Stationarity and ergodic properties for some observation-driven models in random environments. arXiv preprint. Available at arXiv:2007.07623.

[14] Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *J. Time Series Anal.* **27** 923–942. MR2328548 https://doi.org/10.1111/j.1467-9892.2006.00496.x

[15] Fernández-Fontelo, A., Cabaña, A., Puig, P. and Moriña, D. (2016). Under-reported data analysis with INAR-hidden Markov chains. *Stat. Med.* **35** 4875–4890. MR3554999 https://doi.org/10.1002/sim.7026

[16] Fokianos, K. and Kedem, B. (2003). Regression theory for categorical time series. *Statist. Sci.* **18** 357–376. MR2061915 https://doi.org/10.1214/ss/1076102425

[17] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression. *J. Amer. Statist. Assoc.* **104** 1430–1439. MR2596998 https://doi.org/10.1198/jasa.2009.tm08270

[18] Fokianos, K., Støve, B., Tjøstheim, D. and Doukhan, P. (2020). Multivariate count autoregression. *Bernoulli* **26** 471–499. MR4036041 https://doi.org/10.3150/19-BEJ1132

[19] Fokianos, K. and Tjøstheim, D. (2011). Log-linear Poisson autoregression. *J. Multivariate Anal.* **102** 563–578. MR2755016 https://doi.org/10.1016/j.jmva.2010.11.002

[20] Fokianos, K. and Truquet, L. (2018). On categorical time series models with covariates. *Stochastic Processes and Their Applications*.

[21] Francq, C. and Thieu, L.Q. (2019). QML inference for volatility models with covariates. *Econometric Theory* **35** 37–72. MR3904171 https://doi.org/10.1017/S0266466617000512

[22] Francq, C., Wintenberger, O. and Zakoïan, J.-M. (2013). GARCH models without positivity constraints: Exponential or log GARCH? *J. Econometrics* **177** 34–46. MR3103914 https://doi.org/10.1016/j.jeconom.2013.05.004

[23] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: Wiley. MR3186556 https://doi.org/10.1002/9780470670057

[24] Genest, C. and Nešlehová, J. (2007). A primer on copulas for count data. *Astin Bull.* **37** 475–515. MR2422797 https://doi.org/10.2143/AST.37.2.2024077

[25] Han, H. and Kristensen, D. (2014). Asymptotic theory for the QMLE in GARCH-X models with stationary and nonstationary covariates. *J. Bus. Econom. Statist.* **32** 416–429. MR3238595 https://doi.org/10.1080/07350015.2014.897954

[26] Holmes, E.E., Ward, E.J. and Scheuerell, M.D. (2014). Analysis of multivariate time-series using the marss package. *NOAA Fish. Northwest Fish. Sci. Cent.* **2725** 98112.

[27] Kauppi, H. and Saikkonen, P. (2008). Predicting us recessions with dynamic binary response models. *Rev. Econ. Stat.* **90** 777–791.

[28] Lu, S. and Huang, S. (2020). Segmentation of multivariate industrial time series data based on dynamic latent variable predictability. *IEEE Access* **8** 112092–112103.

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

[30] Manner, H., Türk, D. and Eichler, M. (2016). Modeling and forecasting multivariate electricity price spikes. *Energy Econ.* **60** 255–265.

[31] Marbac, M., Biernacki, C. and Vandewalle, V. (2017). Model-based clustering of Gaussian copulas for mixed data. *Comm. Statist. Theory Methods* **46** 11635–11656. MR3735322 https://doi.org/10.1080/03610926.2016.1277753

[32] Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electron. J. Stat.* **6** 1517–1549. MR2988457 https://doi.org/10.1214/12-EJS721

[33] Moysiadis, T. and Fokianos, K. (2014). On binary and categorical time series models with feedback. *J. Multivariate Anal.* **131** 209–228. MR3252645 https://doi.org/10.1016/j.jmva.2014.07.004

[34] Omranian, N., Mueller-Roeber, B. and Nikoloski, Z. (2015). Segmentation of biological multivariate time-series data. *Sci. Rep.* **5** 1–6.

[35] Robin, S., Rodolphe, F. and Schbath, S. (2005). *DNA, Words and Models: Statistics of Exceptional Words*. Cambridge: Cambridge Univ. Press. MR2214200

[36] Russell, J.R. and Engle, R.F. (2005). A discrete-state continuous-time model of financial transactions prices and times: The autoregressive conditional multinomial—autoregressive conditional duration model. *J. Bus. Econom. Statist.* **23** 166–180. MR2157268 https://doi.org/10.1198/073500104000000541

[37] Song, P.X.-K., Li, M. and Yuan, Y. (2009). Joint regression analysis of correlated data using Gaussian copulas. *Biometrics* **65** 60–68. MR2665846 https://doi.org/10.1111/j.1541-0420.2008.01058.x

[38] Straumann, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models*. *Lecture Notes in Statistics* **181**. Berlin: Springer. MR2142271

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

[40] Truquet, L. (2020). Coupling and perturbation techniques for categorical time series. *Bernoulli* **26** 3249–3279. MR4140544 https://doi.org/10.3150/20-BEJ1225

[41] Tsay, R.S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR3236787

[42] Yang, Y. and Kang, J. (2010). Joint analysis of mixed Poisson and continuous longitudinal data with nonignorable missing values. *Comput. Statist. Data Anal.* **54** 193–207. MR2558470 https://doi.org/10.1016/j.csda.2009.08.006

Bernoulli **29**(1), 2023, 696–724
https://doi.org/10.3150/22-BEJ1475

Bootstrap percolation on the stochastic block model

GIOVANNI LUCA TORRISI[1,a], MICHELE GARETTO[2,b] and EMILIO LEONARDI[3,c]

[1] *CNR-IAC, Roma, Italy.* [a] *giovanniluca.torrisi@cnr.it*
[2] *Università di Torino, Torino, Italy.* [b] *michele.garetto@unito.it*
[3] *Politecnico di Torino, Torino, Italy.* [c] *emilio.leonardi@polito.it*

We analyze the bootstrap percolation process on the stochastic block model (SBM), a natural extension of the Erdős–Rényi random graph that incorporates the community structure observed in many real systems. In the SBM, nodes are partitioned into two subsets, which represent different communities, and pairs of nodes are independently connected with a probability that depends on the communities they belong to. Under mild assumptions on the system parameters, we prove the existence of a sharp phase transition for the final number of active nodes and characterize the sub-critical and the super-critical regimes in terms of the number of initially active nodes, which are selected uniformly at random in each community.

Keywords: Bootstrap percolation; random graphs; stochastic block model

References

[1] Abbe, E., Bandeira, A.S. and Hall, G. (2016). Exact recovery in the stochastic block model. *IEEE Trans. Inf. Theory* **62** 471–487. MR3447993 https://doi.org/10.1109/TIT.2015.2490670

[2] Abbe, E. and Sandon, C. (2015). Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science—FOCS 2015* 670–688. Los Alamitos, CA: IEEE Computer Soc. MR3473334 https://doi.org/10.1109/FOCS.2015.47

[3] Abdullah, M.A. and Fountoulakis, N. (2018). A phase transition in the evolution of bootstrap percolation processes on preferential attachment graphs. *Random Structures Algorithms* **52** 379–418. MR3783205 https://doi.org/10.1002/rsa.20744

[4] Airoldi, E.M. et al. (2008). Mixed membership stochastic block models. *J. Mach. Learn. Res.* **9** 1981–2014.

[5] Amini, H. (2010). Bootstrap percolation and diffusion in random graphs with given vertex degrees. *Electron. J. Combin.* **17** Research Paper 25, 20 pp. MR2595485

[6] Amini, H. and Fountoulakis, N. (2014). Bootstrap percolation in power-law random graphs. *J. Stat. Phys.* **155** 72–92. MR3180970 https://doi.org/10.1007/s10955-014-0946-6

[7] Amini, H., Fountoulakis, N. and Panagiotou, K. (2014). Bootstrap percolation in inhomogeneous random graphs. Preprint. Available at arXiv:1402.2815.

