BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Six Number One February 2020 ISSN: 1350-7265

CONTENTS

HAUSER, R.A. and MATZINGER, H. 1
Microscopic path structure of optimally aligned random sequences

CHENG, D., CAMMAROTA, V., FANTAYE, Y., MARINUCCI, D. and SCHWARTZMAN, A. 31
Multiple testing of local maxima for detection of peaks on the (celestial) sphere

BERNHARD, H. and DAS, B. 61
Heavy-tailed random walks, buffered queues and hidden large deviations

GÖTZE, F. and SAMBALE, H. 93
Second order concentration via logarithmic Sobolev inequalities

CASTILLO, I. and SZABÓ, B. 127
Spike and slab empirical Bayes sparse credible sets

GEISS, C., LABART, C. and LUOTO, A. 159
Random walk approximation of BSDEs with Hölder continuous terminal condition

BELITSER, E. and NURUSHEV, N. 191
Needles and straw in a haystack: Robust confidence for possibly sparse sequences

MAZUR, S., OTRYAKHIN, D. and PODOLSKIJ, M. 226
Estimation of the linear fractional stable motion

BOGERD, K., CASTRO, R.M. and VAN DER HOFSTAD, R. 253
Cliques in rank-1 random graphs: The role of inhomogeneity

GAYNANOVA, I. 286
Prediction and estimation consistency of sparse multi-class penalized optimal scoring

ZWINGMANN, T. and HOLZMANN, H. 323
Weak convergence of quantile and expectile processes under general assumptions

GIORDANO, L.M., JOLIS, M. and QUER-SARDANYONS, L. 352
SPDEs with fractional noise in space: Continuity in law with respect to the Hurst index

DING, X. 387
High dimensional deformed rectangular matrices with applications in matrix denoising

SHI, C. and TANG, B. 418
Construction results for strong orthogonal arrays of strength three

(continued)

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

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Six Number One February 2020 ISSN: 1350-7265

CONTENTS

(continued)

MA, Z. and LI, X. 432
Subspace perspective on canonical correlation analysis: Dimension reduction and minimax rates

FOKIANOS, K., STØVE, B., TJØSTHEIM, D. and DOUKHAN, P. 471
Multivariate count autoregression

SHEN, Y. and WANG, Y. 500
Operator-scaling Gaussian random fields via aggregation

DAOUIA, A., GIRARD, S. and STUPFLER, G. 531
Tail expectile process and risk assessment

BEUTNER, E., BORDES, L. and DOYEN, L. 557
Consistent semiparametric estimators for recurrent event times models with application to virtual age models

PRIVAULT, N. and SERAFIN, G. 587
Normal approximation for sums of weighted U-statistics – application to Kolmogorov bounds in random subgraph counting

YANO, K. and KATO, K. 616
On frequentist coverage errors of Bayesian credible sets in moderately high dimensions

SCHNURR, A. 642
The fourth characteristic of a semimartingale

LIANG, M., SCHILLING, R.L. and WANG, J. 664
A unified approach to coupling SDEs driven by Lévy noise and some applications

MINSKER, S. and WEI, X. 694
Robust modifications of U-statistics and applications to covariance estimation problems

ISSOGLIO, E. and RUSSO, F. 728
A Feynman–Kac result via Markov BSDEs with generalised drivers

LEI, J. 767
Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces

BERNOULLI Volume 26 Number 1 February 2020 Pages 1–798 ISI/BS

Volume 26 Number 1   February 2020   ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope

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

Bernoulli Society for Mathematical Statistics and Probability

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

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

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

Executive Committee

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

In 2020 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **26**(1), 2020, 1–30
https://doi.org/10.3150/18-BEJ1053

Microscopic path structure of optimally aligned random sequences

RAPHAEL ANDREAS HAUSER[1,2] and HEINRICH MATZINGER[3]

[1]*Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom. E-mail: hauser@maths.ox.ac.uk; url: www.maths.ox.ac.uk*
[2]*Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, United Kingdom*
[3]*School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, USA. E-mail: matzi@math.gatech.edu; url: http://www.math.gatech.edu*

Considering optimal alignments of two i.i.d. random sequences of length n, we show that for Lebesgue-almost all scoring functions, almost surely the empirical distribution of aligned letter pairs in all optimal alignments converges to a unique limiting distribution as n tends to infinity. This result helps understanding the microscopic path structure of a special type of last-passage percolation problem with correlated weights, an area of long-standing open problems. Characterizing the microscopic path structure also yields robust alternatives to the use of optimal alignment scores alone for testing the homology of genetic sequences.

Keywords: convex geometry; large deviations; percolation theory; sequence alignment

References

[1] Alexander, K.S. (1994). The rate of convergence of the mean length of the longest common subsequence. *Ann. Appl. Probab.* **4** 1074–1082. MR1304773

[2] Amsalu, S., Hauser, R. and Matzinger, H. (2014). A Monte Carlo approach to the fluctuation problem in optimal alignments of random strings. *Markov Process. Related Fields* **20** 107–144. MR3185558

[3] Baxevanis, A.D. and Ouellette, B.F.F. (2005). *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins*. New York: Wiley.

[4] Bertsekas, D.P. and Yu, H. (2011). A unifying polyhedral approximation framework for convex optimization. *SIAM J. Optim.* **21** 333–360. MR2783219 https://doi.org/10.1137/090772204

[5] Borwein, J.M. and Lewis, A.S. (2000). *Convex Analysis and Nonlinear Optimization: Theory and Examples*. *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC* **3**. New York: Springer. MR1757448 https://doi.org/10.1007/978-1-4757-9859-3

[6] Boutet de Monvel, J. (1999). Extensive simulations for longest common subsequences. *Eur. Phys. J. B* **7** 293–308.

[7] Carathéodory, C. (1911). Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen. *Rend. Circ. Mat. Palermo* (2) **32** 193–217.

[8] Chvatal, V. and Sankoff, D. (1975). Longest common subsequences of two random sequences. *J. Appl. Probab.* **12** 306–315. MR0405531 https://doi.org/10.2307/3212444

[9] Durringer, C., Hauser, R. and Matzinger, H. (2008). Approximation to the mean curve in the LCS problem. *Stochastic Process. Appl.* **118** 629–648. MR2394846 https://doi.org/10.1016/j.spa.2007.05.010

[10] Fekete, M. (1923). Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Math. Z.* **17** 228–249. MR1544613 https://doi.org/10.1007/BF01504345

1350-7265

[11] Gruber, P.M. (1993). History of convexity. In *Handbook of Convex Geometry*, *Vol*. *A*, *B* (P.M. Gruber, and J.M. Wills, eds.) 1–15. Amsterdam: North-Holland. MR1242974

[12] Hausdorff, F. (1949). *Grundzüge der Mengenlehre*. New York, NY: Chelsea Publishing Company. MR0031025

[13] Hauser, R., Martínez, S. and Matzinger, H. (2006). Large deviations-based upper bounds on the expected relative length of longest common subsequences. *Adv. in Appl*. *Probab*. **38** 827–852. MR2256880 https://doi.org/10.1239/aap/1158685004

[14] Hauser, R. and Matzinger, H. (2013). Letter change bias and local uniqueness in optimal sequence alignments. *J*. *Stat*. *Phys*. **153** 512–529. MR3107656 https://doi.org/10.1007/s10955-013-0819-4

[15] Hauser, R., Matzinger, H. and Popescu, I. (2018). An upper bound on the convergence rate of a second functional in optimal sequence alignment. *Bernoulli* **24** 971–992. MR3706783 https://doi.org/10.3150/16-BEJ823

[16] Hauser, R.A. and Matzinger, H. (2020). Supplement to "Microscopic path structure of optimally aligned random sequences." https://doi.org/10.3150/18-BEJ1053SUPP.

[17] Hirmo, E., Lember, J. and Matzinger, H. (2012). Detecting the homology of DNA-sequences based on the variety of optimal alignments: A case study. Available at arXiv:1210.3771.

[18] Howard, C.D. (2004). Probability on discrete structures. In *Encyclopaedia of Mathematical Sciences* (H. Kesten, ed.) **110**. Berlin: Springer. MR2023649 https://doi.org/10.1007/978-3-662-09444-0

[19] Jerrum, M. (2003). *Counting*, *Sampling and Integrating*: *Algorithms and Complexity*. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR1960003 https://doi.org/10.1007/978-3-0348-8005-3

[20] Karlin, S. and Altschul, S.F. (1990). Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **87** 2264–2268.

[21] Kingman, J.F.C. (1973). Subadditive ergodic theory. *Ann*. *Probab*. **1** 883–909. MR0356192 https://doi.org/10.1214/aop/1176996798

[22] Krein, M. and Milman, D. (1940). On extreme points of regular convex sets. *Studia Math*. **9** 133–138. MR0004990 https://doi.org/10.4064/sm-9-1-133-138

[23] Lember, J. and Matzinger, H. (2009). Standard deviation of the longest common subsequence. *Ann*. *Probab*. **37** 1192–1235. MR2537552 https://doi.org/10.1214/08-AOP436

[24] McDiarmid, C. (1989). On the method of bounded differences. In *Surveys in Combinatorics*, 1989 (*Norwich*, 1989). *London Mathematical Society Lecture Note Series* **141** 148–188. Cambridge: Cambridge Univ. Press. MR1036755

[25] Mil'man, D. (1947). Characteristics of extremal points of regularly convex sets. *Dokl*. *Akad*. *Nauk SSSR* **57** 119–122. MR0022313

[26] Minkowski, H. (1968). *Geometrie der Zahlen*. *Bibliotheca Mathematica Teubneriana*, *Band* 40. New York: Johnson Reprint Corp. MR0249269

[27] Needleman, S.B. and Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J*. *Mol*. *Biol*. **48** 443–453.

[28] Rademacher, H. (1919). Über partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale. *Math*. *Ann*. **79** 340–359. MR1511935 https://doi.org/10.1007/BF01498415

[29] Rockafellar, R.T. (1997). *Convex Analysis*. *Princeton Landmarks in Mathematics*. Princeton, NJ: Princeton Univ. Press. MR1451876

[30] Steele, J.M. (1986). An Efron–Stein inequality for nonsymmetric statistics. *Ann*. *Statist*. **14** 753–758. MR0840528 https://doi.org/10.1214/aos/1176349952

[31] Steele, J.M. (1989). Kingman's subadditive ergodic theorem. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **25** 93–98. MR0995293

[32] Steele, J.M. (1997). *Probability Theory and Combinatorial Optimization*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **69**. Philadelphia, PA: SIAM. MR1422018 https://doi.org/10.1137/1.9781611970029
[33] Waterman, M. (1994). Estimating statistical significance of sequence alignments. *Philos*. *Trans*. *R*. *Soc*. *Lond*. *B*, *Biol*. *Sci*. **344** 383–390.
[34] Waterman, M. (1995). *Introduction to Computational Biology*. London: Chapman & Hall.

Bernoulli **26**(1), 2020, 31–60
https://doi.org/10.3150/18-BEJ1068

Multiple testing of local maxima for detection of peaks on the (celestial) sphere

DAN CHENG[1], VALENTINA CAMMAROTA[2], YABEBAL FANTAYE[3], DOMENICO MARINUCCI[4] and ARMIN SCHWARTZMAN[5]

[1]*School of Mathematical and Statistical Sciences, Arizona State University, Tempe, USA. E-mail: cheng.stats@gmail.com*
[2]*Department of Statistics, Sapienza University of Rome, Rome, Italy. E-mail: valentina.cammarota@uniroma1.it*
[3]*African Institute for Mathematical Sciences and Department of Mathematics, University of Stellenbosch, Stellenbosch, South Africa. E-mail: yabi@aims.ac.za*
[4]*Department of Mathematics, University of Rome Tor Vergata, Rome, Italy. E-mail: marinucc@mat.uniroma2.it*
[5]*Division of Biostatistics, University of California, San Diego, USA. E-mail: armins@ucsd.edu*

We present a topological multiple testing scheme for detecting peaks on the sphere under isotropic Gaussian noise, where tests are performed at local maxima of the observed field filtered by the spherical needlet transform. Our setting is different from the standard Euclidean large domain asymptotic framework, yet highly relevant to realistic experimental circumstances for some important areas of application in astronomy, namely point-source detection in cosmic Microwave Background radiation (CMB) data. Motivated by this application, we shall focus on cases where a single realization of a smooth isotropic Gaussian random field on the sphere is observed, and a number of well-localized signals are superimposed on such background field. The proposed algorithms, combined with the Benjamini–Hochberg procedure for thresholding p-values, provide asymptotic control of the False Discovery Rate (FDR) and power consistency as the signal strength and the frequency of the needlet transform get large.

Keywords: CMB; false discovery rate; Gaussian random fields; height distribution; needlet transform; overshoot distribution; power; sphere

References

[1] Baldi, P., Kerkyacharian, G., Marinucci, D. and Picard, D. (2009). Asymptotics for spherical needlets. *Ann. Statist.* **37** 1150–1171. MR2509070 https://doi.org/10.1214/08-AOS601
[2] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57** 289–300. MR1325392
[3] Cammarota, V., Marinucci, D. and Wigman, I. (2016). On the distribution of the critical values of random spherical harmonics. *J. Geom. Anal.* **26** 3252–3324. MR3544960 https://doi.org/10.1007/s12220-015-9668-5
[4] Cheng, D., Cammarota, V., Fantaye, Y., Marinucci, D. and Schwartzman, A. (2020). Supplement to "Multiple testing of local maxima for detection of peaks on the (celestial) sphere." https://doi.org/10.3150/18-BEJ1068SUPP.
[5] Cheng, D. and Schwartzman, A. (2015). Distribution of the height of local maxima of Gaussian random fields. *Extremes* **18** 213–240. MR3351815 https://doi.org/10.1007/s10687-014-0211-z

[6] Cheng, D. and Schwartzman, A. (2017). Multiple testing of local maxima for detection of peaks in random fields. *Ann. Statist.* **45** 529–556. MR3650392 https://doi.org/10.1214/16-AOS1458
[7] Cheng, D. and Schwartzman, A. (2018). Expected number and height distribution of critical points of smooth isotropic Gaussian random fields. *Bernoulli* **24** 3422–3446. MR3788177 https://doi.org/10.3150/17-BEJ964
[8] Delabrouille, J. et al. (2013). The pre-launch Planck Sky Model: A model of sky emission at submillimetre to centimetre wavelengths. *Astron. Astrophys.* **553** A96.
[9] Durrer, R. (2008). *The Cosmic Microwave Background*. Cambridge: Cambridge Univ. Press.
[10] Geller, D. and Mayeli, A. (2009). Continuous wavelets on compact manifolds. *Math. Z.* **262** 895–927. Available at arXiv:0811.4440. MR2511756 https://doi.org/10.1007/s00209-008-0405-7
[11] Geller, D. and Mayeli, A. (2009). Nearly tight frames and space-frequency analysis on compact manifolds. *Math. Z.* **263** 235–264. Available at arXiv:0706.3642. MR2534117 https://doi.org/10.1007/s00209-008-0406-6
[12] Górski, K.M., Hivon, E., Banday, A.J., Wandelt, B.D., Hansen, F.K., Reinecke, M. and Bartelmann, M. (2005). HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *Astrophys. J.* **699** 759–771.
[13] Lan, X. and Marinucci, D. (2009). On the dependence structure of wavelet coefficients for spherical random fields. *Stochastic Process. Appl.* **119** 3749–3766. MR2568294 https://doi.org/10.1016/j.spa.2009.07.005
[14] Loh, W.-L. (2005). Fixed-domain asymptotics for a subclass of Matérn-type Gaussian random fields. *Ann. Statist.* **33** 2344–2394. MR2211089 https://doi.org/10.1214/009053605000000516
[15] Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications. London Mathematical Society Lecture Note Series* **389**. Cambridge: Cambridge Univ. Press. MR2840154 https://doi.org/10.1017/CBO9780511751677
[16] Marinucci, D. and Peccati, G. (2013). Mean-square continuity on homogeneous spaces of compact groups. *Electron. Commun. Probab.* **18** no. 37, 10. MR3064996 https://doi.org/10.1214/ECP.v18-2400
[17] Marinucci, D., Pietrobon, D., Balbi, A., Baldi, P., Cabella, P., Kerkyacharian, G., Natoli, P., Picard, D. and Vittorio, N. (2008). Spherical needlets for CMB data analysis. *Mon. Not. R. Astron. Soc.* **383** 539–545.
[18] Mayeli, A. (2010). Asymptotic uncorrelation for Mexican needlets. *J. Math. Anal. Appl.* **363** 336–344. MR2559069 https://doi.org/10.1016/j.jmaa.2009.07.044
[19] Narcowich, F.J., Petrushev, P. and Ward, J.D. (2006). Localized tight frames on spheres. *SIAM J. Math. Anal.* **38** 574–594. MR2237162 https://doi.org/10.1137/040614359
[20] Planck Collaboration (2016). Planck 2015 results. I. Overview of products and scientific results. *Astron. Astrophys.* **594** A1. Available at arXiv:1507.02704.
[21] Planck Collaboration (2016). Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *Astron. Astrophys.* **594** A11. Available at arXiv:1507.02704.
[22] Planck Collaboration (2016). Planck 2015. XXVI. The second Planck catalogue of compact sources. *Astron. Astrophys.* **594** A26.
[23] Schwartzman, A., Gavrilov, Y. and Adler, R.J. (2011). Multiple testing of local maxima for detection of peaks in 1D. *Ann. Statist.* **39** 3290–3319. MR3012409 https://doi.org/10.1214/11-AOS943
[24] Scodeller, S., Hansen, F.K. and Marinucci, D. (2012). Detection of new point sources in WMAP 7 year data using internal templates and needlets. *Astrophys. J.* **753** 27. Available at arXiv:1201.5852.
[25] Scodeller, S., Rudjord, O., Hansen, F.K., Marinucci, D., Geller, D. and Mayeli, A. (2011). Introducing Mexican needlets for CMB analysis: Issues for practical applications and comparison with standard needlets. *Astrophys. J.* **733** 121.

Bernoulli **26**(1), 2020, 61–92
https://doi.org/10.3150/18-BEJ1081

Heavy-tailed random walks, buffered queues and hidden large deviations

HARALD BERNHARD* and BIKRAMJIT DAS**

Singapore University of Technology and Design, Pillar of Engineering Systems and Design, 8 Somapah Road, Singapore 487372. E-mail: *bernhard.harald@gmail.com; **bikram@sutd.edu.sg

It is well-known that large deviations of random walks driven by independent and identically distributed heavy-tailed random variables are governed by the so-called principle of one large jump. We note that further subtleties hold for such random walks in the large deviations scale which we call hidden large deviations. Our results are illustrated using two examples. First, we apply this idea in the context of queueing processes with heavy-tailed service times and study approximations of probabilities of severe congestion times for (buffered) queues. We exhibit our techniques by using limit measures from different large deviation regimes to provide a unified estimate of rare event probabilities in a simulated queue. Furthermore, we use our result to provide probability estimates of rare events governed by more than one jump in case the innovations of a random walk have infinite mean.

Keywords: buffered queues; heavy-tails; large deviations; regular variation

References

[1] Andersen, L.N., Asmussen, S., Glynn, P.W. and Pihlsgård, M. (2015). Lévy processes with two-sided reflection. In *Lévy Matters*. *V*. *Lecture Notes in Math*. **2149** 67–182. Cham: Springer. MR3467345 https://doi.org/10.1007/978-3-319-23138-9_2

[2] Anderson, P.L. and Meerschaert, M.M. (1998). Modeling river flows with heavy tails. *Water Resour*. *Res*. **34** 2271–2280.

[3] Asmussen, S. (2003). *Applied Probability and Queues*: *Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. New York: Springer. MR1978607

[4] Asmussen, S., Binswanger, K. and Højgaard, B. (2000). Rare events simulation for heavy-tailed distributions. *Bernoulli* **6** 303–322. MR1748723 https://doi.org/10.2307/3318578

[5] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J*. *Amer*. *Statist*. *Assoc*. **57** 33–45.

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

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR1015093

[8] Blanchet, J.H. and Liu, J. (2008). State-dependent importance sampling for regularly varying random walks. *Adv*. *in Appl*. *Probab*. **40** 1104–1128. MR2488534

[9] Bollobás, B., Borgs, C., Chayes, J. and Riordan, O. (2003). Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (*Baltimore*, *MD*, 2003) 132–139. New York: ACM. MR1974912

1350-7265

[10] Boxma, O.J. and Cohen, J.W. (1999). The single server queue: Heavy tails and heavy traffic. In *Self-Similar Network Traffic and Performance Evaluation* (K. Park and W. Willinger, eds.) Wiley, New York.

[11] Crovella, M., Bestavros, A. and Taqqu, M.S. (1999). Heavy-tailed probability distributions in the world wide web. In *A Practical Guide to Heavy Tails*: *Statistical Techniques for Analysing Heavy Tailed Distributions* (M.S. Taqqu, R. Adler and R. Feldman, eds.). Boston: Birkhäuser.

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

[13] Das, B., Mitra, A. and Resnick, S. (2013). Living on the multidimensional edge: Seeking hidden risks using regular variation. *Adv. in Appl. Probab.* **45** 139–163. MR3077544

[14] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory*: *An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3

[15] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Berlin: Springer. Corrected reprint of the second (1998) edition. MR2571413 https://doi.org/10.1007/978-3-642-03311-7

[16] Denisov, D., Dieker, A.B. and Shneer, V. (2008). Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.* **36** 1946–1991. MR2440928 https://doi.org/10.1214/07-AOP382

[17] Durrett, R. (2010). *Random Graph Dynamics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **20**. Cambridge: Cambridge Univ. Press. MR2656427

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

[19] Ganesh, A., O'Connell, N. and Wischik, D. (2004). *Big Queues*. *Lecture Notes in Math*. **1838**. Berlin: Springer. MR2045489 https://doi.org/10.1007/b95197

[20] Hult, H. and Lindskog, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process*. *Appl*. **115** 249–274. MR2111194 https://doi.org/10.1016/j.spa.2004.09.003

[21] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ*. *Inst*. *Math*. (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910 https://doi.org/10.2298/PIM0694121H

[22] Hult, H., Lindskog, F., Mikosch, T. and Samorodnitsky, G. (2005). Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.* **15** 2651–2680. MR2187307 https://doi.org/10.1214/105051605000000502

[23] Ibragimov, R., Jaffee, D. and Walden, J. (2011). Diversification disasters. *J. Financ. Econ.* **99** 333–348.

[24] Janssen, A. and Drees, H. (2016). A stochastic volatility model with flexible extremal dependence structure. *Bernoulli* **22** 1448–1490. MR3474822 https://doi.org/10.3150/15-BEJ699

[25] Jelenković, P. and Momčilović, P. (2003). Asymptotic loss probability in a finite buffer fluid queue with heterogeneous heavy-tailed on-off processes. *Ann. Appl. Probab.* **13** 576–603. MR1970278 https://doi.org/10.1214/aoap/1050689595

[26] Jelenković, P.R. (1999). Subexponential loss rates in a $GI/GI/1$ queue with applications. *Queueing Syst*. **33** 91–123. MR1748640 https://doi.org/10.1023/A:1019167927407

[27] Lindley, D.V. (1952). The theory of queues with a single server. *Proc*. *Camb*. *Philos*. *Soc*. **48** 277–289. MR0046597 https://doi.org/10.1017/s0305004100027638

[28] Lindskog, F., Resnick, S.I. and Roy, J. (2014). Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probab. Surv.* **11** 270–314. MR3271332 https://doi.org/10.1214/14-PS231

[29] Mansfield, P., Rachev, S.T. and Samorodnitsky, G. (2001). Long strange segments of a stochastic process. *Ann. Appl. Probab.* **11** 878–921. MR1865027 https://doi.org/10.1214/aoap/1015345352
[30] Mikosch, T. and Wintenberger, O. (2013). Precise large deviations for dependent regularly varying sequences. *Probab. Theory Related Fields* **156** 851–887. MR3078288 https://doi.org/10.1007/s00440-012-0445-0
[31] Nagaev, A.V. (1969). Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. I. *Teor. Veroyatn. Primen.* **14** 51–63. MR0247651
[32] Nagaev, A.V. (1969). Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. II. *Teor. Veroyatn. Primen.* **14** 203–216. MR0247652
[33] Nagaev, A.V. (1969). Limit theorems that take into account large deviations when Cramér's condition is violated. *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk* **13** 17–22. MR0282396
[34] Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5** 303–336. MR2002121 https://doi.org/10.1023/A:1025148622954
[35] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424
[36] Rhee, C.-H., Blanchet, J. and Zwart, B. (2017). Sample path large deviations for heavy-tailed Lévy processes and random walks. ArXiv E-prints arXiv:1606.02795v2.
[37] Smith, R.L. (2003). Statistics of extremes, with applications in environment, insurance and finance. In *SemStat: Seminaire Europeen de Statistique, Exteme Values in Finance, Telecommunications, and the Environment* (B. Finkenstadt and H. Rootzén, eds.) 1–78. London: Chapman-Hall.
[38] Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues. Springer Series in Operations Research*. New York: Springer. MR1876437
[39] Zwart, A.P. (2000). A fluid queue with a finite buffer and subexponential input. *Adv. in Appl. Probab.* **32** 221–243. MR1765162 https://doi.org/10.1239/aap/1013540031
[40] Zwart, A.P. (2001). Tail asymptotics for the busy period in the $GI/G/1$ queue. *Math. Oper. Res.* **26** 485–493. MR1849881 https://doi.org/10.1287/moor.26.3.485.10584

Bernoulli **26**(1), 2020, 93–126
https://doi.org/10.3150/19-BEJ1118

Second order concentration via logarithmic Sobolev inequalities

FRIEDRICH GÖTZE* and HOLGER SAMBALE**

Department of Mathematics, Bielefeld University, Postbox 100131, 33501 Bielefeld, Germany.
E-mail: *goetze@math.uni-bielefeld.de; **hsambale@math.uni-bielefeld.de

We show sharpened forms of the concentration of measure phenomenon centered at first order stochastic expansions. The bound are based on second order difference operators and second order derivatives. Applications to functions on the discrete cube and stochastic Hoeffding type expansions in mathematical statistics are studied as well as linear eigenvalue statistics in random matrix theory.

