BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Seven Number One February 2021 ISSN: 1350-7265

CONTENTS

BARKHAGEN, M., CHAU, N.H., MOULINES, É., RÁSONYI, M., SABANIS, S. and ZHANG, Y. 1
On stochastic gradient Langevin dynamics with dependent data streams in the logconcave case

MARYNYCH, A. and MOLCHANOV, I. 34
Sieving random iterative function systems

BELLEC, P.C. 66
Adaptive confidence sets in shape restricted regression

CHEN, Q. and GOLDSCHMIDT, C. 93
Parking on a random rooted plane tree

NKURUNZIZA, S. 107
Inference problem in generalized fractional Ornstein–Uhlenbeck processes with change-point

DUARTE, E., MARIGLIANO, O. and STURMFELS, B. 135
Discrete statistical models with rational maximum likelihood estimator

QIAO, W. 155
Nonparametric estimation of surface integrals on level sets

KUCHIBHOTLA, A.K., MUKHERJEE, S. and BANERJEE, D. 192
High-dimensional CLT: Improvements, non-uniform extensions and large deviations

BLADT, M., MIDER, M. and SØRENSEN, M. 218
Corrigendum to “Simple simulation of diffusion bridges with application to likelihood inference for diffusions”

GAO, F. and ZHU, L. 221
Precise deviations for Hawkes processes

DE UÑA-ÁLVAREZ, J. and VAN KEILEGOM, I. 249
Efron–Petrosian integrals for doubly truncated data with covariates: An asymptotic analysis

JIANG, D. and BAI, Z. 274
Generalized four moment theorem and an application to CLT for spiked eigenvalues of high-dimensional covariance matrices

FOONDUN, M. and NUALART, E. 295
The Osgood condition for stochastic partial differential equations

(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 Seven Number One February 2021 ISSN: 1350-7265

CONTENTS

(continued)

BOSSY, M., JABIR, J.-F. and MARTÍNEZ, K. 312
On the weak convergence rate of an exponential Euler scheme for SDEs governed by coefficients with superlinear growth

LAZIĆ, P. and SANDRIĆ, N. 348
On sub-geometric ergodicity of diffusion processes

SUNDARARAJAN, R.R., PIPIRAS, V. and POURAHMADI, M. 381
Stationary subspace analysis of nonstationary covariance processes: Eigenstructure description and testing

KIM, I. 419
Comparing a large number of multivariate distributions

CHEN, L.H.Y., RAIČ, M. and THÀNH, L.V. 442
On the error bound in the normal approximation for Jack measures

VAN DELFT, A. and DETTE, H. 469
A similarity measure for second order properties of non-stationary functional time series with applications to clustering and testing

UEDA, Y. 502
Max-convolution semigroups and extreme values in limit theorems for the free multiplicative convolution

WOLFER, G. and KONTOROVICH, A. 532
Statistical estimation of ergodic Markov chain kernel over discrete state space

WANG, D., YU, Y. and RINALDO, A. 554
Optimal covariance change point localization in high dimensions

GUASONI, P., MAYERHOFER, E. and ZHAO, M. 576
Minimal $\mathcal{L}^p$-densities with prescribed marginals

CHEN, X. and YANG, Y. 586
Hanson–Wright inequality in Hilbert spaces with application to K-means clustering for non-Euclidean data

SUN, R., SWART, J.M. and YU, J. 615
An invariance principle for biased voter model interfaces

NIU, Y., PATI, D. and MALLICK, B.K. 637
Bayesian graph selection consistency under model misspecification

NAGY, S., HELANDER, S., VAN BEVER, G., VIITASAARI, L. and ILMONEN, P. 673
Flexible integrated functional depths

BERTI, P., DREASSI, E., PRATELLI, L. and RIGO, P. 702
A class of models for Bayesian predictive inference

BERNOULLI Volume 27 Number 1 February 2021 Pages 1–726 ISI/BS

Volume 27 Number 1   February 2021   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 2021 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **27**(1), 2021, 1–33
https://doi.org/10.3150/19-BEJ1187

On stochastic gradient Langevin dynamics with dependent data streams in the logconcave case

M. BARKHAGEN[1,*], N.H. CHAU[2], É. MOULINES[3,4], M. RÁSONYI[5], S. SABANIS[1,6,†] and Y. ZHANG[1,‡]

[1]*School of Mathematics, The University of Edinburgh, UK.*
E-mail: [*]*n.barkhagen@ed.ac.uk;* [†]*s.sabanis@ed.ac.uk;* [‡]*ying.zhang@ed.ac.uk*
[2]*Center for Mathematical Modeling and Data Science, Osaka University, Japan.*
E-mail: chau@sigmath.es.osaka-u.ac.jp
[3]*Centre de Mathématiques Appliquées, UMR 7641, École Polytechnique, France.*
E-mail: eric.moulines@polytechnique.edu
[4]*HSE University, Russian Federation*
[5]*Alfred Renyi Institue of Mathematics, Hungary.*
E-mail: rasonyi.miklos@renyi.hu
[6]*The Alan Turing Institute, UK*

We study the problem of sampling from a probability distribution π on $\mathbb{R}^d$ which has a density w.r.t. the Lebesgue measure known up to a normalization factor $x \mapsto \mathrm{e}^{-U(x)} / \int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\,\mathrm{d}y$. We analyze a sampling method based on the Euler discretization of the Langevin stochastic differential equations under the assumptions that the potential U is continuously differentiable, ∇U is Lipschitz, and U is strongly concave. We focus on the case where the gradient of the log-density cannot be directly computed but unbiased estimates of the gradient from possibly dependent observations are available. This setting can be seen as a combination of a stochastic approximation (here stochastic gradient) type algorithms with discretized Langevin dynamics. We obtain an upper bound of the Wasserstein-2 distance between the law of the iterates of this algorithm and the target distribution π with constants depending explicitly on the Lipschitz and strong convexity constants of the potential and the dimension of the space. Finally, under weaker assumptions on U and its gradient but in the presence of independent observations, we obtain analogous results in Wasserstein-2 distance.

Keywords: L-mixing; Langevin diffusion; Monte Carlo methods; stochastic approximation

References

[1] Aliprantis, C.D. and Border, K.C. (2006). *Infinite Dimensional Analysis*: *A Hitchhiker's Guide*, 3rd ed. Berlin: Springer. MR2378491

[2] Ambrosio, L., Savaré, G. and Zambotti, L. (2009). Existence and stability for Fokker–Planck equations with log-concave reference measure. *Probab. Theory Related Fields* **145** 517–564. MR2529438 https://doi.org/10.1007/s00440-008-0177-3

[3] Chatterji, N., Flammarion, N., Ma, Y., Bartlett, P. and Jordan, M. (2018). On the theory of variance reduction for stochastic gradient Monte Carlo. Preprint, arXiv:1802.05431.

[4] Chau, H.N., Kumar, C., Rásonyi, M. and Sabanis, S. (2019). On fixed gain recursive estimators with discontinuity in the parameters. *ESAIM Probab. Stat.* **23** 217–244. MR3945579 https://doi.org/10.1051/ps/2018019

[5] Chau, H.N., Moulines, É., Rásonyi, M., Sabanis, S. and Zhang, Y. (2019). On stochastic gradient Langevin dynamics with dependent data streams: The fully non-convex case. Preprint, arXiv:1905.13142.

[6] Dalalyan, A. (2017). Further and stronger analogy between sampling and optimization: Langevin Monte Carlo and gradient descent. In *Proceedings of the* 2017 *Conference on Learning Theory* (S. Kale and O. Shamir, eds.). *Proceedings of Machine Learning Research* **65** 678–689.

1350-7265

[7] Dalalyan, A.S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183
[8] Dalalyan, A.S. and Karagulyan, A. (2019). User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient. *Stochastic Process. Appl.* **129** 5278–5311. MR4025705 https://doi.org/10.1016/j.spa.2019.02.016
[9] Durmus, A. and Moulines, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479 https://doi.org/10.1214/16-AAP1238
[10] Durmus, A. and Moulines, É. (2019). High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli* **25** 2854–2882. MR4003567 https://doi.org/10.3150/18-BEJ1073
[11] Gerencsér, B. and Rásonyi, M. (2019). On the ergodic properties of certain Markov chains in random environments. arXiv:1807.03568v3.
[12] Gerencsér, L. (1989). On a class of mixing processes. *Stoch. Stoch. Rep.* **26** 165–191. MR1018543 https://doi.org/10.1080/17442508908833555
[13] Gerencsér, L. (1992). AR(∞) estimation and nonparametric stochastic complexity. *IEEE Trans. Inf. Theory* **38** 1768–1778. MR1187818 https://doi.org/10.1109/18.165449
[14] Gerencsér, L. (1993). Strong approximation of the recursive prediction error estimator of the parameters of an ARMA process. *Systems Control Lett.* **21** 347–351. MR1241415 https://doi.org/10.1016/0167-6911(93)90078-K
[15] Gerencsér, L. (1994). On Rissanen's predictive stochastic complexity for stationary ARMA processes. *J. Statist. Plann. Inference* **41** 303–325. MR1309616 https://doi.org/10.1016/0378-3758(94)90026-4
[16] Gerencsér, L. (2006). A representation theorem for the error of recursive estimators. *SIAM J. Control Optim.* **44** 2123–2188. MR2248178 https://doi.org/10.1137/S0363012991217421
[17] 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
[18] Laruelle, S. (2011). Analyse d'Algorithmes Stochastiques Appliqués à la Finance. PhD Thesis, Université Paris VI.
[19] Majka, M.B., Mijatović, A. and Szpruch, L. (2018). Non-asymptotic bounds for sampling algorithms without log-concavity. arXiv:1808.07105.
[20] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266 https://doi.org/10.1016/S0304-4149(02)00150-3
[21] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** 518–548. MR1234295 https://doi.org/10.2307/1427522
[22] Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization: A Basic Course*. *Applied Optimization* **87**. Boston, MA: Kluwer Academic. MR2142598 https://doi.org/10.1007/978-1-4419-8853-9
[23] Neveu, J. (1975). *Discrete-Parameter Martingales*, Revised ed. *North-Holland Mathematical Library* **10**. Amsterdam-Oxford: North-Holland; New York: American Elsevier Publishing Co., Inc. Translated from the French by T.P. Speed. MR0402915
[24] Raginsky, M., Rakhlin, A. and Telgarsky, M. (2017). Non-convex learning via stochastic gradient Langevin dynamics: A nonasymptotic analysis. *Proc. Mach. Learn. Res.* **65** 1674–1703.
[25] Rásonyi, M. (2010). On the statistical analysis of quantized Gaussian AR(1) processes. *Internat. J. Adapt. Control Signal Process.* **24** 490–507. MR2666031 https://doi.org/10.1002/acs.1145
[26] Robbins, H. (1955). A remark on Stirling's formula. *Amer. Math. Monthly* **62** 26–29. MR0069328 https://doi.org/10.2307/2308012
[27] Roberts, G.O. and Tweedie, R.L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* **2** 341–363. MR1440273 https://doi.org/10.2307/3318418
[28] Teh, Y.W., Thiery, A.H. and Vollmer, S.J. (2016). Consistency and fluctuations for stochastic gradient Langevin dynamics. *J. Mach. Learn. Res.* **17** Paper No. 7, 33. MR3482927
[29] Villani, C. (2009). *Optimal Transport. Old an New*. Springer.
[30] Welling, M. and Teh, Y.W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the* 28*th International Conference on Machine Learning* 681–688.
[31] Xu, P., Chen, J., Zhou, D. and Gu, Q. (2018). Global convergence of Langevin dynamics-based algorithms for nonconvex optimization. arXiv:1707.06618.

[32] Zhuang, C. (2008). Stochastic approximation methods and applications in financial optimization problems. PhD Thesis, University of Georgia, Athens, Georgia.

Bernoulli **27**(1), 2021, 34–65
https://doi.org/10.3150/20-BEJ1221

Sieving random iterative function systems

ALEXANDER MARYNYCH[1] and ILYA MOLCHANOV[2]

[1]*Faculty of Computer Science and Cybernetics, Taras Shevchenko National University of Kyiv, Volodymyrska 60, 01601 Kyiv, Ukraine.*
E-mail: marynych@unicyb.kiev.ua
[2]*Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, CH-3012 Bern, Switzerland.*
E-mail: ilya.molchanov@stat.unibe.ch

It is known that backward iterations of independent copies of a contractive random Lipschitz function converge almost surely under mild assumptions. By a sieving (or thinning) procedure based on adding to the functions time and space components, it is possible to construct a scale invariant stochastic process. We study its distribution and paths properties. In particular, we show that it is càdlàg and has finite total variation. We also provide examples and analyse various properties of particular sieved iterative function systems including perpetuities and infinite Bernoulli convolutions, iterations of maximum, and random continued fractions.

Keywords: infinite Bernoulli convolutions; iteration; perpetuity; random Lipschitz function; scale invariant process; sieving; thinning

References

[1] Akiyama, S., Feng, D.-J., Kempton, T. and Persson, T. (2018). On the Hausdorff dimension of Bernoulli convolutions. *Int. Math. Res. Not.* rny209. https://doi.org/10.1093/imrn/rny209

[2] Alsmeyer, G. and Fuh, C. (2002). Corrigendum to: "Limit theorems for iterated random functions by regenerative methods". *Stochastic Process. Appl.* **97** 341–345. MR1875338 https://doi.org/10.1016/S0304-4149(01)00144-2

[3] Alsmeyer, G. and Fuh, C.-D. (2001). Limit theorems for iterated random functions by regenerative methods. *Stochastic Process. Appl.* **96** 123–142. MR1856683 https://doi.org/10.1016/S0304-4149(01)00104-1

[4] Alsmeyer, G., Iksanov, A. and Rösler, U. (2009). On distributional properties of perpetuities. *J. Theoret. Probab.* **22** 666–682. MR2530108 https://doi.org/10.1007/s10959-008-0156-8

[5] Alsmeyer, G., Kabluchko, Z. and Marynych, A. (2016). Leader election using random walks. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 1095–1122. MR3580090 https://doi.org/10.30757/alea.v13-39

[6] Alsmeyer, G., Kabluchko, Z. and Marynych, A. (2017). A leader-election procedure using records. *Ann. Probab.* **45** 4348–4388. MR3737913 https://doi.org/10.1214/16-AOP1167

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

[8] Chamayou, J.-F. and Letac, G. (1991). Explicit stationary distributions for compositions of random functions and products of random matrices. *J. Theoret. Probab.* **4** 3–36. MR1088391 https://doi.org/10.1007/BF01046992

[9] Covo, S. (2009). On approximations of small jumps of subordinators with particular emphasis on a Dickman-type limit. *J. Appl. Probab.* **46** 732–755. MR2562319 https://doi.org/10.1239/jap/1253279849

[10] DeGroot, M.H. and Rao, M.M. (1963). Stochastic give-and-take. *J. Math. Anal. Appl.* **7** 489–498. MR0157423 https://doi.org/10.1016/0022-247X(63)90071-4

[11] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.* **41** 45–76. MR1669737 https://doi.org/10.1137/S0036144598338446

[12] Dubins, L.E. and Freedman, D.A. (1967). Random distribution functions. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (*Berkeley, Calif.*, 1965/66) 183–214. Berkeley, CA: Univ. California Press. MR0214109

[13] Duflo, M. (1997). *Random Iterative Models. Applications of Mathematics* (*New York*) **34**. Berlin: Springer. Translated from the 1990 French original by Stephen S. Wilson and revised by the author. MR1485774 https://doi.org/10.1007/978-3-662-12880-0

[14] Dümbgen, L. (2017). Empirische Prozesse. Lecture Notes. Available at https://ilias.unibe.ch/goto_ilias3_unibe_cat_915738.html.

[15] Erdös, P. (1939). On a family of symmetric Bernoulli convolutions. *Amer. J. Math.* **61** 974–976. MR0000311 https://doi.org/10.2307/2371641

[16] Erdös, P. (1940). On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.* **62** 180–186. MR0000858 https://doi.org/10.2307/2371446

[17] Feng, D.-J. (2019). Dimension of invariant measures for affine iterated function systems. Preprint available at https://arxiv.org/abs/1901.01691.

[18] Gnedin, A.V. (2008). Corners and records of the Poisson process in quadrant. *Electron. Commun. Probab.* **13** 187–193. MR2399280 https://doi.org/10.1214/ECP.v13-1351

[19] Goldie, C.M. and Maller, R.A. (2000). Stability of perpetuities. *Ann. Probab.* **28** 1195–1218. MR1797309 https://doi.org/10.1214/aop/1019160331

[20] Grincevičius, A.K. (1981). A random difference equation. *Lith. Math. J.* **21** 302–306.

[21] Hochman, M. and Rapaport, A. (2019). Hausdorff dimension of planar self-affine sets and measures with overlaps. Preprint avialable at https://arxiv.org/abs/1904.09812.

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

[23] Jordan, T., Pollicott, M. and Simon, K. (2007). Hausdorff dimension for randomly perturbed self affine attractors. *Comm. Math. Phys.* **270** 519–544. MR2276454 https://doi.org/10.1007/s00220-006-0161-7

[24] Kesten, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248. MR0440724 https://doi.org/10.1007/BF02392040

[25] Khinchin, A.Ya. (1997). *Continued Fractions*, Russian ed. Mineola, NY: Dover. With a preface by B. V. Gnedenko, Reprint of the 1964 translation. MR1451873

[26] Letac, G. (1986). A contraction principle for certain Markov chains and its applications. In *Random Matrices and Their Applications* (*Brunswick, Maine*, 1984). *Contemp. Math.* **50** 263–273. Providence, RI: Amer. Math. Soc. MR0841098 https://doi.org/10.1090/conm/050/841098

[27] Letac, G. and Seshadri, V. (1983). A characterization of the generalized inverse Gaussian distribution by continued fractions. *Z. Wahrsch. Verw. Gebiete* **62** 485–489. MR0690573 https://doi.org/10.1007/BF00534200

[28] Penrose, M.D. and Wade, A.R. (2004). Random minimal directed spanning trees and Dickman-type distributions. *Adv. in Appl. Probab.* **36** 691–714. MR2079909 https://doi.org/10.1239/aap/1093962229

[29] Peres, Y., Schlag, W. and Solomyak, B. (2000). Sixty years of Bernoulli convolutions. In *Fractal Geometry and Stochastics, II* (*Greifswald/Koserow*, 1998). *Progress in Probability* **46** 39–65. Basel: Birkhäuser. MR1785620

[30] Solomyak, B. (1995). On the random series $\sum \pm\lambda^n$ (an Erdős problem). *Ann. of Math.* (2) **142** 611–625. MR1356783 https://doi.org/10.2307/2118556

[31] Vapnik, V.N. (1998). *Statistical Learning Theory. Adaptive and Learning Systems for Signal Processing, Communications, and Control*. New York: Wiley. A Wiley-Interscience Publication. MR1641250

[32] Varjú, P.P. (2019). On the dimension of Bernoulli convolutions for all transcendental parameters. *Ann. of Math.* (2) **189** 1001–1011. MR3961088 https://doi.org/10.4007/annals.2019.189.3.9

Bernoulli **27**(1), 2021, 66–92
https://doi.org/10.3150/20-BEJ1223

Adaptive confidence sets in shape restricted regression

PIERRE C. BELLEC

Department of Statistics, Hill Center, Busch Campus, Rutgers University, Piscataway, NJ 08854, USA.
E-mail: pierre.bellec@rutgers.edu

A simple construction of adaptive confidence sets is proposed in isotonic, convex and unimodal regression. In univariate isotonic regression, the proposed confidence set enjoys uniform coverage over all non-decreasing regression functions. Furthermore, the diameter of the proposed confidence set automatically adapts to the unknown number of pieces of the true parameter, in the sense that the diameter is bounded from above by the minimax risk over the class of k-piecewise constant functions. The diameter of the confidence set is a simple increasing function of the number of jumps of the isotonic least-squares estimate.