[8] Angel, O. and Kolesnik, B. (2018). Large deviations for subcritical bootstrap percolation on the random graph. Preprint. Available at arXiv:1705.06815v2.

[9] Balogh, J. and Bollobás, B. (2006). Bootstrap percolation on the hypercube. *Probab. Theory Related Fields* **134** 624–648. MR2214907 https://doi.org/10.1007/s00440-005-0451-6

[10] Balogh, J., Peres, Y. and Pete, G. (2006). Bootstrap percolation on infinite trees and non-amenable groups. *Combin. Probab. Comput.* **15** 715–730. MR2248323 https://doi.org/10.1017/S0963548306007619

[11] Balogh, J. and Pittel, B.G. (2007). Bootstrap percolation on the random regular graph. *Random Structures Algorithms* **30** 257–286. MR2283230 https://doi.org/10.1002/rsa.20158

[12] Bollobás, B., Gunderson, K., Holmgren, C., Janson, S. and Przykucki, M. (2014). Bootstrap percolation on Galton–Watson trees. *Electron. J. Probab.* **19** Paper No. 13, 27 pp. MR3164766 https://doi.org/10.1214/EJP.v19-2758

[13] Bollobás, B., Riordan, O., Slivken, E. and Smith, P. (2017). The threshold for jigsaw percolation on random graphs. *Electron. J. Combin.* **24** Paper No. 2.36, 14 pp. MR3665569 https://doi.org/10.37236/6102

[14] Bradonjić, M. and Saniee, I. (2014). Bootstrap percolation on random geometric graphs. *Probab. Engrg. Inform. Sci.* **28** 169–181. MR3244950 https://doi.org/10.1017/S0269964813000405

[15] Brummitt, C.D., Chatterjee, S., Dey, P.S. and Sivakoff, D. (2015). Jigsaw percolation: What social networks can collaboratively solve a puzzle? *Ann. Appl. Probab.* **25** 2013–2038. MR3349000 https://doi.org/10.1214/14-AAP1041

[16] Chalupa, J., Leath, P.L. and Reich, G.R. (1979). Bootstrap percolation on a Bethe lattice. *J. Phys. C* **12** 31–35.

[17] Coja-Oghlan, A. and Lanka, A. (2010). Finding planted partitions in random graphs with general degree distributions. *SIAM J. Discrete Math.* **23** 1682–1714. MR2570199 https://doi.org/10.1137/070699354

[18] Feige, U., Krivelevich, M. and Reichman, D. (2017). Contagious sets in random graphs. *Ann. Appl. Probab.* **27** 2675–2697. MR3719944 https://doi.org/10.1214/16-AAP1254

[19] Fortunato, S. (2010). Community detection in graphs. *Phys. Rep.* **486** 75–174. MR2580414 https://doi.org/10.1016/j.physrep.2009.11.002

[20] Fountoulakis, N., Kang, M., Koch, C. and Makai, T. (2018). A phase transition regarding the evolution of bootstrap processes in inhomogeneous random graphs. *Ann. Appl. Probab.* **28** 990–1051. MR3784494 https://doi.org/10.1214/17-AAP1324

[21] Girvan, M. and Newman, M.E.J. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **99** 7821–7826. MR1908073 https://doi.org/10.1073/pnas.122653799

[22] Gopalan, P.K. and Blei, D.M. (2013). Efficient discovery of overlapping communities in massive networks. *Proc. Natl. Acad. Sci. USA* **110** 14534–14539. MR3105375 https://doi.org/10.1073/pnas.1221839110

[23] Holmgren, C., Juškevičius, T. and Kettle, N. (2017). Majority bootstrap percolation on $G(n,p)$. *Electron. J. Combin.* **24** Paper No. 1.1, 32 pp. MR3609171 https://doi.org/10.37236/6000

[24] Janson, S., Łuczak, T., Turova, T. and Vallier, T. (2012). Bootstrap percolation on the random graph $G_{n,p}$. *Ann. Appl. Probab.* **22** 1989–2047. MR3025687 https://doi.org/10.1214/11-AAP822

[25] Janson, S. et al. (2019). A modified bootstrap percolation on a random graph coupled with a lattice. *Discrete Appl. Math.* **258** 152–165.

[26] Karrer, B. and Newman, M.E.J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E (3)* **83** 016107, 10 pp. MR2788206 https://doi.org/10.1103/PhysRevE.83.016107

[27] Kempe, D., Kleinberg, J. and Tardos, E. (2003). Maximizing the spread of influence through a social network. In *ACM KDD*.

[28] Massoulié, L. Community detection thresholds and the weak Ramanujan property. In *STOC 2004*.

[29] Newman, M.E.J. (2002). Assortative mixing in networks. *Phys. Rev. Lett.* **89** 208701.

[30] Scalia-Tomba, G.-P. (1985). Asymptotic final-size distribution for some chain-binomial processes. *Adv. in Appl. Probab.* **17** 477–495. MR0798872 https://doi.org/10.2307/1427116

[31] Shrestha, M. and Moore, C. (2014). Message-passing approach for threshold models of behavior in networks. *Phys. Rev. E* **89** 022805.

[32] Torrisi, G.L., Garetto, M. and Leonardi, E. (2019). A large deviation approach to super-critical bootstrap percolation on the random graph $G_{n,p}$. *Stochastic Process. Appl.* **129** 1873–1902. MR3958417 https://doi.org/10.1016/j.spa.2018.06.006

[33] Torrisi, G.L., Garetto, M. and Leonardi, E. (2020). Boostrap percolation on the stochastic block model with k communities. Preprint. Available at arXiv:1812.09107.

[34] Torrisi, G.L., Garetto, M. and Leonardi, E. (2023). Supplement to "Bootstrap percolation on the stochastic block model." https://doi.org/10.3150/22-BEJ1475SUPP

[35] Turova, T.S. and Vallier, T. (2015). Bootstrap percolation on a graph with random and local connections. *J. Stat. Phys.* **160** 1249–1276. MR3375589 https://doi.org/10.1007/s10955-015-1294-x

[36] Watts, D.J. (2002). A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. USA* **99** 5766–5771. MR1896072 https://doi.org/10.1073/pnas.082090499

Bernoulli **29**(1), 2023, 725–756
https://doi.org/10.3150/22-BEJ1476

A trajectorial approach to relative entropy dissipation of McKean–Vlasov diffusions: Gradient flows and HWBI inequalities

BERTRAM TSCHIDERER[1,a] and LANE CHUN YEUNG[2,b]

[1] *Faculty of Mathematics, University of Vienna, Austria,* [a] *bertram.tschiderer@univie.ac.at*
[2] *Department of Industrial Engineering and Operations Research, Columbia University, New York, USA,* [b] *l.yeung@columbia.edu*

We formulate a trajectorial version of the relative entropy dissipation identity for McKean–Vlasov diffusions, extending recent results which apply to non-interacting diffusions. Our stochastic analysis approach is based on time-reversal of diffusions and Lions' differential calculus over Wasserstein space. It allows us to compute explicitly the rate of relative entropy dissipation along every trajectory of the underlying diffusion via the semimartingale decomposition of the corresponding relative entropy process. As a first application, we obtain a new interpretation of the gradient flow structure for the granular media equation, generalizing a formulation developed recently for the linear Fokker–Planck equation. Secondly, we show how the trajectorial approach leads to a new derivation of the HWBI inequality, which relates relative entropy (H), Wasserstein distance (W), barycenter (B) and Fisher information (I).