Keywords: bootstrap approximation; concentration of measure phenomenon; functions on the discrete cube; Hoeffding decomposition; logarithmic Sobolev inequalities

References

[1] Adamczak, R. and Wolff, P. (2015). Concentration inequalities for non-Lipschitz functions with bounded derivatives of higher order. *Probab. Theory Related Fields* **162** 531–586. MR3383337 https://doi.org/10.1007/s00440-014-0579-3

[2] Aida, S., Masuda, T. and Shigekawa, I. (1994). Logarithmic Sobolev inequalities and exponential integrability. *J. Funct. Anal.* **126** 83–101. MR1305064 https://doi.org/10.1006/jfan.1994.1142

[3] Arcones, M.A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426

[4] Bickel, P.J., Götze, F. and van Zwet, W.R. (1997). Resampling fewer than n observations: Gains, losses, and remedies for losses. *Statist. Sinica* **7** 1–31. MR1441142

[5] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2017). Second-order concentration on the sphere. *Commun. Contemp. Math.* **19** 1650058, 20. MR3670792 https://doi.org/10.1142/S0219199716500589

[6] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772 https://doi.org/10.1006/jfan.1998.3326

[7] Bobkov, S.G. and Götze, F. (2007). Concentration inequalities and limit theorems for randomized sums. *Probab. Theory Related Fields* **137** 49–81. MR2278452 https://doi.org/10.1007/s00440-006-0500-9

[8] Bobkov, S.G. and Götze, F. (2010). Concentration of empirical distribution functions with applications to non-i.i.d. models. *Bernoulli* **16** 1385–1414. MR2759184 https://doi.org/10.3150/10-BEJ254

[9] Bobkov, S.G., Götze, F. and Sambale, H. (2018). Higher order concentration of measure. *Commun. Contemp. Math.*, online first. https://doi.org/10.1142/S0219199718500438.

[10] Bobkov, S.G. and Ledoux, M. (2000). From Brunn–Minkowski to Brascamp–Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.* **10** 1028–1052. MR1800062 https://doi.org/10.1007/PL00001645

[11] Borodin, A.N. and Ibragimov, I.A. (1994). Limit theorems for functionals of random walks. *Tr. Mat. Inst. Steklova* **195** 286. MR1368394

[12] Boucheron, S., Lugosi, G. and Massart, P. (2003). Concentration inequalities using the entropy method. *Ann*. *Probab*. **31** 1583–1614. MR1989444 https://doi.org/10.1214/aop/1055425791
[13] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities*. Oxford: Oxford Univ. Press. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001
[14] Cordero-Erausquin, D., Fradelizi, M. and Maurey, B. (2004). The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems. *J*. *Funct*. *Anal*. **214** 410–427. MR2083308 https://doi.org/10.1016/j.jfa.2003.12.001
[15] de la Peña, V.H. and Giné, E. (1999). *Decoupling*: *From Dependence to Independence*. *Probability and Its Applications* (*New York*). New York: Springer. MR1666908 https://doi.org/10.1007/978-1-4612-0537-1
[16] Götze, F. (1993). IMS Bulletin.
[17] Götze, F. and Venker, M. (2014). Local universality of repulsive particle systems and random matrices. *Ann*. *Probab*. **42** 2207–2242. MR3265167 https://doi.org/10.1214/13-AOP844
[18] Guionnet, A. and Zeitouni, O. (2000). Concentration of the spectral measure for large matrices. *Electron*. *Commun*. *Probab*. **5** 119–136. MR1781846 https://doi.org/10.1214/ECP.v5-1026
[19] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann*. *Math*. *Stat*. **19** 293–325. MR0026294 https://doi.org/10.1214/aoms/1177730196
[20] Johansson, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math*. *J*. **91** 151–204. MR1487983 https://doi.org/10.1215/S0012-7094-98-09108-6
[21] Khorunzhy, A.M., Khoruzhenko, B.A. and Pastur, L.A. (1996). Asymptotic properties of large random matrices with independent entries. *J*. *Math*. *Phys*. **37** 5033–5060. MR1411619 https://doi.org/10.1063/1.531589
[22] Ledoux, M. (1995/97). On Talagrand's deviation inequalities for product measures. *ESAIM Probab*. *Stat*. **1** 63–87. MR1399224 https://doi.org/10.1051/ps:1997103
[23] Ledoux, M. (1999). Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités*, *XXXIII*. *Lecture Notes in Math*. **1709** 120–216. Berlin: Springer. MR1767995 https://doi.org/10.1007/BFb0096511
[24] Ledoux, M. (2001). *The Concentration of Measure Phenomenon*. *Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347
[25] Lytova, A. and Pastur, L. (2008). On asymptotic behavior of multilinear eigenvalue statistics of random matrices. *J*. *Stat*. *Phys*. **133** 871–882. MR2461187 https://doi.org/10.1007/s10955-008-9644-6
[26] Major, P. (2013). *On the Estimation of Multiple Random Integrals and U-Statistics*. *Lecture Notes in Math*. **2079**. Heidelberg: Springer. MR3087566 https://doi.org/10.1007/978-3-642-37617-7
[27] Politis, D.N. and Romano, J.P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Ann*. *Statist*. **22** 2031–2050. MR1329181 https://doi.org/10.1214/aos/1176325770
[28] Robins, J.M., Li, L., Tchetgen, E.T. and van der Vaart, A. (2016). Asymptotic normality of quadratic estimators. *Stochastic Process*. *Appl*. **126** 3733–3759. MR3565475 https://doi.org/10.1016/j.spa.2016.04.005
[29] Sambale, H. (2015). Second order concentration for functions of independent random variables. Ph.D. thesis, Bielefeld Univ.
[30] Sinai, Y. and Soshnikov, A. (1998). Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol*. *Soc*. *Bras*. *Mat*. **29** 1–24. MR1620151 https://doi.org/10.1007/BF01245866
[31] Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publ*. *Math*. *Inst*. *Hautes Études Sci*. **81** 73–205. MR1361756
[32] Talagrand, M. (1996). New concentration inequalities in product spaces. *Invent*. *Math*. **126** 505–563. MR1419006 https://doi.org/10.1007/s002220050108

Bernoulli **26**(1), 2020, 127–158
https://doi.org/10.3150/19-BEJ1119

Spike and slab empirical Bayes sparse credible sets

ISMAËL CASTILLO[1] and BOTOND SZABÓ[2]

[1]*Sorbonne Université, Laboratoire Probabilités, Statistique et Modélisation; 4, place Jussieu, 75005 Paris, France. E-mail: ismael.castillo@upmc.fr*
[2]*Leiden University, The Netherlands. E-mail: b.t.szabo@math.leidenuniv.nl*

In the sparse normal means model, coverage of adaptive Bayesian posterior credible sets associated to spike and slab prior distributions is considered. The key sparsity hyperparameter is calibrated via marginal maximum likelihood empirical Bayes. First, adaptive posterior contraction rates are derived with respect to d_q-type-distances for $q \leq 2$. Next, under a type of so-called excessive-bias conditions, credible sets are constructed that have coverage of the true parameter at prescribed $1-\alpha$ confidence level and at the same time are of optimal diameter. We also prove that the previous conditions cannot be significantly weakened from the minimax perspective.

Keywords: convergence rates of posterior distributions; credible sets; empirical Bayes; spike and slab prior distributions

References

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

[2] Baraud, Y. (2002). Non-asymptotic minimax rates of testing in signal detection. *Bernoulli* **8** 577–606. MR1935648

[3] Belitser, E. and Ghosal, S. (2019). Empirical Bayes oracle uncertainty quantification for regression. *Ann. Stasist.* To appear.

[4] Belitser, E. and Nurushev, N. (2019). Needles and straw in a haystack: Robust empirical Bayes confidence for possibly sparse sequences. *Bernoull.* To appear.

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

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

[7] Bull, A.D. and Nickl, R. (2013). Adaptive confidence sets in L^2. *Probab. Theory Related Fields* **156** 889–919. MR3078289 https://doi.org/10.1007/s00440-012-0446-z

[8] Carvalho, C.M., Polson, N.G. and Scott, J.G. (2010). The horseshoe estimator for sparse signals. *Biometrika* **97** 465–480. MR2650751 https://doi.org/10.1093/biomet/asq017

[9] Castillo, I. and Mismer, R. (2018). Empirical Bayes analysis of spike and slab posterior distributions. *Electron. J. Stat.* **12** 3953–4001. MR3885271 https://doi.org/10.1214/18-EJS1494

[10] Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. MR3127856 https://doi.org/10.1214/13-AOS1133

[11] Castillo, I. and Szabo, B. (2020). Supplement to "Spike and slab empirical Bayes sparse credible sets." https://doi.org/10.3150/19-BEJ1119SUPP.

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

[13] Donoho, D.L., Johnstone, I.M., Hoch, J.C. and Stern, A.S. (1992). Maximum entropy and the nearly black object. *J. Roy. Statist. Soc. Ser. B* **54** 41–81. With discussion and a reply by the authors. MR1157714

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

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

[16] Golubev, G.K. (2002). Reconstruction of sparse vectors in white Gaussian noise. *Problemy Peredachi Informatsii* **38** 75–91. MR2101314 https://doi.org/10.1023/A:1020098307781

[17] Jiang, W. and Zhang, C.-H. (2009). General maximum likelihood empirical Bayes estimation of normal means. *Ann. Statist.* **37** 1647–1684. MR2533467 https://doi.org/10.1214/08-AOS638

[18] Johnson, V.E. and Rossell, D. (2012). Bayesian model selection in high-dimensional settings. *J. Amer. Statist. Assoc.* **107** 649–660. MR2980074 https://doi.org/10.1080/01621459.2012.682536

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

[20] Johnstone, I.M. and Silverman, B.W. (2005). EbayesThresh: R programs for Empirical Bayes thresholding. *J. Stat. Softw.* **12** .

[21] Johnstone, I.M. and Silverman, B.W. (2005). Empirical Bayes selection of wavelet thresholds. *Ann. Statist.* **33** 1700–1752. MR2166560 https://doi.org/10.1214/009053605000000345

[22] Low, M.G. (1997). On nonparametric confidence intervals. *Ann. Statist.* **25** 2547–2554. MR1604412 https://doi.org/10.1214/aos/1030741084

[23] Martin, R. and Walker, S.G. (2014). Asymptotically minimax empirical Bayes estimation of a sparse normal mean vector. *Electron. J. Stat.* **8** 2188–2206. MR3273623 https://doi.org/10.1214/14-EJS949

[24] Nickl, R. and van de Geer, S. (2013). Confidence sets in sparse regression. *Ann. Statist.* **41** 2852–2876. MR3161450 https://doi.org/10.1214/13-AOS1170

[25] Robins, J. and van der Vaart, A. (2006). Adaptive nonparametric confidence sets. *Ann. Statist.* **34** 229–253. MR2275241 https://doi.org/10.1214/009053605000000877

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

[27] Ročková, V. and George, E.I. (2018). The spike-and-slab LASSO. *J. Amer. Statist. Assoc.* **113** 431–444. MR3803476 https://doi.org/10.1080/01621459.2016.1260469

[28] Rousseau, J. and Szabó, B. (2019). Asymptotic frequentist coverage properties of Bayesian credible sets for sieve priors in general settings. *Ann. Statist.* To appear.

[29] Szabó, B., van der Vaart, A. and van Zanten, H. (2015). Honest Bayesian confidence sets for the L^2-norm. *J. Statist. Plann. Inference* **166** 36–51. MR3390132 https://doi.org/10.1016/j.jspi.2014.06.005

[30] Szabó, B., van der Vaart, A.W. and van Zanten, J.H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428. MR3357861 https://doi.org/10.1214/14-AOS1270

[31] van der Pas, S., Szabó, B. and van der Vaart, A. (2017). Adaptive posterior contraction rates for the horseshoe. *Electron. J. Stat.* **11** 3196–3225. MR3705450 https://doi.org/10.1214/17-EJS1316

[32] van der Pas, S., Szabó, B. and van der Vaart, A. (2017). Uncertainty quantification for the horseshoe (with discussion). *Bayesian Anal*. **12** 1221–1274. MR3724985 https://doi.org/10.1214/17-BA1065

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

[34] van der Pas, S.L., Salomond, J.-B. and Schmidt-Hieber, J. (2016). Conditions for posterior contraction in the sparse normal means problem. *Electron*. *J*. *Stat*. **10** 976–1000. MR3486423 https://doi.org/10.1214/16-EJS1130

Bernoulli **26**(1), 2020, 159–190
https://doi.org/10.3150/19-BEJ1120

Random walk approximation of BSDEs with Hölder continuous terminal condition

CHRISTEL GEISS[1,*], CÉLINE LABART[2] and ANTTI LUOTO[1,**]

[1]*Department of Mathematics and Statistics, University of Jyvaskyla, P.O. Box 35 (MaD), FI-40014, Finland. E-mail:* [*]*christel.geiss@jyu.fi;* [**]*antti.k.luoto@jyu.fi*
[2]*Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France. E-mail:* *celine.labart@univ-smb.fr*

In this paper, we consider the random walk approximation of the solution of a Markovian BSDE whose terminal condition is a locally Hölder continuous function of the Brownian motion. We state the rate of the L_2-convergence of the approximated solution to the true one. The proof relies in part on growth and smoothness properties of the solution u of the associated PDE. Here we improve existing results by showing some properties of the second derivative of u in space.

Keywords: backward stochastic differential equations; numerical scheme; random walk approximation; speed of convergence

References

[1] Alanko, S. (2015). Regression-based Monte Carlo methods for solving nonlinear PDEs. Ph.D. thesis, New York University.

[2] Bally, V. (1997). Approximation scheme for solutions of BSDE. In *Backward Stochastic Differential Equations* (*Paris*, 1995–1996). *Pitman Res*. *Notes Math*. *Ser*. **364** 177–191. Harlow: Longman. MR1752682

[3] Bally, V. and Pagès, G. (2003). A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli* **9** 1003–1049. MR2046816 https://doi.org/10.3150/bj/1072215199

[4] Bender, C. and Zhang, J. (2008). Time discretization and Markovian iteration for coupled FBSDEs. *Ann*. *Appl*. *Probab*. **18** 143–177. MR2380895 https://doi.org/10.1214/07-AAP448

[5] Borodin, A.N. and Salminen, P. (2015). *Handbook of Brownian Motion*: *Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Basel: Birkhäuser.

[6] Bouchard, B. and Touzi, N. (2004). Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process*. *Appl*. **111** 175–206. MR2056536 https://doi.org/10.1016/j.spa.2004.01.001

[7] Briand, Ph., Delyon, B. and Mémin, J. (2001). Donsker-type theorem for BSDEs. *Electron*. *Commun*. *Probab*. **6** 1–14. MR1817885 https://doi.org/10.1214/ECP.v6-1030

[8] Briand, Ph., Delyon, B., Hu, Y., Pardoux, E. and Stoica, L. (2003). L^p solutions of backward stochastic differential equations. *Stochastic Process*. *Appl*. **108** 109–129. MR2008603 https://doi.org/10.1016/S0304-4149(03)00089-9

[9] Briand, Ph. and Labart, C. (2014). Simulation of BSDEs by Wiener chaos expansion. *Ann*. *Appl*. *Probab*. **24** 1129–1171. MR3199982 https://doi.org/10.1214/13-AAP943

1350-7265

[10] Chassagneux, J.-F. (2014). Linear multistep schemes for BSDEs. *SIAM J. Numer. Anal.* **52** 2815–2836. MR3284573 https://doi.org/10.1137/120902951
[11] Chassagneux, J.-F. and Crisan, D. (2014). Runge–Kutta schemes for backward stochastic differential equations. *Ann. Appl. Probab.* **24** 679–720. MR3178495 https://doi.org/10.1214/13-AAP933
[12] Chassagneux, J.-F., Crisan, D. and Delarue, F. (2017). Numerical method for FBSDEs of McKean–Vlasov type. Available at arXiv:1703.02007.
[13] Chassagneux, J.-F. and Garcia Trillos, C.A. (2017). Cubature methods to solve BSDEs: Error expansion and complexity control. Available at arXiv:1702.00999.
[14] Chassagneux, J.-F. and Richou, A. (2015). Numerical stability analysis of the Euler scheme for BSDEs. *SIAM J. Numer. Anal.* **53** 1172–1193. MR3338675 https://doi.org/10.1137/140977047
[15] Chassagneux, J.-F. and Richou, A. (2016). Rate of convergence for discrete-time approximation of reflected BSDEs arising in switching problems. Available at arXiv:1602.00015.
[16] Chaudru de Raynal, P.E. and Garcia Trillos, C.A. (2015). A cubature based algorithm to solve decoupled McKean–Vlasov forward-backward stochastic differential equations. *Stochastic Process. Appl.* **125** 2206–2255. MR3322862 https://doi.org/10.1016/j.spa.2014.11.018
[17] Chevance, D. (1997). Résolution numérique des équations différentielles stochastiques rétrogrades. Ph.D. thesis, Université de Provence-Aix-Marseille I, Marseille.
[18] Coquet, F., Mackevičius, V. and Mémin, J. (1999). Corrigendum to: "Stability in $\mathbb{D}$ of martingales and backward equations under discretization of filtration" [*Stochastic Processes Appl.* **75** (1998), no. 2, 235–248; MR1632205 (99f:60109)]. *Stochastic Process. Appl.* **82** 335–338. MR1700013 https://doi.org/10.1016/S0304-4149(99)00020-4
[19] Crisan, D. and Delarue, F. (2012). Sharp derivative bounds for solutions of degenerate semi-linear partial differential equations. *J. Funct. Anal.* **263** 3024–3101. MR2973334 https://doi.org/10.1016/j.jfa.2012.07.015
[20] Crisan, D., Manolarakis, K. and Touzi, N. (2010). On the Monte Carlo simulation of BSDEs: An improvement on the Malliavin weights. *Stochastic Process. Appl.* **120** 1133–1158. MR2639741 https://doi.org/10.1016/j.spa.2010.03.015
[21] Delarue, F. and Menozzi, S. (2006). A forward-backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.* **16** 140–184. MR2209339 https://doi.org/10.1214/105051605000000674
[22] Geiss, C., Geiss, S. and Gobet, E. (2012). Generalized fractional smoothness and L_p-variation of BSDEs with non-Lipschitz terminal condition. *Stochastic Process. Appl.* **122** 2078–2116. MR2921973 https://doi.org/10.1016/j.spa.2012.02.006
[23] Geiss, C. and Labart, C. (2016). Simulation of BSDEs with jumps by Wiener chaos expansion. *Stochastic Process. Appl.* **126** 2123–2162. MR3483749 https://doi.org/10.1016/j.spa.2016.01.006
[24] Gobet, E. and Labart, C. (2007). Error expansion for the discretization of backward stochastic differential equations. *Stochastic Process. Appl.* **117** 803–829. MR2330720 https://doi.org/10.1016/j.spa.2006.10.007
[25] Gobet, E., Lemor, J.-P. and Warin, X. (2005). A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* **15** 2172–2202. MR2152657 https://doi.org/10.1214/105051605000000412
[26] Henry-Labordère, P., Tan, X. and Touzi, N. (2014). A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl.* **124** 1112–1140. MR3138609 https://doi.org/10.1016/j.spa.2013.10.005
[27] Ma, J., Protter, P., San Martín, J. and Torres, S. (2002). Numerical method for backward stochastic differential equations. *Ann. Appl. Probab.* **12** 302–316. MR1890066 https://doi.org/10.1214/aoap/1015961165
[28] Ma, J. and Zhang, J. (2002). Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.* **12** 1390–1418. MR1936598 https://doi.org/10.1214/aoap/1037125868

[29] Martínez, M., San Martín, J. and Torres, S. (2011). Numerical method for reflected backward stochastic differential equations. *Stoch. Anal. Appl.* **29** 1008–1032. MR2847333 https://doi.org/10.1080/07362994.2011.610162
[30] Mémin, J., Peng, S. and Xu, M. (2008). Convergence of solutions of discrete reflected backward SDE's and simulations. *Acta Math. Appl. Sin. Engl. Ser.* **24** 1–18. MR2385005 https://doi.org/10.1007/s10255-006-6005-6
[31] Peng, S. and Xu, M. (2011). Numerical algorithms for backward stochastic differential equations with 1-D Brownian motion: Convergence and simulations. *ESAIM Math. Model. Numer. Anal.* **45** 335–360. MR2804642 https://doi.org/10.1051/m2an/2010059
[32] Sundar, P. and Yin, H. (2009). Existence and uniqueness of solutions to the backward 2D stochastic Navier–Stokes equations. *Stochastic Process. Appl.* **119** 1216–1234. MR2508571 https://doi.org/10.1016/j.spa.2008.06.007
[33] Walsh, J.B. (2003). The rate of convergence of the binomial tree scheme. *Finance Stoch.* **7** 337–361. MR1994913 https://doi.org/10.1007/s007800200094
[34] Weinan, E., Hutzenthaler, M., Jentzen, A. and Kruse, T. (2017). On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations. Available at arXiv:1708.03223.
[35] Zhang, J. (2001). Some fine properties of backward stochastic differential equations, with applications. Ph.D. thesis, Purdue University.
[36] Zhang, J. (2004). A numerical scheme for BSDEs. *Ann. Appl. Probab.* **14** 459–488. MR2023027 https://doi.org/10.1214/aoap/1075828058
[37] Zhang, J. (2005). Representation of solutions to BSDEs associated with a degenerate FSDE. *Ann. Appl. Probab.* **15** 1798–1831. MR2152246 https://doi.org/10.1214/105051605000000232

Bernoulli **26**(1), 2020, 191–225
https://doi.org/10.3150/19-BEJ1122

Needles and straw in a haystack: Robust confidence for possibly sparse sequences

EDUARD BELITSER[1] and NURZHAN NURUSHEV[1,2]

[1] *Department of Mathematics, VU Amsterdam, The Netherlands. E-mail: e.n.belitser@vu.nl*
[2] *Korteweg-de Vries Institute for Mathematics, University of Amsterdam, The Netherlands. E-mail: n.nurushev@uva.nl*

In the general *signal+noise* (allowing non-normal, non-independent observations) model, we construct an empirical Bayes posterior which we then use for *uncertainty quantification* for the unknown, possibly sparse, signal. We introduce a novel *excessive bias restriction* (EBR) condition, which gives rise to a new slicing of the entire space that is suitable for uncertainty quantification. Under EBR and some mild *exchangeable exponential moment condition* on the noise, we establish the local (oracle) optimality of the proposed confidence ball. Without EBR, we propose another confidence ball of full coverage, but its radius contains an additional $\sigma n^{1/4}$-term. In passing, we also get the local optimal results for *estimation*, *posterior contraction* problems, and the problem of *weak recovery of sparsity structure*. Adaptive minimax results (also for the estimation and posterior contraction problems) over various sparsity classes follow from our local results.

Keywords: confidence set; empirical Bayes posterior; local radial rate

References

[1] Abramovich, F., Grinshtein, V. and Pensky, M. (2007). On optimality of Bayesian testimation in the normal means problem. *Ann. Statist.* **35** 2261–2286. MR2363971 https://doi.org/10.1214/009053607000000226

[2] Babenko, A. and Belitser, E. (2010). Oracle convergence rate of posterior under projection prior and Bayesian model selection. *Math. Methods Statist.* **19** 219–245. MR2742927 https://doi.org/10.3103/S1066530710030026

[3] Baraud, Y. (2004). Confidence balls in Gaussian regression. *Ann. Statist.* **32** 528–551. MR2060168 https://doi.org/10.1214/009053604000000085

[4] Belitser, E. (2017). On coverage and local radial rates of credible sets. *Ann. Statist.* **45** 1124–1151. MR3662450 https://doi.org/10.1214/16-AOS1477

[5] Belitser, E. and Ghosal, S. (2018). Empirical Bayes oracle uncertainty quantification for regression. Preprint.

[6] Belitser, E. and Nurushev, N. (2020). Supplement to "Needles and straw in a haystack: Robust confidence for possibly sparse sequences." https://doi.org/10.3150/19-BEJ1122SUPP.