A similar construction is proposed in convex regression where the true regression function is convex and piecewise affine. Here, the confidence set enjoys uniform coverage and its diameter automatically adapts to the number of affine pieces of the true regression function. The diameter of the confidence set is an increasing function of the number of affine pieces of the convex least-squares estimate.

We explain how to extend this technique to a non-convex set by proposing a similar adaptive confidence set in unimodal regression. The confidence set automatically adapts to the number of jumps of the true unimodal regression function and its diameter is an increasing function of the number of jumps of the unimodal least-squares estimate.

Keywords: adaptive confidence set; convex regresion; isotonic regression; piecewise affine; piecewise constant; shape constraints; unimodal regression

References

[1] Amelunxen, D., Lotz, M., McCoy, M.B. and Tropp, J.A. (2014). Living on the edge: Phase transitions in convex programs with random data. *Inf. Inference* **3** 224–294. MR3311453 https://doi.org/10.1093/imaiai/iau005

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

[3] Bellec, P.C. (2018). Sharp oracle inequalities for least squares estimators in shape restricted regression. *Ann. Statist.* **46** 745–780. MR3782383 https://doi.org/10.1214/17-AOS1566

[4] Bellec, P.C. and Tsybakov, A.B. (2015). Sharp oracle bounds for monotone and convex regression through aggregation. *J. Mach. Learn. Res.* **16** 1879–1892. MR3417801

[5] 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

[6] Cai, T.T. and Low, M.G. (2006). Adaptive confidence balls. *Ann. Statist.* **34** 202–228. MR2275240 https://doi.org/10.1214/009053606000000146

[7] Cai, T.T., Low, M.G. and Xia, Y. (2013). Adaptive confidence intervals for regression functions under shape constraints. *Ann. Statist.* **41** 722–750. MR3099119 https://doi.org/10.1214/12-AOS1068

[8] Carpentier, A. and Verzelen, N. (2019). Optimal sparsity testing in linear regression model. ArXiv preprint. Available at arXiv:1901.08802.

[9] Carpentier, A. and Verzelen, N. (2019). Adaptive estimation of the sparsity in the Gaussian vector model. *Ann. Statist.* **47** 93–126. MR3909928 https://doi.org/10.1214/17-AOS1680

1350-7265

[10] Chandrasekaran, V., Recht, B., Parrilo, P.A. and Willsky, A.S. (2012). The convex geometry of linear inverse problems. *Found. Comput. Math.* **12** 805–849. MR2989474 https://doi.org/10.1007/s10208-012-9135-7
[11] Chatterjee, S., Guntuboyina, A. and Sen, B. (2015). On risk bounds in isotonic and other shape restricted regression problems. *Ann. Statist.* **43** 1774–1800. MR3357878 https://doi.org/10.1214/15-AOS1324
[12] Chatterjee, S., Guntuboyina, A. and Sen, B. (2018). On matrix estimation under monotonicity constraints. *Bernoulli* **24** 1072–1100. MR3706788 https://doi.org/10.3150/16-BEJ865
[13] Chatterjee, S. and Lafferty, J. (2019). Adaptive risk bounds in unimodal regression. *Bernoulli* **25** 1–25. MR3892309 https://doi.org/10.3150/16-bej922
[14] Chen, X., Lin, Q. and Sen, B. (2020). On Degrees of Freedom of Projection Estimators With Applications to Multivariate Nonparametric Regression. *J. Amer. Statist. Assoc.* **115** 173–186. MR4078455 https://doi.org/10.1080/01621459.2018.1537917
[15] Dümbgen, L. (2003). Optimal confidence bands for shape-restricted curves. *Bernoulli* **9** 423–449. MR1997491 https://doi.org/10.3150/bj/1065444812
[16] Flammarion, N., Mao, C. and Rigollet, P. (2019). Optimal rates of statistical seriation. *Bernoulli* **25** 623–653. MR3892331 https://doi.org/10.3150/17-bej1000
[17] Gao, C., Han, F. and Zhang, C.-H. (2017). Minimax risk bounds for piecewise constant models. ArXiv preprint. Available at arXiv:1705.06386.
[18] Guntuboyina, A., Lieu, D., Chatterjee, S. and Sen, B. (2020). Adaptive risk bounds in univariate total variation denoising and trend filtering. *Ann. Statist.* **48** 205–229. MR4065159 https://doi.org/10.1214/18-AOS1799
[19] Guntuboyina, A. and Sen, B. (2015). Global risk bounds and adaptation in univariate convex regression. *Probab. Theory Related Fields* **163** 379–411. MR3405621 https://doi.org/10.1007/s00440-014-0595-3
[20] Guntuboyina, A. and Sen, B. (2018). Nonparametric shape-restricted regression. *Statist. Sci.* **33** 568–594. MR3881209 https://doi.org/10.1214/18-STS665
[21] Han, Q., Wang, T., Chatterjee, S. and Samworth, R.J. (2019). Isotonic regression in general dimensions. *Ann. Statist.* **47** 2440–2471. MR3988762 https://doi.org/10.1214/18-AOS1753
[22] Ingster, Y.I., Tsybakov, A.B. and Verzelen, N. (2010). Detection boundary in sparse regression. *Electron. J. Stat.* **4** 1476–1526. MR2747131 https://doi.org/10.1214/10-EJS589
[23] Matey, N. (2019). Gaussian regression with convex constraints. In *Proceedings of Machine Learning Research* (K. Chaudhuri and M. Sugiyama, eds.). *Proceedings of Machine Learning Research* **89** 31–38. Available at http://proceedings.mlr.press/v89/neykov19b.html.
[24] McCoy, M.B. and Tropp, J.A. (2014). From Steiner formulas for cones to concentration of intrinsic volumes. *Discrete Comput. Geom.* **51** 926–963. MR3216671 https://doi.org/10.1007/s00454-014-9595-4
[25] Meyer, M. and Woodroofe, M. (2000). On the degrees of freedom in shape-restricted regression. *Ann. Statist.* **28** 1083–1104. MR1810920 https://doi.org/10.1214/aos/1015956708
[26] 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
[27] Oymak, S. and Hassibi, B. (2016). Sharp MSE bounds for proximal denoising. *Found. Comput. Math.* **16** 965–1029. MR3529131 https://doi.org/10.1007/s10208-015-9278-4
[28] Stout, Q.F. (2008). Unimodal regression via prefix isotonic regression. *Comput. Statist. Data Anal.* **53** 289–297. MR2649085 https://doi.org/10.1016/j.csda.2008.08.005
[29] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794
[30] Verzelen, N. (2012). Minimax risks for sparse regressions: Ultra-high dimensional phenomenons. *Electron. J. Stat.* **6** 38–90. MR2879672 https://doi.org/10.1214/12-EJS666
[31] Zhang, C.-H. (2002). Risk bounds in isotonic regression. *Ann. Statist.* **30** 528–555. MR1902898 https://doi.org/10.1214/aos/1021379864

Bernoulli **27**(1), 2021, 93–106
https://doi.org/10.3150/20-BEJ1227

Parking on a random rooted plane tree

QIZHAO CHEN[1] and CHRISTINA GOLDSCHMIDT[2]

[1]*The Queen's College, University of Oxford, Oxford, OX1 4AW, UK.*
E-mail: qizhaochen@g.harvard.edu
[2]*Department of Statistics and Lady Margaret Hall, University of Oxford, Oxford, OX2 6QA, UK.*
E-mail: goldschm@stats.ox.ac.uk

In this paper, we investigate a parking process on a uniform random rooted plane tree with n vertices. Every vertex of the tree has a parking space for a single car. Cars arrive at independent uniformly random vertices of the tree. If the parking space at a vertex is unoccupied when a car arrives there, it parks. If not, the car drives towards the root and parks in the first empty space it encounters (if there is one). We are interested in asymptotics of the probability of the event that all cars can park when $\lfloor \alpha n \rfloor$ cars arrive, for $\alpha > 0$. We observe that there is a phase transition at $\alpha_c := \sqrt{2} - 1$: if $\alpha < \alpha_c$ then the event has positive limiting probability, whereas for $\alpha > \alpha_c$ its probability tends to 0. Analogous results have been proved by Lackner and Panholzer (*J. Combin. Theory Ser. A* **142** (2016) 1–28), Goldschmidt and Przykucki (*Combin. Probab. Comput.* **28** (2019) 23–45) and Jones (*J. Appl. Probab.* **56** (2019) 1065–1085) for different underlying random tree models.

Keywords: parking; phase transition; random trees

References

[1] Aldous, D. and Steele, J.M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures*. *Encyclopaedia Math. Sci.* **110** 1–72. Berlin: Springer. MR2023650 https://doi.org/10.1007/978-3-662-09444-0_1
[2] Brown, M., Peköz, E.A. and Ross, S.M. (2010). Some results for skip-free random walk. *Probab. Engrg. Inform. Sci.* **24** 491–507. MR2725345 https://doi.org/10.1017/S0269964810000136
[3] Curien, N. and Hénard, O. (2019). The phase transition for parking on Galton–Watson trees. Available at arXiv:1912.06012 [math.PR].
[4] Goldschmidt, C. and Przykucki, M. (2019). Parking on a random tree. *Combin. Probab. Comput.* **28** 23–45. MR3917904 https://doi.org/10.1017/S0963548318000457
[5] Jones, O.D. (2019). Runoff on rooted trees. *J. Appl. Probab.* **56** 1065–1085. MR4041449 https://doi.org/10.1017/jpr.2019.61
[6] Kesten, H. (1986). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** 425–487. MR0871905
[7] Konheim, A. and Weiss, B. (1966). An occupancy discipline and applications. *SIAM J. Appl. Math.* **14** 1266–1274.
[8] Lackner, M.-L. and Panholzer, A. (2016). Parking functions for mappings. *J. Combin. Theory Ser. A* **142** 1–28. MR3499489 https://doi.org/10.1016/j.jcta.2016.03.001
[9] Luczak, M. and Winkler, P. (2004). Building uniformly random subtrees. *Random Structures Algorithms* **24** 420–443. MR2060629 https://doi.org/10.1002/rsa.20011

Bernoulli **27**(1), 2021, 107–134
https://doi.org/10.3150/20-BEJ1230

Inference problem in generalized fractional Ornstein–Uhlenbeck processes with change-point

SÉVÉRIEN NKURUNZIZA

Mathematics and Statistics Department, University of Windsor, 401 Sunset Avenue, Windsor, ON, N9B 3P4, Canada.
E-mail: severien@uwindsor.ca

In this paper, we study an inference problem in generalized fractional Ornstein–Uhlenbeck (O–U) processes with an unknown change-point when the drift parameter is suspected to satisfy some constraints. The constraint considered is very general and, the testing problem studied generalizes a very recent inference problem in generalized O–U processes. We derive the unrestricted estimator (UE) and the restricted estimator (RE) and we establish the asymptotic properties of the UE and RE. We also propose some shrinkage-type estimators (SEs) as well as a test for testing the constraint. Finally, we derive the asymptotic power of the proposed test and we study the relative risk dominance of the proposed estimators.

Keywords: ADR; change-point; drift-parameter; fractional mean-reverting process; fractional Ornstein–Uhlenbeck process; fractional SDE; shrinkage estimators; testing

References

[1] Aitchison, J. and Silvey, S.D. (1958). Maximum-likelihood estimation of parameters subject to restraints. *Ann. Math. Stat.* **29** 813–828. MR0094873 https://doi.org/10.1214/aoms/1177706538
[2] Bajja, S., Es-Sebaiy, K. and Viitasaari, L. (2017). Least squares estimator of fractional Ornstein–Uhlenbeck processes with periodic mean. *J. Korean Statist. Soc.* **46** 608–622. MR3718150 https://doi.org/10.1016/j.jkss.2017.06.002
[3] Barboza, L.A. and Viens, F.G. (2017). Parameter estimation of Gaussian stationary processes using the generalized method of moments. *Electron. J. Stat.* **11** 401–439. MR3611508 https://doi.org/10.1214/17-EJS1230
[4] Berzin, C., Latour, A. and León, J.R. (2014). *Inference on the Hurst Parameter and the Variance of Diffusions Driven by Fractional Brownian Motion. Lecture Notes in Statistics* **216**. Cham: Springer. With a foreword by Aline Bonami. MR3289986 https://doi.org/10.1007/978-3-319-07875-5
[5] Chen, F., Mamon, R. and Nkurunziza, S. (2018). Inference for a change-point problem under a generalised Ornstein–Uhlenbeck setting. *Ann. Inst. Statist. Math.* **70** 807–853. MR3830289 https://doi.org/10.1007/s10463-017-0610-4
[6] Chen, F. and Nkurunziza, S. (2015). Optimal method in multiple regression with structural changes. *Bernoulli* **21** 2217–2241. MR3378465 https://doi.org/10.3150/14-BEJ642
[7] Chronopoulou, A. and Viens, F.G. (2012). Estimation and pricing under long-memory stochastic volatility. *Ann. Finance* **8** 379–403. MR2922802 https://doi.org/10.1007/s10436-010-0156-4
[8] Dehling, H., Franke, B. and Kott, T. (2010). Drift estimation for a periodic mean reversion process. *Stat. Inference Stoch. Process.* **13** 175–192. MR2729648 https://doi.org/10.1007/s11203-010-9045-8
[9] Dehling, H., Franke, B., Kott, T. and Kulperger, R. (2014). Change point testing for the drift parameters of a periodic mean reversion process. *Stat. Inference Stoch. Process.* **17** 1–18. MR3178011 https://doi.org/10.1007/s11203-014-9092-7
[10] Dehling, H., Franke, B. and Woerner, J.H.C. (2017). Estimating drift parameters in a fractional Ornstein Uhlenbeck process with periodic mean. *Stat. Inference Stoch. Process.* **20** 1–14. MR3619570 https://doi.org/10.1007/s11203-016-9136-2

[11] Gombay, E. (2010). Change detection in linear regression with time series errors. *Canad*. *J*. *Statist*. **38** 65–79. MR2676930 https://doi.org/10.1002/cjs.10043

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

[13] Mathai, A.M. and Provost, S.B. (1992). *Quadratic Forms in Random Variables*: *Theory and Applications*. *Statistics*: *Textbooks and Monographs* **126**. New York: Dekker. MR1192786

[14] Nkurunziza, S. (2012). Shrinkage strategies in some multiple multi-factor dynamical systems. *ESAIM Probab*. *Stat*. **16** 139–150. MR2946124 https://doi.org/10.1051/ps/2010015

[15] Nkurunziza, S. (2020). Supplement to "Inference problem in generalized fractional Ornstein–Uhlenbeck processes with change-point." https://doi.org/10.3150/20-BEJ1230SUPP

[16] Nkurunziza, S. and Zhang, P.P. (2018). Estimation and testing in generalized mean-reverting processes with change-point. *Stat*. *Inference Stoch*. *Process*. **21** 191–215. MR3769837 https://doi.org/10.1007/s11203-016-9151-3

[17] Roy, S., Atchadé, Y. and Michailidis, G. (2017). Change point estimation in high dimensional Markov random-field models. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **79** 1187–1206. MR3689314 https://doi.org/10.1111/rssb.12205

[18] Saleh, A.K.Md.E. (2006). *Theory of Preliminary Test and Stein-Type Estimation with Applications*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR2218139 https://doi.org/10.1002/0471773751

[19] Woody, J. and Lund, R. (2014). A linear regression model with persistent level shifts: An alternative to infill asymptotics. *Statist*. *Probab*. *Lett*. **95** 118–124. MR3262958 https://doi.org/10.1016/j.spl.2014.08.018

[20] Xiao, W., Zhang, W. and Xu, W. (2011). Parameter estimation for fractional Ornstein–Uhlenbeck processes at discrete observation. *Appl*. *Math*. *Model*. **35** 4196–4207. MR2801946 https://doi.org/10.1016/j.apm.2011.02.047

Bernoulli **27**(1), 2021, 135–154
https://doi.org/10.3150/20-BEJ1231

Discrete statistical models with rational maximum likelihood estimator

ELIANA DUARTE[1], ORLANDO MARIGLIANO[2,*] and BERND STURMFELS[2,†]

[1]*Fakultät für Mathematik, Otto-von-Guericke Universität Magdeburg, 39106 Magdeburg, Germany.*
E-mail: eliana.duarte@ovgu.de
[2]*Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstraße 22, 04103 Leipzig, Germany.*
E-mail: [*]orlando.marigliano@mis.mpg.de; [†]bernd@mis.mpg.de

A discrete statistical model is a subset of a probability simplex. Its maximum likelihood estimator (MLE) is a retraction from that simplex onto the model. We characterize all models for which this retraction is a rational function. This is a contribution via real algebraic geometry which rests on results on Horn uniformization due to Huh and Kapranov. We present an algorithm for constructing models with rational MLE, and we demonstrate it on a range of instances. Our focus lies on models familiar to statisticians, like Bayesian networks, decomposable graphical models and staged trees.