Keywords: Relative entropy dissipation; gradient flow; McKean–Vlasov diffusion; granular media equation; HWBI inequality

References

[1] Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W. and Temme, J. (2013). A trajectorial interpretation of Doob's martingale inequalities. *Ann. Appl. Probab.* **23** 1494–1505. MR3098440 https://doi.org/10.1214/12-aap878

[2] Agueh, M., Ghoussoub, N. and Kang, X. (2004). Geometric inequalities via a general comparison principle for interacting gases. *Geom. Funct. Anal.* **14** 215–244. MR2053603 https://doi.org/10.1007/s00039-004-0455-x

[3] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2401600

[4] Backhoff, J., Conforti, G., Gentil, I. and Léonard, C. (2020). The mean field Schrödinger problem: Ergodic behavior, entropy estimates and functional inequalities. *Probab. Theory Related Fields* **178** 475–530. MR4146543 https://doi.org/10.1007/s00440-020-00977-8

[5] Beiglböck, M. and Siorpaes, P. (2015). Pathwise versions of the Burkholder-Davis-Gundy inequality. *Bernoulli* **21** 360–373. MR3322322 https://doi.org/10.3150/13-BEJ570

[6] Benachour, S., Roynette, B., Talay, D. and Vallois, P. (1998). Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* **75** 173–201. MR1632193 https://doi.org/10.1016/S0304-4149(98)00018-0

[7] Benedetto, D., Caglioti, E., Carrillo, J.A. and Pulvirenti, M. (1998). A non-Maxwellian steady distribution for one-dimensional granular media. *J. Stat. Phys.* **91** 979–990. MR1637274 https://doi.org/10.1023/A:1023032000560

[8] Bolley, F., Gentil, I. and Guillin, A. (2013). Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.* **208** 429–445. MR3035983 https://doi.org/10.1007/s00205-012-0599-z

[9] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402

[10] Carmona, R. and Delarue, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. I. Mean Field FBSDEs, Control, and Games*. *Probability Theory and Stochastic Modelling* **83**. Cham: Springer. MR3752669

[11] Carrillo, J.A., Gvalani, R.S., Pavliotis, G.A. and Schlichting, A. (2020). Long-time behaviour and phase transitions for the McKean-Vlasov equation on the torus. *Arch. Ration. Mech. Anal.* **235** 635–690. MR4062483 https://doi.org/10.1007/s00205-019-01430-4

[12] Carrillo, J.A., Jüngel, A., Markowich, P.A., Toscani, G. and Unterreiter, A. (2001). Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.* **133** 1–82. MR1853037 https://doi.org/10.1007/s006050170032

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

[14] Carrillo, J.A., McCann, R.J. and Villani, C. (2006). Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.* **179** 217–263. MR2209130 https://doi.org/10.1007/s00205-005-0386-1

[15] Cattiaux, P. (2004). A pathwise approach of some classical inequalities. *Potential Anal.* **20** 361–394. MR2032116 https://doi.org/10.1023/B:POTA.0000009847.84908.6f

[16] Cattiaux, P. and Guillin, A. (2014). Semi log-concave Markov diffusions. In *Séminaire de Probabilités XLVI*. *Lecture Notes in Math.* **2123** 231–292. Cham: Springer. MR3330820 https://doi.org/10.1007/978-3-319-11970-0_9

[17] Cattiaux, P., Guillin, A. and Malrieu, F. (2008). Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** 19–40. MR2357669 https://doi.org/10.1007/s00440-007-0056-3

[18] Cordero-Erausquin, D., Gangbo, W. and Houdré, C. (2004). Inequalities for generalized entropy and optimal transportation. In *Recent Advances in the Theory and Applications of Mass Transport*. *Contemp. Math.* **353** 73–94. Providence, RI: Amer. Math. Soc. MR2079071 https://doi.org/10.1090/conm/353/06433

[19] Léonard, C. (2014). Some Properties of Path Measures. In *Sémin. Probab. XLVI*, (C. Donati-Martin, A. Lejay and A. Rouault, eds.). *Lecture Notes in Math.* **2123** 207–230. Switzerland: Springer International Publishing.

[20] Duong, M.H. and Tugaut, J. (2016). Stationary solutions of the Vlasov-Fokker-Planck equation: Existence, characterization and phase-transition. *Appl. Math. Lett.* **52** 38–45. MR3416384 https://doi.org/10.1016/j.aml.2015.08.003

[21] Duong, M.H. and Tugaut, J. (2018). The Vlasov-Fokker-Planck equation in non-convex landscapes: Convergence to equilibrium. *Electron. Commun. Probab.* **23** Paper No. 19, 10. MR3779816 https://doi.org/10.1214/18-ECP116

[22] Eldan, R., Lehec, J. and Shenfeld, Y. (2020). Stability of the logarithmic Sobolev inequality via the Föllmer process. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2253–2269. MR4116725 https://doi.org/10.1214/19-AIHP1038

[23] Eldan, R. and Mikulincer, D. (2020). Stability of the Shannon-Stam inequality via the Föllmer process. *Probab. Theory Related Fields* **177** 891–922. MR4126934 https://doi.org/10.1007/s00440-020-00967-w

[24] Erbar, M., Fathi, M., Laschos, V. and Schlichting, A. (2016). Gradient flow structure for McKean-Vlasov equations on discrete spaces. *Discrete Contin. Dyn. Syst.* **36** 6799–6833. MR3567821 https://doi.org/10.3934/dcds.2016096

[25] Feng, J. and Kurtz, T.G. (2006). *Large Deviations for Stochastic Processes*. *Mathematical Surveys and Monographs* **131**. Providence, RI: Amer. Math. Soc. MR2260560 https://doi.org/10.1090/surv/131

[26] Föllmer, H. (1985). An entropy approach to the time reversal of diffusion processes. In *Stochastic Differential Systems (Marseille-Luminy, 1984)*. *Lect. Notes Control Inf. Sci.* **69** 156–163. Berlin: Springer. MR0798318 https://doi.org/10.1007/BFb0005070

[27] Föllmer, H. (1986). Time reversal on Wiener space. In *Stochastic Processes—Mathematics and Physics (Bielefeld, 1984)*. *Lecture Notes in Math.* **1158** 119–129. Berlin: Springer. MR0838561 https://doi.org/10.1007/BFb0080212

[28] Föllmer, H. (1988). Random fields and diffusion processes. In *École D'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*. *Lecture Notes in Math.* **1362** 101–203. Berlin: Springer. MR0983373 https://doi.org/10.1007/BFb0086180

[29] Fontbona, J. and Jourdain, B. (2016). A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. *Ann. Probab.* **44** 131–170. MR3456334 https://doi.org/10.1214/14-AOP969

[30] Gardiner, C. (2009). *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed. *Springer Series in Synergetics*. Berlin: Springer. MR2676235

[31] Gentil, I., Léonard, C., Ripani, L. and Tamanini, L. (2020). An entropic interpolation proof of the HWI inequality. *Stochastic Process. Appl.* **130** 907–923. MR4046525 https://doi.org/10.1016/j.spa.2019.04.002

[32] Haussmann, U.G. and Pardoux, É. (1986). Time reversal of diffusions. *Ann. Probab.* **14** 1188–1205. MR0866342

[33] Herrmann, S. and Tugaut, J. (2010). Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Process. Appl.* **120** 1215–1246. MR2639745 https://doi.org/10.1016/j.spa.2010.03.009

[34] Hu, K., Ren, Z., Šiška, D. and Szpruch, Ł. (2021). Mean-field Langevin dynamics and energy landscape of neural networks. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 2043–2065. MR4328560 https://doi.org/10.1214/20-aihp1140

[35] Jordan, R. and Kinderlehrer, D. (1996). An extended variational principle. In *Partial Differential Equations and Applications*. *Lecture Notes in Pure and Applied Mathematics* **177** 187–200. New York: Dekker. MR1371591 https://doi.org/10.5006/1.3292113

[36] Jordan, R., Kinderlehrer, D. and Otto, F. (1998). The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** 1–17. MR1617171 https://doi.org/10.1137/S0036141096303359

[37] Karatzas, I. and Kardaras, C. (2021). *Portfolio Theory and Arbitrage: A Course in Mathematical Finance*. *Grad. Stud. Math.* **214**. Providence, Rhode Island: Amer. Math. Soc.