[7] Bhattacharya, A., Dunson, D.B., Pati, D. and Pillai, N.S. (2016). Sub-optimality of some continuous shrinkage priors. *Stochastic Process. Appl.* **126** 3828–3842. MR3565480 https://doi.org/10.1016/j.spa.2016.08.007

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

1350-7265

[9] Bull, A.D. (2012). Honest adaptive confidence bands and self-similar functions. *Electron. J. Stat.* **6** 1490–1516. MR2988456 https://doi.org/10.1214/12-EJS720
[10] Bull, A.D. and Nickl, R. (2013). Adaptive confidence sets in L^2. *Probab. Theory Related Fields* **156** 889–919. MR3078289 https://doi.org/10.1007/s00440-012-0446-z
[11] Cai, T.T. and Low, M.G. (2004). An adaptation theory for nonparametric confidence intervals. *Ann. Statist.* **32** 1805–1840. MR2102494 https://doi.org/10.1214/009053604000000049
[12] Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *Ann. Statist.* **43** 1986–2018. MR3375874 https://doi.org/10.1214/15-AOS1334
[13] Castillo, I. and Szabó, B. (2018). Spike and slab empirical Bayes sparse credible sets. Available at arXiv:1808.07721.
[14] Castillo, I. and van der Vaart, A. (2012). Needles and straw in a haystack: Posterior concentration for possibly sparse sequences. *Ann. Statist.* **40** 2069–2101. MR3059077 https://doi.org/10.1214/12-AOS1029
[15] Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425–455. MR1311089 https://doi.org/10.1093/biomet/81.3.425
[16] Donoho, D.L. and Johnstone, I.M. (1994). Minimax risk over l_p-balls for l_q-error. *Probab. Theory Related Fields* **99** 277–303. MR1278886 https://doi.org/10.1007/BF01199026
[17] Donoho, D.L., Johnstone, I.M., Hoch, J.C. and Stern, A.S. (1992). Maximum entropy and the nearly black object (with discussion). *J. Roy. Statist. Soc. Ser. B* **54** 41–81. MR1157714
[18] Johnstone, I. (2017). Gaussian estimation: Sequence and wavelet models. Book draft.
[19] Johnstone, I.M. and Silverman, B.W. (2004). Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences. *Ann. Statist.* **32** 1594–1649. MR2089135 https://doi.org/10.1214/009053604000000030
[20] Li, K.-C. (1989). Honest confidence regions for nonparametric regression. *Ann. Statist.* **17** 1001–1008. MR1015135 https://doi.org/10.1214/aos/1176347253
[21] Martin, R. and Walker, S.G. (2014). Asymptotically minimax empirical Bayes estimation of a sparse normal mean vector. *Electron. J. Stat.* **8** 2188–2206. MR3273623 https://doi.org/10.1214/14-EJS949
[22] Nickl, R. and van de Geer, S. (2013). Confidence sets in sparse regression. *Ann. Statist.* **41** 2852–2876. MR3161450 https://doi.org/10.1214/13-AOS1170
[23] Picard, D. and Tribouley, K. (2000). Adaptive confidence interval for pointwise curve estimation. *Ann. Statist.* **28** 298–335. MR1762913 https://doi.org/10.1214/aos/1016120374
[24] Robins, J. and van der Vaart, A. (2006). Adaptive nonparametric confidence sets. *Ann. Statist.* **34** 229–253. MR2275241 https://doi.org/10.1214/009053605000000877
[25] Ročková, V. (2018). Bayesian estimation of sparse signals with a continuous spike-and-slab prior. *Ann. Statist.* **46** 401–437. MR3766957 https://doi.org/10.1214/17-AOS1554
[26] Rousseau, J. and Szabó, B. (2016). Asymptotic frequentist coverage properties of Bayesian credible sets for sieve priors. Available at arXiv:1609.05067.
[27] Rousseau, J. and Szabo, B. (2017). Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator. *Ann. Statist.* **45** 833–865. MR3650402 https://doi.org/10.1214/16-AOS1469
[28] Szabó, B., van der Vaart, A.W. and van Zanten, J.H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428. MR3357861 https://doi.org/10.1214/14-AOS1270
[29] van der Pas, S., Szabó, B. and van der Vaart, A. (2017). Uncertainty quantification for the horseshoe (with discussion). *Bayesian Anal.* **12** 1221–1274. MR3724985 https://doi.org/10.1214/17-BA1065
[30] van der Pas, S.L., Kleijn, B.J.K. and van der Vaart, A.W. (2014). The horseshoe estimator: Posterior concentration around nearly black vectors. *Electron. J. Stat.* **8** 2585–2618. MR3285877 https://doi.org/10.1214/14-EJS962

Bernoulli **26**(1), 2020, 226–252
https://doi.org/10.3150/19-BEJ1124

Estimation of the linear fractional stable motion

STEPAN MAZUR[1], DMITRY OTRYAKHIN[2,*] and MARK PODOLSKIJ[2,**]

[1] *Unit of Statistics, Örebro University School of Business, Fakultetsgatan 1, 70281 Örebro, Sweden. E-mail: stepan.mazur@oru.se*

[2] *Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus, Denmark. E-mail: [*] d.otryakhin@math.au.dk; [**] mpodolskij@math.au.dk*

In this paper, we investigate the parametric inference for the linear fractional stable motion in high and low frequency setting. The symmetric linear fractional stable motion is a three-parameter family, which constitutes a natural non-Gaussian analogue of the scaled fractional Brownian motion. It is fully characterised by the scaling parameter $\sigma > 0$, the self-similarity parameter $H \in (0,1)$ and the stability index $\alpha \in (0,2)$ of the driving stable motion. The parametric estimation of the model is inspired by the limit theory for stationary increments Lévy moving average processes that has been recently studied in (*Ann. Probab.* **45** (2017) 4477–4528). More specifically, we combine (negative) power variation statistics and empirical characteristic functions to obtain consistent estimates of (σ, α, H). We present the law of large numbers and some fully feasible weak limit theorems.

Keywords: fractional processes; limit theorems; parametric estimation; stable motion

References

[1] Aït-Sahalia, Y. and Jacob, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** 355–392. MR2332279 https://doi.org/10.1214/009053606000001190
[2] Ayache, A., Cohen, S. and Lévy Véhel, J. (2000). The covariance structure of multifractional Brownian motion, with application to long range dependence (extended version). *ICASSP*, Refereed Conference Contribution.
[3] Ayache, A. and Hamonier, J. (2012). Linear fractional stable motion: A wavelet estimator of the α parameter. *Statist. Probab. Lett.* **82** 1569–1575. MR2930661 https://doi.org/10.1016/j.spl.2012.04.005
[4] Bardet, J.-M. and Surgailis, D. (2013). Nonparametric estimation of the local Hurst function of multifractional Gaussian processes. *Stochastic Process. Appl.* **123** 1004–1045. MR3005013 https://doi.org/10.1016/j.spa.2012.11.009
[5] Basse-O'Connor, A., Lachièze-Rey, R. and Podolskij, M. (2017). Power variation for a class of stationary increments Lévy driven moving averages. *Ann. Probab.* **45** 4477–4528. MR3737916 https://doi.org/10.1214/16-AOP1170
[6] Benassi, A., Cohen, S. and Istas, J. (2004). On roughness indices for fractional fields. *Bernoulli* **10** 357–373. MR2046778 https://doi.org/10.3150/bj/1082380223
[7] Bender, C., Lindner, A. and Schicks, M. (2012). Finite variation of fractional Lévy processes. *J. Theoret. Probab.* **25** 594–612. MR2914443 https://doi.org/10.1007/s10959-010-0339-y
[8] Bender, C. and Marquardt, T. (2008). Stochastic calculus for convoluted Lévy processes. *Bernoulli* **14** 499–518. MR2544099 https://doi.org/10.3150/07-BEJ115

[9] Brouste, A. and Fukasawa, M. (2018). Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist.* **46** 2045–2061. MR3845010 https://doi.org/10.1214/17-AOS1611
[10] Cambanis, S., Hardin, C.D. Jr. and Weron, A. (1987). Ergodic properties of stationary stable processes. *Stochastic Process. Appl.* **24** 1–18. MR0883599 https://doi.org/10.1016/0304-4149(87)90024-X
[11] Cambanis, S. and Maejima, M. (1989). Two classes of self-similar stable processes with stationary increments. *Stochastic Process. Appl.* **32** 305–329. MR1014456 https://doi.org/10.1016/0304-4149(89)90082-3
[12] Coeurjolly, J.-F. and Istas, J. (2001). Cramèr–Rao bounds for fractional Brownian motions. *Statist. Probab. Lett.* **53** 435–447. MR1856169 https://doi.org/10.1016/S0167-7152(00)00197-8
[13] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.* **17** 1749–1766. MR1026311 https://doi.org/10.1214/aos/1176347393
[14] Dang, T.T.N. and Istas, J. (2017). Estimation of the Hurst and the stability indices of a H-self-similar stable process. *Electron. J. Stat.* **11** 4103–4150. MR3715823 https://doi.org/10.1214/17-EJS1357
[15] Grahovac, D., Leonenko, N.N. and Taqqu, M.S. (2015). Scaling properties of the empirical structure function of linear fractional stable motion and estimation of its parameters. *J. Stat. Phys.* **158** 105–119. MR3296276 https://doi.org/10.1007/s10955-014-1126-4
[16] Istas, J. and Lang, G. (1997). Quadratic variations and estimation of the local Hölder index of a Gaussian process. *Ann. Inst. Henri Poincaré Probab. Stat.* **33** 407–436. MR1465796 https://doi.org/10.1016/S0246-0203(97)80099-4
[17] Lebovits, J. and Podolskij, M. (2017). Estimation of the global regularity of a multifractional Brownian motion. *Electron. J. Stat.* **11** 78–98. MR3597564 https://doi.org/10.1214/16-EJS1221
[18] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR0242239 https://doi.org/10.1137/1010093
[19] Peltier, R. and Lévy Véhel, J. (1995). Multifractional Brownian motion: Definition and preliminary results. Rapport de recherche de l'INRIA, number 2645.
[20] Pipiras, V. and Taqqu, M.S. (2002). The structure of self-similar stable mixed moving averages. *Ann. Probab.* **30** 898–932. MR1905860 https://doi.org/10.1214/aop/1023481011
[21] Pipiras, V. and Taqqu, M.S. (2003). Central limit theorems for partial sums of bounded functionals of infinite-variance moving averages. *Bernoulli* **9** 833–855. MR2047688 https://doi.org/10.3150/bj/1066418880
[22] Pipiras, V., Taqqu, M.S. and Abry, P. (2007). Bounds for the covariance of functions of infinite variance stable random variables with applications to central limit theorems and wavelet-based estimation. *Bernoulli* **13** 1091–1123. MR2364228 https://doi.org/10.3150/07-BEJ6143
[23] Rajput, B.S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields* **82** 451–487. MR1001524 https://doi.org/10.1007/BF00339998
[24] Resnick, S. and Greenwood, P. (1979). A bivariate stable characterization and domains of attraction. *J. Multivariate Anal.* **9** 206–221. MR0538402 https://doi.org/10.1016/0047-259X(79)90079-4
[25] Rosiński, J. (1995). On the structure of stationary stable processes. *Ann. Probab.* **23** 1163–1187. MR1349166
[26] Samorodnitsky, G. (2005). Null flows, positive flows and the structure of stationary symmetric stable processes. *Ann. Probab.* **33** 1782–1803. MR2165579 https://doi.org/10.1214/009117905000000305
[27] Mazur, S., Otryakhin, D. and Podolskij, M. (2020). Supplement to "Estimation of the linear fractional stable motion." https://doi.org/10.3150/19-BEJ1124SUPP.
[28] Stoev, S., Pipiras, V. and Taqqu, M.S. (2002). Estimation of the self-similarity parameter in linear fractional stable motion. *Signal Process.* **82** 1873–1901.
[29] Stoev, S. and Taqqu, M.S. (2004). Simulation methods for linear fractional stable motion and FARIMA using the fast Fourier transform. *Fractals* **12** 95–121. MR2052735 https://doi.org/10.1142/S0218348X04002379

[30] Stoev, S.A. and Taqqu, M.S. (2006). How rich is the class of multifractional Brownian motions? *Stochastic Process. Appl.* **116** 200–221. MR2197974 https://doi.org/10.1016/j.spa.2005.09.007

Bernoulli **26**(1), 2020, 253–285
https://doi.org/10.3150/19-BEJ1125

Cliques in rank-1 random graphs: The role of inhomogeneity

KAY BOGERD*, RUI M. CASTRO** and REMCO VAN DER HOFSTAD†

Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
E-mail: *k.m.bogerd@tue.nl; **rmcastro@tue.nl; †r.w.v.d.hofstad@tue.nl

We study the asymptotic behavior of the clique number in rank-1 inhomogeneous random graphs, where edge probabilities between vertices are roughly proportional to the product of their vertex weights. We show that the clique number is concentrated on at most two consecutive integers, for which we provide an expression. Interestingly, the order of the clique number is primarily determined by the overall edge density, with the inhomogeneity only affecting multiplicative constants or adding at most a $\log\log(n)$ multiplicative factor. For sparse enough graphs the clique number is always bounded and the effect of inhomogeneity completely vanishes.

Keywords: clique number; Erdős–Rényi random graphs; inhomogeneous random graphs

References

[1] Alon, N., Krivelevich, M. and Sudakov, B. (1998). Finding a large hidden clique in a random graph. *Random Structures Algorithms* **13** 457–466. MR1662795 https://doi.org/10.1002/(SICI)1098-2418(199810/12)13:3/4<457::AID-RSA14>3.3.CO;2-K

[2] Arias-Castro, E. and Verzelen, N. (2014). Community detection in dense random networks. *Ann. Statist.* **42** 940–969. MR3210992 https://doi.org/10.1214/14-AOS1208

[3] Bianconi, G. and Marsili, M. (2006). Emergence of large cliques in random scale-free networks. *Europhys. Lett.* **74** 740–746. MR2240156 https://doi.org/10.1209/epl/i2005-10574-3

[4] Bianconi, G. and Marsili, M. (2006). Number of cliques in random scale-free network ensembles. *Phys. D* **224** 1–6. MR2301503 https://doi.org/10.1016/j.physd.2006.09.013

[5] Bollobás, B., Janson, S. and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. MR2337396 https://doi.org/10.1002/rsa.20168

[6] Britton, T., Deijfen, M. and Martin-Löf, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124** 1377–1397. MR2266448 https://doi.org/10.1007/s10955-006-9168-x

[7] Chung, F. and Lu, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6** 125–145. MR1955514 https://doi.org/10.1007/PL00012580

[8] Chung, F. and Lu, L. (2003). The average distance in a random graph with given expected degrees. *Internet Math.* **1** 91–113. MR2076728

[9] Chung, F. and Lu, L. (2006). The volume of the giant component of a random graph with given expected degrees. *SIAM J. Discrete Math.* **20** 395–411. MR2257269 https://doi.org/10.1137/050630106

[10] Corless, R.M., Gonnet, G.H., Hare, D.E.G., Jeffrey, D.J. and Knuth, D.E. (1996). On the Lambert *W* function. *Adv. Comput. Math.* **5** 329–359. MR1414285 https://doi.org/10.1007/BF02124750

1350-7265

[11] Dekel, Y., Gurel-Gurevich, O. and Peres, Y. (2014). Finding hidden cliques in linear time with high probability. *Combin. Probab. Comput.* **23** 29–49. MR3197965 https://doi.org/10.1017/S096354831300045X
[12] Devroye, L. and Fraiman, N. (2014). Connectivity of inhomogeneous random graphs. *Random Structures Algorithms* **45** 408–420. MR3252923 https://doi.org/10.1002/rsa.20490
[13] Doležal, M., Hladký, J. and Máthé, A. (2017). Cliques in dense inhomogeneous random graphs. *Random Structures Algorithms* **51** 275–314. MR3683364 https://doi.org/10.1002/rsa.20715
[14] Feige, U. and Krauthgamer, R. (2000). Finding and certifying a large hidden clique in a semirandom graph. *Random Structures Algorithms* **16** 195–208. MR1742351 https://doi.org/10.1002/(SICI)1098-2418(200003)16:2<195::AID-RSA5>3.3.CO;2-1
[15] Fraiman, N. and Mitsche, D. (2018). The diameter of inhomogeneous random graphs. *Random Structures Algorithms* **53** 308–326.
[16] Friedrich, T. and Krohmer, A. (2015). Parameterized clique on inhomogeneous random graphs. *Discrete Appl. Math.* **184** 130–138. MR3320708 https://doi.org/10.1016/j.dam.2014.10.018
[17] Grimmett, G.R. and McDiarmid, C.J.H. (1975). On colouring random graphs. *Math. Proc. Cambridge Philos. Soc.* **77** 313–324. MR0369129 https://doi.org/10.1017/S0305004100051124
[18] Janson, S., Łuczak, T. and Norros, I. (2010). Large cliques in a power-law random graph. *J. Appl. Probab.* **47** 1124–1135. MR2752885 https://doi.org/10.1239/jap/1294170524
[19] Lovász, L. and Szegedy, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085 https://doi.org/10.1016/j.jctb.2006.05.002
[20] Matula, D.W. (1972). The employee party problem. *Notices Amer. Math. Soc.* **19** 89–156.
[21] Matula, D.W. (1976). The largest clique size in a random graph. Technical Report CS 7608, Dept. of Comp. Sci. Southern Methodist University, Dallas.
[22] McKinley, G. (2019). Super-logarithmic cliques in dense inhomogeneous random graphs. arXiv:1903.01495.
[23] Norros, I. and Reittu, H. (2006). On a conditionally Poissonian graph process. *Adv. in Appl. Probab.* **38** 59–75. MR2213964 https://doi.org/10.1239/aap/1143936140
[24] Spencer, J.H. and Florescu, L. (2014). *Asymptopia*. *Student Mathematical Library* **71**. Providence, RI: Amer. Math. Soc. MR3185739 https://doi.org/10.1090/stml/071
[25] van der Hofstad, R. (2017). *Random Graphs and Complex Networks*. *Vol.* 1. *Cambridge Series in Statistical and Probabilistic Mathematics* **43**. Cambridge: Cambridge Univ. Press. MR3617364 https://doi.org/10.1017/9781316779422
[26] van der Hofstad, R. (2019+). *Random Graphs and Complex Networks* **2**.
[27] Verzelen, N. and Arias-Castro, E. (2015). Community detection in sparse random networks. *Ann. Appl. Probab.* **25** 3465–3510. MR3404642 https://doi.org/10.1214/14-AAP1080

Bernoulli **26**(1), 2020, 286–322
https://doi.org/10.3150/19-BEJ1126

Prediction and estimation consistency of sparse multi-class penalized optimal scoring

IRINA GAYNANOVA

Department of Statistics, Texas A&M University, MS 3143, College Station, TX 77843, USA.
E-mail: irinag@stat.tamu.edu

Sparse linear discriminant analysis via penalized optimal scoring is a successful tool for classification in high-dimensional settings. While the variable selection consistency of sparse optimal scoring has been established, the corresponding prediction and estimation consistency results have been lacking. We bridge this gap by providing probabilistic bounds on out-of-sample prediction error and estimation error of multi-class penalized optimal scoring allowing for diverging number of classes.

Keywords: classification; high-dimensional regression; lasso; linear discriminant analysis

References

[1] 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

[2] Bien, J., Gaynanova, I., Lederer, J. and Müller, C.L. (2019). Prediction error bounds for linear regression with the TREX. *TEST* **28** 451–474. MR3962062 https://doi.org/10.1007/s11749-018-0584-4

[3] Bouveyron, C., Latouche, P. and Mattei, P.-A. (2018). Bayesian variable selection for globally sparse probabilistic PCA. *Electron. J. Stat.* **12** 3036–3070. MR3856168 https://doi.org/10.1214/18-EJS1450

[4] 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

[5] Bunea, F., Tsybakov, A. and Wegkamp, M. (2007). Sparsity oracle inequalities for the Lasso. *Electron. J. Stat.* **1** 169–194. MR2312149 https://doi.org/10.1214/07-EJS008

[6] Cai, T. and Liu, W. (2011). A direct estimation approach to sparse linear discriminant analysis. *J. Amer. Statist. Assoc.* **106** 1566–1577. MR2896857 https://doi.org/10.1198/jasa.2011.tm11199

[7] Chatterjee, S. (2013). Assumptionless consistency of the Lasso. ArXiv.org.

[8] Clemmensen, L., Hastie, T., Witten, D. and Ersbøll, B. (2011). Sparse discriminant analysis. *Technometrics* **53** 406–413. MR2850472 https://doi.org/10.1198/TECH.2011.08118

[9] Dalalyan, A.S., Hebiri, M. and Lederer, J. (2017). On the prediction performance of the Lasso. *Bernoulli* **23** 552–581. MR3556784 https://doi.org/10.3150/15-BEJ756

[10] Gaynanova, I., Booth, J.G. and Wells, M.T. (2016). Simultaneous sparse estimation of canonical vectors in the $p \gg N$ setting. *J. Amer. Statist. Assoc.* **111** 696–706. MR3538698 https://doi.org/10.1080/01621459.2015.1034318

[11] Gaynanova, I. and Kolar, M. (2015). Optimal variable selection in multi-group sparse discriminant analysis. *Electron. J. Stat.* **9** 2007–2034. MR3393602 https://doi.org/10.1214/15-EJS1064

[12] Gaynanova, I. and Wang, T. (2019). Sparse quadratic classification rules via linear dimension reduction. *J. Multivariate Anal.* **169** 278–299. MR3875600 https://doi.org/10.1016/j.jmva.2018.09.011

1350-7265

[13] Greenshtein, E. and Ritov, Y. (2004). Persistence in high-dimensional linear predictor selection and the virtue of overparametrization. *Bernoulli* **10** 971–988. MR2108039 https://doi.org/10.3150/bj/1106314846
[14] Han, F., Zhao, T. and Liu, H. (2013). CODA: High dimensional copula discriminant analysis. *J. Mach. Learn. Res.* **14** 629–671. MR3033343
[15] Hastie, T., Buja, A. and Tibshirani, R. (1995). Penalized discriminant analysis. *Ann. Statist.* **23** 73–102. MR1331657 https://doi.org/10.1214/aos/1176324456
[16] Hastie, T., Tibshirani, R. and Buja, A. (1994). Flexible discriminant analysis by optimal scoring. *J. Amer. Statist. Assoc.* **89** 1255–1270. MR1310220
[17] Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations. Monographs on Statistics and Applied Probability* **143**. Boca Raton, FL: CRC Press. MR3616141
[18] Hsu, D., Kakade, S.M. and Zhang, T. (2012). A tail inequality for quadratic forms of subgaussian random vectors. *Electron. Commun. Probab.* **17** no. 52, 6. MR2994877 https://doi.org/10.1214/ECP.v17-2079
[19] Kolar, M. and Liu, H. (2015). Optimal feature selection in high-dimensional discriminant analysis. *IEEE Trans. Inform. Theory* **61** 1063–1083. MR3332765 https://doi.org/10.1109/TIT.2014.2381241
[20] Lederer, J., Yu, L. and Gaynanova, I. (2019). Oracle inequalities for high-dimensional prediction. *Bernoulli* **25** 1225–1255. MR3920371 https://doi.org/10.3150/18-bej1019
[21] Li, Y. and Jia, J. (2017). L1 least squares for sparse high-dimensional LDA. *Electron. J. Stat.* **11** 2499–2518. MR3659945 https://doi.org/10.1214/17-EJS1288
[22] Lounici, K., Pontil, M., van de Geer, S. and Tsybakov, A.B. (2011). Oracle inequalities and optimal inference under group sparsity. *Ann. Statist.* **39** 2164–2204. MR2893865 https://doi.org/10.1214/11-AOS896
[23] Mai, Q., Zou, H. and Yuan, M. (2012). A direct approach to sparse discriminant analysis in ultra-high dimensions. *Biometrika* **99** 29–42. MR2899661 https://doi.org/10.1093/biomet/asr066
[24] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis*. London–New York–Toronto, Ont.: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0560319
[25] McLachlan, G.J. (1992). *Discriminant Analysis and Statistical Pattern Recognition. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1190469 https://doi.org/10.1002/0471725293
[26] Merchante, L.F.S., Grandvalet, Y. and Govaert, G. (2012). An efficient approach to sparse linear discriminant analysis. In *Proceedings of the* 29*th International Conference on Machine Learning*, *ICML* 2012 1167–1174. Universite de Technologie de Compiegne, Compiegne, France.
[27] 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
[28] Neykov, M., Ning, Y., Liu, J.S. and Liu, H. (2018). A unified theory of confidence regions and testing for high-dimensional estimating equations. *Statist. Sci.* **33** 427–443. MR3843384 https://doi.org/10.1214/18-STS661
[29] Obozinski, G., Wainwright, M.J. and Jordan, M.I. (2011). Support union recovery in high-dimensional multivariate regression. *Ann. Statist.* **39** 1–47. MR2797839 https://doi.org/10.1214/09-AOS776
[30] Pan, Y., Mai, Q. and Zhang, X. (2018). Covariate-adjusted tensor classification in high dimensions. *J. Amer. Statist. Assoc.* 1–41.
[31] Raskutti, G., Wainwright, M.J. and Yu, B. (2010). Restricted eigenvalue properties for correlated Gaussian designs. *J. Mach. Learn. Res.* **11** 2241–2259. MR2719855
[32] Rigollet, P. and Tsybakov, A. (2011). Exponential screening and optimal rates of sparse estimation. *Ann. Statist.* **39** 731–771. MR2816337 https://doi.org/10.1214/10-AOS854

[33] Rudelson, M. and Zhou, S. (2013). Reconstruction from anisotropic random measurements. *IEEE Trans. Inform. Theory* **59** 3434–3447. MR3061256 https://doi.org/10.1109/TIT.2013.2243201
[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.A. and Bühlmann, P. (2009). On the conditions used to prove oracle results for the Lasso. *Electron. J. Stat.* **3** 1360–1392. MR2576316 https://doi.org/10.1214/09-EJS506
[36] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170
[37] Wainwright, M.J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). *IEEE Trans. Inform. Theory* **55** 2183–2202. MR2729873 https://doi.org/10.1109/TIT.2009.2016018
[38] Witten, D.M. and Tibshirani, R. (2011). Penalized classification using Fisher's linear discriminant. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 753–772. MR2867457 https://doi.org/10.1111/j.1467-9868.2011.00783.x
[39] Wu, Y., Wipf, D.P. and Yun, J.M. (2015). Understanding and evaluating sparse linear discriminant analysis. *AISTATS*.
[40] 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
[41] Zhang, T. (2009). Some sharp performance bounds for least squares regression with L_1 regularization. *Ann. Statist.* **37** 2109–2144. MR2543687 https://doi.org/10.1214/08-AOS659
[42] Zhao, P. and Yu, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.* **7** 2541–2563. MR2274449
[43] Zhou, S. (2009). Restricted eigenvalue conditions on subgaussian random matrices.