Keywords: algebraic statistics; discrete statistical models; graphical models; likelihood geometry; maximum likelihood estimator; real algebraic geometry

References

[1] Ay, N., Jost, J., Lê, H.V. and Schwachhöfer, L. (2017). *Information Geometry*. Cham: Springer. MR3701408 https://doi.org/10.1007/978-3-319-56478-4
[2] Clarke, P. and Cox, D.A. (2020). Moment maps, strict linear precision, and maximum likelihood degree one. *Adv. Math.* **370** 107233. MR4103774 https://doi.org/10.1016/j.aim.2020.107233
[3] Collazo, R.A., Görgen, C. and Smith, J.Q. (2018). *Chain Event Graphs*. *Chapman & Hall/CRC Computer Science and Data Analysis Series*. Boca Raton, FL: CRC Press. MR3752634
[4] Drton, M., Sturmfels, B. and Sullivant, S. (2009). *Lectures on Algebraic Statistics*. *Oberwolfach Seminars* **39**. Basel: Birkhäuser. MR2723140 https://doi.org/10.1007/978-3-7643-8905-5
[5] Duarte, E. and Görgen, C. (2020). Equations defining probability tree models. *J. Symbolic Comput.* **99** 127–146. MR4051945 https://doi.org/10.1016/j.jsc.2019.04.001
[6] Garcia-Puente, L.D. and Sottile, F. (2010). Linear precision for parametric patches. *Adv. Comput. Math.* **33** 191–214. MR2659586 https://doi.org/10.1007/s10444-009-9126-7
[7] Gelfand, I.M., Kapranov, M.M. and Zelevinsky, A.V. (1994). *Discriminants, Resultants, and Multidimensional Determinants*. *Mathematics: Theory & Applications*. Boston, MA: Birkhäuser, Inc. MR1264417 https://doi.org/10.1007/978-0-8176-4771-1
[8] Görgen, C. and Smith, J.Q. (2018). Equivalence classes of staged trees. *Bernoulli* **24** 2676–2692. MR3779698 https://doi.org/10.3150/17-BEJ940
[9] Grayson, D. and Stillman, M. Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/.
[10] Huh, J. (2014). Varieties with maximum likelihood degree one. *J. Algebr. Stat.* **5** 1–17. MR3279951 https://doi.org/10.18409/jas.v5i1.22
[11] Huh, J. and Sturmfels, B. (2014). Likelihood geometry. In *Combinatorial Algebraic Geometry*. *Lecture Notes in Math.* **2108** 63–117. Cham: Springer. MR3329087 https://doi.org/10.1007/978-3-319-04870-3_3
[12] Kapranov, M.M. (1991). A characterization of A-discriminantal hypersurfaces in terms of the logarithmic Gauss map. *Math. Ann.* **290** 277–285. MR1109634 https://doi.org/10.1007/BF01459245
[13] Krasauskas, R. (2002). Toric surface patches. *Adv. Comput. Math.* **17** 89–113. MR1902537 https://doi.org/10.1023/A:1015289823859

[14] Lauritzen, S.L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. New York: Oxford University Press. MR1419991

[15] Silander, T. and Leong, T.-Y. (2013). A dynamic programming algorithm for learning chain event graphs. In *Discovery Science* (J. Fürnkranz, E. Hüllermeier and T. Higuchi, eds.). *Lecture Notes in Computer Science* **8140** 201–216. Berlin: Springer.

[16] Smith, J.Q. and Anderson, P.E. (2008). Conditional independence and chain event graphs. *Artificial Intelligence* **172** 42–68. MR2388535 https://doi.org/10.1016/j.artint.2007.05.004

[17] Sullivant, S. (2018). *Algebraic Statistics*. *Graduate Studies in Mathematics* **194**. Providence, RI: Amer. Math. Soc. MR3838364

[18] https://github.com/emduart2/DiscreteStatisticalModelsWithRationalMLE.

Bernoulli **27**(1), 2021, 155–191
https://doi.org/10.3150/20-BEJ1232

Nonparametric estimation of surface integrals on level sets

WANLI QIAO

Department of Statistics, George Mason University, 4400 University Drive, MS 4A7, Fairfax, VA 22030, USA.
E-mail: *wqiao@gmu.edu*

Surface integrals on density level sets often appear in asymptotic results of nonparametric level set estimation, such as for confidence regions and bandwidth selection. Also surface integrals can be used to describe the shape of level sets. Assuming the integrands are known, we consider three estimators of the surface integrals on density level sets, one as a direct plug-in estimator, and the other two based on different neighborhoods of level sets. For all the three estimators, we derive the rates of convergence and asymptotic distributions.

Keywords: curvatures; kernel density estimation; level sets; positive reach; surface integrals

References

[1] Ambrosio, L., Colesanti, A. and Villa, E. (2008). Outer Minkowski content for some classes of closed sets. *Math*. *Ann*. **342** 727–748. MR2443761 https://doi.org/10.1007/s00208-008-0254-z
[2] Arias-Castro, E., Mason, D. and Pelletier, B. (2016). On the estimation of the gradient lines of a density and the consistency of the mean-shift algorithm. *J*. *Mach*. *Learn*. *Res*. **17** Paper No. 43, 28. MR3491137
[3] Arias-Castro, E. and Rodríguez-Casal, A. (2017). On estimating the perimeter using the alpha-shape. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **53** 1051–1068. MR3689960 https://doi.org/10.1214/16-AIHP747
[4] Armendáriz, I., Cuevas, A. and Fraiman, R. (2009). Nonparametric estimation of boundary measures and related functionals: Asymptotic results. *Adv*. *in Appl*. *Probab*. **41** 311–322. MR2541178 https://doi.org/10.1239/aap/1246886612
[5] Baddeley, A. and Jensen, E.B.V. (2005). *Stereology for Statisticians*. *Monographs on Statistics and Applied Probability* **103**. Boca Raton, FL: CRC Press/CRC. MR2107000
[6] Baddeley, A.J., Gundersen, H.J.G. and Cruz-Orive, L.M. (1986). Estimation of surface area from vertical sections. *J*. *Microsc*. **142** 259–276.
[7] Baíllo, A. (2003). Total error in a plug-in estimator of level sets. *Statist*. *Probab*. *Lett*. **65** 411–417. MR2039885 https://doi.org/10.1016/j.spl.2003.08.007
[8] Baíllo, A., Cuevas, A. and Justel, A. (2000). Set estimation and nonparametric detection. *Canad*. *J*. *Statist*. **28** 765–782. MR1821433 https://doi.org/10.2307/3315915
[9] Cadre, B. (2006). Kernel estimation of density level sets. *J*. *Multivariate Anal*. **97** 999–1023. MR2256570 https://doi.org/10.1016/j.jmva.2005.05.004
[10] Cadre, B., Pelletier, B. and Pudlo, P. (2009). Clustering by estimation of density level sets at a fixed probability. Available at http://hal.archives-ouvertes.fr/docs/00/39/74/37/PDF/tlevel.pdf.
[11] Cannings, T.I., Berrett, T.B. and Samworth, R.J. (2017). Local nearest neighbour classification with applications to semi-supervised learning. Preprint. Available at arXiv:1704.00642.
[12] Canzonieri, V. and Carbone, A. (1998). Clinical and biological applications of image analysis in non-Hodgkin's lymphomas. *Hematol*. *Oncol*. **16** 15–28. https://doi.org/10.1002/(sici)1099-1069(1998030)16:1<15::aid-hon618>3.0.co;2-j
[13] Caselles, V., Haro, G., Sapiro, G. and Verdera, J. (2008). On geometric variational models for inpainting surface holes. *Comput*. *Vis*. *Image Underst*. **111** 351–373.
[14] Chacón, J.E. and Duong, T. (2018). *Multivariate Kernel Smoothing and Its Applications*. *Monographs on Statistics and Applied Probability* **160**. Boca Raton, FL: CRC Press. MR3822372 https://doi.org/10.1201/9780429485572

[15] Chazal, F. (2013). An upper bound for the volume of geodesic balls in submanifolds of Euclidean spaces. Technical report, INRIA.
[16] Chazal, F., Lieutier, A. and Rossignac, J. (2007). Normal-map between normal-compatible manifolds. *Internat. J. Comput. Geom. Appl.* **17** 403–421. MR2362411 https://doi.org/10.1142/S0218195907002422
[17] Chen, Y.-C., Genovese, C.R. and Wasserman, L. (2017). Density level sets: Asymptotics, inference, and visualization. *J. Amer. Statist. Assoc.* **112** 1684–1696. MR3750891 https://doi.org/10.1080/01621459.2016.1228536
[18] Cuevas, A., Fraiman, R. and Pateiro-López, B. (2012). On statistical properties of sets fulfilling rolling-type conditions. *Adv. in Appl. Probab.* **44** 311–329. MR2977397 https://doi.org/10.1239/aap/1339878713
[19] Cuevas, A., Fraiman, R. and Rodríguez-Casal, A. (2007). A nonparametric approach to the estimation of lengths and surface areas. *Ann. Statist.* **35** 1031–1051. MR2341697 https://doi.org/10.1214/009053606000001532
[20] Cuevas, A., González-Manteiga, W. and Rodríguez-Casal, A. (2006). Plug-in estimation of general level sets. *Aust. N. Z. J. Stat.* **48** 7–19. MR2234775 https://doi.org/10.1111/j.1467-842X.2006.00421.x
[21] Cuevas, A. and Pateiro-López, B. (2018). Polynomial volume estimation and its applications. *J. Statist. Plann. Inference* **196** 174–184. MR3772767 https://doi.org/10.1016/j.jspi.2017.11.005
[22] Evans, L.C. and Gariepy, R.F. (1992). *Measure Theory and Fine Properties of Functions*. *Studies in Advanced Mathematics*. Boca Raton, FL: CRC Press. MR1158660
[23] Federer, H. (1959). Curvature measures. *Trans. Amer. Math. Soc.* **93** 418–491. MR0110078 https://doi.org/10.2307/1993504
[24] García Trillos, N., Slepčev, D. and von Brecht, J. (2017). Estimating perimeter using graph cuts. *Adv. in Appl. Probab.* **49** 1067–1090. MR3732187 https://doi.org/10.1017/apr.2017.34
[25] Giné, E. and Guillou, A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 907–921. MR1955344 https://doi.org/10.1016/S0246-0203(02)01128-7
[26] Gokhale, A.M. (1990). Unbiased estimation of curve length in 3-D using vertical slices. *J. Microsc.* **159** 133–141.
[27] Gray, A. (2004). *Tubes*, 2nd ed. *Progress in Mathematics* **221**. Basel: Birkhäuser. With a preface by Vicente Miquel. MR2024928 https://doi.org/10.1007/978-3-0348-7966-8
[28] Guillemin, V. and Pollack, A. (2010). *Differential Topology*. Providence, RI: AMS Chelsea Publishing. MR2680546 https://doi.org/10.1090/chel/370
[29] Hall, P. and Kang, K.-H. (2005). Bandwidth choice for nonparametric classification. *Ann. Statist.* **33** 284–306. MR2157804 https://doi.org/10.1214/009053604000000959
[30] Hall, P. and Murison, R.D. (1993). Correcting the negativity of high-order kernel density estimators. *J. Multivariate Anal.* **47** 103–122. MR1239108 https://doi.org/10.1006/jmva.1993.1073
[31] Hartigan, J.A. (1987). Estimation of a convex density contour in two dimensions. *J. Amer. Statist. Assoc.* **82** 267–270. MR0883354
[32] Jiménez, R. and Yukich, J.E. (2011). Nonparametric estimation of surface integrals. *Ann. Statist.* **39** 232–260. MR2797845 https://doi.org/10.1214/10-AOS837
[33] Kerscher, M. (2000). Statistical analysis of large-scale structure in the Universe. In *Statistical Physics and Spatial Statistics*: *The Art of Analyzing and Modeling Spatial Structures and Pattern Formation*. *Lecture Notes in Physics* **554** 36–71. Berlin: Springer. MR1870944 https://doi.org/10.1007/3-540-45043-2_3
[34] Kindlmann, G., Whitaker, R., Tasdizen, T. and Möller, T. (2003). Curvature-based transfer functions for direct volume rendering: Methods and applications. In *Proc. IEEE Visualization* 2003 513–520.
[35] Lee, J.M. (1997). *Riemannian Manifolds*: *An Introduction to Curvature*. *Graduate Texts in Mathematics* **176**. New York: Springer. MR1468735 https://doi.org/10.1007/b98852
[36] Lim, P.H., Bagci, U. and Bai, L. (2013). Introducing Willmore flow into level set segmentation of spinal vertebrae. *IEEE Trans. Biomed. Eng.* **60** 115–122.
[37] Mammen, E. and Polonik, W. (2013). Confidence regions for level sets. *J. Multivariate Anal.* **122** 202–214. MR3189318 https://doi.org/10.1016/j.jmva.2013.07.017
[38] Mammen, E. and Tsybakov, A.B. (1999). Smooth discrimination analysis. *Ann. Statist.* **27** 1808–1829. MR1765618 https://doi.org/10.1214/aos/1017939240

[39] Martínez, V.J. and Saar, E. (2001). *Statistics of the Galaxy Distribution*. Roca Raton, FL: CRC Press/CRC.
[40] Mason, D.M. and Polonik, W. (2009). Asymptotic normality of plug-in level set estimates. *Ann. Appl. Probab.* **19** 1108–1142. MR2537201 https://doi.org/10.1214/08-AAP569
[41] Pateiro-López, B. and Rodríguez-Casal, A. (2008). Length and surface area estimation under smoothness restrictions. *Adv. in Appl. Probab.* **40** 348–358. MR2431300
[42] Pateiro-López, B. and Rodríguez-Casal, A. (2009). Surface area estimation under convexity type assumptions. *J. Nonparametr. Stat.* **21** 729–741. MR2549435 https://doi.org/10.1080/10485250902971732
[43] Polonik, W. (1995). Measuring mass concentrations and estimating density contour clusters—An excess mass approach. *Ann. Statist.* **23** 855–881. MR1345204 https://doi.org/10.1214/aos/1176324626
[44] Pranav, P., Edelsbrunner, H., van de Weygaert, R., Vegter, G., Kerber, M., Jones, B.J.T. and Wintraecken, M. (2017). The topology of the cosmic web in terms of persistent Betti numbers. *Mon. Not. R. Astron. Soc.* **465** 4281–4310.
[45] Qiao, W. (2020). Asymptotics and optimal bandwidth for nonparametric estimation of density level sets. *Electron. J. Stat.* **14** 302–344. MR4048601 https://doi.org/10.1214/19-EJS1668
[46] Qiao, W. (2020). Asymptotic confidence regions for density ridges. Preprint. Available at arXiv:2004.11354.
[47] Qiao, W. and Polonik, W. (2019). Nonparametric confidence regions for level sets: Statistical properties and geometry. *Electron. J. Stat.* **13** 985–1030. MR3934621 https://doi.org/10.1214/19-EJS1543
[48] Rigollet, P. and Vert, R. (2009). Optimal rates for plug-in estimators of density level sets. *Bernoulli* **15** 1154–1178. MR2597587 https://doi.org/10.3150/09-BEJ184
[49] Rinaldo, A. and Wasserman, L. (2010). Generalized density clustering. *Ann. Statist.* **38** 2678–2722. MR2722453 https://doi.org/10.1214/10-AOS797
[50] Salas, W.A., Boles, S.H., Frolking, S., Xiao, X. and Li, C. (2003). The perimeter/area ratio as an index of misregistration bias in land cover change estimates. *Int. J. Remote Sens.* **24** 1165–1170.
[51] Samworth, R.J. (2012). Optimal weighted nearest neighbour classifiers. *Ann. Statist.* **40** 2733–2763. MR3097618 https://doi.org/10.1214/12-AOS1049
[52] Schmalzing, J. and Górski, K.M. (1998). Minkowski functionals used in the morphological analysis of cosmic microwave background anisotropy maps. *Mon. Not. R. Astron. Soc.* **297** 355–365.
[53] Seifert, U. (1997). Configurations of fluid membranes and vesicles. *Adv. Phys.* **46** 13–137.
[54] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*. London: CRC Press. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9
[55] Singh, A., Scott, C. and Nowak, R. (2009). Adaptive Hausdorff estimation of density level sets. *Ann. Statist.* **37** 2760–2782. MR2541446 https://doi.org/10.1214/08-AOS661
[56] Sriperumbudur, B. and Steinwart, I. (2012). Consistency and rates for clustering with DBSCAN. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*. *PMLR* **22** 1090–1098.
[57] Steinwart, I. (2015). Fully adaptive density-based clustering. *Ann. Statist.* **43** 2132–2167. MR3396981 https://doi.org/10.1214/15-AOS1331
[58] Steinwart, I., Hush, D. and Scovel, C. (2005). A classification framework for anomaly detection. *J. Mach. Learn. Res.* **6** 211–232. MR2249820
[59] Troutt, M.D., Pang, W.K. and Hou, S.H. (2004). *Vertical Density Representation and Its Applications*. River Edge, NJ: World Scientific Co., Inc. MR2051604 https://doi.org/10.1142/9789812562616
[60] Tsybakov, A.B. (1997). On nonparametric estimation of density level sets. *Ann. Statist.* **25** 948–969. MR1447735 https://doi.org/10.1214/aos/1069362732
[61] Walther, G. (1997). Granulometric smoothing. *Ann. Statist.* **25** 2273–2299. MR1604445 https://doi.org/10.1214/aos/1030741072
[62] Weyl, H. (1939). On the volume of tubes. *Amer. J. Math.* **61** 461–472. MR1507388 https://doi.org/10.2307/2371513
[63] Willmore, T.J. (1965). Note on embedded surfaces. *An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat.* (*N.S.*) **11B** 493–496. MR0202066

Bernoulli **27**(1), 2021, 192–217
https://doi.org/10.3150/20-BEJ1233

High-dimensional CLT: Improvements, non-uniform extensions and large deviations

ARUN KUMAR KUCHIBHOTLA[*], SOMABHA MUKHERJEE[†] and DEBAPRATIM BANERJEE[‡]

Department of Statistics, The Wharton School, University of Pennsylvania, 3730 Walnut Street, Jon M. Huntsman Hall, Philadelphia, PA 19104, USA.
E-mail: [*]*arunku@upenn.edu;* [†]*somabha@upenn.edu;* [‡]*dban@upenn.edu*

Central limit theorems (CLTs) for high-dimensional random vectors with dimension possibly growing with the sample size have received a lot of attention in the recent times. Chernozhukov et al. (*Ann. Probab.* **45** (2017) 2309–2352) proved a Berry–Esseen type result for high-dimensional averages for the class of sparsely convex sets including hyperrectangles as a special case and they proved that the rate of convergence can be upper bounded by $n^{-1/6}$ up to a polynomial factor of $\log p$ (where n represents the sample size and p denotes the dimension). Convergence to zero of the bound requires $\log^7 p = o(n)$. We improve upon their result, for hyperrectangles, which only requires $\log^4 p = o(n)$ (in the best case). This improvement is made possible by a sharper dimension-free anti-concentration inequality for Gaussian process on a compact metric space. In addition, we prove two non-uniform variants of the high-dimensional CLT based on the large deviation and non-uniform CLT results for random variables in a Banach space by Bentkus, Račkauskas, and Paulauskas. We apply our results in the context of post-selection inference in linear regression and of empirical processes.