[38] Karatzas, I., Maas, J. and Schachermayer, W. (2021). Trajectorial dissipation and gradient flow for the relative entropy in Markov chains. *Commun. Inf. Syst.* **21** 481–536. MR4273512 https://doi.org/10.4310/CIS.2021.v21.n4.a1

[39] Karatzas, I., Schachermayer, W. and Tschiderer, B. (2020). Trajectorial Otto calculus. arXiv:1811.08686.

[40] Karatzas, I., Schachermayer, W. and Tschiderer, B. (2021). A trajectorial approach to the gradient flow properties of Langevin-Smoluchowski diffusions. *Teor. Veroyatn. Primen.* **66** 839–888. MR4331222 https://doi.org/10.4213/tvp55

[41] Lehec, J. (2013). Representation formula for the entropy and functional inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 885–899. MR3112438 https://doi.org/10.1214/11-aihp464

[42] Léonard, C. (2014). A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.* **34** 1533–1574. MR3121631 https://doi.org/10.3934/dcds.2014.34.1533

[43] Lions, P.L. (2007–2008). Théorie des jeux de champ moyen et applications. Lectures at the Collège de France. https://www.college-de-france.fr/site/pierre-louis-lions/course-2007-2008.htm.

[44] Liptser, R.S. and Shiryaev, A.N. (2001). *Statistics of Random Processes. I: General Theory*, expanded ed. *Applications of Mathematics (New York)* **5**. Berlin: Springer. MR1800857

[45] Malrieu, F. (2003). Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13** 540–560. MR1970276 https://doi.org/10.1214/aoap/1050689593

[46] McCann, R.J. (1997). A convexity principle for interacting gases. *Adv. Math.* **128** 153–179. MR1451422 https://doi.org/10.1006/aima.1997.1634

[47] McKean, H.P. Jr. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907

[48] Mei, S., Montanari, A. and Nguyen, P.-M. (2018). A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA* **115** E7665–E7671. MR3845070 https://doi.org/10.1073/pnas.1806579115

[49] Otto, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations* **26** 101–174. MR1842429 https://doi.org/10.1081/PDE-100002243

[50] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620 https://doi.org/10.1006/jfan.1999.3557

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

[52] Risken, H. (1984). *The Fokker-Planck Equation: Methods of Solution and Applications*. *Springer Series in Synergetics* **18**. Berlin: Springer. MR0749386 https://doi.org/10.1007/978-3-642-96807-5

[53] Rogers, L.C.G. (1985). Smooth transition densities for one-dimensional diffusions. *Bull. Lond. Math. Soc.* **17** 157–161. MR0806242 https://doi.org/10.1112/blms/17.2.157

[54] Tugaut, J. (2013). Convergence to the equilibria for self-stabilizing processes in double-well landscape. *Ann. Probab.* **41** 1427–1460. MR3098681 https://doi.org/10.1214/12-AOP749

[55] Tugaut, J. (2014). Phase transitions of McKean-Vlasov processes in double-wells landscape. *Stochastics* **86** 257–284. MR3180036 https://doi.org/10.1080/17442508.2013.775287

[56] Tugaut, J. (2014). Self-stabilizing processes in multi-wells landscape in $\mathbb{R}^d$-invariant probabilities. *J. Theoret. Probab.* **27** 57–79. MR3174216 https://doi.org/10.1007/s10959-012-0435-2

[57] Veretennikov, A.Yu. (2006). On ergodic measures for McKean-Vlasov stochastic equations. In *Monte Carlo and Quasi-Monte Carlo Methods 2004* 471–486. Berlin: Springer. MR2208726 https://doi.org/10.1007/3-540-31186-6_29

[58] Villani, C. (2003). *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Providence, RI: Amer. Math. Soc. MR1964483 https://doi.org/10.1090/gsm/058

[59] Villani, C. (2006). Mathematics of granular materials. *J. Stat. Phys.* **124** 781–822. MR2264625 https://doi.org/10.1007/s10955-006-9038-6

Bernoulli **29**(1), 2023, 757–784
https://doi.org/10.3150/22-BEJ1478

SDEs with critical time dependent drifts: Weak solutions

MICHAEL RÖCKNER[1,2,a] and GUOHUAN ZHAO[3,b]

[1]*Department of Mathematics, Bielefeld University, Germany*
[2]*Academy of Mathematics and Systems Science, Chinese Academy of Sciences (CAS), Beijing, P.R.China,*
[a]*roeckner@math.uni-bielefeld.de*
[3]*Department of Mathematics, Bielefeld University, Germany,* [b]*zhaoguohuan@gmail.com*

For $d \geq 3$, we prove that time-inhomogeneous stochastic differential equations driven by additive noises with drifts in critical Lebesgue space $L^q([0,T]; L^p(\mathbb{R}^d))$, where $(p,q) \in (d,\infty] \times [2,\infty)$ and $d/p + 2/q = 1$, or $(p,q) = (d,\infty)$ and $\operatorname{div} b \in L^\infty([0,T]; L^{d/2+\varepsilon}(\mathbb{R}^d))$, are well-posed. The weak uniqueness is obtained by solving corresponding Kolmogorov backward equations in some second-order Sobolev spaces, which is analytically interesting in itself.

Keywords: Weak solutions; Ladyzhenskaya–Prodi–Serrin condition; Kolmogorov equations; De Giorgi's method

References

[1] Bahouri, H., Chemin, J.-Y. and Danchin, R. (2011). *Fourier Analysis and Nonlinear Partial Differential Equations. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **343**. Heidelberg: Springer. MR2768550 https://doi.org/10.1007/978-3-642-16830-7

[2] Beck, L., Flandoli, F., Gubinelli, M. and Maurelli, M. (2019). Stochastic ODEs and stochastic linear PDEs with critical drift: Regularity, duality and uniqueness. *Electron. J. Probab.* **24** Paper No. 136, 72. MR4040996 https://doi.org/10.1214/19-ejp379

[3] Caffarelli, L.A. and Vasseur, A.F. (2010). The De Giorgi method for regularity of solutions of elliptic equations and its applications to fluid dynamics. *Discrete Contin. Dyn. Syst. Ser. S* **3** 409–427. MR2660718 https://doi.org/10.3934/dcdss.2010.3.409

[4] Cherny, A.S. (2002). On the uniqueness in law and the pathwise uniqueness for stochastic differential equations. *Theory Probab. Appl.* **46** 406–419.

[5] Constantin, P. and Iyer, G. (2008). A stochastic Lagrangian representation of the three-dimensional incompressible Navier–Stokes equations. *Comm. Pure Appl. Math.* **61** 330–345. MR2376844 https://doi.org/10.1002/cpa.20192

[6] Dong, H. and Du, D. (2009). The Navier–Stokes equations in the critical Lebesgue space. *Comm. Math. Phys.* **292** 811–827. MR2551795 https://doi.org/10.1007/s00220-009-0852-y

[7] Escauriaza, L., Seregin, G. and Šverák, V. (2003). Backward uniqueness for parabolic equations. *Arch. Ration. Mech. Anal.* **169** 147–157. MR2005639 https://doi.org/10.1007/s00205-003-0263-8

[8] Fedrizzi, E. and Flandoli, F. (2011). Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics* **83** 241–257. MR2810591 https://doi.org/10.1080/17442508.2011.553681

[9] Gallagher, I., Koch, G.S. and Planchon, F. (2013). A profile decomposition approach to the $L_t^\infty(L_x^3)$ Navier–Stokes regularity criterion. *Math. Ann.* **355** 1527–1559. MR3037023 https://doi.org/10.1007/s00208-012-0830-0

[10] Grafakos, L. (2008). *Classical Fourier Analysis*, 2nd ed. *Graduate Texts in Mathematics* **249**. New York: Springer. MR2445437

[11] Han, Q. and Lin, F. (2011). *Elliptic Partial Differential Equations* **1**. Providence: Am. Math. Soc.