Bernoulli **26**(1), 2020, 323–351
https://doi.org/10.3150/19-BEJ1127

Weak convergence of quantile and expectile processes under general assumptions

TOBIAS ZWINGMANN* and HAJO HOLZMANN**

Fachbreich Mathematik und Informatik, Philipps-Universität Marburg, Germany.
*E-mail: *zwingmann@mathematik.uni-marburg.de; **holzmann@mathematik.uni-marburg.de*

We show weak convergence of quantile and expectile processes to Gaussian limit processes in the space of bounded functions endowed with an appropriate semimetric which is based on the concepts of epi- and hypo- convergence as introduced in A. Bücher, J. Segers and S. Volgushev (2014), '*When Uniform Weak Convergence Fails: Empirical Processes for Dependence Functions and Residuals via Epi- and Hypographs*', *Annals of Statistics* **42**. We impose assumptions for which it is known that weak convergence with respect to the supremum norm generally fails to hold. For quantiles, we consider stationary observations, where the marginal distribution function is assumed to be strictly increasing and continuous except for finitely many points and to admit strictly positive – possibly infinite – left- and right-sided derivatives. For expectiles, we focus on independent and identically distributed (i.i.d.) observations. Only a finite second moment and continuity at the boundary points but no further smoothness properties of the distribution function are required. We also show consistency of the bootstrap for this mode of convergence in the i.i.d. case for quantiles and expectiles.

Keywords: epi- and hypo convergence; expectile process; quantile process; weak convergence

References

[1] Bellini, F., Klar, B., Müller, A. and Rosazza Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.* **54** 41–48. MR3145849 https://doi.org/10.1016/j.insmatheco.2013.10.015

[2] Bücher, A., Segers, J. and Volgushev, S. (2014). When uniform weak convergence fails: Empirical processes for dependence functions and residuals via epi- and hypographs. *Ann. Statist.* **42** 1598–1634. MR3262462 https://doi.org/10.1214/14-AOS1237

[3] Bühlmann, P. (2002). Bootstraps for time series. *Statist. Sci.* **17** 52–72. MR1910074 https://doi.org/10.1214/ss/1023798998

[4] Dehling, H., Durieu, O. and Volny, D. (2009). New techniques for empirical processes of dependent data. *Stochastic Process. Appl.* **119** 3699–3718. MR2568292 https://doi.org/10.1016/j.spa.2009.07.003

[5] Holzmann, H. and Klar, B. (2016). Expectile asymptotics. *Electron. J. Stat.* **10** 2355–2371. MR3544290 https://doi.org/10.1214/16-EJS1173

[6] Knight, K. (1998). Bootstrapping sample quantiles in non-regular cases. *Statist. Probab. Lett.* **37** 259–267. MR1614966 https://doi.org/10.1016/S0167-7152(97)00125-9

[7] Knight, K. (2002). What are the limiting distributions of quantile estimators? In *Statistical Data Analysis Based on the L_1-Norm and Related Methods* (*Neuchâtel*, 2002) (Y. Dodge , ed.). *Stat. Ind. Technol.* 47–65. Basel: Birkhäuser. MR2001304

[8] Koenker, R. (2005). *Quantile Regression. Econometric Society Monographs* **38**. Cambridge: Cambridge Univ. Press. MR2268657 https://doi.org/10.1017/CBO9780511754098

[9] McNeil, A.J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management*: *Concepts*, *Techniques and Tools*, Revised ed. *Princeton Series in Finance*. Princeton, NJ: Princeton Univ. Press. MR3445371

[10] Molchanov, I. (2005). *Theory of Random Sets*. *Probability and Its Applications* (*New York*). London: Springer. MR2132405

[11] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847. MR0906565 https://doi.org/10.2307/1911031

[12] Rio, E. (2013, September). Inequalities and limit theorems for weakly dependent sequences. Lecture.

[13] 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

[14] 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

[15] Ziegel, J.F. (2016). Coherence and elicitability. *Math*. *Finance* **26** 901–918. MR3551510 https://doi.org/10.1111/mafi.12080

[16] Zwingmann, T. and Holzmann, H. (2020). Supplement to “Weak convergence of quantile and expectile processes under general assumptions.” https://doi.org/10.3150/19-BEJ1127SUPP.

Bernoulli **26**(1), 2020, 352–386
https://doi.org/10.3150/19-BEJ1128

SPDEs with fractional noise in space: Continuity in law with respect to the Hurst index

LUCA M. GIORDANO[1,2], MARIA JOLIS[2,*] and
LLUÍS QUER-SARDANYONS[2,**]

[1]*Department of Mathematics, University of Milano, Via C. Saldini 50, 20133 Milano, Italy.*
E-mail: luca.giordano@unimi.it
[2]*Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Catalonia, Spain.*
*E-mail: *mjolis@mat.uab.cat; **quer@mat.uab.cat*

In this article, we consider the quasi-linear stochastic wave and heat equations on the real line and with an additive Gaussian noise which is white in time and behaves in space like a fractional Brownian motion with Hurst index $H \in (0, 1)$. The drift term is assumed to be globally Lipschitz. We prove that the solution of each of the above equations is continuous in terms of the index H, with respect to the convergence in law in the space of continuous functions.

Keywords: fractional noise; stochastic heat equation; stochastic wave equation; weak convergence

References

[1] Ait Ouahra, M. and Sghir, A. (2017). Continuity in law with respect to the Hurst index of some additive functionals of sub-fractional Brownian motion. *Stoch. Anal. Appl.* **35** 677–690. MR3651137 https://doi.org/10.1080/07362994.2017.1309979

[2] Balan R.M., Quer-Sardanyons, L. and Song, J. (2018). Existence of density for the stochastic wave equation with space-time homogeneous Gaussian noise. Available at https://arxiv.org/abs/1805.06936.

[3] Balan, R.M. (2012). Linear SPDEs driven by stationary random distributions. *J. Fourier Anal. Appl.* **18** 1113–1145. MR3000977 https://doi.org/10.1007/s00041-012-9240-7

[4] Balan, R.M., Jolis, M. and Quer-Sardanyons, L. (2015). SPDEs with affine multiplicative fractional noise in space with index $1/4 < H < 1/2$. *Electron. J. Probab.* **20** 54. MR3354614 https://doi.org/10.1214/EJP.v20-3719

[5] Balan, R.M., Jolis, M. and Quer-Sardanyons, L. (2017). Intermittency for the hyperbolic Anderson model with rough noise in space. *Stochastic Process. Appl.* **127** 2316–2338. MR3652415 https://doi.org/10.1016/j.spa.2016.10.009

[6] Bertola, G. and Jolis, M. (2015). Weak convergence for the stochastic wave equation driven by Gausisan white noise. Working draft.

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

[8] Chen, L. and Dalang, R.C. (2015). Moment bounds and asymptotics for the stochastic wave equation. *Stochastic Process. Appl.* **125** 1605–1628. MR3310358 https://doi.org/10.1016/j.spa.2014.11.009

1350-7265

[9] 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

[10] Dalang, R.C. and Quer-Sardanyons, L. (2011). Stochastic integrals for spde's: A comparison. *Expo. Math.* **29** 67–109. MR2785545 https://doi.org/10.1016/j.exmath.2010.09.005

[11] Hu, Y., Huang, J., Lê, K., Nualart, D. and Tindel, S. (2017). Stochastic heat equation with rough dependence in space. *Ann. Probab.* **45** 4561–4616. MR3737918 https://doi.org/10.1214/16-AOP1172

[12] Hu, Y., Huang, J., Nualart, D. and Tindel, S. (2015). Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.* **20** 55. MR3354615 https://doi.org/10.1214/EJP.v20-3316

[13] Jolis, M. and Viles, N. (2007). Continuity in law with respect to the Hurst parameter of the local time of the fractional Brownian motion. *J. Theoret. Probab.* **20** 133–152. MR2324522 https://doi.org/10.1007/s10959-007-0054-5

[14] Jolis, M. and Viles, N. (2007). Continuity with respect to the Hurst parameter of the laws of the multiple fractional integrals. *Stochastic Process. Appl.* **117** 1189–1207. MR2343935 https://doi.org/10.1016/j.spa.2006.12.005

[15] Jolis, M. and Viles, N. (2010). Continuity in the Hurst parameter of the law of the Wiener integral with respect to the fractional Brownian motion. *Statist. Probab. Lett.* **80** 566–572. MR2595131 https://doi.org/10.1016/j.spl.2009.12.011

[16] Jolis, M. and Viles, N. (2010). Continuity in the Hurst parameter of the law of the symmetric integral with respect to the fractional Brownian motion. *Stochastic Process. Appl.* **120** 1651–1679. MR2673969 https://doi.org/10.1016/j.spa.2010.05.002

[17] Pipiras, V. and Taqqu, M.S. (2000). Integration questions related to fractional Brownian motion. *Probab. Theory Related Fields* **118** 251–291. MR1790083 https://doi.org/10.1007/s440-000-8016-7

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

[19] Wu, D. and Xiao, Y. (2009). Continuity in the Hurst index of the local times of anisotropic Gaussian random fields. *Stochastic Process. Appl.* **119** 1823–1844. MR2519346 https://doi.org/10.1016/j.spa.2008.09.001

[20] Yor, M. (1983). Le drap brownien comme limite en loi de temps locaux linéaires. In *Seminar on Probability, XVII. Lecture Notes in Math.* **986** 89–105. Berlin: Springer. MR0770400 https://doi.org/10.1007/BFb0068303

Bernoulli **26**(1), 2020, 387–417
https://doi.org/10.3150/19-BEJ1129

High dimensional deformed rectangular matrices with applications in matrix denoising

XIUCAI DING

Department of Statistics, University of Toronto, Toronto, ON, M5S 3G3, Canada.
E-mail: xiucai.ding@mail.utoronto.ca

We consider the recovery of a low rank $M \times N$ matrix S from its noisy observation $\tilde{S}$ in the high dimensional framework when M is comparable to N. We propose two efficient estimators for S under two different regimes. Our analysis relies on the local asymptotics of the eigenstructure of large dimensional rectangular matrices with finite rank perturbation. We derive the convergent limits and rates for the singular values and vectors for such matrices.

Keywords: matrix denoising; random matrices; rotation invariant estimation; singular value decomposition

References

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

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

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

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

[5] Bun, J., Allez, R., Bouchaud, J.-P. and Potters, M. (2016). Rotational invariant estimator for general noisy matrices. *IEEE Trans. Inform. Theory* **62** 7475–7490. MR3599095 https://doi.org/10.1109/TIT.2016.2616132

[6] Bun, J., Bouchaud, J.-P. and Potters, M. (2017). Cleaning large correlation matrices: Tools from random matrix theory. *Phys. Rep.* **666** 1–109. MR3590056 https://doi.org/10.1016/j.physrep.2016.10.005

[7] Ding, X. (2017). Asymptotics of empirical eigen-structure of general covariance matrices. Available at arXiv:1708.06296.

[8] Ding, X. (2020). Supplement to "High dimensional deformed rectangular matrices with applications in matrix denoising." https://doi.org/10.3150/19-BEJ1129SUPP.

[9] Ding, X. (2019). Singular vector distribution of sample covariance matrices. *Adv. in Appl. Probab.* **51** 236–267. MR3984017 https://doi.org/10.1017/apr.2019.10

1350-7265

[10] Ding, X. and Yang, F. (2018). A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices. *Ann. Appl. Probab.* **28** 1679–1738. MR3809475 https://doi.org/10.1214/17-AAP1341

[11] Donoho, D.L. (1995). De-noising by soft-thresholding. *IEEE Trans. Inform. Theory* **41** 613–627. MR1331258 https://doi.org/10.1109/18.382009

[12] Elad, M. (2010). *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer. MR2677506 https://doi.org/10.1007/978-1-4419-7011-4

[13] Gavish, M. and Donoho, D.L. (2014). The optimal hard threshold for singular values is $4/\sqrt{3}$. *IEEE Trans. Inform. Theory* **60** 5040–5053. MR3245370 https://doi.org/10.1109/TIT.2014.2323359

[14] Gavish, M. and Donoho, D.L. (2017). Optimal shrinkage of singular values. *IEEE Trans. Inform. Theory* **63** 2137–2152. MR3626861 https://doi.org/10.1109/TIT.2017.2653801

[15] Golub, G.H. and Van Loan, C.F. (1996). *Matrix Computations*, 3rd ed. *Johns Hopkins Studies in the Mathematical Sciences*. Baltimore, MD: Johns Hopkins Univ. Press. MR1417720

[16] James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013). *An Introduction to Statistical Learning*. *Springer Texts in Statistics* **103**. New York: Springer. MR3100153 https://doi.org/10.1007/978-1-4614-7138-7

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

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

[19] Knowles, A. and Yin, J. (2017). Anisotropic local laws for random matrices. *Probab. Theory Related Fields* **169** 257–352. MR3704770 https://doi.org/10.1007/s00440-016-0730-4

[20] Laloux, L., Cizeau, P., Potters, M. and Bouchaud, J. (2000). Random matrix theory and financial correlations. *Int. J. Theor. Appl. Finance* **3** 391–397.

[21] Lam, C. and Yao, Q. (2012). Factor modeling for high-dimensional time series: Inference for the number of factors. *Ann. Statist.* **40** 694–726. MR2933663 https://doi.org/10.1214/12-AOS970

[22] Ledoit, O. and Péché, S. (2011). Eigenvectors of some large sample covariance matrix ensembles. *Probab. Theory Related Fields* **151** 233–264. MR2834718 https://doi.org/10.1007/s00440-010-0298-3

[23] Lee, M., Shen, H., Huang, J.Z. and Marron, J.S. (2010). Biclustering via sparse singular value decomposition. *Biometrics* **66** 1087–1095. MR2758496 https://doi.org/10.1111/j.1541-0420.2010.01392.x

[24] Marčenko, V.A. and Pastur, L.A. (1967). Distribution of eigenvalues for some sets of random matrices. *Math. USSR, Sb.* **1** 457.

[25] Nadakuditi, R.R. (2014). OptShrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage. *IEEE Trans. Inform. Theory* **60** 3002–3018. MR3200641 https://doi.org/10.1109/TIT.2014.2311661

[26] Nadakuditi, R.R. and Edelman, A. (2008). Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Trans. Signal Process.* **56** 2625–2638. MR1500236 https://doi.org/10.1109/TSP.2008.917356

[27] Passemier, D. and Yao, J. (2014). Estimation of the number of spikes, possibly equal, in the high-dimensional case. *J. Multivariate Anal.* **127** 173–183. MR3188885 https://doi.org/10.1016/j.jmva.2014.02.017

[28] Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865

[29] Pillai, N.S. and Yin, J. (2014). Universality of covariance matrices. *Ann. Appl. Probab.* **24** 935–1001. MR3199978 https://doi.org/10.1214/13-AAP939

[30] Pizzo, A., Renfrew, D. and Soshnikov, A. (2013). On finite rank deformations of Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 64–94. MR3060148 https://doi.org/10.1214/11-AIHP459

[31] Pontes, B., Giráldez, R. and Aguilar-Ruiz, J. (2015). Biclustering on expression data: A review. *BMC Bioinform*. **57** 163–180.

[32] Renfrew, D. and Soshnikov, A. (2013). On finite rank deformations of Wigner matrices II: Delocalized perturbations. *Random Matrices Theory Appl*. **2** 1250015, 36. MR3039820 https://doi.org/10.1142/S2010326312500153

[33] Silverstein, J.W. (2009). The Stieltjes transform and its role in eigenvalue behavior of large dimensional random matrices. In *Random Matrix Theory and Its Applications*. *Lect*. *Notes Ser. Inst*. *Math*. *Sci*. *Natl*. *Univ. Singap*. **18** 1–25. Hackensack, NJ: World Sci. Publ. MR2603192 https://doi.org/10.1142/9789814273121_0001

[34] Tao, T. (2012). *Topics in Random Matrix Theory*. *Graduate Studies in Mathematics* **132**. Providence, RI: Amer. Math. Soc. MR2906465 https://doi.org/10.1090/gsm/132

[35] Tracy, C.A. and Widom, H. (1996). On orthogonal and symplectic matrix ensembles. *Comm*. *Math*. *Phys*. **177** 727–754. MR1385083

[36] Tufts, D. and Shah, A. (1993). Estimation of a signal waveform from noisy data using low-rank approximation to a data matrix. *IEEE Trans*. *Signal Process*. **41** 7475–7490.

[37] Yang, D., Ma, Z. and Buja, A. (2016). Rate optimal denoising of simultaneously sparse and low rank matrices. *J*. *Mach*. *Learn*. *Res*. **17** Paper No. 92, 27. MR3543498

Bernoulli **26**(1), 2020, 418–431
https://doi.org/10.3150/19-BEJ1130

Construction results for strong orthogonal arrays of strength three

CHENLU SHI* and BOXIN TANG**

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada. E-mail: *chenlus@sfu.ca; **boxint@sfu.ca

Strong orthogonal arrays were recently introduced as a class of space-filling designs for computer experiments. The most attractive are those of strength three for their economical run sizes. Although the existence of strong orthogonal arrays of strength three has been completely characterized, the construction of these arrays has not been explored. In this paper, we provide a systematic and comprehensive study on the construction of these arrays, with the aim at better space-filling properties. Besides various characterizing results, three families of strength-three strong orthogonal arrays are presented. One of these families deserves special mention, as the arrays in this family enjoy almost all of the space-filling properties of strength-four strong orthogonal arrays, and do so with much more economical run sizes than the latter. The theory of maximal designs and their doubling constructions plays a crucial role in many of theoretical developments.

Keywords: computer experiment; doubling and projection; maximal design; second order saturated design; space-filling design

References

[1] Bates, R.A., Buck, R.J., Riccomagno, E. and Wynn, H.P. (1996). Experimental design and observation for large systems. *J. Roy. Statist. Soc. Ser. B* **58** 77–94, 95–111. MR1379235

[2] Bingham, D., Sitter, R.R. and Tang, B. (2009). Orthogonal and nearly orthogonal designs for computer experiments. *Biometrika* **96** 51–65. MR2482134 https://doi.org/10.1093/biomet/asn057

[3] Butler, N.A. (2003). Some theory for constructing minimum aberration fractional factorial designs. *Biometrika* **90** 233–238. MR1966563 https://doi.org/10.1093/biomet/90.1.233

[4] Chen, H.H. and Cheng, C.-S. (2006). Doubling and projection: A method of constructing two-level designs of resolution IV. *Ann. Statist.* **34** 546–558. MR2275253 https://doi.org/10.1214/009053605000000813

[5] Cheng, C.S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments*. Boca Raton: CRC Press.

[6] Dey, A. and Mukerjee, R. (1999). *Fractional Factorial Plans. Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1679441 https://doi.org/10.1002/9780470316986

[7] Fang, K.-T., Li, R. and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments. Chapman & Hall/CRC Computer Science and Data Analysis Series*. Boca Raton, FL: CRC Press/CRC. MR2223960

[8] Fang, K.-T., Lin, D.K.J., Winker, P. and Zhang, Y. (2000). Uniform design: Theory and application. *Technometrics* **42** 237–248. MR1801031 https://doi.org/10.2307/1271079

[9] Georgiou, S.D. and Efthimiou, I. (2014). Some classes of orthogonal Latin hypercube designs. *Statist. Sinica* **24** 101–120. MR3183676

1350-7265

[10] Haaland, B. and Qian, P.Z.G. (2010). An approach to constructing nested space-filling designs for multi-fidelity computer experiments. *Statist*. *Sinica* **20** 1063–1075. MR2729853
[11] He, Y. and Tang, B. (2013). Strong orthogonal arrays and associated Latin hypercubes for computer experiments. *Biometrika* **100** 254–260. MR3034340 https://doi.org/10.1093/biomet/ass065
[12] He, Y. and Tang, B. (2014). A characterization of strong orthogonal arrays of strength three. *Ann*. *Statist*. **42** 1347–1360. MR3226159 https://doi.org/10.1214/14-AOS1225
[13] Hedayat, A.S., Sloane, N.J.A. and Stufken, J. (1999). *Orthogonal Arrays*: *Theory and Applications*. *Springer Series in Statistics*. New York: Springer. MR1693498 https://doi.org/10.1007/978-1-4612-1478-6
[14] Johnson, M.E., Moore, L.M. and Ylvisaker, D. (1990). Minimax and maximin distance designs. *J*. *Statist*. *Plann*. *Inference* **26** 131–148. MR1079258 https://doi.org/10.1016/0378-3758(90)90122-B
[15] Lawrence, K.M. (1996). A combinatorial characterization of (t, m, s)-nets in base b. *J*. *Combin*. *Des*. **4** 275–293. MR1391810 https://doi.org/10.1002/(SICI)1520-6610(1996)4:4<275::AID-JCD5>3.0.CO;2-C
[16] Lin, C.D. and Kang, L. (2016). A general construction for space-filling Latin hypercubes. *Statist*. *Sinica* **26** 675–690. MR3497766
[17] Liu, H. and Liu, M.-Q. (2015). Column-orthogonal strong orthogonal arrays and sliced strong orthogonal arrays. *Statist*. *Sinica* **25** 1713–1734. MR3409089
[18] McKay, M.D., Beckman, R.J. and Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21** 239–245. MR0533252 https://doi.org/10.2307/1268522
[19] Moon, H., Dean, A. and Santner, T. (2011). Algorithms for generating maximin Latin hypercube and orthogonal designs. *J*. *Stat*. *Theory Pract*. **5** 81–98. MR2829824 https://doi.org/10.1080/15598608.2011.10412052
[20] Mullen, G.L. and Schmid, W.C. (1996). An equivalence between (t, m, s)-nets and strongly orthogonal hypercubes. *J*. *Combin*. *Theory Ser. A* **76** 164–174. MR1406000 https://doi.org/10.1006/jcta.1996.0098
[21] 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
[22] Owen, A.B. (1992). Orthogonal arrays for computer experiments, integration and visualization. *Statist*. *Sinica* **2** 439–452. MR1187952
[23] Owen, A.B. (1997). Scrambled net variance for integrals of smooth functions. *Ann*. *Statist*. **25** 1541–1562. MR1463564 https://doi.org/10.1214/aos/1031594731
[24] Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer. MR2160708 https://doi.org/10.1007/978-1-4757-3799-8
[25] Tang, B. (1993). Orthogonal array-based Latin hypercubes. *J*. *Amer. Statist*. *Assoc*. **88** 1392–1397. MR1245375
[26] Wang, L., Xiao, Q. and Xu, H. (2018). Optimal maximin L_1-distance Latin hypercube designs based on good lattice point designs. *Ann*. *Statist*. **46** 3741–3766. MR3852667 https://doi.org/10.1214/17-AOS1674
[27] Wu, C.-F.J. (1989). Construction of $2^m 4^n$ designs via a grouping scheme. *Ann*. *Statist*. **17** 1880–1885. MR1026317 https://doi.org/10.1214/aos/1176347399
[28] Ye, K.Q. (1998). Orthogonal column Latin hypercubes and their application in computer experiments. *J*. *Amer. Statist*. *Assoc*. **93** 1430–1439. MR1666638 https://doi.org/10.2307/2670057

Bernoulli **26**(1), 2020, 432–470
https://doi.org/10.3150/19-BEJ1131

Subspace perspective on canonical correlation analysis: Dimension reduction and minimax rates

ZHUANG MA[1] and XIAODONG LI[2]

[1] *Department of Statistics, the Wharton School, University of Pennsylvania, 3730 Walnut street, Suite 400, Philadelphia, PA 19104, USA. E-mail: kop.mazhuang@gmail.com*
[2] *Department of Statistics, University of California, Davis, Davis, CA 95616, USA. E-mail: xdgli@ucdavis.edu*

Canonical correlation analysis (CCA) is a fundamental statistical tool for exploring the correlation structure between two sets of random variables. In this paper, motivated by the recent success of applying CCA to learn low dimensional representations of high dimensional objects, we propose two losses based on the principal angles between the model spaces spanned by the sample canonical variates and their population correspondents, respectively. We further characterize the non-asymptotic error bounds for the estimation risks under the proposed error metrics, which reveal how the performance of sample CCA depends adaptively on key quantities including the dimensions, the sample size, the condition number of the covariance matrices and particularly the population canonical correlation coefficients. The optimality of our uniform upper bounds is also justified by lower-bound analysis based on stringent and localized parameter spaces. To the best of our knowledge, for the first time our paper separates p_1 and p_2 for the first order term in the upper bounds without assuming the residual correlations are zeros. More significantly, our paper derives $(1-\lambda_k^2)(1-\lambda_{k+1}^2)/(\lambda_k-\lambda_{k+1})^2$ for the first time in the non-asymptotic CCA estimation convergence rates, which is essential to understand the behavior of CCA when the leading canonical correlation coefficients are close to 1.