Keywords: anti-concentration; Cramér type large deviation; empirical processes; nonuniform CLT; Orlicz norms; post-selection inference

References

[1] Bachoc, F., Blanchard, G. and Neuvial, P. (2018). On the post selection inference constant under restricted isometry properties. *Electron. J. Stat.* **12** 3736–3757. MR3878579 https://doi.org/10.1214/18-ejs1490
[2] Banerjee, D., Kuchibhotla, A.K. and Mukherjee, S. (2018). Cramér-type large deviation and non-uniform central limit theorems in high dimensions. Preprint. Available at arXiv:1806.06153v1.
[3] Baraud, Y. (2016). Bounding the expectation of the supremum of an empirical process over a (weak) VC-major class. *Electron. J. Stat.* **10** 1709–1728. MR3522658 https://doi.org/10.1214/15-EJS1055
[4] Belloni, A., Chernozhukov, V., Chetverikov, D., Hansen, C. and Kato, K. (2018). High-dimensional econometrics and regularized GMM. Preprint. Available at arXiv:1806.01888.
[5] Bentkus, V. (1990). Smooth approximations of the norm and differentiable functions with bounded support in Banach space l_∞^k. *Lith. Math. J.* **30** 223–230.
[6] Bentkus, V. (2004). A Lyapunov type bound in $\mathbf{R}^d$. *Teor. Veroyatn. Primen.* **49** 400–410. MR2144310 https://doi.org/10.1137/S0040585X97981123
[7] Bentkus, V., Götze, F., Paulauskas, V. and Račkauskas, A. (2000). The accuracy of Gaussian approximation in Banach spaces. In *Limit Theorems of Probability Theory* 25–111. Berlin: Springer.
[8] Bentkus, V. and Račkauskas, A. (1990). On probabilities of large deviations in Banach spaces. *Probab. Theory Related Fields* **86** 131–154. MR1065276 https://doi.org/10.1007/BF01474639
[9] Bentkus, V.Y. (1987). Large deviations in Banach spaces. *Theory Probab. Appl.* **31** 627–632.
[10] Bentkus, V.Yu. (1985). Lower bounds for the rate of convergence in the central limit theorem in Banach spaces. *Liet. Mat. Rink.* **25** 10–21. MR0823198
[11] Berk, R., Brown, L., Buja, A., Zhang, K. and Zhao, L. (2013). Valid post-selection inference. *Ann. Statist.* **41** 802–837. MR3099122 https://doi.org/10.1214/12-AOS1077

1350-7265

[12] 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
[13] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42** 1564–1597. MR3262461 https://doi.org/10.1214/14-AOS1230
[14] Chernozhukov, V., Chetverikov, D. and Kato, K. (2015). Comparison and anti-concentration bounds for maxima of Gaussian random vectors. *Probab. Theory Related Fields* **162** 47–70. MR3350040 https://doi.org/10.1007/s00440-014-0565-9
[15] Chernozhukov, V., Chetverikov, D. and Kato, K. (2017). Detailed proof of Nazarov's inequality. Preprint. Available at arXiv:1711.10696.
[16] 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
[17] Chernozhukov, V., Chetverikov, D., Kato, K. and Koike, Y. (2019). Improved central limit theorem and bootstrap approximations in high dimensions. Preprint. Available at arXiv:1912.10529.
[18] Giné, E. (1976). Bounds for the speed of convergence in the central limit theorem in $C(S)$. *Z. Wahrsch. Verw. Gebiete* **36** 317–331. MR0418190 https://doi.org/10.1007/BF00532697
[19] Han, Q. (2019). Global empirical risk minimizers with "shape constraints" are rate optimal in general dimensions. Preprint. Available at arXiv:1905.12823.
[20] Han, Q. and Wellner, J.A. (2019). Convergence rates of least squares regression estimators with heavy-tailed errors. *Ann. Statist.* **47** 2286–2319. MR3953452 https://doi.org/10.1214/18-AOS1748
[21] Koike, Y. (2019). High-dimensional central limit theorems for homogeneous sums. Preprint. Available at arXiv:1902.03809.
[22] Koike, Y. (2019). Notes on the dimension dependence in high-dimensional central limit theorems for hyperrectangles. Preprint. Available at arXiv:1911.00160.
[23] Kuchibhotla, A.K., Brown, L.D., Buja, A., George, E.I. and Zhao, L. (2018). A model free perspective for linear regression: Uniform-in-model bounds for post selection inference. Preprint. Available at arXiv:1802.05801.
[24] Kuchibhotla, A.K., Brown, L.D., Buja, A., George, E.I. and Zhao, L. (2018). Valid post-selection inference in assumption-lean linear regression. Preprint. Available at arXiv:1806.04119.
[25] Kuchibhotla, A.K., Mukherjee, S. and Banerjee, D. (2020). Supplement to "High-dimensional CLT: Improvements, non-uniform extensions and large deviations." https://doi.org/10.3150/20-BEJ1233SUPP
[26] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces*: *Isoperimetry and Processes*. *Classics in Mathematics*. Berlin: Springer. MR2814399
[27] Lopes, M.E., Lin, Z. and Müller, H.-G. (2020). Bootstrapping max statistics in high dimensions: Near-parametric rates under weak variance decay and application to functional and multinomial data. *Ann. Statist.* **48** 1214–1229. MR4102694 https://doi.org/10.1214/19-AOS1844
[28] 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
[29] Norvaiša, R. and Paulauskas, V. (1991). Rate of convergence in the central limit theorem for empirical processes. *J. Theoret. Probab.* **4** 511–534. MR1115160 https://doi.org/10.1007/BF01210322
[30] Paulauskas, V. and Račkauskas, A. (1989). *Approximation Theory in the Central Limit Theorem*: *Exact Results in Banach Spaces*. *Mathematics and Its Applications* (*Soviet Series*) **32**. Dordrecht: Kluwer Academic. MR1015294 https://doi.org/10.1007/978-94-011-7798-6
[31] Paulauskas, V. and Račkauskas, A. (1991). Nonuniform estimates in the central limit theorem in Banach spaces. *Liet. Mat. Rink.* **31** 483–496. MR1162240 https://doi.org/10.1007/BF00973059
[32] Paulauskas, V. and Račkauskas, A. (2012). *Approximation Theory in the Central Limit Theorem*: *Exact Results in Banach Spaces*. *Mathematics and Its Applications* **32**. Berlin: Springer.
[33] Petrov, V.V. (1995). *Limit Theorems of Probability Theory*: *Sequences of Independent Random Variables*. *Oxford Studies in Probability* **4**. New York: The Clarendon Press. MR1353441
[34] Saulis, L. and Statulevičius, V.A. (1991). *Limit Theorems for Large Deviations*. *Mathematics and Its Applications* (*Soviet Series*) **73**. Dordrecht: Kluwer Academic. MR1171883 https://doi.org/10.1007/978-94-011-3530-6

[35] Sazonov, V.V. (1981). *Normal Approximation – Some Recent Advances*. *Lecture Notes in Math*. **879**. Berlin: Springer. MR0643968

[36] Sazonov, V.V. and Ul'yanov, V.V. (1982). On the accuracy of normal approximation. *J. Multivariate Anal.* **12** 371–384. MR0666012 https://doi.org/10.1016/0047-259X(82)90072-0

[37] Sun, Q. (2020). Gaussian approximations for maxima of random vectors under $(2+\iota)$-th moments. *Statist. Probab. Lett*. **158** Art. ID 108523. MR4027763 https://doi.org/10.1016/j.spl.2019.05.022

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

[39] van de Geer, S. and Wainwright, M.J. (2017). On concentration for (regularized) empirical risk minimization. *Sankhya A* **79** 159–200. MR3707417 https://doi.org/10.1007/s13171-017-0111-9

[40] 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

Bernoulli **27**(1), 2021, 218–220
https://doi.org/10.3150/20-BEJ1234

Corrigendum to "Simple simulation of diffusion bridges with application to likelihood inference for diffusions"

MOGENS BLADT[1,*], MARCIN MIDER[2] and MICHAEL SØRENSEN[1,†]

[1]*Dept. of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark.*
E-mail: [*]*bladt@math.ku.dk;* [†]*michael@math.ku.dk*
[2]*Max Planck Institute for Mathematics in the Sciences, Inselstraße 22, 04103 Leipzig, Germany.*
E-mail: marcin.mider@mis.mpg.de

We correct an error in Theorem 2.1 in Bladt and Sørensen (*Bernoulli* **20** (2014) 645–675), where the initial distribution of an auxiliary diffusion process that is used to describe the distribution of the proposed approximate diffusion bridge is wrong. As a consequence, we also correct the pseudo marginal Metropolis-Hastings algorithm that has an exact diffusion bridge as its target distribution. The same auxiliary diffusion plays a central role in the algorithm.

Keywords: diffusion bridge; discretely sampled diffusions; pseudo-marginal MCMC

References

[1] Bladt, M. and Sørensen, M. (2014). Simple simulation of diffusion bridges with application to likelihood inference for diffusions. *Bernoulli* **20** 645–675. MR3178513 https://doi.org/10.3150/12-BEJ501

Bernoulli **27**(1), 2021, 221–248
https://doi.org/10.3150/20-BEJ1235

Precise deviations for Hawkes processes

FUQING GAO[1] and LINGJIONG ZHU[2]

[1]*School of Mathematics and Statistics, Wuhan University, Wuhan 430072, People's Republic of China.*
E-mail: fqgao@whu.edu.cn
[2]*Department of Mathematics, Florida State University, 1017 Academic Way, Tallahassee, FL-32306, United States of America.*
E-mail: zhu@math.fsu.edu

Hawkes process is a class of simple point processes with self-exciting and clustering properties. Hawkes process has been widely applied in finance, neuroscience, social networks, criminology, seismology, and many other fields. In this paper, we study precise deviations for Hawkes processes for large time asymptotics, that strictly extends and improves the existing results in the literature. Numerical illustrations will also be provided.

Keywords: Hawkes processes; large deviations; precise deviations

References

[1] Abergel, F. and Jedidi, A. (2015). Long-time behavior of a Hawkes process-based limit order book. *SIAM J. Financial Math*. **6** 1026–1043. MR3418223 https://doi.org/10.1137/15M1011469
[2] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2015). Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ*. **117** 585–606.
[3] Aït-Sahalia, Y., Laeven, R.J.A. and Pelizzon, L. (2014). Mutual excitation in Eurozone sovereign CDS. *J. Econometrics* **183** 151–167. MR3276385 https://doi.org/10.1016/j.jeconom.2014.05.006
[4] Alfonsi, A. and Blanc, P. (2016). Dynamic optimal execution in a mixed-market-impact Hawkes price model. *Finance Stoch*. **20** 183–218. MR3441291 https://doi.org/10.1007/s00780-015-0282-y
[5] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* **13** 65–77. MR3005350 https://doi.org/10.1080/14697688.2011.647054
[6] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl*. **123** 2475–2499. MR3054533 https://doi.org/10.1016/j.spa.2013.04.007
[7] Bauwens, L. and Hautsch, N. (2009). *Handbook of Financial Time Series: Modelling Financial High Frequency Data Using Point Processes*. Berlin: Springer.
[8] Bordenave, C. and Torrisi, G.L. (2007). Large deviations of Poisson cluster processes. *Stoch. Models* **23** 593–625. MR2362700 https://doi.org/10.1080/15326340701645959
[9] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab*. **24** 1563–1588. MR1411506 https://doi.org/10.1214/aop/1065725193
[10] Chavez-Demoulin, V., Davison, A.C. and McNeil, A.J. (2005). Estimating value-at-risk: A point process approach. *Quant. Finance* **5** 227–234. MR2240248 https://doi.org/10.1080/14697680500039613
[11] Chu, S.C. and Metcalf, F.T. (1967). On Gronwall's inequality. *Proc. Amer. Math. Soc*. **18** 439–440. MR0212529 https://doi.org/10.2307/2035473
[12] Crane, R. and Sornette, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci. USA* **105** 15649–15653. https://doi.org/10.1073/pnas.0803685105
[13] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics (New York)* **38**. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
[14] Errais, E., Giesecke, K. and Goldberg, L.R. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math*. **1** 642–665. MR2719785 https://doi.org/10.1137/090771272

[15] Féray, V., Méliot, P.-L. and Nikeghbali, A. (2016). *Mod-ϕ Convergence*: *Normality Zones and Precise Deviations*. *SpringerBriefs in Probability and Mathematical Statistics*. Cham: Springer. MR3585777 https://doi.org/10.1007/978-3-319-46822-8

[16] Gao, X., Zhou, X. and Zhu, L. (2018). Transform analysis for Hawkes processes with applications in dark pool trading. *Quant*. *Finance* **18** 265–282. MR3750735 https://doi.org/10.1080/14697688.2017.1403151

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

[18] Gao, X. and Zhu, L. (2018). Limit theorems for Markovian Hawkes processes with a large initial intensity. *Stochastic Process*. *Appl*. **128** 3807–3839. MR3860011 https://doi.org/10.1016/j.spa.2017.12.001

[19] Gao, X. and Zhu, L. (2018). Large deviations and applications for Markovian Hawkes processes with a large initial intensity. *Bernoulli* **24** 2875–2905. MR3779705 https://doi.org/10.3150/17-BEJ948

[20] Giot, P. (2005). Market risk models for intraday data. *Eur. J. Finance* **11** 309–324.

[21] Gusto, G. and Schbath, S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes' model. *Stat*. *Appl*. *Genet*. *Mol*. *Biol*. **4** Art. 24, 28. MR2170440 https://doi.org/10.2202/1544-6115.1119

[22] Hainaut, D. (2016). A bivariate Hawkes process for interest rates modelling. *Econ*. *Model*. **57** 180–196.

[23] Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. MR0278410 https://doi.org/10.1093/biomet/58.1.83

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

[25] Jaisson, T. and Rosenbaum, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Ann*. *Appl*. *Probab*. **25** 600–631. MR3313750 https://doi.org/10.1214/14-AAP1005

[26] Jaisson, T. and Rosenbaum, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Ann*. *Appl*. *Probab*. **26** 2860–2882. MR3563196 https://doi.org/10.1214/15-AAP1164

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

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

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

[30] Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Ann*. *Inst*. *Statist*. *Math*. **50** 379–402.

[31] Pernice, V., Staude, B., Cardanobile, S. and Rotter, S. (2011). How structure determines correlations in neuronal networks. *PLoS Comput*. *Biol*. **7** e1002059, 14. MR2821638 https://doi.org/10.1371/journal.pcbi.1002059

[32] Porter, M.D. and White, G. (2012). Self-exciting hurdle models for terrorist activity. *Ann*. *Appl*. *Stat*. **6** 106–124. MR2951531 https://doi.org/10.1214/11-AOAS513

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

[34] Sokol, A. and Hansen, N.R. (2015). Exponential martingales and changes of measure for counting processes. *Stoch*. *Anal*. *Appl*. **33** 823–843. MR3378040 https://doi.org/10.1080/07362994.2015.1040890

[35] Stabile, G. and Torrisi, G.L. (2010). Risk processes with non-stationary Hawkes claims arrivals. *Methodol*. *Comput*. *Appl*. *Probab*. **12** 415–429. MR2665268 https://doi.org/10.1007/s11009-008-9110-6

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

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

[38] Zhang, X., Blanchet, J., Giesecke, K. and Glynn, P.W. (2009). Rare event simulation for a generalized Hawkes process. In *Proceedings of the* 2009 *Winter Simulation Conference*.

[39] Zhang, X., Blanchet, J., Giesecke, K. and Glynn, P.W. (2015). Affine point processes: Approximation and efficient simulation. *Math*. *Oper. Res*. **40** 797–819. MR3423739 https://doi.org/10.1287/moor.2014.0696
[40] Zhao, Q., Erdogdu, M.A., He, H.Y., Rajaraman, A. and Jure, L. (2015). SEISMIC: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the* 21*th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1513–1522.
[41] Zhu, L. (2013). Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance Math*. *Econom*. **53** 544–550. MR3130449 https://doi.org/10.1016/j.insmatheco.2013.08.008
[42] Zhu, L. (2013). Moderate deviations for Hawkes processes. *Statist*. *Probab*. *Lett*. **83** 885–890. MR3040318 https://doi.org/10.1016/j.spl.2012.12.011
[43] Zhu, L. (2013). Central limit theorem for nonlinear Hawkes processes. *J*. *Appl*. *Probab*. **50** 760–771. MR3102513 https://doi.org/10.1239/jap/1378401234
[44] Zhu, L. (2014). Process-level large deviations for nonlinear Hawkes point processes. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **50** 845–871. MR3224291 https://doi.org/10.1214/12-AIHP532
[45] Zhu, L. (2015). Large deviations for Markovian nonlinear Hawkes processes. *Ann*. *Appl*. *Probab*. **25** 548–581. MR3313748 https://doi.org/10.1214/14-AAP1003

Bernoulli **27**(1), 2021, 249–273
https://doi.org/10.3150/20-BEJ1236

Efron–Petrosian integrals for doubly truncated data with covariates: An asymptotic analysis

JACOBO DE UÑA-ÁLVAREZ[1] and INGRID VAN KEILEGOM[2]

[1]*Department of Statistics and OR, SiDOR research group & CINBIO, Universidade de Vigo, Spain.*
E-mail: jacobo@uvigo.es
[2]*Research Centre for Operations Research and Statistics (ORSTAT), KU Leuven, Belgium.*
E-mail: ingrid.vankeilegom@kuleuven.be

In survival analysis, epidemiology and related fields there exists an increasing interest in statistical methods for doubly truncated data. Double truncation appears with interval sampling and other sampling schemes, and refers to situations in which the target variable is subject to two (left and right) random observation limits. Doubly truncated data require specific corrections for the observational bias, and this affects a variety of settings including the estimation of marginal and multivariate distributions, regression problems, and multi-state models. In this work multivariate Efron–Petrosian integrals for doubly truncated data are introduced. These integrals naturally arise when the goal is the estimation of the mean of a general transformation which involves the doubly truncated variable and covariates. An asymptotic representation of the Efron–Petrosian integrals as a sum of i.i.d. terms is derived and, from this, consistency and distributional convergence are established. As a by-product, uniform i.i.d. representations for the marginal nonparametric maximum likelihood estimator and its corresponding weighting process are provided. Applications to correlation analysis, regression, and competing risks models are presented. A simulation study is reported too.