[12] Kinzebulatov, D. and Madou, K.R. (2021). Stochastic equations with time-dependent singular drift. Preprint. Available at arXiv:2105.07312.

[13] Kinzebulatov, D. and Semënov, Y.A. (2021). Feller generators and stochastic differential equations with singular (form-bounded) drift. *Osaka J. Math.* **58** 855–883. MR4335377

[14] Kinzebulatov, D. and Semënov, Yu.A. (2020). Brownian motion with general drift. *Stochastic Process. Appl.* **130** 2737–2750. MR4080729 https://doi.org/10.1016/j.spa.2019.08.003

[15] Krylov, N.V. (2001). The heat equation in $L_q((0,T),L_p)$-spaces with weights. *SIAM J. Math. Anal.* **32** 1117–1141. MR1828321 https://doi.org/10.1137/S0036141000372039

[16] Krylov, N.V. (2021). On strong solutions of Itô's equations with $A \in W^1_d$ and $b \in L_d$. *Ann. Probab.* **49** 3142–3167. MR4348687 https://doi.org/10.1214/21-aop1525

[17] Krylov, N.V. (2020). On time inhomogeneous stochastic Itô equations with drift in L_{d+1}. *Ukraïn. Mat. Zh.* **72** 1232–1253. MR4207066 https://doi.org/10.37863/umzh.v72i9.6280

[18] Krylov, N.V. (2021). On diffusion processes with drift in L_d. *Probab. Theory Related Fields* **179** 165–199. MR4221656 https://doi.org/10.1007/s00440-020-01007-3

[19] Krylov, N.V. (2021). On stochastic equations with drift in L_d. *Ann. Probab.* **49** 2371–2398. MR4317707 https://doi.org/10.1214/21-aop1510

[20] Krylov, N.V. (2021). On stochastic Itô processes with drift in L_d. *Stochastic Process. Appl.* **138** 1–25. MR4252191 https://doi.org/10.1016/j.spa.2021.04.005

[21] Krylov, N.V. (2022). Some properties of solutions of Itô equations with drift in L_d. *Stochastic Process. Appl.* **147** 363–387. MR4382079 https://doi.org/10.1016/j.spa.2022.01.021

[22] Krylov, N.V. and Röckner, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** 154–196. MR2117951 https://doi.org/10.1007/s00440-004-0361-z

[23] Kufner, A. and Wannebo, A. (1995). An interpolation inequality involving Hölder norms. *Georgian Math. J.* **2** 603–612. MR1357997 https://doi.org/10.1007/BF02262857

[24] Lee, H. and Trutnau, G. (2021). Existence, uniqueness and ergodic properties for time-homogeneous Itô-SDEs with locally integrable drifts and Sobolev diffusion coefficients. *Tohoku Math. J. (2)* **73** 159–198. MR4278743 https://doi.org/10.2748/tmj.20200218

[25] Menoukeu-Pamen, O., Meyer-Brandis, T., Nilssen, T., Proske, F. and Zhang, T. (2013). A variational approach to the construction and Malliavin differentiability of strong solutions of SDE's. *Math. Ann.* **357** 761–799. MR3096525 https://doi.org/10.1007/s00208-013-0916-3

[26] Mohammed, S.-E.A., Nilssen, T.K. and Proske, F.N. (2015). Sobolev differentiable stochastic flows for SDEs with singular coefficients: Applications to the transport equation. *Ann. Probab.* **43** 1535–1576. MR3342670 https://doi.org/10.1214/14-AOP909

[27] Nam, K. (2020). Stochastic differential equations with critical drifts. *Stochastic Process. Appl.* **130** 5366–5393. MR4127332 https://doi.org/10.1016/j.spa.2020.03.010

[28] Nazarov, A.I. and Ural'tseva, N.N. (2012). The Harnack inequality and related properties for solutions of elliptic and parabolic equations with divergence-free lower-order coefficients. *St. Petersburg Math. J.* **23** 93–115.

[29] Nirenberg, L. (1959). On elliptic partial differential equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (3)* **13** 115–162. MR0109940

[30] Rezakhanlou, F. (2014). Regular flows for diffusions with rough drifts. Preprint. Available at arXiv:1405.5856.

[31] Röckner, M. and Zhao, G. (2021). SDEs with critical time dependent drifts: Strong solutions. Preprint.

[32] Sobolevskiĭ, P.E. (1977). Fractional powers of coercive-positive sums of operators. *Sibirsk. Mat. Zh.* **18** 637–657, 719. MR0454693

[33] Stroock, D.W. and Varadhan, S.R.S. (2007). *Multidimensional Diffusion Processes*. Berlin: Springer.

[34] Tartar, L. (1998). Imbedding theorems of Sobolev spaces into Lorentz spaces. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)* **1** 479–500. MR1662313

[35] Veretennikov, A.Ju. (1980). On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat. Sb.* **153** 434–452.

[36] Veretennikov, A.Ju. and Krylov, N.V. (1976). On explicit formulas for solutions of stochastic equations. *Math. USSR, Sb.* **29** 239–256.

[37] Xia, P., Xie, L., Zhang, X. and Zhao, G. (2020). $L^q(L^p)$-theory of stochastic differential equations. *Stochastic Process. Appl.* **130** 5188–5211. MR4108486 https://doi.org/10.1016/j.spa.2020.03.004

[38] YaWeing, J., Lv, G. and Wu, J.-L. (2017). On weak solutions of stochastic differential equations with sharp drift coefficients. Preprint. Available at arXiv:1711.05058.

[39] Zhang, X. (2005). Strong solutions of SDES with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.* **115** 1805–1818. MR2172887 https://doi.org/10.1016/j.spa.2005.06.003

[40] Zhang, X. (2010). A stochastic representation for backward incompressible Navier–Stokes equations. *Probab. Theory Related Fields* **148** 305–332. MR2653231 https://doi.org/10.1007/s00440-009-0234-6

[41] Zhang, X. (2011). Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.* **16** 1096–1116. MR2820071 https://doi.org/10.1214/EJP.v16-887

[42] Zhang, X. (2016). Stochastic differential equations with Sobolev diffusion and singular drift and applications. *Ann. Appl. Probab.* **26** 2697–2732. MR3563191 https://doi.org/10.1214/15-AAP1159

[43] Zhang, X. and Zhao, G. (2018). Singular Brownian diffusion processes. *Commun. Math. Stat.* **6** 533–581. MR3877717 https://doi.org/10.1007/s40304-018-0164-7

[44] Zhang, X. and Zhao, G. (2021). Stochastic Lagrangian path for Leray's solutions of 3D Navier–Stokes equations. *Comm. Math. Phys.* **381** 491–525. MR4207449 https://doi.org/10.1007/s00220-020-03888-w

[45] Zhao, G. (2019). Stochastic Lagrangian flows for SDEs with rough coefficients. Preprint. Available at arXiv:1911.05562.

[46] Zvonkin, A.K. (1974). A transformation of the phase space of a diffusion process that removes the drift. *Math. USSR, Sb.* **22** 129–149.

Bernoulli **29**(1), 2023, 785–814
https://doi.org/10.3150/22-BEJ1479

Linear multifractional stable sheets in the broad sense: Existence and joint continuity of local times

YUJIA DING[1,a], QIDI PENG[1,b] and YIMIN XIAO[2,c]

[1]*Institute of Mathematical Sciences, Claremont Graduate University, USA,* [a]*yujia.ding@cgu.edu,* [b]*qidi.peng@cgu.edu*
[2]*Department of Statistics and Probability, Michigan State University, USA,* [c]*xiaoy@msu.edu*

We introduce the notion of *linear multifractional stable sheets in the broad sense* (LMSS) with $\alpha \in (0,2]$, to include both linear multifractional Brownian sheets ($\alpha = 2$) and linear multifractional stable sheets ($\alpha < 2$). The purpose of the present paper is to study the existence and joint continuity of the local times of LMSS, and also the local Hölder condition of the local times in the set variable. Among the main results of this paper, Theorem 2.4 provides a *sufficient and necessary condition* for the existence of local times of LMSS; Theorem 3.1 shows a *sufficient condition* for the joint continuity of local times; and Theorem 4.1 proves a sharp local Hölder condition for the local times in the set variable. All these theorems improve significantly the existing results for the local times of multifractional Brownian sheets and linear multifractional stable sheets in the literature.