Keywords: canonical correlation analysis; dimension reduction; minimax rates

References

[1] Anderson, T.W. (1999). Asymptotic theory for canonical correlation analysis. *J. Multivariate Anal*. **70** 1–29. MR1701396 https://doi.org/10.1006/jmva.1999.1810

[2] Arora, R. and Livescu, K. (2013). Multi-view cca-based acoustic features for phonetic recognition across speakers and domains. In *Acoustics*, *Speech and Signal Processing* (*ICASSP*), 2013 *IEEE International Conference on* 7135–7139. IEEE.

[3] Cai, T., Ma, Z. and Wu, Y. (2015). Optimal estimation and rank detection for sparse spiked covariance matrices. *Probab. Theory Related Fields* **161** 781–815. MR3334281 https://doi.org/10.1007/s00440-014-0562-z

[4] Cai, T.T., Ma, Z. and Wu, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist*. **41** 3074–3110. MR3161458 https://doi.org/10.1214/13-AOS1178

[5] Cai, T.T. and Zhang, A. (2018). Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics. *Ann. Statist*. **46** 60–89. MR3766946 https://doi.org/10.1214/17-AOS1541

1350-7265

[6] Chaudhuri, K., Kakade, S.M., Livescu, K. and Sridharan, K. (2009). Multi-view clustering via canonical correlation analysis. In *Proceedings of the* 26*th Annual International Conference on Machine Learning* 129–136. ACM.
[7] Chen, X., Liu, H. and Carbonell, J.G. (2012). Structured sparse canonical correlation analysis. In *International Conference on Artificial Intelligence and Statistics* 199–207.
[8] Dhillon, P.S., Foster, D. and Ungar, L. (2011). Multi-view learning of word embeddings via cca. In *Advances in Neural Information Processing Systems* (*NIPS*), *Vol*. 24.
[9] Faruqui, M. and Dyer, C. (2014). Improving vector space word representations using multilingual correlation. Association for Computational Linguistics.
[10] Foster, D.P., Johnson, R., Kakade, S.M. and Zhang, T. (2008). Multi-view dimensionality reduction via canonical correlation analysis. Technical report.
[11] Friman, O., Borga, M., Lundberg, P. and Knutsson, H. (2003). Adaptive analysis of fmri data. *NeuroImage* **19** 837–845.
[12] Fukumizu, K., Bach, F.R. and Jordan, M.I. (2009). Kernel dimension reduction in regression. *Ann*. *Statist*. **37** 1871–1905. MR2533474 https://doi.org/10.1214/08-AOS637
[13] Gao, C., Ma, Z., Ren, Z. and Zhou, H.H. (2015). Minimax estimation in sparse canonical correlation analysis. *Ann*. *Statist*. **43** 2168–2197. MR3396982 https://doi.org/10.1214/15-AOS1332
[14] Gao, C., Ma, Z. and Zhou, H.H. (2017). Sparse CCA: Adaptive estimation and computational barriers. *Ann*. *Statist*. **45** 2074–2101. MR3718162 https://doi.org/10.1214/16-AOS1519
[15] Gong, Y., Ke, Q., Isard, M. and Lazebnik, S. (2014). A multi-view embedding space for modeling Internet images, tags, and their semantics. *Int*. *J*. *Comput*. *Vis*. **106** 210–233.
[16] Horn, R.A. and Johnson, C.R. (1991). *Topics in Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR1091716 https://doi.org/10.1017/CBO9780511840371
[17] Hotelling, H. (1936). Relations between two sets of variables. *Biometrika* **28** 312–377.
[18] Kakade, S.M. and Foster, D.P. (2007). Multi-view regression via canonical correlation analysis. In *Learning Theory*. *Lecture Notes in Computer Science* **4539** 82–96. Berlin: Springer. MR2397580 https://doi.org/10.1007/978-3-540-72927-3_8
[19] Kim, T.-K., Wong, S.-F. and Cipolla, R. (2007). Tensor canonical correlation analysis for action classification. In *Computer Vision and Pattern Recognition*, 2007. *CVPR'*07. *IEEE Conference on* 1–8. IEEE.
[20] Ma, Z. and Li, X. (2020). Supplement to "Subspace perspective on canonical correlation analysis: Dimension reduction and minimax rates." https://doi.org/10.3150/19-BEJ1131SUPP.
[21] Mathias, R. (1993). The Hadamard operator norm of a circulant and applications. *SIAM J*. *Matrix Anal*. *Appl*. **14** 1152–1167. MR1238930 https://doi.org/10.1137/0614080
[22] Rasiwasia, N., Costa Pereira, J., Coviello, E., Doyle, G., Lanckriet, G.R., Levy, R. and Vasconcelos, N. (2010). A new approach to cross-modal multimedia retrieval. In *Proceedings of the* 18*th ACM International Conference on Multimedia* 251–260. ACM.
[23] Rudelson, M. and Vershynin, R. (2013). Hanson–Wright inequality and sub-Gaussian concentration. *Electron*. *Commun*. *Probab*. **18** no. 82, 9. MR3125258 https://doi.org/10.1214/ECP.v18-2865
[24] Sridharan, K. and Kakade, S.M. (2008). An information theoretic framework for multi-view learning. In *COLT* (R.A. Servedio and T. Zhang, eds.) 403–414. Omnipress.
[25] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170
[26] Vu, V.Q. and Lei, J. (2013). Minimax sparse principal subspace estimation in high dimensions. *Ann*. *Statist*. **41** 2905–2947. MR3161452 https://doi.org/10.1214/13-AOS1151
[27] Wang, W., Arora, R., Livescu, K. and Bilmes, J. (2015). On deep multi-view representation learning. In *Proceedings of the* 32*nd International Conference on Machine Learning* (*ICML*-15) 1083–1092.

[28] Wedin, P. (1972). Perturbation bounds in connection with singular value decomposition. *BIT* **12** 99–111. MR0309968 https://doi.org/10.1007/bf01932678
[29] Wedin, P.Å. (1983). On angles between subspaces of a finite dimensional inner product space. In *Matrix Pencils* 263–285. Springer.
[30] Witten, D.M., Tibshirani, R. and Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics* kxp008.

Bernoulli **26**(1), 2020, 471–499
https://doi.org/10.3150/19-BEJ1132

Multivariate count autoregression

KONSTANTINOS FOKIANOS[1], BÅRD STØVE[2,*], DAG TJØSTHEIM[2,**] and PAUL DOUKHAN[3]

[1]*Department of Mathematics & Statistics, Lancaster University, Lancaster, LA1 4YF, UK.*
E-mail: k.fokianos@lancaster.ac.uk
[2]*Department of Mathematics, University of Bergen, Postbox 7802, 5020 Bergen, Norway.*
E-mail: [*]Bard.Stove@math.uib.no; [**]Dag.Tjostheim@math.uib.no
[3]*AGM UMR 8088, University of Cergy-Pontoise, Department of Mathematics, 2 av. Adolphe Chauvin, 95302, Cergy-Pontoise, CEDEX, France.*
E-mail: doukhan@u-cergy.fr

We are studying linear and log-linear models for multivariate count time series data with Poisson marginals. For studying the properties of such processes we develop a novel conceptual framework which is based on copulas. Earlier contributions impose the copula on the joint distribution of the vector of counts by employing a continuous extension methodology. Instead we introduce a copula function on a vector of associated continuous random variables. This construction avoids conceptual difficulties related to the joint distribution of counts yet it keeps the properties of the Poisson process marginally. Furthermore, this construction can be employed for modeling multivariate count time series with other marginal count distributions. We employ Markov chain theory and the notion of weak dependence to study ergodicity and stationarity of the models we consider. Suitable estimating equations are suggested for estimating unknown model parameters. The large sample properties of the resulting estimators are studied in detail. The work concludes with some simulations and a real data example.

Keywords: autocorrelation; copula; ergodicity; generalized linear models; perturbation; prediction; stationarity; volatility

References

[1] Ahmad, A. (2016). Contributions à l'éconemétrie des séries temporelles à valeurs entières. Ph.D. thesis, Univ. Charles De Gaulle-Lille III, France.
[2] Ahmad, A. and Francq, C. (2016). Poisson QMLE of count time series models. *J. Time Series Anal.* **37** 291–314. MR3512959 https://doi.org/10.1111/jtsa.12167
[3] Andreassen, C.M. (2013). Models and inference for correlated count data. Ph.D. thesis, Aarhus Univ., Denmark.
[4] Andrews, D.W.K. (1984). Nonstrong mixing autoregressive processes. *J. Appl. Probab.* **21** 930–934. MR0766830 https://doi.org/10.2307/3213710
[5] Basawa, I.V. and Prakasa Rao, B.L.S. (1980). *Statistical Inference for Stochastic Processes*. London-New York: Academic Press. MR0586053
[6] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1
[7] Christou, V. and Fokianos, K. (2014). Quasi-likelihood inference for negative binomial time series models. *J. Time Series Anal.* **35** 55–78. MR3148248 https://doi.org/10.1111/jtsa.12050

1350-7265

[8] Cox, D.R. (1981). Statistical analysis of time series: Some recent developments. *Scand. J. Stat.* **8** 93–115. MR0623586
[9] Davis, R.A., Dunsmuir, W.T.M. and Wang, Y. (2000). On autocorrelation in a Poisson regression model. *Biometrika* **87** 491–505. MR1789805 https://doi.org/10.1093/biomet/87.3.491
[10] 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
[11] Dedecker, J., Doukhan, P., Lang, G., José, R., León, R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence: With Examples and Applications. Lecture Notes in Statistics* **190**. New York: Springer. MR2338725
[12] Denuit, M. and Lambert, P. (2005). Constraints on concordance measures in bivariate discrete data. *J. Multivariate Anal.* **93** 40–57. MR2119763 https://doi.org/10.1016/j.jmva.2004.01.004
[13] 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
[14] Douc, R., Fokianos, K. and Moulines, E. (2017). Asymptotic properties of quasi-maximum likelihood estimators in observation-driven time series models. *Electron. J. Stat.* **11** 2707–2740. MR3679907 https://doi.org/10.1214/17-EJS1299
[15] Doukhan, P., Fokianos, K. and Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statist. Probab. Lett.* **82** 942–948. MR2910041 https://doi.org/10.1016/j.spl.2012.01.015
[16] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** 313–342. MR1719345 https://doi.org/10.1016/S0304-4149(99)00055-1
[17] Doukhan, P. and Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Process. Appl.* **118** 1997–2013. MR2462284 https://doi.org/10.1016/j.spa.2007.12.004
[18] Efron, B. (1986). Double exponential families and their use in generalized linear regression. *J. Amer. Statist. Assoc.* **81** 709–721. MR0860505
[19] 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
[20] Fokianos, K. (2016). Statistical analysis of count time series models: A GLM perspective. In *Handbook of Discrete-Valued Time Series* (R. Davis, S. Holan, R. Lund and N. Ravishanker, eds.). *Chapman & Hall/CRC Handb. Mod. Stat. Methods* 3–27. Boca Raton, FL: CRC Press. MR3699398
[21] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Poisson autoregression (complete vesrion). Available at http://pubs.amstat.org/toc/jasa/104/488.
[22] 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
[23] 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
[24] Fokianos, K., Støve, B., Tjøstheim, D. and Doukhan, P. (2020). Supplement to "Multivariate count autoregression." https://doi.org/10.3150/19-BEJ1132SUPP.
[25] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: Wiley. MR3186556 https://doi.org/10.1002/9780470670057
[26] Franke, J. and Subba Rao, T. (1995). Multivariate first-order integer values autoregressions. Technical report, Department of Mathematics, UMIST.
[27] 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
[28] Heinen, A. and Rengifo, E. (2008). Multivariate reduced rank regression in non-Gaussian contexts, using copulas. *Comput. Statist. Data Anal.* **52** 2931–2944. MR2424769 https://doi.org/10.1016/j.csda.2007.08.012

[29] Heyde, C.C. (1997). *Quasi-Likelihood and Its Application*: *A General Approach to Optimal Parameter Estimation*. *Springer Series in Statistics*. New York: Springer. MR1461808 https://doi.org/10.1007/b98823

[30] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. *Monographs on Statistics and Applied Probability* **73**. London: CRC Press. MR1462613 https://doi.org/10.1201/b13150

[31] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. *Wiley Series in Probability and Statistics*: *Applied Probability and Statistics*. New York: Wiley. MR1429617

[32] Jørgensen, B., Lundbye-Christensen, S., Song, P.X.-K. and Sun, L. (1996). State-space models for multivariate longitudinal data of mixed types. *Canad*. *J*. *Statist*. **24** 385–402. MR1421864 https://doi.org/10.2307/3315747

[33] Jung, R.C., Liesenfeld, R. and Richard, J.-F. (2011). Dynamic factor models for multivariate count data: An application to stock-market trading activity. *J*. *Bus*. *Econom*. *Statist*. **29** 73–85. MR2789392 https://doi.org/10.1198/jbes.2009.08212

[34] Karlis, D. (2016). Models for multivariate count time series. In *Handbook of Discrete-Valued Time Series*. *Chapman & Hall/CRC Handb*. *Mod*. *Stat*. *Methods* 407–424. Boca Raton, FL: CRC Press. MR3699415

[35] Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann*. *Statist*. **6** 629–642. MR0494770

[36] Latour, A. (1997). The multivariate GINAR(p) process. *Adv. in Appl*. *Probab.* **29** 228–248. MR1432938 https://doi.org/10.2307/1427868

[37] Lee, Y., Lee, S. and Tjøstheim, D. (2018). Asymptotic normality and parameter change test for bivariate Poisson INGARCH models. *TEST* **27** 52–69. MR3764023 https://doi.org/10.1007/s11749-016-0510-6

[38] Liang, K.Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73** 13–22. MR0836430 https://doi.org/10.1093/biomet/73.1.13

[39] Liu, H. (2012). Some models for time series of count. PhD thesis, Columbia Univ., USA.

[40] Livsey, J., Lund, R., Kechagias, S. and Pipiras, V. (2018). Multivariate integer-valued time series with flexible autocovariances and their application to major hurricane counts. *Ann*. *Appl*. *Stat*. **12** 408–431. MR3773399 https://doi.org/10.1214/17-AOAS1098

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

[42] McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*, 2nd edn. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR3223057 https://doi.org/10.1007/978-1-4899-3242-6

[43] Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. *Communications and Control Engineering Series*. London: Springer London, Ltd. MR1287609 https://doi.org/10.1007/978-1-4471-3267-7

[44] Neumann, M.H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17** 1268–1284. MR2854772 https://doi.org/10.3150/10-BEJ313

[45] Nikoloulopoulos, A.K. (2013). On the estimation of normal copula discrete regression models using the continuous extension and simulated likelihood. *J*. *Statist*. *Plann*. *Inference* **143** 1923–1937. MR3095082 https://doi.org/10.1016/j.jspi.2013.06.015

[46] Paul, M., Held, L. and Toschke, A.M. (2008). Multivariate modelling of infectious disease surveillance data. *Stat*. *Med*. **27** 6250–6267. MR2522320 https://doi.org/10.1002/sim.3440

[47] Pedeli, X. and Karlis, D. (2013). On composite likelihood estimation of a multivariate INAR(1) model. *J*. *Time Series Anal*. **34** 206–220. MR3028367 https://doi.org/10.1111/jtsa.12003

[48] Pedeli, X. and Karlis, D. (2013). Some properties of multivariate INAR(1) processes. *Comput*. *Statist*. *Data Anal*. **67** 213–225. MR3079598 https://doi.org/10.1016/j.csda.2013.05.019

[49] Ravishanker, N., Serhiyenko, V. and Willig, M.R. (2014). Hierarchical dynamic models for multivariate times series of counts. *Stat. Interface* **7** 559–570. MR3302382 https://doi.org/10.4310/SII.2014.v7.n4.a11
[50] Ravishanker, N., Venkatesan, R. and Hu, S. (2016). Dynamic models for time series of counts with a marketing application. In *Handbook of Discrete-Valued Time Series* (R. Davis, S. Holan R. Lund and N. Ravishanker, eds.). *Chapman & Hall/CRC Handb. Mod. Stat. Methods* 425–445. Boca Raton, FL: CRC Press. MR3699416
[51] Rydberg, T.H. and Shephard, N. (2000). A modeling framework for the prices and times of trades on the New York stock exchange. In *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerlad, R.L. Smith, A.T. Walden and P.C. Young, eds.) 217–246. Cambridge: Isaac Newton Institute and Cambridge University Press.
[52] Seber, G.A.F. (2008). *A Matrix Handbook for Statisticians*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR2365265
[53] Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series*. *Springer Series in Statistics*. New York: Springer. MR1785484 https://doi.org/10.1007/978-1-4612-1162-4
[54] Tjøstheim, D. (2012). Some recent theory for autoregressive count time series. *TEST* **21** 413–438. MR2983224 https://doi.org/10.1007/s11749-012-0296-0
[55] Tjøstheim, D. (2015). Count time series with observation-driven autoregressive parameter dynamics. In *Handbook of Discrete-Valued Time Series* (R. Davis, S. Holan, R. Lund and N. Ravishanker, eds.). *Handbooks of Modern Statistical Methods* 77–100. London: Chapman & Hall.
[56] Wang, C., Liu, H., Yao, J., Davis, R.A. and Li, W.K. (2014). Self-excited threshold Poisson autoregression. *J. Amer. Statist. Assoc.* **109** 777–787. MR3223749 https://doi.org/10.1080/01621459.2013.872994
[57] Woodard, D.B., Matteson, D.S. and Henderson, S.G. (2011). Stationarity of generalized autoregressive moving average models. *Electron. J. Stat.* **5** 800–828. MR2824817 https://doi.org/10.1214/11-EJS627
[58] Wu, W.B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4** 207–226. MR2812816 https://doi.org/10.4310/SII.2011.v4.n2.a15

Bernoulli **26**(1), 2020, 500–530
https://doi.org/10.3150/19-BEJ1133

Operator-scaling Gaussian random fields via aggregation

YI SHEN[1] and YIZAO WANG[2]

[1]*Department of Statistics and Actuarial Science, University of Waterloo, Mathematics 3 Building, 200 University Avenue West Waterloo, ON N2L 3G1, Canada. E-mail: yi.shen@uwaterloo.ca*
[2]*Department of Mathematical Sciences, University of Cincinnati, 2815 Commons Way, ML-0025, Cincinnati, OH 45221-0025, USA. E-mail: yizao.wang@uc.edu*

We propose an aggregated random-field model, and investigate the scaling limits of the aggregated partial-sum random fields. In this model, each copy in the aggregation is a ± 1-valued random field built from two correlated one-dimensional random walks, the law of each determined by a random persistence parameter. A flexible joint distribution of the two parameters is introduced, and given the parameters the two correlated random walks are conditionally independent. For the aggregated random field, when the persistence parameters are independent, the scaling limit is a fractional Brownian sheet. When the persistence parameters are tail-dependent, characterized in the framework of multivariate regular variation, the scaling limit is more delicate, and in particular depends on the growth rates of the underlying rectangular region along two directions: at different rates different operator-scaling Gaussian random fields appear as the region area tends to infinity. In particular, at the so-called critical speed, a large family of Gaussian random fields with long-range dependence arise in the limit. We also identify four different regimes at non-critical speed where fractional Brownian sheets arise in the limit.

Keywords: aggregation; fractional Brownian sheet; functional central limit theorem; Gaussian random field; long-range dependence; operator-scaling property

References

[1] Benson, D.A., Meerschaert, M.M., Baeumer, B. and Scheffler, H.-P. (2006). Aquifer operator scaling and the effect on solute mixing and dispersion. *Water Resour. Res.* **42**.

[2] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Heidelberg: Springer. MR3075595 https://doi.org/10.1007/978-3-642-35512-7

[3] Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Stat.* **42** 1656–1670. MR0383482 https://doi.org/10.1214/aoms/1177693164

[4] Biermé, H. and Durieu, O. (2014). Invariance principles for self-similar set-indexed random fields. *Trans. Amer. Math. Soc.* **366** 5963–5989. MR3256190 https://doi.org/10.1090/S0002-9947-2014-06135-7

[5] Biermé, H., Durieu, O. and Wang, Y. (2017). Invariance principles for operator-scaling Gaussian random fields. *Ann. Appl. Probab.* **27** 1190–1234. MR3655864 https://doi.org/10.1214/16-AAP1229

[6] Biermé, H., Meerschaert, M.M. and Scheffler, H.-P. (2007). Operator scaling stable random fields. *Stochastic Process. Appl.* **117** 312–332. MR2290879 https://doi.org/10.1016/j.spa.2006.07.004

1350-7265

[7] Biermé, H., Richard, F., Rachidi, M. and Benhamou, C.-L. (2009). Anisotropic texture modeling and applications to medical image analysis. In *Mathematical Methods for Imaging and Inverse Problems*. *ESAIM Proc*. **26** 100–122. Les Ulis: EDP Sci. MR2498142 https://doi.org/10.1051/proc/2009008

[8] Durieu, O. and Wang, Y. (2016). From infinite urn schemes to decompositions of self-similar Gaussian process. *Electron*. *J*. *Probab*. **21** Paper No. 43, 23 pp. MR3530320 https://doi.org/10.1214/16-EJP4492

[9] Durieu, O. and Wang, Y. (2019). From random partitions to fractional Brownian sheets. *Bernoulli* **25** 1412–1450. MR3920377 https://doi.org/10.3150/18-bej1025

[10] Enriquez, N. (2004). A simple construction of the fractional Brownian motion. *Stochastic Process*. *Appl*. **109** 203–223. MR2031768 https://doi.org/10.1016/j.spa.2003.10.008

[11] Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. *J*. *Econometrics* **14** 227–238. MR0597259 https://doi.org/10.1016/0304-4076(80)90092-5

[12] Hammond, A. and Sheffield, S. (2013). Power law Pólya's urn and fractional Brownian motion. *Probab*. *Theory Related Fields* **157** 691–719. MR3129801 https://doi.org/10.1007/s00440-012-0468-6

[13] Kaj, I. and Taqqu, M.S. (2008). Convergence to fractional Brownian motion and to the Telecom process: The integral representation approach. In *In and Out of Equilibrium*. 2. *Progress in Probability* **60** 383–427. Basel: Birkhäuser. MR2477392 https://doi.org/10.1007/978-3-7643-8786-0_19

[14] Lavancier, F. (2006). Long memory random fields. In *Dependence in Probability and Statistics*. *Lect*. *Notes Stat*. **187** 195–220. New York: Springer. MR2283256 https://doi.org/10.1007/0-387-36062-X_9

[15] Lavancier, F. (2007). Invariance principles for non-isotropic long memory random fields. *Stat*. *Inference Stoch*. *Process*. **10** 255–282. MR2321311 https://doi.org/10.1007/s11203-006-9001-9

[16] Lavancier, F. (2011). Aggregation of isotropic autoregressive fields [corrigendum to MR2523650]. *J*. *Statist*. *Plann*. *Inference* **141** 3862–3866. MR2823655 https://doi.org/10.1016/j.jspi.2011.06.003

[17] Leonenko, N. and Taufer, E. (2013). Disaggregation of spatial autoregressive processes. *Spat*. *Stat*. **3** 1–20.

[18] Li, Y., Wang, W. and Xiao, Y. (2015). Exact moduli of continuity for operator-scaling Gaussian random fields. *Bernoulli* **21** 930–956. MR3338652 https://doi.org/10.3150/13-BEJ593

[19] Lopes, R. and Betrouni, N. (2009). Fractal and multifractal analysis: A review. *Med*. *Image Anal*. **13** 634–649.

[20] Meerschaert, M.M., Dogan, M., Dam, R.L., Hyndman, D.W. and Benson, D.A. (2013). Hydraulic conductivity fields: Gaussian or not? *Water Resour*. *Res*. **49** 4730–4737.