Keywords: Donsker class; double truncation; interval sampling; survival analysis; weak convergence

References

[1] Aalen, O.O. and Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. *Scand. J. Statist.* **5** 141–150. MR0509450

[2] Austin, M.D., Simon, D.K. and Betensky, R.A. (2014). Computationally simple estimation and improved efficiency for special cases of double truncation. *Lifetime Data Anal.* **20** 335–354. MR3217541 https://doi.org/10.1007/s10985-013-9287-z

[3] Bai, Z.D., Rao, C.R. and Wu, Y. (1992). M-estimation of multivariate linear regression parameters under a convex discrepancy function. *Statist. Sinica* **2** 237–254. MR1152307

[4] de Uña-Álvarez, J. (2003). Large sample results under biased sampling when covariables are present. *Statist. Probab. Lett.* **63** 287–293. MR1986328 https://doi.org/10.1016/S0167-7152(03)00093-2

[5] de Uña-Álvarez, J. (2020). Nonparametric estimation of the cumulative incidences of competing risks under double truncation. *Biom J.* **62** 852–867. https://doi.org/10.1002/bimj.201800323

[6] Du, J., Zhang, Z. and Xie, T. (2018). A weighted M-estimator for linear regression models with randomly truncated data. *Statist. Probab. Lett.* **138** 90–94. MR3788723 https://doi.org/10.1016/j.spl.2018.02.055

[7] Efron, B. and Petrosian, V. (1999). Nonparametric methods for doubly truncated data. *J. Amer. Statist. Assoc.* **94** 824–834. MR1723343 https://doi.org/10.2307/2669997

[8] Emura, T., Hu, Y.-H. and Konno, Y. (2017). Asymptotic inference for maximum likelihood estimators under the special exponential family with double-truncation. *Statist. Papers* **58** 877–909. MR3686855 https://doi.org/10.1007/s00362-015-0730-y

[9] Frank, G., Chae, M. and Kim, Y. (2019). Additive time-dependent hazard model with doubly truncated data. *J. Korean Statist. Soc.* **48** 179–193. MR3958707 https://doi.org/10.1016/j.jkss.2018.10.005

1350-7265

[10] Gürler, Ü. (1996). Bivariate estimation with right-truncated data. *J. Amer. Statist. Assoc.* **91** 1152–1165. MR1424615 https://doi.org/10.2307/2291734
[11] Gürler, Ü., Stute, W. and Wang, J.-L. (1993). Weak and strong quantile representations for randomly truncated data with applications. *Statist. Probab. Lett.* **17** 139–148. MR1223536 https://doi.org/10.1016/0167-7152(93)90008-7
[12] Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.* **40** 633–643. MR0238419 https://doi.org/10.1214/aoms/1177697731
[13] Kantorovich, L.V. and Akilov, G.P. (1982). *Functional Analysis*, 2nd ed. Oxford-Elmsford, NY: Pergamon Press. MR0664597
[14] Mandel, M., de Uña-Álvarez, J., Simon, D.K. and Betensky, R.A. (2018). Inverse probability weighted Cox regression for doubly truncated data. *Biometrics* **74** 481–487. MR3825334 https://doi.org/10.1111/biom.12771
[15] Moreira, C. and de Uña-Álvarez, J. (2010). Bootstrapping the NPMLE for doubly truncated data. *J. Nonparametr. Stat.* **22** 567–583. MR2682208 https://doi.org/10.1080/10485250903556102
[16] Moreira, C. and de Uña-Álvarez, J. (2010). A semiparametric estimator of survival for doubly truncated data. *Stat. Med.* **29** 3147–3159. MR2758709 https://doi.org/10.1002/sim.3938
[17] Moreira, C., de Uña-Álvarez, J. and Braekers, R. (2018). Nonparametric estimation of a distribution function from doubly truncated data under dependence. Available at arXiv:1802.08579v1.
[18] Moreira, C., de Uña-Álvarez, J. and Crujeiras, R. (2010). DTDA: An R package to analyze randomly truncated data. *J. Stat. Softw.* **37** 1–20.
[19] Moreira, C., de Uña-Álvarez, J. and Van Keilegom, I. (2014). Goodness-of-fit tests for a semiparametric model under random double truncation. *Comput. Statist.* **29** 1365–1379. MR3266063 https://doi.org/10.1007/s00180-014-0496-z
[20] Pollard, D. (1991). Asymptotics for least absolute deviation regression estimators. *Econometric Theory* **7** 186–199. MR1128411 https://doi.org/10.1017/S0266466600004394
[21] Rennert, L. and Xie, S.X. (2018). Cox regression model with doubly truncated data. *Biometrics* **74** 725–733. MR3825359 https://doi.org/10.1111/biom.12809
[22] Sánchez Sellero, C., González Manteiga, W. and Van Keilegom, I. (2005). Uniform representation of product-limit integrals with applications. *Scand. J. Stat.* **32** 563–581. MR2232343 https://doi.org/10.1111/j.1467-9469.2005.00453.x
[23] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. MR0595165
[24] Shen, P. (2010). Nonparametric analysis of doubly truncated data. *Ann. Inst. Statist. Math.* **62** 835–853. MR2669740 https://doi.org/10.1007/s10463-008-0192-2
[25] Shen, P. and Liu, Y. (2019). Pseudo maximum likelihood estimation for the Cox model with doubly truncated data. *Statist. Papers* **60** 1207–1224. MR4008684 https://doi.org/10.1007/s00362-016-0870-8
[26] Shen, P.-S. (2010). Semiparametric analysis of doubly truncated data. *Comm. Statist. Theory Methods* **39** 3178–3190. MR2755432 https://doi.org/10.1080/03610920903219272
[27] Stute, W. (1993). Almost sure representations of the product-limit estimator for truncated data. *Ann. Statist.* **21** 146–156. MR1212170 https://doi.org/10.1214/aos/1176349019
[28] Stute, W. (1999). Nonlinear censored regression. *Statist. Sinica* **9** 1089–1102. MR1744826
[29] van der Laan, M.J. (1996). Nonparametric estimation of the bivariate survival function with truncated data. *J. Multivariate Anal.* **58** 107–131. MR1408131 https://doi.org/10.1006/jmva.1996.0042
[30] 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
[31] 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
[32] van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942 https://doi.org/10.1214/074921707000000382
[33] Woodroofe, M. (1985). Estimating a distribution function with truncated data. *Ann. Statist.* **13** 163–177. MR0773160 https://doi.org/10.1214/aos/1176346584

[34] Xiao, J. and Hudgens, M.G. (2019). On nonparametric maximum likelihood estimation with double truncation. *Biometrika* **106** 989–996. MR4031212 https://doi.org/10.1093/biomet/asz038
[35] Ying, Z., Yu, W., Zhao, Z. and Zheng, M. (2020). Regression analysis of doubly truncated data. *J. Amer. Statist. Assoc.* **115** 810–821. MR4107681 https://doi.org/10.1080/01621459.2019.1585252
[36] Zhu, H. and Wang, M.-C. (2012). Analysing bivariate survival data with interval sampling and application to cancer epidemiology. *Biometrika* **99** 345–361. MR2931258 https://doi.org/10.1093/biomet/ass009
[37] Zhu, H. and Wang, M.-C. (2014). Nonparametric inference on bivariate survival data with interval sampling: Association estimation and testing. *Biometrika* **101** 519–533. MR3254898 https://doi.org/10.1093/biomet/asu005

Bernoulli **27**(1), 2021, 274–294
https://doi.org/10.3150/20-BEJ1237

Generalized four moment theorem and an application to CLT for spiked eigenvalues of high-dimensional covariance matrices

DANDAN JIANG[1] and ZHIDONG BAI[2]

[1]*School of Mathematics and Statistics, Xi'an Jiaotong University, No.28, Xianning West Road, Xi'an 710049, China.*
E-mail: *jiangdd@xjtu.edu.cn*
[2]*KLASMOE and School of Mathematics and Statistics, Northeast Normal University, No. 5268 People's Street, Changchun 130024, China.*
E-mail: *baizd@nenu.edu.cn*

We consider a more generalized spiked covariance matrix, which is a general non-negative definite matrix with the spiked eigenvalues scattered into spaces of a few bulks and the largest ones allowed to tend to infinity. The study is split into two cases by whether the maximum absolute value of the eigenvector of the corresponding spikes tends to zero or not. On one hand, if it is zero, a Generalized Four Moment Theorem (G4MT) is proposed by relaxing the matching of the 3rd and the 4th moment to the tail probability decaying with certain rate, which shows the universality of the asymptotic law for the spiked eigenvalues of the generalized spiked covariance model. On the other hand, if it is not zero, the matches of the third and fourth moments in usual four moment theorem are weakened to only requiring the match of the 4th moment. Moreover, by applying the results to the Central Limit Theorem (CLT) for the spiked eigenvalues of the generalized spiked covariance model, we successively remove the restrictive condition of block wise diagonal assumption on the population covariance matrix in the previous works. This condition implies an unrealistic fact that the spiked eigenvalues and bulked eigenvalues are generated by independent variables, respectively. Thus, the new CLT will have much better application domain.

Keywords: central limit theorem; generalized four moment theorem; high-dimensional covariance matrix; random matrix theory; spiked model

References

[1] Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* **70** 191–221. MR1926259 https://doi.org/10.1111/1468-0262.00273
[2] Bai, Z., Choi, K.P. and Fujikoshi, Y. (2018). Consistency of AIC and BIC in estimating the number of significant components in high-dimensional principal component analysis. *Ann. Statist.* **46** 1050–1076. MR3797996 https://doi.org/10.1214/17-AOS1577
[3] Bai, Z. and Ding, X. (2012). Estimation of spiked eigenvalues in spiked models. *Random Matrices Theory Appl.* **1** 1150011, 21. MR2934717 https://doi.org/10.1142/S2010326311500110
[4] Bai, Z. and Yao, J. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474. MR2451053 https://doi.org/10.1214/07-AIHP118
[5] Bai, Z. and Yao, J. (2012). On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.* **106** 167–177. MR2887686 https://doi.org/10.1016/j.jmva.2011.10.009
[6] Bai, Z.D., Miao, B.Q., Rao and Radbakrisbna, C. (1991). Estimation of directions of arrival of signals: Asymptotic results. In *Advances in Spectrum Analysis and Array Processing*, *Vol. I* 327–347.
[7] Bai, Z.D. and Silverstein, J.W. (1999). Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.* **27** 1536–1555. MR1733159 https://doi.org/10.1214/aop/1022677458

[8] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845
[9] 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
[10] Baik, J. and Silverstein, J.W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680 https://doi.org/10.1016/j.jmva.2005.08.003
[11] Ben Arous, G. and Péché, S. (2005). Universality of local eigenvalue statistics for some sample covariance matrices. *Comm. Pure Appl. Math.* **58** 1316–1357. MR2162782 https://doi.org/10.1002/cpa.20070
[12] Cai, T.T., Han, X. and Pan, G.M. (2019). Limiting laws for divergent spiked eigenvalues and largest non-spiked eigenvalue of sample covariance matrices. *Ann. Statist.* To appear. http://arxiv.org/abs/1711.00217v2.
[13] Dyson, F.J. (1970). Correlations between eigenvalues of a random matrix. *Comm. Math. Phys.* **19** 235–250. MR0278668
[14] Erdős, L., Péché, S., Ramírez, J.A., Schlein, B. and Yau, H.-T. (2010). Bulk universality for Wigner matrices. *Comm. Pure Appl. Math.* **63** 895–925. MR2662426 https://doi.org/10.1002/cpa.20317
[15] Erdős, L., Ramírez, J.A., Schlein, B. and Yau, H.-T. (2010). Universality of sine-kernel for Wigner matrices with a small Gaussian perturbation. *Electron. J. Probab.* **15** 526–603. MR2639734 https://doi.org/10.1214/EJP.v15-768
[16] Hoyle, D.C. and Rattray, M. (2004). Principal-component-analysis eigenvalue spectra from data with symmetry-breaking structure. *Phys. Rev. E* **69** 026124.
[17] Hu, J. and Bai, Z.D. (2014). Estimation of directions of arrival of signals: Asymptotic results. *Sci. China Math.* **57**. https://doi.org/10.1007/s11425-014-4855-6
[18] Jiang, D. and Bai, Z. (2020). Supplement to "Generalized four moment theorem and an application to CLT for spiked eigenvalues of high-dimensional covariance matrices." https://doi.org/10.3150/20-BEJ1237SUPP
[19] Johansson, K. (2001). Universality of the local spacing distribution in certain ensembles of Hermitian Wigner matrices. *Comm. Math. Phys.* **215** 683–705. MR1810949 https://doi.org/10.1007/s002200000328
[20] Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961 https://doi.org/10.1214/aos/1009210544
[21] Mehta, M.L. (1967). *Random Matrices and the Statistical Theory of Energy Levels*. New York: Academic Press. MR0220494
[22] Onatski, A. (2009). Testing hypotheses about the numbers of factors in large factor models. *Econometrica* **77** 1447–1479. MR2561070 https://doi.org/10.3982/ECTA6964
[23] Passemier, D. and Yao, J.-F. (2012). On determining the number of spikes in a high-dimensional spiked population model. *Random Matrices Theory Appl.* **1** 1150002, 19. MR2930380 https://doi.org/10.1142/S201032631150002X
[24] Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica* **17** 1617–1642. MR2399865
[25] Skorohod, A.V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290. MR0084897
[26] Soshnikov, A. (1999). Universality at the edge of the spectrum in Wigner random matrices. *Comm. Math. Phys.* **207** 697–733. MR1727234 https://doi.org/10.1007/s002200050743
[27] Tao, T. and Vu, V. (2015). Random matrices: Universality of local spectral statistics of non-Hermitian matrices. *Ann. Probab.* **43** 782–874. MR3306005 https://doi.org/10.1214/13-AOP876
[28] Wang, W. and Fan, J. (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. *Ann. Statist.* **45** 1342–1374. MR3662457 https://doi.org/10.1214/16-AOS1487
[29] Wigner, E.P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. of Math.* (2) **67** 325–327. MR0095527 https://doi.org/10.2307/1970008

Bernoulli **27**(1), 2021, 295–311
https://doi.org/10.3150/20-BEJ1240

The Osgood condition for stochastic partial differential equations

MOHAMMUD FOONDUN[1] and EULALIA NUALART[2]

[1] *University of Strathclyde, Department of Mathematics and Statistics, 26 Richmond Street, Glasgow G1 1XH, Scotland.*
E-mail: mohammud.foondun@strath.ac.uk
[2] *Universitat Pompeu Fabra and Barcelona Graduate School of Economics, Department of Economics and Business, Ramón Trias Fargas 25-27, 08005 Barcelona, Spain.*
E-mail: eulalia.nualart@upf.edu

We study the following equation

$$\frac{\partial u(t,x)}{\partial t} = \Delta u(t,x) + b\big(u(t,x)\big) + \sigma \dot{W}(t,x), \quad t > 0,$$

where σ is a positive constant and $\dot{W}$ is a space–time white noise. The initial condition $u(0,x) = u_0(x)$ is assumed to be a nonnegative and continuous function. We first study the problem on $[0,1]$ with homogeneous Dirichlet boundary conditions. Under some suitable conditions, together with a theorem of Bonder and Groisman in (*Phys. D* **238** (2009) 209–215), our first result shows that the solution blows up in finite time if and only if for some $a > 0$,

$$\int_a^\infty \frac{1}{b(s)}\,\mathrm{d}s < \infty,$$

which is the well-known Osgood condition. We also consider the same equation on the whole line and show that the above condition is sufficient for the nonexistence of global solutions. Various other extensions are provided; we look at equations with fractional Laplacian and spatial colored noise in $\mathbf{R}^d$.

Keywords: fractional stochastic heat equation; space–time white noise; spatial colored noise

References

[1] Bonder, J.F. and Groisman, P. (2009). Time–space white noise eliminates global solutions in reaction-diffusion equations. *Phys. D* **238** 209–215. MR2516340 https://doi.org/10.1016/j.physd.2008.09.005
[2] Chen, Z.-Q. and Song, R. (1997). Intrinsic ultracontractivity and conditional gauge for symmetric stable processes. *J. Funct. Anal.* **150** 204–239. MR1473631 https://doi.org/10.1006/jfan.1997.3104
[3] Dalang, R., Khoshnevisan, D., Mueller, C., Nualart, D. and Xiao, Y. (2009). *A Minicourse on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1962**. Berlin: Springer. Held at the University of Utah, Salt Lake City, UT, May 8–19, 2006, Edited by Khoshnevisan and Firas Rassoul-Agha. MR1500166
[4] Dalang, R.C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** no. 6, 29 pp. MR1684157 https://doi.org/10.1214/EJP.v4-43
[5] Dalang, R.C., Khoshnevisan, D. and Nualart, E. (2007). Hitting probabilities for systems of non-linear stochastic heat equations with additive noise. *ALEA Lat. Am. J. Probab. Math. Stat.* **3** 231–271. MR2365643
[6] Dalang, R.C., Khoshnevisan, D. and Zhang, T. (2019). Global solutions to stochastic reaction-diffusion equations with super-linear drift and multiplicative noise. *Ann. Probab.* **47** 519–559. MR3909975 https://doi.org/10.1214/18-AOP1270
[7] Fujita, H. (1966). On the blowing up of solutions of the Cauchy problem for $u_t = \Delta u + u^{1+\alpha}$. *J. Fac. Sci., Univ. Tokyo, Sect. 1A, Math.* **13** 109–124. MR0214914

[8] Hayakawa, K. (1973). On nonexistence of global solutions of some semilinear parabolic differential equations. *Proc*. *Japan Acad*. *Ser. A Math*. *Sci*. **49** 503–505. MR0338569

[9] Houdré, C. and Villa, J. (2003). An example of infinite dimensional quasi-helix. In *Stochastic Models* (*Mexico City*, 2002). *Contemp*. *Math*. **336** 195–201. Providence, RI: Amer. Math. Soc. MR2037165 https://doi.org/10.1090/conm/336/06034

[10] 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

[11] Kobayashi, K., Sirao, T. and Tanaka, H. (1977). On the growing up problem for semilinear heat equations. *J*. *Math*. *Soc*. *Japan* **29** 407–424. MR0450783 https://doi.org/10.2969/jmsj/02930407

[12] Kolokoltsov, V. (2000). Symmetric stable laws and stable-like jump-diffusions. *Proc*. *Lond*. *Math*. *Soc*. (3) **80** 725–768. MR1744782 https://doi.org/10.1112/S0024611500012314

[13] Lei, P. and Nualart, D. (2009). A decomposition of the bifractional Brownian motion and some applications. *Statist*. *Probab*. *Lett*. **79** 619–624. MR2499385 https://doi.org/10.1016/j.spl.2008.10.009

[14] León, J.A. and Villa, J. (2011). An Osgood criterion for integral equations with applications to stochastic differential equations with an additive noise. *Statist*. *Probab*. *Lett*. **81** 470–477. MR2765167 https://doi.org/10.1016/j.spl.2010.12.001

[15] Quittner, P. and Souplet, P. (2007). *Superlinear Parabolic Problems*: *Blow-up*, *Global Existence and Steady States*. *Birkhäuser Advanced Texts*: *Basler Lehrbücher*. [*Birkhäuser Advanced Texts*: *Basel Textbooks*]. Basel: Birkhäuser. MR2346798

[16] Sanz-Solé, M. and Sarrà, M. (2002). Hölder continuity for the stochastic heat equation with spatially correlated noise. In *Seminar on Stochastic Analysis*, *Random Fields and Applications*, *III* (*Ascona*, 1999). *Progress in Probability* **52** 259–268. Basel: Birkhäuser. MR1958822

[17] Vazquez, J.L. (1999). Domain of existence and blowup for the exponential reaction-diffusion equation. *Indiana Univ. Math*. *J*. **48** 677–709. MR1722813 https://doi.org/10.1512/iumj.1999.48.1581

[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

Bernoulli **27**(1), 2021, 312–347
https://doi.org/10.3150/20-BEJ1241

On the weak convergence rate of an exponential Euler scheme for SDEs governed by coefficients with superlinear growth

MIREILLE BOSSY[1], JEAN-FRANÇOIS JABIR[2,3] and KERLYNS MARTÍNEZ[1,4]

[1] *Université Côte d'Azur, Inria, France.*
E-mail: mireille.bossy@inria.fr
[2] *School of Mathematics, University of Edinburgh, United Kingdom.*
E-mail: jjabir@ed.ac.uk
[3] *National Research University Higher School of Economics, Moscow, Russia.*
E-mail: jjabir@hse.ru
[4] *Instituto de Ingeniería Matemática and CIMFAV, Facultad de Ingeniería, Universidad de Valparaíso, Valparaíso, Chile.*
E-mail: kerlynsmartinez@gmail.com

We consider the problem of designing robust numerical integration scheme of the solution of a one-dimensional SDE with non-globally Lipschitz drift and diffusion coefficients behaving as x^α, with $\alpha > 1$. We propose an (semi-explicit) exponential-Euler scheme for which we obtain a theoretical convergence rate for the weak error. To this aim, we analyze the $C^{1,4}$ regularity of the solution of the associated backward Kolmogorov PDE using its Feynman–Kac representation and the flow derivative of the involved processes. Under some suitable hypotheses on the parameters of the model, we prove a rate of weak convergence of order one for the proposed exponential Euler scheme, and illustrate it with some numerical experiments.