Keywords: Linear multifractional Brownian sheets; linear multifractional stable sheets; local times; joint continuity

References

[1] Ayache, A., Roueff, F. and Xiao, Y. (2007). Joint continuity of the local times of linear fractional stable sheets. *C. R. Math. Acad. Sci. Paris* **344** 635–640. MR2334075 https://doi.org/10.1016/j.crma.2007.03.028

[2] Ayache, A., Roueff, F. and Xiao, Y. (2007). Local and asymptotic properties of linear fractional stable sheets. *C. R. Math. Acad. Sci. Paris* **344** 389–394. MR2310675 https://doi.org/10.1016/j.crma.2007.01.017

[3] Ayache, A., Roueff, F. and Xiao, Y. (2009). Linear fractional stable sheets: Wavelet expansion and sample path properties. *Stochastic Process. Appl.* **119** 1168–1197. MR2508569 https://doi.org/10.1016/j.spa.2008.06.004

[4] Ayache, A., Wu, D. and Xiao, Y. (2008). Joint continuity of the local times of fractional Brownian sheets. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 727–748. MR2446295 https://doi.org/10.1214/07-AIHP131

[5] Ayache, A. and Xiao, Y. (2005). Asymptotic properties and Hausdorff dimensions of fractional Brownian sheets. *J. Fourier Anal. Appl.* **11** 407–439. MR2169474 https://doi.org/10.1007/s00041-005-4048-3

[6] Berman, S.M. (1970). Gaussian processes with stationary increments: Local times and sample function properties. *Ann. Math. Stat.* **41** 1260–1272. MR0272035 https://doi.org/10.1214/aoms/1177696901

[7] Boufoussi, B., Dozzi, M. and Guerbaz, R. (2006). On the local time of multifractional Brownian motion. *Stochastics* **78** 33–49. MR2219711 https://doi.org/10.1080/17442500600578073

[8] Dozzi, M. (2003). Occupation density and sample path properties of N-parameter processes. In *Topics in Spatial Stochastic Processes (Martina Franca, 2001). Lecture Notes in Math.* **1802** 127–166. Berlin: Springer. MR1975519 https://doi.org/10.1007/978-3-540-36259-3_4

[9] Ehm, W. (1981). Sample function properties of multiparameter stable processes. *Z. Wahrsch. Verw. Gebiete* **56** 195–228. MR0618272 https://doi.org/10.1007/BF00535741

[10] Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* **8** 1–67. Available at https://doi.org/10.1214/aop/1176994824.

[11] Khoshnevisan, D. (2002). *Multiparameter Processes*. *Springer Monographs in Mathematics: An Introduction to Random Fields*. New York: Springer. MR1914748 https://doi.org/10.1007/b97363

[12] Meerschaert, M., Wu, D. and Xiao, Y. (2008). Local times of multifractional Brownian sheets. *Bernoulli* **14** 865–898. MR2537815 https://doi.org/10.3150/08-BEJ126

[13] Nolan, J.P. (1989). Local nondeterminism and local times for stable processes. *Probab. Theory Related Fields* **82** 387–410. MR1001520 https://doi.org/10.1007/BF00339994

[14] Pitt, L.D. (1978). Local times for Gaussian vector fields. *Indiana Univ. Math. J.* **27** 309–330. MR0471055 https://doi.org/10.1512/iumj.1978.27.27024

[15] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes*. *Stochastic Modeling*. New York: CRC Press. MR1280932

[16] Shen, G., Yu, Q. and Li, Y. (2020). Local times of linear multifractional stable sheets. *Appl. Math. J. Chinese Univ. Ser. B* **35** 1–15. MR4078814 https://doi.org/10.1007/s11766-020-3548-x

[17] Stoev, S. and Taqqu, M.S. (2004). Stochastic properties of the linear multifractional stable motion. *Adv. in Appl. Probab.* **36** 1085–1115. MR2119856 https://doi.org/10.1239/aap/1103662959

[18] Stoev, S. and Taqqu, M.S. (2005). Path properties of the linear multifractional stable motion. *Fractals* **13** 157–178. MR2151096 https://doi.org/10.1142/S0218348X05002775

[19] Xiao, Y. (1997). Hölder conditions for the local times and the Hausdorff measure of the level sets of Gaussian random fields. *Probab. Theory Related Fields* **109** 129–157. MR1469923 https://doi.org/10.1007/s004400050128

[20] Xiao, Y. (1999). The Hausdorff dimension of the level sets of stable processes in random scenery. *Acta Sci. Math. (Szeged)* **65** 385–395. MR1702175

[21] 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 https://doi.org/10.1007/978-3-540-85994-9_5

[22] Xiao, Y. (2011). Properties of strong local nondeterminism and local times of stable random fields. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*. *Progress in Probability* **63** 279–308. Basel: Birkhäuser/Springer Basel AG. MR2857032 https://doi.org/10.1007/978-3-0348-0021-1_18

[23] Xiao, Y. and Zhang, T. (2002). Local times of fractional Brownian sheets. *Probab. Theory Related Fields* **124** 204–226. MR1936017 https://doi.org/10.1007/s004400200210

Bernoulli **29**(1), 2023, 815–838
https://doi.org/10.3150/22-BEJ1480

Non-homogeneous Poisson process intensity modeling and estimation using measure transport

TIN LOK JAMES NG[a] and ANDREW ZAMMIT-MANGION[b]

School of Computer Science and Statistics, Trinity College Dublin, Ireland, [a]*ngja@tcd.ie*
School of Mathematics and Applied Statistics, University of Wollongong, Australia, [b]*azm@uow.edu.au*

Non-homogeneous Poisson processes are used in a wide range of scientific disciplines, ranging from the environmental sciences to the health sciences. Often, the central object of interest in a point process is the underlying intensity function. Here, we present a general model for the intensity function of a non-homogeneous Poisson process using measure transport. The model is built from a flexible bijective mapping that maps from the underlying intensity function of interest to a simpler reference intensity function. We enforce bijectivity by modeling the map as a composition of multiple bijective maps that have increasing triangular structure, and show that the model exhibits an important approximation property. Estimation of the flexible mapping is accomplished within an optimization framework, wherein computations are efficiently done using tools originally designed to facilitate deep learning, and a graphics processing unit. Point process simulation and uncertainty quantification are straightforward to do with the proposed model. We demonstrate the potential benefits of our proposed method over conventional approaches to intensity modeling through various simulation studies. We also illustrate the use of our model on a real data set containing the locations of seismic events near Fiji since 1964.

Keywords: Poisson point process; intensity estimation; measure transport; deep neural network

References

Adams, R.P., Murray, I. and MacKay, D.J. C. (2009). Tractable nonparametric Bayesian inference in Poisson processes with Gaussian process intensities. In *Proceedings of the 26th Annual International Conference on Machine Learning* 9–16.

Bachl, F.E., Lindgren, F., Borchers, D.L. and Illian, J.B. (2019). Inlabru: An R package for Bayesian spatial modelling from ecological survey data. *Methods Ecol. Evol.* **10** 760–766.

Barron, A.R. (1994). Approximation and estimation bounds for artificial neural networks. *Mach. Learn.* **14** 115–133.