[21] 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

[22] Mikosch, T., Resnick, S., Rootzén, H. and Stegeman, A. (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann*. *Appl*. *Probab*. **12** 23–68. MR1890056 https://doi.org/10.1214/aoap/1015961155

[23] Mikosch, T. and Samorodnitsky, G. (2007). Scaling limits for cumulative input processes. *Math*. *Oper*. *Res*. **32** 890–918. MR2363203 https://doi.org/10.1287/moor.1070.0267

[24] Pilipauskaitė, V. and Surgailis, D. (2017). Scaling transition for nonlinear random fields with long-range dependence. *Stochastic Process*. *Appl*. **127** 2751–2779. MR3660890 https://doi.org/10.1016/j.spa.2016.12.011

[25] Pinsky, M.A. (2002). *Introduction to Fourier Analysis and Wavelets*. *Brooks/Cole Series in Advanced Mathematics*. Pacific Grove, CA: Brooks/Cole. MR2100936

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

[27] Puplinskaitė, D. and Surgailis, D. (2015). Scaling transition for long-range dependent Gaussian random fields. *Stochastic Process. Appl.* **125** 2256–2271. MR3322863 https://doi.org/10.1016/j.spa.2014.12.011

[28] Puplinskaitė, D. and Surgailis, D. (2016). Aggregation of autoregressive random fields and anisotropic long-range dependence. *Bernoulli* **22** 2401–2441. MR3498033 https://doi.org/10.3150/15-BEJ733

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

[30] Robinson, P.M. (1978). Statistical inference for a random coefficient autoregressive model. *Scand. J. Stat.* **5** 163–168. MR0509453

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

[32] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Stochastic Modeling*. New York: CRC Press. MR1280932

[33] Shen, Y. and Wang, Y. (2020). Supplement to "Operator-scaling Gaussian random fields via aggregation." https://doi.org/10.3150/19-BEJ1133SUPP.

[34] Wang, Y. (2014). An invariance principle for fractional Brownian sheets. *J. Theoret. Probab.* **27** 1124–1139. MR3278934 https://doi.org/10.1007/s10959-013-0483-2

[35] 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

[36] Xiao, Y. (2013). Recent developments on fractal properties of Gaussian random fields. In *Further Developments in Fractals and Related Fields. Trends Math.* 255–288. New York: Birkhäuser/Springer. MR3184196 https://doi.org/10.1007/978-0-8176-8400-6_13

Bernoulli **26**(1), 2020, 531–556
https://doi.org/10.3150/19-BEJ1137

Tail expectile process and risk assessment

ABDELAATI DAOUIA[1], STÉPHANE GIRARD[2] and GILLES STUPFLER[3]

[1]*Toulouse School of Economics, University of Toulouse Capitole, France.*
E-mail: abdelaati.daouia@tse-fr.eu
[2]*Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.*
E-mail: stephane.girard@inria.fr
[3]*School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK.*
E-mail: gilles.stupfler@nottingham.ac.uk

Expectiles define a least squares analogue of quantiles. They are determined by tail expectations rather than tail probabilities. For this reason and many other theoretical and practical merits, expectiles have recently received a lot of attention, especially in actuarial and financial risk management. Their estimation, however, typically requires to consider non-explicit asymmetric least squares estimates rather than the traditional order statistics used for quantile estimation. This makes the study of the tail expectile process a lot harder than that of the standard tail quantile process. Under the challenging model of heavy-tailed distributions, we derive joint weighted Gaussian approximations of the tail empirical expectile and quantile processes. We then use this powerful result to introduce and study new estimators of extreme expectiles and the standard quantile-based expected shortfall, as well as a novel expectile-based form of expected shortfall. Our estimators are built on general weighted combinations of both top order statistics and asymmetric least squares estimates. Some numerical simulations and applications to actuarial and financial data are provided.

Keywords: asymmetric least squares; coherent risk measures; expected shortfall; expectile; extrapolation; extremes; heavy tails; tail index

References

[1] Abdous, B. and Rémillard, B. (1995). Relating quantiles and expectiles under weighted-symmetry. *Ann. Inst. Statist. Math.* **47** 371–384. MR1345429 https://doi.org/10.1007/BF00773468
[2] Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Bank. Financ.* **26** 1505–1518.
[3] Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *J. Bank. Financ.* **26** 1487–1503.
[4] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Math. Finance* **9** 203–228. MR1850791 https://doi.org/10.1111/1467-9965.00068
[5] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications. Wiley Series in Probability and Statistics*. Chichester: Wiley. MR2108013 https://doi.org/10.1002/0470012382
[6] Bellini, F. and Di Bernardino, E. (2017). Risk management with expectiles. *Eur. J. Finance* **23** 487–506.
[7] Bellini, F., Klar, B., Müller, A. and Gianin, E.R. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.* **54** 41–48. MR3145849 https://doi.org/10.1016/j.insmatheco.2013.10.015
[8] Breckling, J. and Chambers, R. (1988). M-quantiles. *Biometrika* **75** 761–771. MR0995118 https://doi.org/10.2307/2336317

1350-7265

[9] Cai, J.-J., Einmahl, J.H.J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: The mean when a related variable is extreme. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 417–442. MR3310533 https://doi.org/10.1111/rssb.12069
[10] Csörgő, M. and Horváth, L. (1987). Approximation of intermediate quantile processes. *J. Multivariate Anal.* **21** 250–262. MR0884099 https://doi.org/10.1016/0047-259X(87)90004-2
[11] Daníelsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O. and Shin, H.S. (2001). An Academic Response to Basel II. Special paper no. 130, Financial Markets Group, London School of Economics.
[12] Daouia, A., Gijbels, I. and Stupfler, G. (2018). Extremiles: A new perspective on asymmetric least squares. *J. Amer. Statist. Assoc.* To appear. Available at https://www.tandfonline.com/doi/full/10.1080/01621459.2018.1498348.
[13] Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 263–292. MR3763692 https://doi.org/10.1111/rssb.12254
[14] Daouia, A., Girard, S. and Stupfler, G. (2019). Extreme M-quantiles as risk measures: From L^1 to L^p optimization. *Bernoulli* **25** 264–309. MR3892320 https://doi.org/10.3150/17-bej987
[15] Daouia, A., Girard, S. and Stupfler, G. (2020). Supplement to "Tail expectile process and risk assessment." https://doi.org/10.3150/19-BEJ1137SUPP.
[16] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3
[17] de Haan, L., Mercadier, C. and Zhou, C. (2016). Adapting extreme value statistics to financial time series: Dealing with bias and serial dependence. *Finance Stoch.* **20** 321–354. MR3479324 https://doi.org/10.1007/s00780-015-0287-6
[18] Dietrich, D., de Haan, L. and Hüsler, J. (2002). Testing extreme value conditions. *Extremes* **5** 71–85. MR1947789 https://doi.org/10.1023/A:1020934126695
[19] Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Stat.* **25** 187–210. MR1614276 https://doi.org/10.1111/1467-9469.00097
[20] Drees, H., de Haan, L. and Li, D. (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions. *J. Statist. Plann. Inference* **136** 3498–3538. MR2284668 https://doi.org/10.1016/j.jspi.2005.02.017
[21] Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 505–562. MR3506792 https://doi.org/10.1111/rssb.12154
[22] Einmahl, J.H.J. and Mason, D.M. (1988). Strong limit theorems for weighted quantile processes. *Ann. Probab.* **16** 1623–1643. MR0958207
[23] El Methni, J., Gardes, L. and Girard, S. (2014). Non-parametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scand. J. Stat.* **41** 988–1012. MR3277034 https://doi.org/10.1111/sjos.12078
[24] El Methni, J. and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions. *Statist. Sinica* **27** 907–930. MR3675036
[25] El Methni, J. and Stupfler, G. (2018). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions. *Econ. Stat.* **6** 129–148. MR3797979 https://doi.org/10.1016/j.ecosta.2017.03.002
[26] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613 https://doi.org/10.1007/978-3-642-33483-2
[27] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174. MR0378204

[28] Holzmann, H. and Klar, B. (2016). Expectile asymptotics. *Electron. J. Stat.* **10** 2355–2371. MR3544290 https://doi.org/10.1214/16-EJS1173
[29] Hüsler, J. and Li, D. (2006). On testing extreme value conditions. *Extremes* **9** 69–86. MR2327882 https://doi.org/10.1007/s10687-006-0025-8
[30] Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50. MR0474644 https://doi.org/10.2307/1913643
[31] Krätschmer, V. and Zähle, H. (2017). Statistical inference for expectile-based risk measures. *Scand. J. Stat.* **44** 425–454. MR3658521
[32] Kuan, C.-M., Yeh, J.-H. and Hsu, Y.-C. (2009). Assessing value at risk with CARE, the conditional autoregressive expectile models. *J. Econometrics* **150** 261–270. MR2535521 https://doi.org/10.1016/j.jeconom.2008.12.002
[33] Mao, T., Ng, K.W. and Hu, T. (2015). Asymptotic expansions of generalized quantiles and expectiles for extreme risks. *Probab. Engrg. Inform. Sci.* **29** 309–327. MR3355606 https://doi.org/10.1017/S0269964815000017
[34] Mao, T. and Yang, F. (2015). Risk concentration based on expectiles for extreme risks under FGM copula. *Insurance Math. Econom.* **64** 429–439. MR3394666 https://doi.org/10.1016/j.insmatheco.2015.06.009
[35] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847. MR0906565 https://doi.org/10.2307/1911031
[36] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424
[37] Rockafellar, R.T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *J. Risk* **2** 21–42.
[38] Rockafellar, R.T. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *J. Bank. Financ.* **26** 1443–1471.
[39] Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression. *Comput. Statist. Data Anal.* **56** 755–767. MR2888723 https://doi.org/10.1016/j.csda.2010.11.015
[40] Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles. *J. Financ. Econom.* **6** 231–252.
[41] Vandewalle, B. and Beirlant, J. (2006). On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance Math. Econom.* **38** 441–459. MR2229509 https://doi.org/10.1016/j.insmatheco.2005.11.002
[42] Wang, S.S. (1996). Premium calculation by transforming the layer premium density. *Astin Bull.* **26** 71–92.
[43] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73** 812–815. MR0521329
[44] Wirch, J.L. and Hardy, M.R. (1999). A synthesis of risk measures for capital adequacy. *Insurance Math. Econom.* **25** 337–347.
[45] Ziegel, J.F. (2016). Coherence and elicitability. *Math. Finance* **26** 901–918. MR3551510 https://doi.org/10.1111/mafi.12080
[46] Zwingmann, T. and Holzmann, H. (2016). Asymptotics for the expected shortfall. Preprint.

Bernoulli **26**(1), 2020, 557–586
https://doi.org/10.3150/19-BEJ1140

Consistent semiparametric estimators for recurrent event times models with application to virtual age models

ERIC BEUTNER[1], LAURENT BORDES[2] and LAURENT DOYEN[3]

[1] *Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. E-mail: e.a.beutner@vu.nl*
[2] *Laboratoire de Mathématiques et de leurs Applications – IPRA, Univ. Pau and Pays de l'Adour, CNRS, F-64000 Pau, France. E-mail: laurent.bordes@univ-pau.fr*
[3] *Laboratoire Jean Kuntzmann, Univ. Grenoble Alpes, CNRS, F-38000 Grenoble, France. E-mail: laurent.doyen@univ-grenoble-alpes.fr*

Virtual age models are very useful to analyse recurrent events. Among the strengths of these models is their ability to account for treatment (or intervention) effects after an event occurrence. Despite their flexibility for modeling recurrent events, the number of applications is limited. This seems to be a result of the fact that in the semiparametric setting all the existing results assume the virtual age function that describes the treatment (or intervention) effects to be known. This shortcoming can be overcome by considering semiparametric virtual age models with parametrically specified virtual age functions. Yet, fitting such a model is a difficult task. Indeed, it has recently been shown that for these models the standard profile likelihood method fails to lead to consistent estimators. Here we show that consistent estimators can be constructed by smoothing the profile log-likelihood function appropriately. We show that our general result can be applied to most of the relevant virtual age models of the literature. Our approach shows that empirical process techniques may be a worthwhile alternative to martingale methods for studying asymptotic properties of these inference methods. A simulation study is provided to illustrate our consistency results together with an application to real data.

Keywords: effective age process; recurrent event data; semiparametric inference; smoothed profile likelihood; virtual age process

References

[1] Adekpedjou, A. and Stocker, R. (2015). A general class of semiparametric models for recurrent event data. *J. Statist. Plann. Inference* **156** 48–63. MR3267512 https://doi.org/10.1016/j.jspi.2014.07.012

[2] 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

[3] Beutner, E., Bordes, L. and Doyen, L. (2017). The failure of the profile likelihood method for a large class of semi-parametric models. *Bernoulli* **23** 3650–3684. MR3654819 https://doi.org/10.3150/16-BEJ861

[4] Beutner, E., Bordes, L. and Doyen, L. (2020). Supplement to "Consistent semiparametric estimators for recurrent event times models with application to virtual age models." https://doi.org/10.3150/19-BEJ1140SUPPA, https://doi.org/10.3150/19-BEJ1140SUPPB, https://

1350-7265

doi.org/10.3150/19-BEJ1140SUPPC, https://doi.org/10.3150/19-BEJ1140SUPPD, https://doi.org/10.3150/19-BEJ1140SUPPE, https://doi.org/10.3150/19-BEJ1140SUPPF, https://doi.org/10.3150/19-BEJ1140SUPPG.

[5] Cook, R.J. and Lawless, J.F. (2007). *The Statistical Analysis of Recurrent Events*. *Statistics for Biology and Health*. New York: Springer. MR3822124

[6] Dorado, C., Hollander, M. and Sethuraman, J. (1997). Nonparametric estimation for a general repair model. *Ann*. *Statist*. **25** 1140–1160. MR1447744 https://doi.org/10.1214/aos/1069362741

[7] Doyen, L. and Gaudoin, O. (2004). Classes of imperfect repair models based on reduction of failure intensity or virtual age. *Reliab*. *Eng*. *Syst*. *Saf*. **84** 45–56.

[8] 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

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

[10] Guerra de Toledo, M.L., Freitas, M.A., Colosimo, E.A. and Gilardoni, G.L. (2016). Optimal periodic maintenance policy under imperfect repair: A case study of off-road engines. *IIE Trans*. **48** 747–758.

[11] Jacod, J. (1974/75). Multivariate point processes: Predictable projection, Radon–Nikodým derivatives, representation of martingales. *Z*. *Wahrsch*. *Verw*. *Gebiete* **31** 235–253. MR0380978 https://doi.org/10.1007/BF00536010

[12] Kijima, M. (1989). Some results for repairable systems with general repair. *J*. *Appl*. *Probab*. **26** 89–102. MR0981254 https://doi.org/10.1017/s0021900200041826

[13] Kijima, M., Morimura, H. and Suzuki, Y. (1988). Periodical replacement problem without assuming minimal repair. *European J*. *Oper*. *Res*. **37** 194–203. MR0963925 https://doi.org/10.1016/0377-2217(88)90329-3

[14] Lawless, J.F. (1995). The analysis of recurrent events for multiple subjects. *J*. *R*. *Stat*. *Soc*. *Ser*. *C*. *Appl*. *Stat*. **44** 487–498.

[15] 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

[16] Murphy, S.A. and van der Vaart, A.W. (2000). On profile likelihood. *J*. *Amer*. *Statist*. *Assoc*. **95** 449–485. With comments and a rejoinder by the authors. MR1803168 https://doi.org/10.2307/2669386

[17] Peña, E.A. (2006). Dynamic modeling and statistical analysis of event times. *Statist*. *Sci*. **21** 487–500. MR2369983 https://doi.org/10.1214/088342306000000349

[18] Peña, E.A. (2016). Asymptotics for a class of dynamic recurrent event models. *J*. *Nonparametr*. *Stat*. **28** 716–735. MR3555454 https://doi.org/10.1080/10485252.2016.1225733

[19] Peña, E.A., Slate, E.H. and González, J.R. (2007). Semiparametric inference for a general class of models for recurrent events. *J*. *Statist*. *Plann*. *Inference* **137** 1727–1747. MR2323859 https://doi.org/10.1016/j.jspi.2006.05.004

[20] Peña, E.A., Strawderman, R.L. and Hollander, M. (2001). Nonparametric estimation with recurrent event data. *J*. *Amer*. *Statist*. *Assoc*. **96** 1299–1315. MR1946578 https://doi.org/10.1198/016214501753381922

[21] Sellke, T. (1988). Weak convergence of the Aalen estimator for a censored renewal process. In *Statistical Decision Theory and Related Topics*, *IV*, *Vol*. 2 (*West Lafayette*, *Ind*., 1986) 183–194. New York: Springer. MR0927132

[22] Sellke, T. and Siegmund, D. (1983). Sequential analysis of the proportional hazards model. *Biometrika* **70** 315–326. MR0712020 https://doi.org/10.1093/biomet/70.2.315

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

[24] 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

[25] 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

[26] Zeng, D. and Lin, D.Y. (2007). Efficient estimation for the accelerated failure time model. *J*. *Amer*. *Statist*. *Assoc*. **102** 1387–1396. MR2412556 https://doi.org/10.1198/016214507000001085

[27] Zeng, D. and Lin, D.Y. (2010). A general asymptotic theory for maximum likelihood estimation in semiparametric regression models with censored data. *Statist*. *Sinica* **20** 871–910. MR2682647

Bernoulli **26**(1), 2020, 587–615
https://doi.org/10.3150/19-BEJ1141

Normal approximation for sums of weighted U-statistics – application to Kolmogorov bounds in random subgraph counting

NICOLAS PRIVAULT[1] and GRZEGORZ SERAFIN[2]

[1]*School of Physical and Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link, Singapore 637371, Singapore. E-mail: nprivault@ntu.edu.sg*
[2]*Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Ul. Wybrzeże Wyspiańskiego 27, 54-129 Wrocław, Poland. E-mail: grzegorz.serafin@pwr.edu.pl*

We derive normal approximation bounds in the Kolmogorov distance for sums of discrete multiple integrals and weighted U-statistics made of independent Bernoulli random variables. Such bounds are applied to normal approximation for the renormalized subgraph counts in the Erdős–Rényi random graph. This approach completely solves a long-standing conjecture in the general setting of arbitrary graph counting, while recovering recent results obtained for triangles and improving other bounds in the Wasserstein distance.

Keywords: Berry–Esseen bound; central limit theorem; Kolmogorov distance; Malliavin–Stein method; normal approximation; random graph; Stein–Chen method; subgraph count

References

[1] Barbour, A.D., Holst, L. and Janson, S. (1992). *Poisson Approximation*. *Oxford Studies in Probability* **2**. Oxford University Press: New York. MR1163825

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

[3] Chen, L.H.Y. and Shao, Q.-M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** 581–599. MR2331265 https://doi.org/10.3150/07-BEJ5164

[4] Döbler, C. and Krokowski, K. (2019). On the fourth moment condition for Rademacher chaos. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 61–97. MR3901641 https://doi.org/10.1214/17-aihp876

[5] Erdős, P. and Rényi, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. MR0120167

[6] Erdős, P. and Rényi, A. (1961). On the evolution of random graphs. *Bull. Inst. Internat. Statist.* **38** 343–347. MR0148055

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

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

[9] Krokowski, K. (2017). Poisson approximation of Rademacher functionals by the Chen–Stein method and Malliavin calculus. *Commun. Stoch. Anal.* **11** 195–222. MR3687774 https://doi.org/10.31390/cosa.11.2.05

[10] Krokowski, K., Reichenbachs, A. and Thäle, C. (2016). Berry–Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 763–803. MR3498009 https://doi.org/10.1214/14-AIHP652

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

[12] Lachièze-Rey, R. and Peccati, G. (2013). Fine Gaussian fluctuations on the Poisson space, I: Contractions, cumulants and geometric random graphs. *Electron. J. Probab.* **18** no. 32, 32. MR3035760 https://doi.org/10.1214/EJP.v18-2104

[13] Lachièze-Rey, R. and Peccati, G. (2013). Fine Gaussian fluctuations on the Poisson space II: Rescaled kernels, marked processes and geometric U-statistics. *Stochastic Process. Appl.* **123** 4186–4218. MR3096352 https://doi.org/10.1016/j.spa.2013.06.004

[14] Last, G., Peccati, G. and Schulte, M. (2016). Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields* **165** 667–723. MR3520016 https://doi.org/10.1007/s00440-015-0643-7

[15] Mantei, A. (2015). Stochastisches Kalkül in diskreter Zeit. Masterarbeit, Institut für Mathematik, Universität Potsdam.

[16] 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

[17] Nourdin, I., Peccati, G. and Reinert, G. (2010). Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.* **15** 1703–1742. MR2735379 https://doi.org/10.1214/EJP.v15-843

[18] Privault, N. (2008). Stochastic analysis of Bernoulli processes. *Probab. Surv.* **5** 435–483. MR2476738 https://doi.org/10.1214/08-PS139

[19] Privault, N. (2009). *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales. Lecture Notes in Math.* **1982**. Berlin: Springer. MR2531026 https://doi.org/10.1007/978-3-642-02380-4

[20] Privault, N. and Torrisi, G.L. (2015). The Stein and Chen–Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **12** 309–356. MR3357130

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

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

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

[24] Ruciński, A. and Vince, A. (1986). Strongly balanced graphs and random graphs. *J. Graph Theory* **10** 251–264. MR0890231 https://doi.org/10.1002/jgt.3190100214

Bernoulli **26**(1), 2020, 616–641
https://doi.org/10.3150/19-BEJ1142

On frequentist coverage errors of Bayesian credible sets in moderately high dimensions

KEISUKE YANO[1] and KENGO KATO[2]

[1]*Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: yano@mist.i.u-tokyo.ac.jp*
[2]*Department of Statistics and Data Science, Cornell University, 1194 Comstock Hall, Ithaca, NY 14853, USA. E-mail: kk976@cornell.edu*

In this paper, we study frequentist coverage errors of Bayesian credible sets for an approximately linear regression model with (moderately) high dimensional regressors, where the dimension of the regressors may increase with but is smaller than the sample size. Specifically, we consider quasi-Bayesian inference on the slope vector under the quasi-likelihood with Gaussian error distribution. Under this setup, we derive finite sample bounds on frequentist coverage errors of Bayesian credible rectangles. Derivation of those bounds builds on a novel Berry–Esseen type bound on quasi-posterior distributions and recent results on high-dimensional CLT on hyperrectangles. We use this general result to quantify coverage errors of Castillo–Nickl and L^{∞}-credible bands for Gaussian white noise models, linear inverse problems, and (possibly non-Gaussian) nonparametric regression models. In particular, we show that Bayesian credible bands for those nonparametric models have coverage errors decaying polynomially fast in the sample size, implying advantages of Bayesian credible bands over confidence bands based on extreme value theory.

Keywords: Castillo–Nickl band; credible rectangle; sieve prior

References

[1] Abramovich, F. and Silverman, B.W. (1998). Wavelet decomposition approaches to statistical inverse problems. *Biometrika* **85** 115–129. MR1627226 https://doi.org/10.1093/biomet/85.1.115
[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] Atchadé, Y.A. (2017). On the contraction properties of some high-dimensional quasi-posterior distributions. *Ann. Statist.* **45** 2248–2273. MR3718168 https://doi.org/10.1214/16-AOS1526
[4] Belloni, A., Chernozhukov, V., Chetverikov, D. and Kato, K. (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *J. Econometrics* **186** 345–366. MR3343791 https://doi.org/10.1016/j.jeconom.2015.02.014
[5] Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. MR0348906
[6] Bontemps, D. (2011). Bernstein–von Mises theorems for Gaussian regression with increasing number of regressors. *Ann. Statist.* **39** 2557–2584. MR2906878 https://doi.org/10.1214/11-AOS912
[7] Boucheron, S. and Gassiat, E. (2009). A Bernstein–von Mises theorem for discrete probability distributions. *Electron. J. Stat.* **3** 114–148. MR2471588 https://doi.org/10.1214/08-EJS262
[8] Castillo, I. (2014). On Bayesian supremum norm contraction rates. *Ann. Statist.* **42** 2058–2091. MR3262477 https://doi.org/10.1214/14-AOS1253

[9] Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. MR3127856 https://doi.org/10.1214/13-AOS1133

[10] 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

[11] Castillo, I. and Rousseau, J. (2015). A Bernstein–von Mises theorem for smooth functionals in semiparametric models. *Ann. Statist.* **43** 2353–2383. MR3405597 https://doi.org/10.1214/15-AOS1336

[12] Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *Ann. Statist.* **43** 1986–2018. MR3375874 https://doi.org/10.1214/15-AOS1334

[13] 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

[14] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Anti-concentration and honest, adaptive confidence bands. *Ann. Statist.* **42** 1787–1818. MR3262468 https://doi.org/10.1214/14-AOS1235

[15] 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

[16] Chernozhukov, V. and Hong, H. (2003). An MCMC approach to classical estimation. *J. Econometrics* **115** 293–346. MR1984779 https://doi.org/10.1016/S0304-4076(03)00100-3

[17] Claeskens, G. and Van Keilegom, I. (2003). Bootstrap confidence bands for regression curves and their derivatives. *Ann. Statist.* **31** 1852–1884. MR2036392 https://doi.org/10.1214/aos/1074290329

[18] Clarke, B. and Ghosal, S. (2010). Reference priors for exponential families with increasing dimension. *Electron. J. Stat.* **4** 737–780. MR2678969 https://doi.org/10.1214/10-EJS569

[19] Cox, D.D. (1993). An analysis of Bayesian inference for nonparametric regression. *Ann. Statist.* **21** 903–923. MR1232525 https://doi.org/10.1214/aos/1176349157

[20] DeVore, R.A. and Lorentz, G.G. (1993). *Constructive Approximation. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **303**. Berlin: Springer. MR1261635 https://doi.org/10.1007/978-3-662-02888-9

[21] Donoho, D.L. (1995). Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmon. Anal.* **2** 101–126. MR1325535 https://doi.org/10.1006/acha.1995.1008

[22] Florens, J.-P. and Simoni, A. (2012). Nonparametric estimation of an instrumental regression: A quasi-Bayesian approach based on regularized posterior. *J. Econometrics* **170** 458–475. MR2970326 https://doi.org/10.1016/j.jeconom.2012.05.016

[23] Freedman, D. (1999). On the Bernstein–von Mises theorem with infinite-dimensional parameters. *Ann. Statist.* **27** 1119–1140. MR1740119 https://doi.org/10.1214/aos/1017938917

[24] Gao, C. and Zhou, H.H. (2016). Bernstein–von Mises theorems for functionals of the covariance matrix. *Electron. J. Stat.* **10** 1751–1806. MR3522660 https://doi.org/10.1214/15-EJS1048

[25] Ghosal, S. (1999). Asymptotic normality of posterior distributions in high-dimensional linear models. *Bernoulli* **5** 315–331. MR1681701 https://doi.org/10.2307/3318438

[26] Ghosal, S. (2000). Asymptotic normality of posterior distributions for exponential families when the number of parameters tends to infinity. *J. Multivariate Anal.* **74** 49–68. MR1790613 https://doi.org/10.1006/jmva.1999.1874

[27] 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

[28] Giné, E. and Nickl, R. (2010). Confidence bands in density estimation. *Ann. Statist.* **38** 1122–1170. MR2604707 https://doi.org/10.1214/09-AOS738

[29] Giné, E. and Nickl, R. (2011). Rates of contraction for posterior distributions in L^r-metrics, $1 \leq r \leq \infty$. *Ann. Statist.* **39** 2883–2911. MR3012395 https://doi.org/10.1214/11-AOS924

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

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

[32] Hoffmann, M., Rousseau, J. and Schmidt-Hieber, J. (2015). On adaptive posterior concentration rates. *Ann. Statist.* **43** 2259–2295. MR3396985 https://doi.org/10.1214/15-AOS1341

[33] Johnstone, I. Gaussian Estimation: Sequence and Wavelet Models. Unpublished manuscript.