Keywords: Feynman–Kac representation; numerical scheme; polynomial coefficients; rate of convergence; stochastic differential equation; weak convergence

References

[1] Ait-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate. *Rev. Financ. Stud.* **9** 385–426.

[2] Alfonsi, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11** 355–384. MR2186814 https://doi.org/10.1163/156939605777438569

[3] Beccari, M., Hutzenthaler, M., Jentzen, A., Kurniawan, R., Lindner, F. and Salimova, D. Strong and weak divergence of exponential and linear-implicit Euler approximations for stochastic partial differential equations with superlinearly growing nonlinearities. Available at arXiv:1903.06066.

[4] Bossy, M. and Diop, A. (2010). Weak convergence analysis of the symmetrized Euler scheme for one dimensional SDEs with diffusion coefficient $|x|^\alpha$, $\alpha \in [\frac{1}{2}, 1)$. Inria Research report no. 5396. Available at arxiv:1508.04573.

[5] Bossy, M., Jabir, J.-F. and Martínez, K. (2020). Supplement to "On the weak convergence rate of an exponential Euler scheme for SDEs governed by coefficients with superlinear growth." https://doi.org/10.3150/20-BEJ1241SUPP.

[6] Bossy, M. and Olivero, H. (2018). Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs. *Bernoulli* **24** 1995–2042. MR3757521 https://doi.org/10.3150/16-BEJ918

[7] Chan, K.C., Karolyi, G.A., Longstaff, F.A. and Sanders, A.B. (1992). An empirical investigation of alternative models of the short-term interest rate. *J. Finance* **47** 1209–1227.

[8] Chassagneux, J.-F., Jacquier, A. and Mihaylov, I. (2016). An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *SIAM J. Financial Math.* **7** 993–1021. MR3582409 https://doi.org/10.1137/15M1017788

[9] Cox, J.C., Ingersoll, J.E. Jr. and Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* **53** 385–407. MR0785475 https://doi.org/10.2307/1911242

[10] Delbaen, F. and Shirakawa, H. (2002). A note on option pricing for the constant elasticity of variance model. *Asia-Pac. Financ. Mark.* **9** 85–99.

[11] Dereich, S., Neuenkirch, A. and Szpruch, L. (2012). An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **468** 1105–1115. MR2898556 https://doi.org/10.1098/rspa.2011.0505

[12] Friedman, A. (1975). *Stochastic Differential Equations and Applications. Vol.* 1. New York: Academic Press. MR0494490

[13] Gyöngy, I. (1998). A note on Euler's approximations. *Potential Anal.* **8** 205–216. MR1625576 https://doi.org/10.1023/A:1008605221617

[14] Higham, D.J., Mao, X. and Stuart, A.M. (2002). Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40** 1041–1063. MR1949404 https://doi.org/10.1137/S0036142901389530

[15] Hochbruck, M. and Ostermann, A. (2010). Exponential integrators. *Acta Numer.* **19** 209–286. MR2652783 https://doi.org/10.1017/S0962492910000048

[16] Hutzenthaler, M. and Jentzen, A. (2020). On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with nonglobally monotone coefficients. *Ann. Probab.* **48** 53–93. MR4079431 https://doi.org/10.1214/19-AOP1345

[17] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E. (2011). Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** 1563–1576. MR2795791 https://doi.org/10.1098/rspa.2010.0348

[18] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E. (2012). Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** 1611–1641. MR2985171 https://doi.org/10.1214/11-AAP803

[19] Karatzas, I. and Shreve, S.E. (1988). *Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics* **113**. New York: Springer. MR0917065 https://doi.org/10.1007/978-1-4684-0302-2

[20] Lamba, H., Mattingly, J.C. and Stuart, A.M. (2007). An adaptive Euler–Maruyama scheme for SDEs: Convergence and stability. *IMA J. Numer. Anal.* **27** 479–506. MR2337577 https://doi.org/10.1093/imanum/drl032

[21] Martínez, K. (2019). Penalized stochastic optimal control problems for singular McKean–Vlasov Dynamics and turbulent kinetic energy modeling with calibration on lagrangian turbulent flow models. Ph.D. thesis in Mathematics, Pontificia Universidad Católica de Valparaíso, Universidad Técnica Federico Santa María and Universidad de Valparaíso. August 2019. https://doi.org/10.13140/RG.2.2.17285.09446

[22] Milstein, G.N. and Tretyakov, M.V. (2005). Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients. *SIAM J. Numer. Anal.* **43** 1139–1154. MR2177799 https://doi.org/10.1137/040612026

[23] Pope, D.A. (1963). An exponential method of numerical integration of ordinary differential equations. *Commun. ACM* **6** 491–493. MR0153118 https://doi.org/10.1145/366707.367592

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

[25] Sabanis, S. (2016). Euler approximations with varying coefficients: The case of superlinearly growing diffusion coefficients. *Ann. Appl. Probab.* **26** 2083–2105. MR3543890 https://doi.org/10.1214/15-AAP1140

[26] Szpruch, L., Mao, X., Higham, D.J. and Pan, J. (2011). Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. *BIT* **51** 405–425. MR2806537 https://doi.org/10.1007/s10543-010-0288-y

[27] Yan, L. (2002). The Euler scheme with irregular coefficients. *Ann. Probab.* **30** 1172–1194. MR1920104 https://doi.org/10.1214/aop/1029867124

Bernoulli **27**(1), 2021, 348–380
https://doi.org/10.3150/20-BEJ1242

On sub-geometric ergodicity of diffusion processes

PETRA LAZIĆ* and NIKOLA SANDRIĆ†

Department of Mathematics, University of Zagreb, 10000 Zagreb, Croatia.
E-mail: *petralaz@math.hr; †nsandric@math.hr

In this article, we discuss ergodicity properties of a diffusion process given through an Itô stochastic differential equation. We identify conditions on the drift and diffusion coefficients which result in sub-geometric ergodicity of the corresponding semigroup with respect to the total variation distance. We also prove sub-geometric contractivity and ergodicity of the semigroup under a class of Wasserstein distances. Finally, we discuss sub-geometric ergodicity of two classes of Markov processes with jumps.

Keywords: asymptotic flatness; diffusion process; sub-geometric ergodicity; total variation distance; Wasserstein distance

References

[1] Abourashchi, N. and Veretennikov, A.Y. (2010). On stochastic averaging and mixing. *Theory Stoch. Process.* **16** 111–129. MR2779833

[2] Albeverio, S., Brzeźniak, Z. and Wu, J.-L. (2010). Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients. *J. Math. Anal. Appl.* **371** 309–322. MR2661009 https://doi.org/10.1016/j.jmaa.2010.05.039

[3] Arapostathis, A., Borkar, V.S. and Ghosh, M.K. (2012). *Ergodic Control of Diffusion Processes*. *Encyclopedia of Mathematics and Its Applications* **143**. Cambridge: Cambridge Univ. Press. MR2884272

[4] Arapostathis, A., Pang, G. and Sandrić, N. (2019). Ergodicity of a Lévy-driven SDE arising from multiclass many-server queues. *Ann. Appl. Probab.* **29** 1070–1126. MR3910024 https://doi.org/10.1214/18-AAP1430

[5] Bass, R.F. and Cranston, M. (1986). The Malliavin calculus for pure jump processes and applications to local time. *Ann. Probab.* **14** 490–532. MR0832021

[6] Bhattacharya, R.N. (1978). Criteria for recurrence and existence of invariant measures for multidimensional diffusions. *Ann. Probab.* **6** 541–553. MR0494525

[7] Blumenthal, R.M. and Getoor, R.K. (1968). *Markov Processes and Potential Theory*. *Pure and Applied Mathematics*, *Vol*. 29. New York: Academic Press. MR0264757

[8] Bolley, F., Gentil, I. and Guillin, A. (2012). Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations. *J. Funct. Anal.* **263** 2430–2457. MR2964689 https://doi.org/10.1016/j.jfa.2012.07.007

[9] Böttcher, B., Schilling, R. and Wang, J. (2013). *Lévy Matters*. *III*. *Lecture Notes in Math.* **2099**. Cham: Springer. MR3156646 https://doi.org/10.1007/978-3-319-02684-8

[10] Butkovsky, O. (2014). Subgeometric rates of convergence of Markov processes in the Wasserstein metric. *Ann. Appl. Probab.* **24** 526–552. MR3178490 https://doi.org/10.1214/13-AAP922

[11] Chen, M. (2000). Explicit bounds of the first eigenvalue. *Sci. China Ser. A* **43** 1051–1059. MR1802148 https://doi.org/10.1007/BF02898239

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

[13] Deng, C.-S., Schilling, R.L. and Song, Y.-H. (2017). Subgeometric rates of convergence for Markov processes under subordination. *Adv. in Appl. Probab.* **49** 162–181. MR3631220 https://doi.org/10.1017/apr.2016.83

[14] Deng, C.-S., Schilling, R.L. and Song, Y.-H. (2018). Correction: "Subgeometric rates of convergence for Markov processes under subordination" [MR3631220]. *Adv. in Appl. Probab.* **50** 1005. MR3877262 https://doi.org/10.1017/apr.2018.44

[15] Douc, R., Fort, G. and Guillin, A. (2009). Subgeometric rates of convergence of f-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** 897–923. MR2499863 https://doi.org/10.1016/j.spa.2008.03.007

[16] Douc, R., Fort, G., Moulines, E. and Soulier, P. (2004). Practical drift conditions for subgeometric rates of convergence. *Ann. Appl. Probab.* **14** 1353–1377. MR2071426 https://doi.org/10.1214/105051604000000323

[17] Douc, R., Moulines, E., Priouret, P. and Soulier, P. (2018). *Markov Chains. Springer Series in Operations Research and Financial Engineering*. Cham: Springer. MR3889011 https://doi.org/10.1007/978-3-319-97704-1

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

[19] Durmus, A., Fort, G. and Moulines, É. (2016). Subgeometric rates of convergence in Wasserstein distance for Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1799–1822. MR3573296 https://doi.org/10.1214/15-AIHP699

[20] Durrett, R. (1996). *Stochastic Calculus. Probability and Stochastics Series*. Boca Raton, FL: CRC Press. MR1398879

[21] Eberle, A. (2011). Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris* **349** 1101–1104. MR2843007 https://doi.org/10.1016/j.crma.2011.09.003

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

[23] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658

[24] Fort, G. and Moulines, E. (2003). Polynomial ergodicity of Markov transition kernels. *Stochastic Process. Appl.* **103** 57–99. MR1947960 https://doi.org/10.1016/S0304-4149(02)00182-5

[25] Fort, G. and Roberts, G.O. (2005). Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15** 1565–1589. MR2134115 https://doi.org/10.1214/105051605000000115

[26] Friedman, A. (1975). *Stochastic Differential Equations and Applications. Vol.* 1. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0494490

[27] Hairer, M. (2016). *Convergence of Markov Processes. Lecture Notes, University of Warwick*. Available at http://www.hairer.org/notes/Convergence.pdf.

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

[29] Has'minskiĭ, R.Z. (1960). Ergodic properties of recurrent diffusion processes and stabilization of the solution of the Cauchy problem for parabolic equations. *Teor. Veroyatn. Primen.* **5** 196–214. MR0133871

[30] Ishikawa, Y. (2001). Density estimate in small time for jump processes with singular Lévy measures. *Tohoku Math. J.* (2) **53** 183–202. MR1829978 https://doi.org/10.2748/tmj/1178207478

[31] Kevei, P. (2018). Ergodic properties of generalized Ornstein–Uhlenbeck processes. *Stochastic Process. Appl.* **128** 156–181. MR3729534 https://doi.org/10.1016/j.spa.2017.04.010

[32] Khasminskii, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Heidelberg: Springer. MR2894052 https://doi.org/10.1007/978-3-642-23280-0

[33] Knopova, V. and Kulik, A. (2018). Parametrix construction of the transition probability density of the solution to an SDE driven by α-stable noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 100–140. MR3765882 https://doi.org/10.1214/16-AIHP796

[34] Knopova, V. and Schilling, R.L. (2012). Transition density estimates for a class of Lévy and Lévy-type processes. *J. Theoret. Probab.* **25** 144–170. MR2886383 https://doi.org/10.1007/s10959-010-0300-0

[35] Knopova, V. and Schilling, R.L. (2013). A note on the existence of transition probability densities of Lévy processes. *Forum Math.* **25** 125–149. MR3010850 https://doi.org/10.1515/form.2011.108

[36] Kolokoltsov, V. (2000). Symmetric stable laws and stable-like jump-diffusions. *Proc. Lond. Math. Soc.* (3) **80** 725–768. MR1744782 https://doi.org/10.1112/S0024611500012314

[37] Kolokoltsov, V.N. (2011). *Markov Processes, Semigroups and Generators. De Gruyter Studies in Mathematics* **38**. Berlin: de Gruyter. MR2780345

[38] Kulik, A. (2015). *Introduction to Ergodic Rates for Markov Chains and Processes. Lectures in Pure and Applied Mathematics* **2**. Potsdam: Potsdam Univ. Press. MR3617049

[39] Kulik, A. (2018). *Ergodic Behavior of Markov Processes. De Gruyter Studies in Mathematics* **67**. Berlin: de Gruyter. MR3791835

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

[41] Kwon, Y. and Lee, C. (1999). Strong Feller property and irreducibility of diffusions with jumps. *Stoch. Stoch. Rep.* **67** 147–157. MR1717795 https://doi.org/10.1080/17442509908834206

[42] Luo, D. and Wang, J. (2016). Exponential convergence in L^p-Wasserstein distance for diffusion processes without uniformly dissipative drift. *Math. Nachr.* **289** 1909–1926. MR3563910 https://doi.org/10.1002/mana.201500351

[43] Majka, M.B. (2016). A note on existence of global solutions and invariant measures for jump SDEs with locally one-sided lipschitz drift. Preprint. Available at https://arxiv.org/pdf/1612.03824.pdf.

[44] 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

[45] Mandl, P. (1968). *Analytical Treatment of One-Dimensional Markov Processes. Die Grundlehren der Mathematischen Wissenschaften, Band* 151. Prague: Academia Publishing House of the Czechoslovak Academy of Sciences; New York: Springer. MR0247667

[46] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010

[47] Masuda, H. (2009). Erratum to: "Ergodicity and exponential β-mixing bound for multidimensional diffusions with jumps" [Stochastic Process. Appl. 117 (2007) 35–56] [MR2287102]. *Stochastic Process. Appl.* **119** 676–678. MR2494009 https://doi.org/10.1016/j.spa.2008.02.010

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

[49] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.* **25** 487–517. MR1234294 https://doi.org/10.2307/1427521

[50] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** 518–548. MR1234295 https://doi.org/10.2307/1427522

[51] Pang, G. and Sandrić, N. (2016). Ergodicity and fluctuations of a fluid particle driven by diffusions with jumps. *Commun. Math. Sci.* **14** 327–362. MR3436243 https://doi.org/10.4310/CMS.2016.v14.n2.a2

[52] Picard, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. MR1402654 https://doi.org/10.1007/BF01191910

[53] Picard, J. (2010). Erratum to: On the existence of smooth densities for jump processes [MR1402654]. *Probab. Theory Related Fields* **147** 711–713. MR2639720 https://doi.org/10.1007/s00440-010-0267-x

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

[55] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol.* 2. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. MR1780932 https://doi.org/10.1017/CBO9781107590120

[56] Sandrić, N. (2016). Ergodicity of Lévy-type processes. *ESAIM Probab. Stat.* **20** 154–177. MR3528622 https://doi.org/10.1051/ps/2016009

[57] Sandrić, N. (2016). Long-time behavior for a class of Feller processes. *Trans. Amer. Math. Soc.* **368** 1871–1910. MR3449227 https://doi.org/10.1090/tran/6371

[58] Sandrić, N. (2017). A note on the Birkhoff ergodic theorem. *Results Math.* **72** 715–730. MR3684455 https://doi.org/10.1007/s00025-017-0681-9

[59] 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

[60] Schilling, R.L., Song, R. and Vondraček, Z. (2010). *Bernstein Functions. De Gruyter Studies in Mathematics* **37**. Berlin: de Gruyter. MR2598208

[61] Stramer, O. and Tweedie, R.L. (1997). Existence and stability of weak solutions to stochastic differential equations with non-smooth coefficients. *Statist*. *Sinica* **7** 577–593. MR1467449

[62] Stroock, D.W. (1975). Diffusion processes associated with Lévy generators. *Z. Wahrsch*. *Verw*. *Gebiete* **32** 209–244. MR0433614 https://doi.org/10.1007/BF00532614

[63] Tuominen, P. and Tweedie, R.L. (1994). Subgeometric rates of convergence of f-ergodic Markov chains. *Adv. in Appl*. *Probab*. **26** 775–798. MR1285459 https://doi.org/10.2307/1427820

[64] Tweedie, R.L. (1994). Topological conditions enabling use of Harris methods in discrete and continuous time. *Acta Appl*. *Math*. **34** 175–188. MR1273853 https://doi.org/10.1007/BF00994264

[65] Veretennikov, A.Y. (1997). On polynomial mixing bounds for stochastic differential equations. *Stochastic Process*. *Appl*. **70** 115–127. MR1472961 https://doi.org/10.1016/S0304-4149(97)00056-2

[66] Veretennikov, A.Y. (1999). On polynomial mixing and the rate of convergence for stochastic differential and difference equations. *Teor. Veroyatn*. *Primen*. **44** 312–327. MR1751475 https://doi.org/10.1137/S0040585X97977550

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

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

[69] Wang, F.-Y. (2011). Coupling for Ornstein–Uhlenbeck processes with jumps. *Bernoulli* **17** 1136–1158. MR2854767 https://doi.org/10.3150/10-BEJ308

[70] Wang, J. (2008). Criteria for ergodicity of Lévy type operators in dimension one. *Stochastic Process*. *Appl*. **118** 1909–1928. MR2454470 https://doi.org/10.1016/j.spa.2007.11.003

[71] Wang, J. (2011). Stability of Markov processes generated by Lévy-type operators. *Chinese Ann*. *Math*. *Ser*. *A* **32** 33–50. MR2663819

[72] 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

[73] Wee, I.-S. (1999). Stability for multidimensional jump-diffusion processes. *Stochastic Process*. *Appl*. **80** 193–209. MR1682247 https://doi.org/10.1016/S0304-4149(98)00078-7

Bernoulli **27**(1), 2021, 381–418
https://doi.org/10.3150/20-BEJ1243

Stationary subspace analysis of nonstationary covariance processes: Eigenstructure description and testing

RAANJU R. SUNDARARAJAN[1], VLADAS PIPIRAS[2] and MOHSEN POURAHMADI[3]

[1]*Department of Statistical Science, Southern Methodist University.*
E-mail: rsundararajan@smu.edu
[2]*Department of Statistics and Operations Research, University of North Carolina.*
E-mail: pipiras@email.unc.edu
[3]*Department of Statistics, Texas A&M University.*
E-mail: pourahm@stat.tamu.edu

Stationary subspace analysis (SSA) searches for linear combinations of the components of nonstationary vector time series that are stationary. These linear combinations and their number define an associated stationary subspace and its dimension. SSA is studied here for zero mean nonstationary covariance processes. We characterize stationary subspaces and their dimensions in terms of eigenvalues and eigenvectors of certain symmetric matrices. This characterization is then used to derive formal statistical tests for estimating dimensions of stationary subspaces. Eigenstructure-based techniques are also proposed to estimate stationary subspaces, without relying on previously used computationally intensive optimization-based methods. Finally, the introduced methodologies are examined on simulated and real data.