Bogachev, V.I., Kolesnikov, A.V. and Medvedev, K.V. (2005). Triangular transformations of measures. *Mat. Sb.* **196** 3–30. MR2144274 https://doi.org/10.1070/SM2005v196n03ABEH000882

Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Math. Control Signals Systems* **2** 303–314. MR1015670 https://doi.org/10.1007/BF02551274

de Miranda, J.C.S. and Morettin, P.A. (2011). Estimation of the intensity of non-homogeneous point processes via wavelets. *Ann. Inst. Statist. Math.* **63** 1221–1246. MR2830858 https://doi.org/10.1007/s10463-010-0283-8

Dias, R., Ferreira, C.S. and Garcia, N.L. (2008). Penalized maximum likelihood estimation for a function of the intensity of a Poisson point process. *Stat. Inference Stoch. Process.* **11** 11–34. MR2357551 https://doi.org/10.1007/s11203-006-9005-5

Diggle, P. (1985). A kernel method for smoothing point process data. *J. Roy. Statist. Soc. Ser. C* **34** 138–147.

Dinh, L., Krueger, D. and Bengio, Y. (2015). NICE: Non-linear independent components estimation. In *Workshop Track Proceedings of the 3rd International Conference on Learning Representations*.

Dinh, L., Sohl-Dickstein, J. and Bengio, S. (2017). Density estimation using real NVP. In *Conference Track Proceedings of the 5th International Conference on Learning Representations*.

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

Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika* **68** 589–599. MR0637776 https://doi.org/10.1093/biomet/68.3.589

Eldan, R. and Shamir, O. (2016). The power of depth for feedforward neural networks. In *Proceedings of the 29th Annual Conference on Learning Theory*. *Proceedings of Machine Learning Research* **49** 907–940.

Fine, T.L. (1999). *Feedforward Neural Network Methodology*. *Statistics for Engineering and Physical Science*. New York: Springer. MR1691898

Fogel, D.B. (1991). An information criterion for optimal neural network selection. *IEEE Trans. Neural Netw.* **2** 490–497.

Germain, M., Gregor, K., Murray, I. and Larochelle, H. (2015). MADE: Masked autoencoder for distribution estimation. In *Proceedings of the 32nd International Conference on Machine Learning* 881–889.

Hong, L.-L. and Guo, S.-W. (1995). Nonstationary Poisson model for earthquake occurrences. *Bull. Seismol. Soc. Amer.* **85** 814–824.

Hornik, K., Stinchcombe, M. and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Netw.* **2** 359–366.

Huang, C.-W., Krueger, D., Lacoste, A. and Courville, A. (2018). Neural autoregressive flows. In *Proceedings of the 35th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* **80** 2078–2087.

Illian, J.B., Sørbye, S.H. and Rue, H. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *Ann. Appl. Stat.* **6** 1499–1530. MR3058673 https://doi.org/10.1214/11-AOAS530

Kingma, D.P., Salimans, T., Jozefowicz, R., Chen, X., Sutskever, I. and Welling, M. (2016). Improved variational inference with inverse autoregressive flow. In *Advances in Neural Information Processing Systems 29* 4743–4751.

Kolaczyk, E.D. (1999). Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds. *Statist. Sinica* **9** 119–135. MR1678884

Letham, B., Letham, L.M. and Rudin, C. (2016). Bayesian inference of arrival rate and substitution behavior from sales transaction data with stockouts. In *Proceedings of the 22nd International Conference on Knowledge Discovery and Data Mining* 1695–1704.

Leung, F.H. F., Lam, H.K., Ling, S.H. and Tam, P.K. S. (2003). Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Trans. Neural Netw.* **14** 79–88.

Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. Roy. Statist. Soc. Ser. B* **73** 423–498.

Lindqvist, B.H. (2006). On the statistical modeling and analysis of repairable systems. *Statist. Sci.* **21** 532–551. MR2369984 https://doi.org/10.1214/088342306000000448

Lloyd, C., Gunter, T., Osborne, M.A. and Roberts, S.J. (2015). Variational inference for Gaussian process modulated Poisson processes. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning* 1814–1822.

Marzouk, Y., Moselhy, T., Parno, M. and Spantini, A. (2017). Sampling via measure transport: An introduction. In *Handbook of Uncertainty Quantification. Vol. 1, 2, 3* 785–825. Cham: Springer. MR3821485

McCann, R.J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80** 309–323. MR1369395 https://doi.org/10.1215/S0012-7094-95-08013-2

Møller, J., Syversveen, A.R. and Waagepetersen, R.P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482. MR1650019 https://doi.org/10.1111/1467-9469.00115

Nestoridis, V. and Stefanopoulos, V. (2007). Universal series and appoximate identities Technical Report.

Papamakarios, G., Pavlakou, T. and Murray, I. (2017). Masked autoregressive flow for density estimation. *Adv. Neural Inf. Process. Syst.* **30** 2338–2347.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L. and Lerer, A. (2017). Automatic differentiation in PyTorch. In *Advances Im Neural Information Processing Systems 30 Workshop on Autodiff*.

Raghu, M., Poole, B., Kleinberg, J., Ganguli, S. and Sohl-Dickstein, J. (2017). On the expressive power of deep neural networks. In *Proceedings of the 34th International Conference on Machine Learning*. *Proceedings of Machine Learning Research* **70** 2847–2854.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9

Taddy, M. and Kottas, A. (2010). Mixture modeling for marked Poisson processes. *Bayesian Anal.* **7** 335–362.

Tsai, J.-T., Chou, J.-H. and Liu, T.-K. (2006). Tuning the structure and parameters of a neural network by using hybrid Taguchi-genetic algorithm. *IEEE Trans. Neural Netw.* **17** 69–80. https://doi.org/10.1109/TNN.2005.860885

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

Weinan, E. and Wang, Q. (2018). Exponential convergence of the deep neural network approximation for analytic functions. *Sci. China Math.* **61** 1733–1740. MR3856963 https://doi.org/10.1007/s11425-018-9387-x

Zammit Mangion, A., Yuan, K., Kadirkamanathan, V., Niranjan, M. and Sanguinetti, G. (2011). Online variational inference for state-space models with point-process observations. *Neural Comput.* **23** 1967–1999. MR2839905 https://doi.org/10.1162/NECO_a_00156

Zhao, M. and Xie, M. (1996). On maximum likelihood estimation for a general non-homogeneous Poisson process. *Scand. J. Stat.* **23** 597–607. MR1439714

Bernoulli **29**(1), 2023, 839–874
https://doi.org/10.3150/22-BEJ1481

General construction and classes of explicit L^1-optimal couplings

GIOVANNI PUCCETTI[1,a] and LUDGER RÜSCHENDORF[2,b]

[1]*University of Milano, Via Conservatorio 7, 20122 Milano, Italy,* [a]*giovanni.puccetti@unimi.it*
[2]*University of Freiburg, Ernst-Zermelo-Straße 1, 79104 Freiburg, Germany,* [b]*ruschen@stochastik.uni-freiburg.de*

The main scope of this paper is to give some explicit classes of examples of L^1-optimal couplings. Optimal transportation w.r.t. the Kantorovich metric ℓ_1 (resp. the Wasserstein metric W_1) between two absolutely continuous measures is known since the basic papers of Kantorovich and Rubinstein (*Dokl. Akad. Nauk SSSR* **115** (1957) 1058–1061) and Sudakov (*Proc. Steklov Inst. Math.* **141** (1979) 1–178) to occur on rays induced by a decomposition of the basic space (and more generally to higher dimensional decompositions in the case of general measures) induced by the corresponding dual potentials. Several papers have given this kind of structural result and established existence and uniqueness of solutions in varying generality. Since the dual problems pose typically too strong challenges to be solved in explicit form, these structural results have so far been applied for the solution of few particular instances.

First, we give a self-contained review of some basic optimal coupling results and we propose and investigate in particular some basic principles for the construction of L^1-optimal couplings given by a reduction principle and some usable forms of the decomposition method. This reduction principle, together with symmetry properties of the reduced measures, gives a hint to the decomposition of the space into sectors and via the non crossing property of optimal transport leads to the choice of transportation rays. The optimality of the induced transports is then a consequence of the characterization results of optimal couplings.