[34] Johnstone, I.M. (2010). High dimensional Bernstein–von Mises: Simple examples. In *Borrowing Strength: Theory Powering Application – a Festschrift for Lawrence D. Brown. Inst. Math. Stat. (IMS) Collect.* **6** 87–98. Beachwood, OH: IMS. MR2798513

[35] Kato, K. (2013). Quasi-Bayesian analysis of nonparametric instrumental variables models. *Ann. Statist.* **41** 2359–2390. MR3127869 https://doi.org/10.1214/13-AOS1150

[36] Knapik, B.T., van der Vaart, A.W. and van Zanten, J.H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.* **39** 2626–2657. MR2906881 https://doi.org/10.1214/11-AOS920

[37] Leahu, H. (2011). On the Bernstein–von Mises phenomenon in the Gaussian white noise model. *Electron. J. Stat.* **5** 373–404. MR2802048 https://doi.org/10.1214/11-EJS611

[38] Nazarov, F. (2003). On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a Gaussian measure. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1807** 169–187. Berlin: Springer. MR2083397 https://doi.org/10.1007/978-3-540-36428-3_15

[39] Nickl, R. (2019). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc. (JEMS)*. To appear.

[40] Nickl, R. and Söhl, J. (2017). Nonparametric Bayesian posterior contraction rates for discretely observed scalar diffusions. *Ann. Statist.* **45** 1664–1693. MR3670192 https://doi.org/10.1214/16-AOS1504

[41] 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

[42] Ray, K. (2017). Adaptive Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **45** 2511–2536. MR3737900 https://doi.org/10.1214/16-AOS1533

[43] Rivoirard, V. and Rousseau, J. (2012). Bernstein–von Mises theorem for linear functionals of the density. *Ann. Statist.* **40** 1489–1523. MR3015033 https://doi.org/10.1214/12-AOS1004

[44] Smirnov, N.V. (1950). On the construction of confidence regions for the density of distribution of random variables. *Dokl. Akad. Nauk SSSR* **74** 189–191. MR0037494

[45] Sniekers, S. and van der Vaart, A. (2015). Credible sets in the fixed design model with Brownian motion prior. *J. Statist. Plann. Inference* **166** 78–86. MR3390135 https://doi.org/10.1016/j.jspi.2014.07.008

[46] Spokoiny, V. Bernstein–von Mises Theorem for growing parameter dimension. Available at arXiv:1302.3430.

[47] Szabó, B., van der Vaart, A. and van Zanten, H. (2015). Honest Bayesian confidence sets for the L^2-norm. *J. Statist. Plann. Inference* **166** 36–51. MR3390132 https://doi.org/10.1016/j.jspi.2014.06.005

[48] Szabó, B., van der Vaart, A.W. and van Zanten, J.H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428. MR3357861 https://doi.org/10.1214/14-AOS1270

[49] 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

[50] Yang, Y., Bhattacharya, A. and Pati, D. Frequentist coverage and sup-norm convergence rate in Gaussian process regression. Available at arXiv:1708.04753.

[51] Yano, K. and Kato, K. (2020). Supplement to "On frequentist coverage errors of Bayesian credible sets in moderately high dimensions." https://doi.org/10.3150/19-BEJ1142SUPP.
[52] Yoo, W., Rousseau, J. and Rivoirard, V. Adaptive supremum norm posterior contraction: Wavelet spike-and-slab and anisotropic Besov spaces. Available at arXiv:1708.01909.
[53] Yoo, W.W. and Ghosal, S. (2016). Supremum norm posterior contraction and credible sets for nonparametric multivariate regression. *Ann. Statist.* **44** 1069–1102. MR3485954 https://doi.org/10.1214/15-AOS1398

Bernoulli **26**(1), 2020, 642–663
https://doi.org/10.3150/19-BEJ1145

The fourth characteristic of a semimartingale

ALEXANDER SCHNURR

Department of Mathematics, University Siegen, Walter-Flex-Street 3, 57068 Siegen, Germany.
E-mail: *schnurr@mathematik.uni-siegen.de*

We extend the class of semimartingales in a natural way. This allows us to incorporate processes having paths that leave the state space $\mathbb{R}^d$. In particular, Markov processes related to sub-Markovian kernels, but also non-Markovian processes with path-dependent behavior. By carefully distinguishing between two killing states, we are able to introduce a fourth semimartingale characteristic which generalizes the fourth part of the Lévy quadruple. Using the probabilistic symbol, we analyze the close relationship between the generators of certain Markov processes with killing and their (now four) semimartingale characteristics.

Keywords: killing; Markov process; semimartingale; symbol

References

[1] Behme, A. and Schnurr, A. (2015). A criterion for invariant measures of Itô processes based on the symbol. *Bernoulli* **21** 1697–1718. MR3352058 https://doi.org/10.3150/14-BEJ618
[2] Berg, C. and Forst, G. (1975). *Potential Theory on Locally Compact Abelian Groups*. New York: Springer. MR0481057
[3] Blumenthal, R.M. and Getoor, R.K. (1968). *Markov Processes and Potential Theory*. New York: Academic Press. MR0264757
[4] Böttcher, B., Schilling, R., Wang, J. and Lévy III, M. (2013). *Lévy-Type Processes*: *Construction*, *Approximation and Sample Path Properties*. *Lecture Notes in Math*. **2099**. Cham: Springer. MR3156646 https://doi.org/10.1007/978-3-319-02684-8
[5] Böttcher, B. and Schilling, R.L. (2009). Approximation of Feller processes by Markov chains with Lévy increments. *Stoch*. *Dyn*. **9** 71–80. MR2502474 https://doi.org/10.1142/S0219493709002555
[6] Chen, L. and Filipović, D. (2005). A simple model for credit migration and spread curves. *Finance Stoch*. **9** 211–231. MR2211125 https://doi.org/10.1007/s00780-004-0140-9
[7] Cherdito, P., Filipović, D. and Kimmel, R.L. (2007). Market price risk specification for affine models: Theory and evidence. *J*. *Financ*. *Econ*. **83** 123–170.
[8] Cheridito, P., Filipović, D. and Yor, M. (2005). Equivalent and absolutely continuous measure changes for jump-diffusion processes. *Ann*. *Appl*. *Probab*. **15** 1713–1732. MR2152242 https://doi.org/10.1214/105051605000000197
[9] Çinlar, E., Jacod, J., Protter, P. and Sharpe, M.J. (1980). Semimartingales and Markov processes. *Z*. *Wahrsch*. *Verw*. *Gebiete* **54** 161–219. MR0597337 https://doi.org/10.1007/BF00531446
[10] Courrège, P. (1965/66). Sur la forme intégro-différentielle des opérateurs de C_k^∞ dans C satisfaisant au principe du maximum. *Sém*. *Théorie du Potentiel*. *Exposé* **2** 38.
[11] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes*: *Characterization and Convergence*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658
[12] Fuglede, B. (1972). *Finely Harmonic Functions*. Berlin: Springer. MR0450590
[13] Hoh, W. (1995). Pseudodifferential operators with negative definite symbols and the martingale problem. *Stoch*. *Stoch*. *Rep*. **55** 225–252. MR1378858

1350-7265

[14] Hoh, W. (1998). Pseudo differential operators generating Markov processes. Habilitationsschrift, Universität Bielefeld.

[15] Ikeda, N. and Watanabe, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. *North-Holland Mathematical Library* **24**. Tokyo: Kodansha. MR0637061

[16] Jacob, N. (1996). *Pseudo-Differential Operators and Markov Processes*. *Mathematical Research* **94**. Berlin: Akademie Verlag. MR1409607

[17] Jacob, N. (1998). Characteristic functions and symbols in the theory of Feller processes. *Potential Anal*. **8** 61–68. MR1608650 https://doi.org/10.1023/A:1017983112289

[18] Jacob, N. (2001). *Pseudo Differential Operators and Markov Processes*. *Vol*. *I*: *Fourier Analysis and Semigroups*. London: Imperial College Press. MR1873235 https://doi.org/10.1142/9781860949746

[19] Jacob, N. (2002). *Pseudo Differential Operators & Markov Processes*. *Vol*. *II*: *Generators and Their Potential Theory*. London: Imperial College Press. MR1917230 https://doi.org/10.1142/9781860949562

[20] Jacob, N. (2005). *Pseudo Differential Operators and Markov Processes*. *Vol*. *III*: *Markov Processes and Applications*. London: Imperial College Press. MR2158336 https://doi.org/10.1142/9781860947155

[21] Jacod, J. and Shiryaev, A.N. (1987). *Limit Theorems for Stochastic Processes*. *Grundlehren der Mathematischen Wissenschaften* **288**. Berlin: Springer. MR0959133 https://doi.org/10.1007/978-3-662-02514-7

[22] Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *J*. *Appl*. *Probab*. **41** 601–622. MR2074811 https://doi.org/10.1017/s0021900200020428

[23] Klüppelberg, C., Maller, R. and Szimayer, A. (2011). The COGARCH: A review, with news on option pricing and statistical inference. In *Surveys in Stochastic Processes* (J. Blath, P. Imkeller and S. Roelly, eds.). *EMS Ser*. *Congr*. *Rep*. 29–58. Zürich: Eur. Math. Soc. MR2883852 https://doi.org/10.4171/072-1/2

[24] Knopova, V. and Schilling, R.L. (2014). On the small-time behaviour of Lévy-type processes. *Stochastic Process*. *Appl*. **124** 2249–2265. MR3188355 https://doi.org/10.1016/j.spa.2014.02.008

[25] Manstavičius, M. and Schnurr, A. (2017). Criteria for the finiteness of the strong p-variation for Lévy-type processes. *Stoch*. *Anal*. *Appl*. **35** 873–899. MR3686473 https://doi.org/10.1080/07362994.2017.1333007

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

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

[28] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[29] Schilling, R.L. (1998). Conservativeness and extensions of Feller semigroups. *Positivity* **2** 239–256. MR1653474 https://doi.org/10.1023/A:1009748105208

[30] Schilling, R.L. (1998). Feller processes generated by pseudo-differential operators: On the Hausdorff dimension of their sample paths. *J*. *Theoret*. *Probab*. **11** 303–330. MR1622573 https://doi.org/10.1023/A:1022678219821

[31] Schilling, R.L. (1998). Growth and Hölder conditions for the sample paths of Feller processes. *Probab*. *Theory Related Fields* **112** 565–611. MR1664705 https://doi.org/10.1007/s004400050201

[32] Schilling, R.L. and Schnurr, A. (2010). The symbol associated with the solution of a stochastic differential equation. *Electron*. *J*. *Probab*. **15** 1369–1393. MR2721050 https://doi.org/10.1214/EJP.v15-807

[33] Schilling, R.L. and Wang, J. (2013). Some theorems on Feller processes: Transience, local times and ultracontractivity. *Trans. Amer. Math. Soc.* **365** 3255–3286. MR3034465 https://doi.org/10.1090/S0002-9947-2012-05738-2

[34] Schnurr, A. (2009). The Symbol of a Markov Semimartingale. Ph.D. thesis, TU Dresden.

[35] Schnurr, A. (2012). On the semimartingale nature of Feller processes with killing. *Stochastic Process. Appl.* **122** 2758–2780. MR2926174 https://doi.org/10.1016/j.spa.2012.04.009

[36] Schnurr, A. (2013). Generalization of the Blumenthal–Getoor index to the class of homogeneous diffusions with jumps and some applications. *Bernoulli* **19** 2010–2032. MR3129042 https://doi.org/10.3150/12-BEJ441

[37] Schnurr, A. (2013). COGARCH: Symbol, generator and characteristics. *Rend. Semin. Mat. Univ. Politec. Torino* **71** 251–260. MR3345055

Bernoulli **26**(1), 2020, 664–693
https://doi.org/10.3150/19-BEJ1148

A unified approach to coupling SDEs driven by Lévy noise and some applications

MINGJIE LIANG[1,2], RENÉ L. SCHILLING[3] and JIAN WANG[4]

[1]*College of Information Engineering, Sanming University, 365004 Sanming, P.R. China. E-mail: liangmingjie@aliyun.com*
[2]*College of Mathematics and Informatics, Fujian Normal University, 350007 Fuzhou, P.R. China*
[3]*TU Dresden, Fakultät Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany. E-mail: rene.schilling@tu-dresden.de*
[4]*College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA), Fujian Normal University, 350007 Fuzhou, P.R. China. E-mail: jianwang@fjnu.edu.cn*

We present a general method to construct couplings of stochastic differential equations driven by Lévy noise in terms of coupling operators. This approach covers both coupling by reflection and refined basic coupling which are often discussed in the literature. As applications, we prove regularity results for the transition semigroups and obtain successful couplings for the solutions to stochastic differential equations driven by additive Lévy noise.

Keywords: coupling by reflection; coupling operator; Lévy process; optimal coupling; refined basic coupling; successful coupling

References

[1] Aldous, D. (1983). Random walks on finite groups and rapidly mixing Markov chains. In *Seminar on Probability*, *XVII*. *Lecture Notes in Math*. **986** 243–297. Berlin: Springer. MR0770418 https://doi.org/10.1007/BFb0068322
[2] Alishahi, K. and Salavati, S. (2016). A sufficient condition for absolute continuity of infinitely divisible distributions. Available at arXiv:1606.07106.
[3] Alishahi, K. and Salavati, S. (2018). Local coupling property for Markov processes with applications to Lévy processes. Available at arXiv:1802.09608.
[4] Banerjee, S. and Kendall, W.S. (2017). Rigidity for Markovian maximal couplings of elliptic diffusions. *Probab. Theory Related Fields* **168** 55–112. MR3651049 https://doi.org/10.1007/s00440-016-0706-4
[5] Böttcher, B. (2017). Markovian maximal coupling of Markov processes. Available at arXiv:1710.09654.
[6] Böttcher, B., Schilling, R. and Wang, J. (2013). *Lévy Matters*. *III*: *Lévy-Type Processes*: *Construction*, *Approximation and Sample Path Properties*. *Lecture Notes in Math*. **2099**. Cham: Springer. MR3156646 https://doi.org/10.1007/978-3-319-02684-8
[7] Böttcher, B., Schilling, R.L. and Wang, J. (2011). Constructions of coupling processes for Lévy processes. *Stochastic Process*. *Appl*. **121** 1201–1216. MR2794973 https://doi.org/10.1016/j.spa.2011.02.007
[8] Burdzy, K. and Kendall, W.S. (2000). Efficient Markovian couplings: Examples and counterexamples. *Ann*. *Appl*. *Probab*. **10** 362–409. MR1768241 https://doi.org/10.1214/aoap/1019487348

1350-7265

[9] Chen, M. (1994). Optimal Markovian couplings and applications. *Acta Math. Sin.* (*Engl. Ser.*) **10** 260–275. MR1415697 https://doi.org/10.1007/BF02560717

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

[11] Chen, M.F. and Li, S.F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.* **17** 151–177. MR0972776

[12] Chen, Z.-Q., Song, R. and Zhang, X. (2018). Stochastic flows for Lévy processes with Hölder drifts. *Rev. Mat. Iberoam.* **34** 1755–1788. MR3896248 https://doi.org/10.4171/rmi/1042

[13] Cranston, M. (1991). Gradient estimates on manifolds using coupling. *J. Funct. Anal.* **99** 110–124. MR1120916 https://doi.org/10.1016/0022-1236(91)90054-9

[14] Dereich, S. (2011). Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *Ann. Appl. Probab.* **21** 283–311. MR2759203 https://doi.org/10.1214/10-AAP695

[15] Dereich, S. and Heidenreich, F. (2011). A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. *Stochastic Process. Appl.* **121** 1565–1587. MR2802466 https://doi.org/10.1016/j.spa.2011.03.015

[16] Ernst, P.A., Kendall, W.S., Roberts, G.O. and Rosenthal, J.S. (2019). MEXIT: Maximal un-coupling times for stochastic processes. *Stochastic Process. Appl.* **129** 355–380. MR3907003 https://doi.org/10.1016/j.spa.2018.03.001

[17] Gangbo, W. and McCann, R.J. (1996). The geometry of optimal transportation. *Acta Math.* **177** 113–161. MR1440931 https://doi.org/10.1007/BF02392620

[18] Griffeath, D. (1974/75). A maximal coupling for Markov chains. *Z. Wahrsch. Verw. Gebiete* **31** 95–106. MR0370771 https://doi.org/10.1007/BF00539434

[19] Hsu, E.P. and Sturm, K.-T. (2013). Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.* **1** 93–104. MR3197874 https://doi.org/10.1007/s40304-013-0007-5

[20] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. *North-Holland Mathematical Library* **24**. Amsterdam: North-Holland. MR1011252

[21] Kühn, F. and Schilling, R.L. (2019). Strong convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs. *Stochastic Process. Appl.* **129** 2654–2680. MR3980140 https://doi.org/10.1016/j.spa.2018.07.018

[22] Liang, M. and Wang, J. (2018). Gradient estimates and ergodicity for SDEs driven multiplcative Lévy noises via coupling. Available at arXiv:1801.05936.

[23] Liggett, T.M. (1985). *Interacting Particle Systems*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **276**. New York: Springer. MR0776231 https://doi.org/10.1007/978-1-4613-8542-4

[24] Lindvall, T. (1992). *Lectures on the Coupling Method*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR1180522

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

[26] Luo, D. and Wang, J. (2019). Coupling by reflection and Hölder regularity for non-local operators of variable order. *Trans. Amer. Math. Soc.* **371** 431–459. MR3885150 https://doi.org/10.1090/tran/7259

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

[28] Majka, M.B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stochastic Process. Appl.* **127** 4083–4125. MR3718107 https://doi.org/10.1016/j.spa.2017.03.020

[29] McCann, R.J. (1999). Exact solutions to the transportation problem on the line. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **455** 1341–1380. MR1701760 https://doi.org/10.1098/rspa.1999.0364

[30] Priola, E. (2012). Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math*. **49** 421–447. MR2945756
[31] Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal*. **236** 244–264. MR2227134 https://doi.org/10.1016/j.jfa.2005.12.010
[32] Rachev, S.T. and Rüschendorf, L. (1998). *Mass Transportation Problems. Vol. II: Applications. Probability and Its Applications (New York)*. New York: Springer. MR1619171
[33] Revuz, D. (1984). *Markov Chains*, 2nd ed. *North-Holland Mathematical Library* **11**. Amsterdam: North-Holland. MR0758799
[34] Roberts, G.O. and Rosenthal, J.S. (2004). General state space Markov chains and MCMC algorithms. *Probab. Surv.* **1** 20–71. MR2095565 https://doi.org/10.1214/154957804100000024
[35] Schilling, R.L., Sztonyk, P. and Wang, J. (2012). Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18** 1128–1149. MR2995789 https://doi.org/10.3150/11-BEJ375
[36] Schilling, R.L. and Wang, J. (2011). On the coupling property of Lévy processes. *Ann. Inst. Henri Poincaré Probab. Stat*. **47** 1147–1159. MR2884228 https://doi.org/10.1214/10-AIHP400
[37] Schilling, R.L. and Wang, J. (2012). On the coupling property and the Liouville theorem for Ornstein–Uhlenbeck processes. *J. Evol. Equ*. **12** 119–140. MR2891204 https://doi.org/10.1007/s00028-011-0126-y
[38] Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration. Probability and Its Applications (New York)*. New York: Springer. MR1741181 https://doi.org/10.1007/978-1-4612-1236-2
[39] Villani, C. (2003). *Topics in Optimal Transportation. Graduate Studies in Mathematics* **58**. Providence, RI: Am. Math. Soc. MR1964483 https://doi.org/10.1007/b12016
[40] Wang, J. (2010). Regularity of semigroups generated by Lévy type operators via coupling. *Stochastic Process. Appl*. **120** 1680–1700. MR2673970 https://doi.org/10.1016/j.spa.2010.04.007
[41] Wang, J. (2016). L^p-Wasserstein distance for stochastic differential equations driven by Lévy processes. *Bernoulli* **22** 1598–1616. MR3474827 https://doi.org/10.3150/15-BEJ705
[42] Zhang, X. (2013). Stochastic differential equations with Sobolev drifts and driven by α-stable processes. *Ann. Inst. Henri Poincaré Probab. Stat*. **49** 1057–1079. MR3127913 https://doi.org/10.1214/12-AIHP476

Bernoulli **26**(1), 2020, 694–727
https://doi.org/10.3150/19-BEJ1149

Robust modifications of U-statistics and applications to covariance estimation problems

STANISLAV MINSKER[1] and XIAOHAN WEI[2]

[1]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA. E-mail: minsker@usc.edu*
[2]*Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA. E-mail: xiaohanw@usc.edu*

Let Y be a d-dimensional random vector with unknown mean μ and covariance matrix Σ. This paper is motivated by the problem of designing an estimator of Σ that admits exponential deviation bounds in the operator norm under minimal assumptions on the underlying distribution, such as existence of only 4th moments of the coordinates of Y. To address this problem, we propose robust modifications of the operator-valued U-statistics, obtain non-asymptotic guarantees for their performance, and demonstrate the implications of these results to the covariance estimation problem under various structural assumptions.

Keywords: covariance estimation; heavy tails; robust estimators; U-statistics

References

[1] Aleksandrov, A.B. and Peller, V.V. (2016). Operator Lipschitz functions. *Russian Math. Surveys* **71** 605. MR3588921 https://doi.org/10.4213/rm9729

[2] Alon, N., Matias, Y. and Szegedy, M. (1996). The space complexity of approximating the frequency moments. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing* (*Philadelphia*, *PA*, 1996) 20–29. New York: ACM. MR1427494 https://doi.org/10.1145/237814.237823

[3] Angluin, D. and Valiant, L.G. (1979). Fast probabilistic algorithms for Hamiltonian circuits and matchings. *J. Comput. System Sci.* **18** 155–193. MR0532174 https://doi.org/10.1016/0022-0000(79)90045-X

[4] Arcones, M.A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426

[5] Bertsekas, D.P. (2009). *Convex Optimization Theory*. Nashua, NH: Athena Scientific. MR2830150

[6] Bhatia, R. (2013). *Matrix Analysis* **169**. Springer Science & Business Media. MR1477662 https://doi.org/10.1007/978-1-4612-0653-8

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

[8] Carlen, E. (2010). Trace inequalities and quantum entropy: An introductory course. In *Entropy and the Quantum*. *Contemp. Math.* **529** 73–140. Providence, RI: Amer. Math. Soc. MR2681769 https://doi.org/10.1090/conm/529/10428

[9] 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

[10] Catoni, O. (2016). PAC-Bayesian bounds for the Gram matrix and least squares regression with a random design. arXiv preprint arXiv:1603.05229.

[11] Catoni, O. and Giulini, I. (2017). Dimension-free PAC-Bayesian bounds for matrices, vectors, and linear least squares regression. arXiv preprint arXiv:1712.02747.

[12] 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

[13] Devroye, L., Lerasle, M., Lugosi, G. and Oliveira, R.I. (2016). Sub-Gaussian mean estimators. *Ann. Statist.* **44** 2695–2725. MR3576558 https://doi.org/10.1214/16-AOS1440

[14] de la Peña, V.H. and Giné, E. (1999). *Decoupling*: *From Dependence to Independence*, *Randomly Stopped Processes*. *U-Statistics and Processes*. *Martingales and Beyond*. *Probability and Its Applications* (*New York*). New York: Springer. MR1666908 https://doi.org/10.1007/978-1-4612-0537-1

[15] Fan, J., Ke, Y., Sun, Q. and Zhou, W.-X. (2017). FARM-Test: factor-adjusted robust multiple testing with false discovery control. arXiv preprint arXiv:1711.05386.