Keywords: dimension test; eigen-decomposition; local and global dimensions; multivariate nonstationarity

References

[1] Blythe, D.A.J., von Bunau, P., Meinecke, F.C. and Muller, K.R. (2012). Feature extraction for change-point detection using stationary subspace analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **23** 631–643.

[2] Bunse-Gerstner, A., Byers, R., Mehrmann, V. and Nichols, N.K. (1991). Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numer. Math.* **60** 1–39. MR1131497 https://doi.org/10.1007/BF01385712

[3] Cardinali, A. and Nason, G.P. (2010). Costationarity of locally stationary time series. *J. Time Ser. Econom.* **2** Art. 1, 33. MR2915633 https://doi.org/10.2202/1941-1928.1074

[4] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620

[5] Dahlhaus, R. (2012). 13 – locally stationary processes. In *Time Series Analysis: Methods and Applications* (Subba Rao, T., Subba Rao, S. and Rao, C.R., eds.). *Handbook of Statistics* **30** 351–413. Amsterdam: Elsevier. MR3295420

[6] de Jong, P. (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Related Fields* **75** 261–277. MR0885466 https://doi.org/10.1007/BF00354037

[7] Donald, S.G., Fortuna, N. and Pipiras, V. (2007). On rank estimation in symmetric matrices: The case of indefinite matrix estimators. *Econometric Theory* **23** 1217–1232. MR2397198 https://doi.org/10.1017/S0266466607070478

[8] Donald, S.G., Fortuna, N. and Pipiras, V. (2011). Local and global rank tests for multivariate varying-coefficient models. *J. Bus. Econom. Statist.* **29** 295–306. MR2807882 https://doi.org/10.1198/jbes.2010.07303

1350-7265

[9] Düker, M., Pipiras, V. and Sundararajan, R. (2019). Cotrending: Testing for common deterministic trends in varying means model. Preprint.
[10] Fortuna, N. (2008). Local rank tests in a multivariate nonparametric relationship. *J. Econometrics* **142** 162–182. MR2408734 https://doi.org/10.1016/j.jeconom.2007.03.001
[11] Golub, G.H. and Van Loan, C.F. (2013). *Matrix Computations*, 4th ed. *Johns Hopkins Studies in the Mathematical Sciences*. Baltimore, MD: Johns Hopkins Univ. Press. MR3024913
[12] Kaplan, A.Y., Fingelkurts, A.A., Fingelkurts, A.A., Borisov, S.V. and Darkhovsky, B.S. (2005). Nonstationary nature of the brain activity as revealed by EEG/MEG: Methodological, practical and conceptual challenges. *Signal Process*. **85** 2190–2212.
[13] Kato, T. (1966). *Perturbation Theory for Linear Operators*. *Die Grundlehren der Mathematischen Wissenschaften*, *Band* 132. New York: Springer. MR0203473
[14] Lotte, F. and Guan, C.T. (2011). Regularizing common spatial patterns to improve BCI designs: Unified theory and new algorithms. *IEEE Trans*. *Biomed*. *Eng*. **58** 355–362.
[15] Magnus, J.R. and Neudecker, H. (1999). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR1698873
[16] Naeem, M., Brunner, C., Leeb, R., Graimann, B. and Pfurtscheller, G. (2006). Seperability of four-class motor imagery data using independent components analysis. *J*. *Neural Eng*. **3** 208.
[17] Ombao, H., von Sachs, R. and Guo, W. (2005). SLEX analysis of multivariate nonstationary time series. *J*. *Amer*. *Statist*. *Assoc*. **100** 519–531. MR2160556 https://doi.org/10.1198/016214504000001448
[18] Pons, P. and Latapy, M. (2005). Computing communities in large networks using random walks. In *Computer and Information Sciences – ISCIS* 2005 (P. Yolum, T. Güngör, F. Gürgen and C. Özturan, eds.) 284–293. Berlin, Heidelberg: Springer.
[19] Robin, J.-M. and Smith, R.J. (2000). Tests of rank. *Econometric Theory* **16** 151–175. MR1763430 https://doi.org/10.1017/S0266466600162012
[20] Sundararajan, R.R., Palma, M.A. and Pourahmadi, M. (2017). Reducing brain signal noise in the prediction of economic choices: A case study in neuroeconomics. *Front*. *Neurosci*. **11** 704.
[21] Sundararajan, R.R., Pipiras, V. and Pourahmadi, M. (2021). Supplement to "Stationary subspace analysis of nonstationary covariance processes: Eigenstructure description and testing." https://doi.org/10.3150/20-BEJ1243SUPP
[22] Sundararajan, R.R. and Pourahmadi, M. (2018). Stationary subspace analysis of nonstationary processes. *J*. *Time Series Anal*. **39** 338–355. MR3796523 https://doi.org/10.1111/jtsa.12274
[23] von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann*. *Math*. *Stat*. **36** 299–303. MR0170407 https://doi.org/10.1214/aoms/1177700291
[24] von Bünau, P., Meinecke, F.C., Király, F.C. and Müller, K.-R. (2009). Finding stationary subspaces in multivariate time series. *Phys*. *Rev*. *Lett*. **103** 214101.
[25] von Bünau, P., Meinecke, F.C., Scholler, S. and Müller, K.R. (2010). Finding stationary brain sources in EEG data. In 2010 *Annual International Conference of the IEEE Engineering in Medicine and Biology* 2810–2813.

Bernoulli **27**(1), 2021, 419–441
https://doi.org/10.3150/20-BEJ1244

Comparing a large number of multivariate distributions

ILMUN KIM

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

In this paper, we propose a test for the equality of multiple distributions based on kernel mean embeddings. Our framework provides a flexible way to handle multivariate data by virtue of kernel methods and allows the number of distributions to increase with the sample size. This is in contrast to previous studies that have been mostly restricted to classical univariate settings with a fixed number of distributions. By building on Cramér-type moderate deviation for degenerate two-sample V-statistics, we derive the limiting null distribution of the test statistic and show that it converges to a Gumbel distribution. The limiting distribution, however, depends on an infinite number of nuisance parameters, which makes it infeasible for use in practice. To address this issue, the proposed test is implemented via the permutation procedure and is shown to be minimax rate optimal against sparse alternatives. During our analysis, an exponential concentration inequality for the permuted test statistic is developed which may be of independent interest.

Keywords: Bobkov's inequality; K-sample test; maximum mean discrepancy; permutation test

References

[1] Adamczak, R., Chafaï, D. and Wolff, P. (2016). Circular law for random matrices with exchangeable entries. *Random Structures Algorithms* **48** 454–479. MR3481269 https://doi.org/10.1002/rsa.20599
[2] Albert, M. (2019). Concentration inequalities for randomly permuted sums. In *High Dimensional Probability VIII* 341–383. Springer.
[3] Anderson, T.W. and Darling, D.A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Stat.* **23** 193–212. MR0050238 https://doi.org/10.1214/aoms/1177729437
[4] Arias-Castro, E., Candès, E.J. and Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *Ann. Statist.* **39** 2533–2556. MR2906877 https://doi.org/10.1214/11-AOS910
[5] Arratia, R., Goldstein, L. and Gordon, L. (1989). Two moments suffice for Poisson approximations: The Chen–Stein method. *Ann. Probab.* **17** 9–25. MR0972770
[6] Baringhaus, L. and Franz, C. (2004). On a new multivariate two-sample test. *J. Multivariate Anal.* **88** 190–206. MR2021870 https://doi.org/10.1016/S0047-259X(03)00079-4
[7] Bobkov, S.G. (2004). Concentration of normalized sums and a central limit theorem for noncorrelated random variables. *Ann. Probab.* **32** 2884–2907. MR2094433 https://doi.org/10.1214/009117904000000720
[8] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. With a foreword by Michel Ledoux. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001
[9] Bouzebda, S., Keziou, A. and Zari, T. (2011). K-sample problem using strong approximations of empirical copula processes. *Math. Methods Statist.* **20** 14–29. MR2811029 https://doi.org/10.3103/S1066530711010029
[10] Burke, M.D. (1979). On the asymptotic power of some k-sample statistics based on the multivariate empirical process. *J. Multivariate Anal.* **9** 183–205. MR0538401 https://doi.org/10.1016/0047-259X(79)90078-2
[11] Cai, T., Liu, W. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.* **108** 265–277. MR3174618 https://doi.org/10.1080/01621459.2012.758041

1350-7265

[12] Cai, T.T., Liu, W. and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 349–372. MR3164870 https://doi.org/10.1111/rssb.12034
[13] Cai, T.T. and Xia, Y. (2014). High-dimensional sparse MANOVA. *J. Multivariate Anal.* **131** 174–196. MR3252643 https://doi.org/10.1016/j.jmva.2014.07.002
[14] Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072 https://doi.org/10.1007/s00440-006-0029-y
[15] Chen, S. and Pokojovy, M. (2018). Modern and classical k-sample omnibus tests. *Wiley Interdiscip. Rev.: Comput. Stat.* **10** e1418, 12. MR3749550 https://doi.org/10.1002/wics.1418
[16] Conover, W.J. (1965). Several k-sample Kolmogorov–Smirnov tests. *Ann. Math. Stat.* **36** 1019–1026. MR0175230 https://doi.org/10.1214/aoms/1177700073
[17] Drton, M., Han, F. and Shi, H. (2018). High dimensional independence testing with maxima of rank correlations. arXiv preprint arXiv:1812.06189.
[18] Fan, J., Liao, Y. and Yao, J. (2015). Power enhancement in high-dimensional cross-sectional tests. *Econometrica* **83** 1497–1541. MR3384226 https://doi.org/10.3982/ECTA12749
[19] Fukumizu, K., Gretton, A., Sun, X. and Schölkopf, B. (2008). Kernel measures of conditional dependence. In *Advances in Neural Information Processing Systems* 489–496.
[20] Gretton, A., Borgwardt, K.M., Rasch, M., Schölkopf, B. and Smola, A.J. (2007). A kernel method for the two-sample problem. In *Advances in Neural Information Processing Systems* 513–520.
[21] Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773. MR2913716
[22] Hall, P. (1991). On convergence rates of suprema. *Probab. Theory Related Fields* **89** 447–455. MR1118558 https://doi.org/10.1007/BF01199788
[23] Han, F., Chen, S. and Liu, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306 https://doi.org/10.1093/biomet/asx050
[24] He, H.Y., Basu, K., Zhao, Q. and Owen, A.B. (2019). Permutation p-value approximation via generalized Stolarsky invariance. *Ann. Statist.* **47** 583–611. MR3909943 https://doi.org/10.1214/18-AOS1702
[25] Hušková, M. and Meintanis, S.G. (2008). Tests for the multivariate k-sample problem based on the empirical characteristic function. *J. Nonparametr. Stat.* **20** 263–277. MR2421770 https://doi.org/10.1080/10485250801948294
[26] Jeng, X.J., Cai, T.T. and Li, H. (2010). Optimal sparse segment identification with application in copy number variation analysis. *J. Amer. Statist. Assoc.* **105** 1156–1166. MR2752611 https://doi.org/10.1198/jasa.2010.tm10083
[27] Jiang, B., Ye, C. and Liu, J.S. (2015). Nonparametric K-sample tests via dynamic slicing. *J. Amer. Statist. Assoc.* **110** 642–653. MR3367254 https://doi.org/10.1080/01621459.2014.920257
[28] Kiefer, J. (1959). K-sample analogues of the Kolmogorov–Smirnov and Cramér–V. Mises tests. *Ann. Math. Stat.* **30** 420–447. MR0102882 https://doi.org/10.1214/aoms/1177706261
[29] Kim, I. (2020). Supplement to "Comparing a large number of multivariate distributions." https://doi.org/10.3150/20-BEJ1244SUPP
[30] Knijnenburg, T.A., Wessels, L.F., Reinders, M.J. and Shmulevich, I. (2009). Fewer permutations, more accurate P-values. *Bioinformatics* **25** i161–i168.
[31] Kontorovich, A. (2014). Concentration in unbounded metric spaces and algorithmic stability. In *International Conference on Machine Learning* 28–36.
[32] Lee, A.J. (1990). *U-Statistics: Theory and Practice. Statistics: Textbooks and Monographs* **110**. New York: Dekker. MR1075417
[33] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR2135927
[34] Lemeshko, B.Y. and Veretelnikova, I.V. (2018). On some new K-samples tests for testing the homogeneity of distribution laws. In 2018 *XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering* (*APEIE*) 153–157. IEEE.
[35] Li, T. and Yuan, M. (2019). On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. arXiv preprint arXiv:1909.03302.
[36] Liu, W. and Li, Y.Q. (2020). Sign-based test for mean vector in high-dimensional and sparse settings. *Acta Math. Sin.* (*Engl. Ser.*) **36** 93–108. MR4049304 https://doi.org/10.1007/s10114-019-8290-z

[37] Martínez-Camblor, P., De Uña-Álvarez, J. and Corral, N. (2008). k-sample test based on the common area of kernel density estimators. *J. Statist. Plann. Inference* **138** 4006–4020. MR2455983 https://doi.org/10.1016/j.jspi.2008.02.008
[38] Massart, P. (1990). The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *Ann. Probab.* **18** 1269–1283. MR1062069
[39] 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
[40] Muandet, K., Fukumizu, K., Sriperumbudur, B. and Schölkopf, B. (2016). Kernel mean embedding of distributions: A review and beyonds. *Stat* **1050** 31.
[41] Mukhopadhyay, S. and Wang, K. (2018). Nonparametric high-dimensional K-sample comparison. arXiv preprint arXiv:1810.01724.
[42] Quessy, J.-F. and Éthier, F. (2012). Cramér–von Mises and characteristic function tests for the two and k-sample problems with dependent data. *Comput. Statist. Data Anal.* **56** 2097–2111. MR2892402 https://doi.org/10.1016/j.csda.2011.12.021
[43] Rizzo, M.L. and Székely, G.J. (2010). DISCO analysis: A nonparametric extension of analysis of variance. *Ann. Appl. Stat.* **4** 1034–1055. MR2758432 https://doi.org/10.1214/09-AOAS245
[44] Scholz, F.-W. and Stephens, M.A. (1987). k-sample Anderson–Darling tests. *J. Amer. Statist. Assoc.* **82** 918–924. MR0910001
[45] Sejdinovic, D., Sriperumbudur, B., Gretton, A. and Fukumizu, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.* **41** 2263–2291. MR3127866 https://doi.org/10.1214/13-AOS1140
[46] Sosthene, A., Balogoun, K., Martial Nkiet, G. and Ogouyandjou, C. (2018). Kernel based method for the k-sample problem. arXiv preprint arXiv:1812.00100.
[47] Sriperumbudur, B.K., Fukumizu, K. and Lanckriet, G.R.G. (2011). Universality, characteristic kernels and RKHS embedding of measures. *J. Mach. Learn. Res.* **12** 2389–2410. MR2825431
[48] Székely, G.J. and Rizzo, M.L. (2004). Testing for equal distributions in high dimension. *InterStat* **5** 1249–1272.
[49] Thas, O. (2010). *Comparing Distributions*. *Springer Series in Statistics*. New York: Springer. MR2547894 https://doi.org/10.1007/b99044
[50] Tolstikhin, I., Sriperumbudur, B.K. and Muandet, K. (2017). Minimax estimation of kernel mean embeddings. *J. Mach. Learn. Res.* **18** Paper No. 86, 47. MR3714249
[51] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. With a foreword by Sara van de Geer. MR3837109 https://doi.org/10.1017/9781108231596
[52] Wyłupek, G. (2010). Data-driven k-sample tests. *Technometrics* **52** 107–123. MR2654991 https://doi.org/10.1198/TECH.2009.08101
[53] Zhan, D. and Hart, J.D. (2014). Testing equality of a large number of densities. *Biometrika* **101** 449–464. MR3215359 https://doi.org/10.1093/biomet/asu002
[54] Zhang, J. and Wu, Y. (2007). k-sample tests based on the likelihood ratio. *Comput. Statist. Data Anal.* **51** 4682–4691. MR2364473 https://doi.org/10.1016/j.csda.2006.08.029

Bernoulli **27**(1), 2021, 442–468
https://doi.org/10.3150/20-BEJ1245

On the error bound in the normal approximation for Jack measures

LOUIS H.Y. CHEN[1], MARTIN RAIČ[2,3,4] and LÊ VĂN THÀNH[5]

[1]*Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, Singapore 119076, Singapore.*
E-mail: matchyl@nus.edu.sg
[2]*University of Ljubljana, Ljubljana, Slovenia.*
E-mail: martin.raic@fmf.uni-lj.si
[3]*University of Primorska, Koper, Slovenia*
[4]*Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia*
[5]*Department of Mathematics, Vinh University, 182 Le Duan, Vinh, Nghe An, Vietnam.*
E-mail: levt@vinhuni.edu.vn

In this paper, we obtain uniform and non-uniform bounds on the Kolmogorov distance in the normal approximation for Jack deformations of the character ratio, by using Stein's method and zero-bias couplings. Our uniform bound comes very close to that conjectured by Fulman (*J. Combin. Theory Ser. A* **108** (2004) 275–296). As a by-product of the proof of the non-uniform bound, we obtain a Rosenthal-type inequality for zero-bias couplings.

Keywords: Jack deformation; Jack measure; Kolmogorov distance; non-uniform bound; rate of convergence; Stein's method; uniform bound; zero-bias coupling

References

[1] Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178. MR1682248 https://doi.org/10.1090/S0894-0347-99-00307-0

[2] Borodin, A., Okounkov, A. and Olshanski, G. (2000). Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13** 481–515. MR1758751 https://doi.org/10.1090/S0894-0347-00-00337-4

[3] Chatterjee, S. and Dey, P.S. (2010). Applications of Stein's method for concentration inequalities. *Ann. Probab.* **38** 2443–2485. MR2683635 https://doi.org/10.1214/10-AOP542

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

[5] Chen, L.H.Y. and Shao, Q.-M. (2001). A non-uniform Berry–Esseen bound via Stein's method. *Probab. Theory Related Fields* **120** 236–254. MR1841329 https://doi.org/10.1007/PL00008782

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

[7] Chen, L.H.Y., Goldstein, L. and Röllin, A. (2020). Stein's method via induction. Available at https://arxiv.org/abs/1903.09319.