Then, we apply these principles to determine in explicit form L^1-optimal couplings for several classes of examples of elliptical distributions. In particular, we give for the first time a general construction of L^1-optimal couplings between two bivariate Gaussian distributions. We also discuss optimality of special constructions like shifts and scalings, and provide an extended class of dual functionals allowing for the closed-form computation of the ℓ_1-metric or of accurate lower bounds of it in a variety of examples.

Keywords: Kantorovich ℓ_1-metric; L^1-Wasserstein distance; optimal mass transportation; optimal couplings; Gaussian distributions; Monge-Kantorovich problem; Kantorovich-Rubinstein Theorem

References

Alfonsi, A. and Jourdain, B. (2014). A remark on the optimal transport between two probability measures sharing the same copula. *Statist. Probab. Lett.* **84** 131–134. MR3131266 https://doi.org/10.1016/j.spl.2013.09.035

Ambrosio, L. (2003). Lecture notes on optimal transport problems. In *Mathematical Aspects of Evolving Interfaces (Funchal, 2000). Lecture Notes in Math.* **1812** 1–52. Berlin: Springer. MR2011032 https://doi.org/10.1007/978-3-540-39189-0_1

Ambrosio, L., Kirchheim, B. and Pratelli, A. (2004). Existence of optimal transport maps for crystalline norms. *Duke Math. J.* **125** 207–241. MR2096672 https://doi.org/10.1215/S0012-7094-04-12521-7

Ambrosio, L. and Pratelli, A. (2003). Existence and stability results in the L^1 theory of optimal transportation. In *Optimal Transportation and Applications (Martina Franca, 2001). Lecture Notes in Math.* **1813** 123–160. Berlin: Springer. MR2006307 https://doi.org/10.1007/978-3-540-44857-0_5

Bianchini, S. and Daneri, S. (2018). On Sudakov's type decomposition of transference plans with norm costs. *Mem. Amer. Math. Soc.* **251** vi+112. MR3748028 https://doi.org/10.1090/memo/1197

Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402

Caffarelli, L.A., Feldman, M. and McCann, R.J. (2002). Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. *J. Amer. Math. Soc.* **15** 1–26. MR1862796 https://doi.org/10.1090/S0894-0347-01-00376-9

Caravenna, L. (2011). A proof of Sudakov theorem with strictly convex norms. *Math. Z.* **268** 371–407. MR2805441 https://doi.org/10.1007/s00209-010-0677-6

Carlier, G., Oberman, A. and Oudet, E. (2015). Numerical methods for matching for teams and Wasserstein barycenters. *ESAIM Math. Model. Numer. Anal.* **49** 1621–1642. MR3423268 https://doi.org/10.1051/m2an/2015033

Cuesta-Albertos, J.A., Rüschendorf, L. and Tuero-Díaz, A. (1993). Optimal coupling of multivariate distributions and stochastic processes. *J. Multivariate Anal.* **46** 335–361. MR1240428 https://doi.org/10.1006/jmva.1993.1064

Dall'Aglio, G. (1956). Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (3)* **10** 35–74. MR0081577

Dietrich, H. (1988). Zur c-Konvexität und c-Subdifferenzierbarkeit von Funktionalen. *Optimization* **19** 355–371. MR0955448 https://doi.org/10.1080/02331938808843352

Dowson, D.C. and Landau, B.V. (1982). The Fréchet distance between multivariate normal distributions. *J. Multivariate Anal.* **12** 450–455. MR0666017 https://doi.org/10.1016/0047-259X(82)90077-X

Eckstein, S. and Kupper, M. (2021). Computation of optimal transport and related hedging problems via penalization and neural networks. *Appl. Math. Optim.* **83** 639–667. MR4239795 https://doi.org/10.1007/s00245-019-09558-1

Evans, L.C. and Gangbo, W. (1999). Differential equations methods for the Monge-Kantorovich mass transfer problem. *Mem. Amer. Math. Soc.* **137** viii+66. MR1464149 https://doi.org/10.1090/memo/0653

Fang, K.W. (2017). *Symmetric Multivariate and Related Distributions*. Boca Raton: CRC Press/CRC.

Gangbo, W. and McCann, R.J. (1995). Optimal maps in Monge's mass transport problem. *C. R. Acad. Sci. Paris Sér. I Math.* **321** 1653–1658. MR1367824

Gangbo, W. and McCann, R.J. (1996). The geometry of optimal transportation. *Acta Math.* **177** 113–161. MR1440931 https://doi.org/10.1007/BF02392620

Kantorovic, L.V. and Rubinstein, G.S. (1957). On a functional space and certain extremum problems. *Dokl. Akad. Nauk SSSR* **115** 1058–1061. MR0094707

Kantorovich, L.V. (1942). On the translocation of masses. *C. R. (Dokl.) Acad. Sci. URSS* **37** 199–201. MR0009619

Kantorovich, L.V. (1948). On a problem of Monge. *Uspekhi Mat. Nauk* **3** 225–226. In Russian.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications*. Cambridge: Cambridge Univ. Press. MR2038227 https://doi.org/10.1017/CBO9780511550683

Olkin, I. and Pukelsheim, F. (1982). The distance between two random vectors with given dispersion matrices. *Linear Algebra Appl.* **48** 257–263. MR0683223 https://doi.org/10.1016/0024-3795(82)90112-4

Orlova, D.Y., Zimmerman, N., Meehan, S., Meehan, C., Waters, J. et al. (2016). Earth mover's distance (EMD): A true metric for comparing biomarker expression levels in cell populations. *PLoS ONE* **11** e0151859.

Peyré, G. and Cuturi, M. (2018). Computational optimal transport. With applications to data sciences. *Found. Trends Mach. Learn.* **11** 1–262.

Puccetti, G. (2017). An algorithm to approximate the optimal expected inner product of two vectors with given marginals. *J. Math. Anal. Appl.* **451** 132–145. MR3619231 https://doi.org/10.1016/j.jmaa.2017.02.003

Rachev, S.T. and Rüschendorf, L. (1998). *Mass Transportation Problems. Vol. I-II*. New York: Springer.

Rubner, Y., Tomasi, C. and Guibas, L.J. (2000). The Earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* **40** 99–121.

Rüschendorf, L. (1991). Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals (Rome, 1990). Math. Appl.* **67** 151–187. Dordrecht: Kluwer Academic. MR1215951

Rüschendorf, L. (1995). Optimal solutions of multivariate coupling problems. *Appl. Math. (Warsaw)* **23** 325–338. MR1360058 https://doi.org/10.4064/am-23-3-325-338

Rüschendorf, L. (1996). Developments on Fréchet-bounds. In *Distributions with Fixed Marginals and Related Topics (Seattle, WA, 1993). Institute of Mathematical Statistics Lecture Notes—Monograph Series* **28** 273–296. Hayward, CA: IMS. MR1485538 https://doi.org/10.1214/lnms/1215452625

Rüschendorf, L. (1998). Wasserstein metric. In *Encyclopaedia of Mathematics. Supplement, I, II, III* (M. Hazewinkel, ed.) Kluwer Academic (1997–2001).

Rüschendorf, L. and Rachev, S.T. (1990). A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.* **32** 48–54. MR1035606 https://doi.org/10.1016/0047-259X(90)90070-X

Ruttenberg, B.E. and Singh, A.K. (2011). Indexing the Earth mover's distance using normal distributions. *Proc. VLDB Endow.* **5** 205–216.

Sudakov, V.N. (1979). Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.* (translation from Tr. Mat. Inst. Steklov 141 (1976)) **141** 1–178.

Trudinger, N.S. and Wang, X.-J. (2001). On the Monge mass transfer problem. *Calc. Var. Partial Differential Equations* **13** 19–31. MR1854255 https://doi.org/10.1007/PL00009922

Uckelmann, L. (1998). Über das Monge-Kantorovich Transportproblem und dessen Verallgemeinerungen Dissertation, Univ. Freiburg.

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