[16] Fan, J., Wang, W. and Zhu, Z. (2016). Robust low-rank matrix recovery. arXiv preprint arXiv:1603.08315.

[17] Giné, E., Latała, R. and Zinn, J. (2000). Exponential and moment inequalities for U-statistics. In *High Dimensional Probability*, *II* (*Seattle*, *WA*, 1999). *Progress in Probability* **47** 13–38. Boston, MA: Birkhäuser. MR1857312

[18] Giulini, I. (2016). Robust principal component analysis in Hilbert spaces. arXiv preprint arXiv:1606.00187.

[19] Han, F. and Liu, H. (2017). Statistical analysis of latent generalized correlation matrix estimation in transelliptical distribution. *Bernoulli* **23** 23–57. MR3556765 https://doi.org/10.3150/15-BEJ702

[20] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325. MR0026294 https://doi.org/10.1214/aoms/1177730196

[21] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR0144363

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

[23] Huber, P.J. (2011). Robust statistics. In *International Encyclopedia of Statistical Science* 1248–1251. Springer.

[24] Hubert, M., Rousseeuw, P.J. and Van Aelst, S. (2008). High-breakdown robust multivariate methods. *Statist. Sci.* **23** 92–119. MR2431867 https://doi.org/10.1214/088342307000000087

[25] Joly, E. and Lugosi, G. (2016). Robust estimation of U-statistics. *Stochastic Process. Appl.* **126** 3760–3773. MR3565476 https://doi.org/10.1016/j.spa.2016.04.021

[26] Kabanava, M. and Rauhut, H. (2017). Masked Toeplitz covariance estimation. arXiv preprint arXiv:1709.09377.

[27] 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

[28] Lepskiĭ, O.V. (1991). Asymptotically minimax adaptive estimation. I: Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36** 682–697. MR1147167 https://doi.org/10.1137/1136085

[29] Lerasle, M. and Oliveira, R.I. (2011). Robust empirical mean estimators. arXiv preprint arXiv:1112.3914.

[30] Levina, E. and Vershynin, R. (2012). Partial estimation of covariance matrices. *Probab. Theory Related Fields* **153** 405–419. MR2948681 https://doi.org/10.1007/s00440-011-0349-4

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

[32] Maronna, R.A. (1976). Robust M-estimators of multivariate location and scatter. *Ann. Statist.* **4** 51–67. MR0388656
[33] 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
[34] Minsker, S. (2018). Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries. *Ann. Statist.* **46** 2871–2903. MR3851758 https://doi.org/10.1214/17-AOS1642
[35] Minsker, S. and Wei, X. Estimation of the covariance structure of heavy-tailed distributions. In *NIPS* 2017.
[36] Nemirovsky, A.S. and Yudin, D.B. (1983). *Problem Complexity and Method Efficiency in Optimization*. *Wiley-Interscience Series in Discrete Mathematics*. New York: Wiley. Translated from the Russian and with a preface by E.R. Dawson. MR0702836
[37] Polyak, B.T. and Khlebnikov, M.V. (2017). Principle component analysis: Robust versions. *Autom. Remote Control* **78** 490–506. MR3666090
[38] Tropp, J.A. (2012). User-friendly tail bounds for sums of random matrices. *Found. Comput. Math.* **12** 389–434. MR2946459 https://doi.org/10.1007/s10208-011-9099-z
[39] Tyler, D.E. (1987). A distribution-free M-estimator of multivariate scatter. *Ann. Statist.* **15** 234–251. MR0885734 https://doi.org/10.1214/aos/1176350263
[40] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170
[41] Wegkamp, M. and Zhao, Y. (2016). Adaptive estimation of the copula correlation matrix for semiparametric elliptical copulas. *Bernoulli* **22** 1184–1226. MR3449812 https://doi.org/10.3150/14-BEJ690
[42] Zhang, T., Cheng, X. and Singer, A. (2016). Marčenko–Pastur law for Tyler's M-estimator. *J. Multivariate Anal.* **149** 114–123. MR3507318 https://doi.org/10.1016/j.jmva.2016.03.010

Bernoulli **26**(1), 2020, 728–766
https://doi.org/10.3150/19-BEJ1150

A Feynman–Kac result via Markov BSDEs with generalised drivers

ELENA ISSOGLIO[1] and FRANCESCO RUSSO[2]

[1]*Department of Mathematics, University of Leeds, Leeds, LS2 9JT, UK. E-mail: e.issoglio@leeds.ac.uk*
[2]*ENSTA Paris, Institut Polytechnique de Paris, Unité de Mathématiques appliquées, 828, bd. des Maréchaux, F-91120 Palaiseau, France. E-mail: francesco.russo@ensta-paris.fr*

In this paper, we investigate BSDEs where the driver contains a distributional term (in the sense of generalised functions) and derive general Feynman–Kac formulae related to these BSDEs. We introduce an integral operator to give sense to the equation and then we show the existence of a strong solution employing results on a related PDE. Due to the irregularity of the driver, the Y-component of a couple (Y, Z) solving the BSDE is not necessarily a semimartingale but a weak Dirichlet process.

Keywords: backward stochastic differential equations (BSDEs); distributional driver; Feynman–Kac formula; generalised and rough coefficients; pointwise product; weak Dirichlet process

References

[1] Barrasso, A. and Russo, F. (2017). Backward stochastic differential equations with no driving martingale, Markov processes and associated pseudo partial differential equations. Part II: Decoupled mild solutions and examples. Preprint. Available at hal-01505974.
[2] Bismut, J.-M. (1973). Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.* **44** 384–404. MR0329726 https://doi.org/10.1016/0022-247X(73)90066-8
[3] Bouleau, N. and Yor, M. (1981). Sur la variation quadratique des temps locaux de certaines semimartingales. *C. R. Acad. Sci., Sér.* 1 *Math.* **292** 491–494. MR0612544
[4] Cheridito, P. (2003). Arbitrage in fractional Brownian motion models. *Finance Stoch.* **7** 533–553. MR2014249 https://doi.org/10.1007/s007800300101
[5] Cheridito, P. and Nam, K. (2017). BSE's, BSDE's and fixed-point problems. *Ann. Probab.* **45** 3795–3828. MR3729615 https://doi.org/10.1214/16-AOP1149
[6] Coviello, R., di Girolami, C. and Russo, F. (2011). On stochastic calculus related to financial assets without semimartingales. *Bull. Sci. Math.* **135** 733–774. MR2838099 https://doi.org/10.1016/j.bulsci.2011.06.008
[7] Davies, E.B. (1989). *Heat Kernels and Spectral Theory*. *Cambridge Tracts in Mathematics* **92**. Cambridge: Cambridge Univ. Press. MR0990239 https://doi.org/10.1017/CBO9780511566158
[8] Delarue, F. and Diel, R. (2016). Rough paths and 1d SDE with a time dependent distributional drift: Application to polymers. *Probab. Theory Related Fields* **165** 1–63. MR3500267 https://doi.org/10.1007/s00440-015-0626-8
[9] Diehl, J. and Zhang, J. (2017). Backward stochastic differential equations with Young drift. *Probab. Uncertain. Quant. Risk* **2** Paper No. 5. MR3663305 https://doi.org/10.1186/s41546-017-0016-5
[10] El Karoui, N., Peng, S. and Quenez, M.C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71. MR1434407 https://doi.org/10.1111/1467-9965.00022

[11] Engelbert, H.-J. and Wolf, J. (1998). Strong Markov local Dirichlet processes and stochastic differential equations. *Teor. Veroyatn. Primen.* **43** 331–348. MR1679006 https://doi.org/10.1137/S0040585X97976829
[12] Erraoui, M., Ouknine, Y. and Sbi, A. (1998). Reflected solutions of backward stochastic differential equations with distribution as terminal condition. *Random Oper. Stoch. Equ.* **6** 1–16. MR1610189 https://doi.org/10.1515/rose.1998.6.1.1
[13] Flandoli, F., Issoglio, E. and Russo, F. (2017). Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.* **369** 1665–1688. MR3581216 https://doi.org/10.1090/tran/6729
[14] Flandoli, F., Russo, F. and Wolf, J. (2003). Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.* **40** 493–542. MR1988703
[15] Flandoli, F., Russo, F. and Wolf, J. (2004). Some SDEs with distributional drift. II. Lyons–Zheng structure, Itô's formula and semimartingale characterization. *Random Oper. Stoch. Equ.* **12** 145–184. MR2065168 https://doi.org/10.1163/156939704323074700
[16] Gozzi, F. and Russo, F. (2006). Weak Dirichlet processes with a stochastic control perspective. *Stochastic Process. Appl.* **116** 1563–1583. MR2269216 https://doi.org/10.1016/j.spa.2006.04.009
[17] Hinz, M., Issoglio, E. and Zähle, M. (2014). Elementary pathwise methods for nonlinear parabolic and transport type stochastic partial differential equations with fractal noise. In *Modern Stochastics and Applications. Springer Optim. Appl.* **90** 123–141. Cham: Springer. MR3236072 https://doi.org/10.1007/978-3-319-03512-3_8
[18] Issoglio, E. (2013). Transport equations with fractal noise – existence, uniqueness and regularity of the solution. *Z. Anal. Anwend.* **32** 37–53. MR3061360 https://doi.org/10.4171/ZAA/1473
[19] Issoglio, E. (2019). A non-linear parabolic PDE with a distributional coefficient and its applications to stochastic analysis. *J. Differential Equations* **267** 5976–6003. MR3996791 https://doi.org/10.1016/j.jde.2019.06.014
[20] Issoglio, E. and Jing, S. (2018). Forward–backward SDEs with distributional coefficients. *Stochastic Process. Appl.*, in press. https://doi.org/10.1016/j.spa.2019.01.001
[21] Issoglio, E. and Zähle, M. (2015). Regularity of the solutions to SPDEs in metric measure spaces. *Stoch. Partial Differ. Equ. Anal. Comput.* **3** 272–289. MR3350453 https://doi.org/10.1007/s40072-015-0048-8
[22] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2
[23] Lunardi, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems. Progress in Nonlinear Differential Equations and Their Applications* **16**. Basel: Birkhäuser. MR1329547 https://doi.org/10.1007/978-3-0348-9234-6
[24] Mathieu, P. (1994). Zero white noise limit through Dirichlet forms, with application to diffusions in a random medium. *Probab. Theory Related Fields* **99** 549–580. MR1288070 https://doi.org/10.1007/BF01206232
[25] Pardoux, É. and Peng, S.G. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Lett.* **14** 55–61. MR1037747 https://doi.org/10.1016/0167-6911(90)90082-6
[26] Pardoux, E. and Răşcanu, A. (2014). *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations. Stochastic Modelling and Applied Probability* **69**. Cham: Springer. MR3308895 https://doi.org/10.1007/978-3-319-05714-9
[27] Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations. Applied Mathematical Sciences* **44**. New York: Springer. MR0710486 https://doi.org/10.1007/978-1-4612-5561-1

[28] Pham, H. (2009). *Continuous-Time Stochastic Control and Optimization with Financial Applications*. *Stochastic Modelling and Applied Probability* **61**. Berlin: Springer. MR2533355 https://doi.org/10.1007/978-3-540-89500-8

[29] Runst, T. and Sickel, W. (1996). *Sobolev Spaces of Fractional Order*, *Nemytskij Operators*, *and Nonlinear Partial Differential Equations*. *De Gruyter Series in Nonlinear Analysis and Applications* **3**. Berlin: de Gruyter. MR1419319 https://doi.org/10.1515/9783110812411

[30] Russo, F. and Trutnau, G. (2007). Some parabolic PDEs whose drift is an irregular random noise in space. *Ann*. *Probab*. **35** 2213–2262. MR2353387 https://doi.org/10.1214/009117906000001178

[31] Russo, F. and Vallois, P. (1995). The generalized covariation process and Itô formula. *Stochastic Process*. *Appl*. **59** 81–104. MR1350257 https://doi.org/10.1016/0304-4149(95)93237-A

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

[33] Russo, F. and Wurzer, L. (2017). Elliptic PDEs with distributional drift and backward SDEs driven by a càdlàg martingale with random terminal time. *Stoch*. *Dyn*. **17** 1750030. MR3646613 https://doi.org/10.1142/S0219493717500307

[34] Seignourel, P. (2000). Discrete schemes for processes in random media. *Probab. Theory Related Fields* **118** 293–322. MR1800534 https://doi.org/10.1007/PL00008743

[35] Triebel, H. (1978). *Interpolation Theory*, *Function Spaces*, *Differential Operators*. *North-Holland Mathematical Library* **18**. Amsterdam: North-Holland. MR0503903

[36] Triebel, H. (2008). *Function Spaces and Wavelets on Domains*. *EMS Tracts in Mathematics* **7**. Zürich: Eur. Math. Soc. MR2455724 https://doi.org/10.4171/019

[37] Triebel, H. (2010). *Bases in Function Spaces*, *Sampling*, *Discrepancy*, *Numerical Integration*. *EMS Tracts in Mathematics* **11**. Zürich: Eur. Math. Soc. MR2667814 https://doi.org/10.4171/085

[38] Üstünel, A.S. (1982). A generalization of Itô's formula. *J. Funct*. *Anal*. **47** 143–152. MR0664333 https://doi.org/10.1016/0022-1236(82)90102-1

Bernoulli **26**(1), 2020, 767–798
https://doi.org/10.3150/19-BEJ1151

Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces

JING LEI

Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: jinglei@andrew.cmu.edu

We provide upper bounds of the expected Wasserstein distance between a probability measure and its empirical version, generalizing recent results for finite dimensional Euclidean spaces and bounded functional spaces. Such a generalization can cover Euclidean spaces with large dimensionality, with the optimal dependence on the dimensionality. Our method also covers the important case of Gaussian processes in separable Hilbert spaces, with rate-optimal upper bounds for functional data distributions whose coordinates decay geometrically or polynomially. Moreover, our bounds of the expected value can be combined with mean-concentration results to yield improved exponential tail probability bounds for the Wasserstein error of empirical measures under Bernstein-type or log Sobolev-type conditions.

Keywords: concentration inequality; empirical measure; empirical process; functional data; Wasserstein distance

References

[1] Airoldi, E.M., Costa, T.B. and Chan, S.H. (2013). Stochastic blockmodel approximation of a graphon: Theory and consistent estimation. In *Advances in Neural Information Processing Systems* 692–700.

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

[3] Arjovsky, M., Chintala, S. and Bottou, L. (2017). Wasserstein GAN. Available at arXiv:1701.07875.

[4] Bickel, P.J. and Chen, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci. USA* **106** 21068–21073.

[5] Bobkov, S. and Ledoux, M. (2014). One-dimensional empirical measures, order statistics and Kantorovich transport distances. Preprint.

[6] Boissard, E. (2011). Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance. *Electron. J. Probab.* **16** 2296–2333. MR2861675 https://doi.org/10.1214/EJP.v16-958

[7] Boissard, E. and Le Gouic, T. (2014). On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 539–563. MR3189084 https://doi.org/10.1214/12-AIHP517

[8] Bolley, F., Guillin, A. and Villani, C. (2007). Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* **137** 541–593. MR2280433 https://doi.org/10.1007/s00440-006-0004-7

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

[10] de la Peña, V.H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** 537–564. MR1681153 https://doi.org/10.1214/aop/1022677271

[11] Dedecker, J. and Fan, X. (2015). Deviation inequalities for separately Lipschitz functionals of iterated random functions. *Stochastic Process. Appl.* **125** 60–90. MR3274692 https://doi.org/10.1016/j.spa.2014.08.001

[12] Dereich, S., Scheutzow, M. and Schottstedt, R. (2013). Constructive quantization: Approximation by empirical measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1183–1203. MR3127919 https://doi.org/10.1214/12-AIHP489

[13] Dudley, R.M. (1968). The speed of mean Glivenko–Cantelli convergence. *Ann. Math. Stat.* **40** 40–50. MR0236977 https://doi.org/10.1214/aoms/1177697802

[14] Dumer, I. (2006). Covering an ellipsoid with equal balls. *J. Combin. Theory Ser. A* **113** 1667–1676. MR2269546 https://doi.org/10.1016/j.jcta.2006.03.021

[15] Dumer, I., Pinsker, M.S. and Prelov, V.V. (2004). On coverings of ellipsoids in Euclidean spaces. *IEEE Trans. Inform. Theory* **50** 2348–2356. MR2097051 https://doi.org/10.1109/TIT.2004.834759

[16] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7

[17] Gao, C., Lu, Y. and Zhou, H.H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43** 2624–2652. MR3405606 https://doi.org/10.1214/15-AOS1354

[18] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* 2672–2680.

[19] Gozlan, N. and Léonard, C. (2010). Transport inequalities. A survey. *Markov Process. Related Fields* **16** 635–736. MR2895086

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

[21] Hoover, D.N. (1982). Row-column exchangeability and a generalized model for probability. In *Exchangeability in Probability and Statistics* (*Rome*, 1981) 281–291. Amsterdam: North-Holland. MR0675982

[22] Kallenberg, O. (1989). On the representation theorem for exchangeable arrays. *J. Multivariate Anal.* **30** 137–154. MR1003713 https://doi.org/10.1016/0047-259X(89)90092-4

[23] Kantorovič, L.V. and Rubinšteĭn, G.Š. (1958). On a space of completely additive functions. *Vestnik Leningrad Univ. Math.* **13** 52–59. MR0102006

[24] Kantorovitch, L. (1942). On the translocation of masses. *C. R.* (*Dokl.*) *Acad. Sci. URSS* **37** 199–201. MR0009619

[25] Kontorovich, A. (2014). Concentration in unbounded metric spaces and algorithmic stability. In *International Conference on Machine Learning* 28–36.

[26] Ledoux, M. (1999). Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXIII. Lecture Notes in Math.* **1709** 120–216. Berlin: Springer. MR1767995 https://doi.org/10.1007/BFb0096511

[27] Lei, J. (2018). Network representation using graph root distributions. Available at arXiv:1802.09684.

[28] Lovász, L. (2012). *Large Networks and Graph Limits. American Mathematical Society Colloquium Publications* **60**. Providence, RI: Amer. Math. Soc. MR3012035 https://doi.org/10.1090/coll/060

[29] Luschgy, H. and Pagès, G. (2004). Sharp asymptotics of the Kolmogorov entropy for Gaussian measures. *J. Funct. Anal.* **212** 89–120. MR2065239 https://doi.org/10.1016/j.jfa.2003.09.004

[30] Monge, G. (1781). Mémoire sur la théorie des déblais des remblais. History of the Royal Academy of Sciences of Paris.

[31] Ramsay, J.O. (2006). *Functional Data Analysis*. New York: Wiley Online Library.
[32] Singh, S. and Póczos, B. (2018). Minimax distribution estimation in Wasserstein distance. Available at arXiv:1802.08855.
[33] Sommerfeld, M. and Munk, A. (2018). Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 219–238. MR3744719 https://doi.org/10.1111/rssb.12236
[34] Talagrand, M. (1994). The transportation cost from the uniform measure to the empirical measure in dimension ≥ 3. *Ann. Probab.* **22** 919–959. MR1288137
[35] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794
[36] van de Geer, S. (1995). Exponential inequalities for martingales, with application to maximum likelihood estimation for counting processes. *Ann. Statist.* **23** 1779–1801. MR1370307 https://doi.org/10.1214/aos/1176324323
[37] 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
[38] Vasershtein, L.N. (1969). Markov processes over denumerable products of spaces describing large system of automata. *Problemy Peredachi Informatsii* **5** 64–72. MR0314115 https://doi.org/10.1016/s0016-0032(33)90010-1
[39] Verger-Gaugry, J.-L. (2005). Covering a ball with smaller equal balls in $\mathbb{R}^n$. *Discrete Comput. Geom.* **33** 143–155. MR2105755 https://doi.org/10.1007/s00454-004-2916-2
[40] 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
[41] Weed, J. and Bach, F. (2019). Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** 2620–2648. MR4003560 https://doi.org/10.3150/18-BEJ1065
[42] Williamson, R.C., Smola, A.J. and Schölkopf, B. (2001). Generalization performance of regularization networks and support vector machines via entropy numbers of compact operators. *IEEE Trans. Inform. Theory* **47** 2516–2532. MR1873936 https://doi.org/10.1109/18.945262
[43] Wolfe, P.J. and Olhede, S.C. (2013). Nonparametric graphon estimation. Available at arXiv:1309.5936.
[44] Ying, Y. McDiarmid's inequalities of Bernstein and Bennett forms. Available at https://www.albany.edu/~yy298919/McDiarmid.pdf.

Bernoulli Forthcoming Papers

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

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

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

CÉNAC, P., DE LOYNES, B., OFFRET, Y. and ROUSSELLE, A.
Recurrence of multidimensional persistent random walks. Fourier and series criteria

FAN, J.Y., HAMZA, K., JAGERS, P. and KLEBANER, F.
Convergence of the age structure of general schemes of population processes

NOURDIN, I., PECCATI, G. and SEURET, S.
Sojourn time dimensions of fractional Brownian motion

HIRSCH, C. and MÖNCH, C.
Distances and large deviations in the spatial preferential attachment model

BONNET, G. and CHENAVIER, N.
The maximal degree in a Poisson–Delaunay graph

DÖRING, L. and WEISSMANN, P.
Stable processes conditioned to hit an interval continuously from the outside

MUKHERJEE, S.
Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics

CONTI, P.L., MARELLA, D., MECATTI, F. and ANDREIS, F.
A unified principled framework for resampling based on pseudo-populations: Asymptotic theory

MARIUCCI, E., RAY, K. and SZABÓ, B.
A Bayesian nonparametric approach to log-concave density estimation

ALETTI, G., CRIMALDI, I. and GHIGLIETTI, A.
Interacting reinforced stochastic processes: Statistical inference based on the weighted empirical means

LIU, Y. and PAGÈS, G.
Characterization of probability distribution convergence in Wasserstein distance by L^p-quantization error function

CORLIN MARCHAND, D. and MANOLESCU, I.
Influence of the seed in *affine* preferential attachment trees

SCHWEINBERGER, M.
Consistent structure estimation of exponential-family random graph models with block structure

Continues

Bernoulli Forthcoming Papers—*Continued*

JIANG, B., CHEN, Z. and LENG, C.
Dynamic linear discriminant analysis in high dimensional space

GANTERT, N., HEYDENREICH, M. and HIRSCHER, T.
Strictly weak consensus in the uniform compass model on $\mathbb{Z}$

DOLERA, E. and FAVARO, S.
Rates of convergence in de Finetti's representation theorem, and Hausdorff moment problem

CHIGANSKY, P., KLEPTSYNA, M. and MARUSHKEVYCH, D.
On the eigenproblem for Gaussian bridges

TALAY, D. and TOMAŠEVIĆ, M.
A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The one-dimensional case

TKACHOV, P.
On stability of traveling wave solutions for integro–differential equations related to branching Markov processes

VAN HALA, M., BANDYOPADHYAY, S., LAHIRI, S.N. and NORDMAN, D.J.
A general frequency domain method for assessing spatial covariance structures

PAVLYUKEVICH, I. and SHEVCHENKO, G.
Stratonovich stochastic differential equation with irregular coefficients: Girsanov's example revisited

WANG, W., SU, Z. and XIAO, Y.
The moduli of non-differentiability for Gaussian random fields with stationary increments

CONFORTI, G. and RIPANI, L.
Around the entropic Talagrand inequality

KOLB, M. and SAVOV, M.
A characterization of the finiteness of perpetual integrals of Lévy processes

BAI, S. and TAQQU, M.S.
Limit theorems for long-memory flows on Wiener chaos

GAO, C.
Robust Regression via Multivariate Regression Depth

KLUSOWSKI, J. M. and WU, Y.
Estimating the Number of Connected Components in a Graph via Subgraph Sampling

BRUNEL, V.-E.
Deviation Inequalities for Random Polytopes in Arbitrary Convex Bodies

BUTUCEA, C., DUBOIS, A., KROLL, M. and SAUMARD, A.
Local differential privacy: Elbow effect in optimal density estimation and adaptation over Besov ellipsoids

Continues

Bernoulli Forthcoming Papers—*Continued*

BIN, X., BUNEAR, F. and WEGKAMP, M.
A fast algorithm with minimax optimal guarantees for topic models with an unknown number of topics

KALBASI, K. and MOUNTFORD, T.
On the probability distribution of the local times of diagonally operator-self-similar Gaussian fields with stationary increments

GADAT, S., GERCHINOVITZ, S. and MARTEAU, C.
Optimal functional supervised classification with separation condition

KERKYACHARIAN, G., GEORGIADIS, A., PETRUSHEV, P., PICARD, D.
Kernel and wavelet density estimators on manifolds and more general metric spaces

SAMBALE, H. and SINULIS, A.
Logarithmic Sobolev inequalities for finite spin systems and applications

NAJAFI, A., MOTAHARI, A. and RABIEE, H.R.
Reliable Clustering of Bernoulli Mixture Models

MAYER, U., ZAHLE, H. and ZHOU, Z.
Functional weak limit theorem for an empirical process of non-stationary time series and its application

OSEKOWSKI, A.
On the best constant in the martingale version of Fefferman's inequality

JAUCH, M., HOFF, P.D. and DUNSON, D.B.
Random orthogonal matrices and the Cayley transform