[8] Diaconis, P. and Shahshahani, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR0626813 https://doi.org/10.1007/BF00535487

[9] Dołęga, M. and Féray, V. (2016). Gaussian fluctuations of Young diagrams and structure constants of Jack characters. *Duke Math. J.* **165** 1193–1282. MR3498866 https://doi.org/10.1215/00127094-3449566

[10] Dołęga, M. and Śniady, P. (2019). Gaussian fluctuations of Jack-deformed random Young diagrams. *Probab. Theory Related Fields* **174** 133–176. MR3947322 https://doi.org/10.1007/s00440-018-0854-9

[11] Eskin, A. and Okounkov, A. (2001). Asymptotics of numbers of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials. *Invent*. *Math*. **145** 59–103. MR1839286 https://doi.org/10.1007/s002220100142
[12] Frobenius, G.F. (1900). Uber die Charaktere der symmetrischen Gruppe. In *Sitz. Konig. Preuss. Akad. Wissen* 516–534. In *Gesammelte Abhandlungen III* 148–166. Heidelberg: Springer-Verlag (1968).
[13] Fulman, J. (2004). Stein's method, Jack measure, and the Metropolis algorithm. *J. Combin*. *Theory Ser. A* **108** 275–296. MR2098845 https://doi.org/10.1016/j.jcta.2004.07.003
[14] Fulman, J. (2006). An inductive proof of the Berry–Esseen theorem for character ratios. *Ann*. *Comb*. **10** 319–332. MR2284273 https://doi.org/10.1007/s00026-006-0290-x
[15] Fulman, J. and Goldstein, L. (2011). Zero biasing and Jack measures. *Combin*. *Probab*. *Comput*. **20** 753–762. MR2825589 https://doi.org/10.1017/S0963548311000241
[16] Goldstein, L. (2005). Berry–Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J*. *Appl*. *Probab*. **42** 661–683. MR2157512 https://doi.org/10.1239/jap/1127322019
[17] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann*. *Appl*. *Probab*. **7** 935–952. MR1484792 https://doi.org/10.1214/aoap/1043862419
[18] Hora, A. (1998). Central limit theorem for the adjacency operators on the infinite symmetric group. *Comm*. *Math*. *Phys*. **195** 405–416. MR1637801 https://doi.org/10.1007/s002200050395
[19] Hora, A. and Obata, N. (2007). *Quantum Probability and Spectral Analysis of Graphs*. *Theoretical and Mathematical Physics*. Berlin: Springer. With a foreword by Luigi Accardi. MR2316893
[20] Ibragimov, R. and Sharakhmetov, Sh. (1997). On an exact constant for the Rosenthal inequality. *Theory Probab*. *Appl*. **42** 294–302. MR1474714
[21] Johansson, K. (2001). Discrete orthogonal polynomial ensembles and the Plancherel measure. *Ann*. *of Math*. (2) **153** 259–296. MR1826414 https://doi.org/10.2307/2661375
[22] Johnson, W.B., Schechtman, G. and Zinn, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann*. *Probab*. **13** 234–253. MR0770640
[23] Kerov, S. (1993). Gaussian limit for the Plancherel measure of the symmetric group. *C*. *R*. *Acad*. *Sci*. *Paris Sér. I Math*. **316** 303–308. MR1204294
[24] Kerov, S.V. (2000). Anisotropic Young diagrams and Jack symmetric functions. *Funct*. *Anal*. *Appl*. **34** 41–51. MR1756734
[25] Latała, R. (1997). Estimation of moments of sums of independent real random variables. *Ann*. *Probab*. **25** 1502–1513. MR1457628 https://doi.org/10.1214/aop/1024404522
[26] Matsumoto, S. (2008). Jack deformations of Plancherel measures and traceless Gaussian random matrices. *Electron*. *J*. *Combin*. **15** Research Paper 149. MR2465773
[27] Rosenthal, H.P. (1970). On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables. *Israel J*. *Math*. **8** 273–303. MR0271721 https://doi.org/10.1007/BF02771562
[28] Sagan, B.E. (2001). *The Symmetric Group*: *Representations*, *Combinatorial Algorithms*, *and Symmetric Functions*, 2nd ed. *Graduate Texts in Mathematics* **203**. New York: Springer. MR1824028 https://doi.org/10.1007/978-1-4757-6804-6
[29] Shao, Q.-M. and Su, Z.-G. (2006). The Berry–Esseen bound for character ratios. *Proc*. *Amer. Math*. *Soc*. **134** 2153–2159. MR2215787 https://doi.org/10.1090/S0002-9939-05-08177-3
[30] Śniady, P. (2006). Gaussian fluctuations of characters of symmetric groups and of Young diagrams. *Probab*. *Theory Related Fields* **136** 263–297. MR2240789 https://doi.org/10.1007/s00440-005-0483-y
[31] Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **7**. Hayward, CA: IMS. MR0882007

Bernoulli **27**(1), 2021, 469–501
https://doi.org/10.3150/20-BEJ1246

A similarity measure for second order properties of non-stationary functional time series with applications to clustering and testing

ANNE VAN DELFT[1] and HOLGER DETTE[2]

[1]*Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA.*
E-mail: anne.vandelft@columbia.edu
[2]*Ruhr-Universität Bochum, Fakultät für Mathematik, 44780 Bochum, Germany.*
E-mail: holger.dette@rub.de

Due to the surge of data storage techniques, the need for the development of appropriate techniques to identify patterns and to extract knowledge from the resulting enormous data sets, which can be viewed as collections of dependent functional data, is of increasing interest in many scientific areas. We develop a similarity measure for spectral density operators of a collection of functional time series, which is based on the aggregation of Hilbert–Schmidt differences of the individual time-varying spectral density operators. Under fairly general conditions, the asymptotic properties of the corresponding estimator are derived and asymptotic normality is established. The introduced statistic lends itself naturally to quantify (dis)-similarity between functional time series, which we subsequently exploit in order to build a spectral clustering algorithm. Our algorithm is the first of its kind in the analysis of non-stationary (functional) time series and enables to discover particular patterns by grouping together 'similar' series into clusters, thereby reducing the complexity of the analysis considerably. The algorithm is simple to implement and computationally feasible. As a further application, we provide a simple test for the hypothesis that the second order properties of two non-stationary functional time series coincide.

Keywords: clustering; functional data; local stationarity; spectral analysis; time series

References

[1] Abraham, C., Cornillon, P.A., Matzner-Løber, E. and Molinari, N. (2003). Unsupervised curve clustering using B-splines. *Scand. J. Stat.* **30** 581–595. MR2002229 https://doi.org/10.1111/1467-9469.00350

[2] Aghabozorgi, S., Sirkhorshidi, A.S. and Wah, Y.W. (2015). Time-series clustering – a decade review. *Inform. Sci.* **53** 16–38.

[3] Aston, J.A.D. and Kirch, C. (2012). Detecting and estimating changes in dependent functional data. *J. Multivariate Anal.* **109** 204–220. MR2922864 https://doi.org/10.1016/j.jmva.2012.03.006

[4] Aue, A. and van Delft, A. (2020). Testing for stationarity of functional time series in the frequency domain. *Ann. Statist.* **48** 2505–2547. https://doi.org/10.1214/19-AOS1895

[5] Bauwens, L. and Rombouts, J.V.K. (2007). Bayesian clustering of many Garch models. *Econometric Rev.* **26** 365–386. MR2364366 https://doi.org/10.1080/07474930701220576

[6] Böhm, H., Ombao, H., von Sachs, R. and Sanes, J. (2010). Classification of multivariate non-stationary signals: The SLEX-shrinkage approach. *J. Statist. Plann. Inference* **140** 3754–3763. MR2674163 https://doi.org/10.1016/j.jspi.2010.04.040

[7] Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and Applications. Lecture Notes in Statistics* **149**. New York: Springer. MR1783138 https://doi.org/10.1007/978-1-4612-1154-9

[8] Caliński, T. and Harabasz, J. (1974). A dendrite method for cluster analysis. *Commun. Stat.* **3** 1–27. MR0375641 https://doi.org/10.1080/03610927408827101

[9] Chamroukhi, F. and Nguyen, H.D. (2019). Model-based clustering and classification of functional data. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **9** e1298. https://doi.org/10.1002/widm.1298

1350-7265

[10] Chandler, G. and Polonik, W. (2006). Discrimination of locally stationary time series based on the excess mass functional. *J. Amer. Statist. Assoc.* **101** 240–253. MR2268042 https://doi.org/10.1198/016214505000000899
[11] Charrad, M., Ghazzali, N., Boiteau, V. and Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *J. Stat. Softw.* **61** 1–36.
[12] Chung, F. and Radcliffe, M. (2011). On the spectra of general random graphs. *Electron. J. Combin.* **18** Paper 215, 14. MR2853072
[13] Chung, F.R.K. (1997). *Spectral Graph Theory. CBMS Regional Conference Series in Mathematics* **92**. Washington, DC: Published for the Conference Board of the Mathematical Sciences; Providence, RI: Amer. Math. Soc. MR1421568
[14] Coates, D.S. and Diggle, P.J. (1986). Tests for comparing two estimated spectral densities. *J. Time Series Anal.* **7** 7–20. MR0832349 https://doi.org/10.1111/j.1467-9892.1986.tb00482.x
[15] Corduas, M. and Piccolo, D. (2008). Time series clustering and classification by the autoregressive metric. *Comput. Statist. Data Anal.* **52** 1860–1872. MR2418476 https://doi.org/10.1016/j.csda.2007.06.001
[16] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620
[17] Davis, C. and Kahan, W.M. (1970). The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* **7** 1–46. MR0264450 https://doi.org/10.1137/0707001
[18] Delaigle, A., Hall, P. and Pham, T. (2019). Clustering functional data into groups by using projections. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 271–304. MR3928143 https://doi.org/10.1111/rssb.12310
[19] Dette, H. (2009). Bootstrapping frequency domain tests in multivariate time series with an application to comparing spectral densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 831–857. MR2750097 https://doi.org/10.1111/j.1467-9868.2009.00709.x
[20] Dette, H. and Hildebrandt, T. (2012). A note on testing hypotheses for stationary processes in the frequency domain. *J. Multivariate Anal.* **104** 101–114. MR2832189 https://doi.org/10.1016/j.jmva.2011.07.002
[21] Eichler, M. (2008). Testing nonparametric and semiparametric hypotheses in vector stationary processes. *J. Multivariate Anal.* **99** 968–1009. MR2405101 https://doi.org/10.1016/j.jmva.2007.06.003
[22] Euán, C., Ombao, H. and Ortega, J. (2018). Spectral synchronicity in brain signals. *Stat. Med.* **37** 2855–2873. MR3832245 https://doi.org/10.1002/sim.7695
[23] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics*. New York: Springer. MR2229687
[24] Floriello, D. and Vitelli, V. (2017). Sparse clustering of functional data. *J. Multivariate Anal.* **154** 1–18. MR3588554 https://doi.org/10.1016/j.jmva.2016.10.008
[25] Fokianos, K. and Promponas, V.J. (2012). Biological applications of time series frequency domain clustering. *J. Time Series Anal.* **33** 744–756. MR2969908 https://doi.org/10.1111/j.1467-9892.2011.00758.x
[26] Frühwirth-Schnatter, S. and Kaufmann, S. (2008). Model-based clustering of multiple time series. *J. Bus. Econom. Statist.* **26** 78–89. MR2422063 https://doi.org/10.1198/073500107000000106
[27] Gordon, A.D. (1999). *Classification*, 2nd ed. London: Chapman and Hall–CRC. MR0637465
[28] Hartigan, J.A. (1975). *Clustering Algorithms. Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR0405726
[29] Harvill, J.L., Kohli, P. and Ravishanker, N. (2017). Clustering nonlinear, nonstationary time series using BSLEX. *Methodol. Comput. Appl. Probab.* **19** 935–955. MR3683978 https://doi.org/10.1007/s11009-016-9528-1
[30] Heard, N.A., Holmes, C.C. and Stephens, D.A. (2006). A quantitative study of gene regulation involved in the immune response of anopheline mosquitoes: An application of Bayesian hierarchical clustering of curves. *J. Amer. Statist. Assoc.* **101** 18–29. MR2252430 https://doi.org/10.1198/016214505000000187
[31] Holan, S.H. and Ravishanker, N. (2018). Time series clustering and classification via frequency domain methods. *Wiley Interdiscip. Rev.: Comput. Stat.* **10** e1444, 15. MR3873674 https://doi.org/10.1002/wics.1444
[32] Horváth, L., Hušková, M. and Rice, G. (2013). Test of independence for functional data. *J. Multivariate Anal.* **117** 100–119. MR3053537 https://doi.org/10.1016/j.jmva.2013.02.005
[33] Huang, H.-Y., Ombao, H. and Stoffer, D.S. (2004). Discrimination and classification of nonstationary time series using the SLEX model. *J. Amer. Statist. Assoc.* **99** 763–774. MR2090909 https://doi.org/10.1198/016214504000001105

[34] Ieva, F., Paganoni, A.M., Pigoli, D. and Vitelli, V. (2013). Multivariate functional clustering for the morphological analysis of electrocardiograph curves. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **62** 401–418. MR3060623 https://doi.org/10.1111/j.1467-9876.2012.01062.x

[35] Jacques, J. and Preda, C. (2014). Model-based clustering for multivariate functional data. *Comput. Statist. Data Anal.* **71** 92–106. MR3131956 https://doi.org/10.1016/j.csda.2012.12.004

[36] Jacques, J. and Preda, C. (2014). Functional data clustering: A survey. *Adv. Data Anal. Classif.* **8** 231–255. MR3253859 https://doi.org/10.1007/s11634-013-0158-y

[37] Jentsch, C. and Pauly, M. (2015). Testing equality of spectral densities using randomization techniques. *Bernoulli* **21** 697–739. MR3338644 https://doi.org/10.3150/13-BEJ584

[38] Juárez, M.A. and Steel, M.F.J. (2010). Model-based clustering of non-Gaussian panel data based on skew-*t* distributions. *J. Bus. Econom. Statist.* **28** 52–66. MR2650600 https://doi.org/10.1198/jbes.2009.07145

[39] Kakizawa, Y., Shumway, R.H. and Taniguchi, M. (1998). Discrimination and clustering for multivariate time series. *J. Amer. Statist. Assoc.* **93** 328–340. MR1614589 https://doi.org/10.2307/2669629

[40] Kalpakis, K., Gada, D. and Puttagunta, V. (2001). Distance measures for effective clustering of arima time-series. In *Proceedings of the* 2001 *IEEE International Conference on Data Mining*, *San Jose* 273–280.

[41] Kaufman, L. and Rousseeuw, P.J. (1990). *Finding Groups in Data*: *An Introduction to Cluster Analysis*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1044997 https://doi.org/10.1002/9780470316801

[42] Krzanowski, W.J. and Lai, Y.T. (1988). A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics* **44** 23–34. MR0931626 https://doi.org/10.2307/2531893

[43] Lei, J. and Rinaldo, A. (2015). Consistency of spectral clustering in stochastic block models. *Ann. Statist.* **43** 215–237. MR3285605 https://doi.org/10.1214/14-AOS1274

[44] Leucht, A., Paporoditis, E. and Sapatinas, T. (2018). Testing equality of spectral density operators for functional linear processes. arXiv:1804.03366.

[45] Milligan, G.W. and Cooper, M.C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika* **50** 159–179.

[46] Ng, A., Jordan, S. and Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems* 14 (T. Dietterich, S. Becker and Z. Ghahramani, eds.) MIT Press.

[47] Ombao, H.C., Raz, J.A., von Sachs, R. and Malow, B.A. (2001). Automatic statistical analysis of bivariate nonstationary time series. *J. Amer. Statist. Assoc.* **96** 543–560. MR1946424 https://doi.org/10.1198/016214501753168244

[48] Paparoditis, E. and Sapatinas, T. (2016). Bootstrap-based testing of equality of mean functions or equality of covariance operators for functional data. *Biometrika* **103** 727–733. MR3551795 https://doi.org/10.1093/biomet/asw033

[49] Peng, J. and Müller, H.-G. (2008). Distance-based clustering of sparsely observed stochastic processes, with applications to online auctions. *Ann. Appl. Stat.* **2** 1056–1077. MR2516804 https://doi.org/10.1214/08-AOAS172

[50] Rohe, K., Chatterjee, S. and Yu, B. (2011). Spectral clustering and the high-dimensional stochastic blockmodel. *Ann. Statist.* **39** 1878–1915. MR2893856 https://doi.org/10.1214/11-AOS887

[51] Sakiyama, K. and Taniguchi, M. (2004). Discriminant analysis for locally stationary processes. *J. Multivariate Anal.* **90** 282–300. MR2081780 https://doi.org/10.1016/j.jmva.2003.08.002

[52] Savvides, A., Promponas, V.J. and Fokianos, K. (2008). Clustering of biological time series by cepstral coefficients based distances. *Pattern Recognit.* **41** 2398–2412.

[53] Shi, J. and Malik, J. (2002). Nomalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 888–905.

[54] Stewart, G.W. and Sun, J.G. (1990). *Matrix Perturbation Theory*. *Computer Science and Scientific Computing*. Boston, MA: Academic Press. MR1061154

[55] Tavakoli, S. and Panaretos, V.M. (2016). Detecting and localizing differences in functional time series dynamics: A case study in molecular biophysics. *J. Amer. Statist. Assoc.* **111** 1020–1035. MR3561926 https://doi.org/10.1080/01621459.2016.1147355

[56] Theodoridis, S. and Koutroumbas, K. (2008). *Pattern Recognition*, 4th ed. New York: Academic Press.

[57] Tibshirani, R., Walther, G. and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 411–423. MR1841503 https://doi.org/10.1111/1467-9868.00293
[58] van Delft, A. (2020). A note on quadratic forms of stationary functional time series under mild conditions. *Stochastic Process. Appl.* **130** 4206–4251. MR4102264 https://doi.org/10.1016/j.spa.2019.12.002
[59] van Delft, A., Characiejus, V. and Dette, H. (2019). A nonparametric test for stationarity in functional time series. *Statist. Sinica*. To appear. https://doi.org/10.5705/ss.202018.0320
[60] van Delft, A. and Dette, H. (2020). Supplement to "A similarity measure for second order properties of non-stationary functional time series with applications to clustering and testing." https://doi.org/10.3150/20-BEJ1246SUPP
[61] van Delft, A. and Eichler, M. (2018). Locally stationary functional time series. *Electron. J. Stat.* **12** 107–170. MR3746979 https://doi.org/10.1214/17-EJS1384
[62] Vlachos, M., Lin, J., Keogh, E. and Gunopulos, D. (2003). A wavelet-based anytime algorithm for k-means clustering of time series. In *Proc. Workshop on Clustering High Dimensionality Data and Its Applications, San Francisco*.
[63] von Luxburg, U. (2007). A tutorial on spectral clustering. *Stat. Comput.* **17** 395–416. MR2409803 https://doi.org/10.1007/s11222-007-9033-z
[64] von Luxburg, U., Belkin, M. and Bousquet, O. (2008). Consistency of spectral clustering. *Ann. Statist.* **36** 555–586. MR2396807 https://doi.org/10.1214/009053607000000640
[65] Zhang, X. and Shao, X. (2015). Two sample inference for the second-order property of temporally dependent functional data. *Bernoulli* **21** 909–929. MR3338651 https://doi.org/10.3150/13-BEJ592

Bernoulli **27**(1), 2021, 502–531
https://doi.org/10.3150/20-BEJ1247

Max-convolution semigroups and extreme values in limit theorems for the free multiplicative convolution

YUKI UEDA

Department of Mathematics, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan.
E-mail: yuuki1114@math.sci.hokudai.ac.jp

We investigate relations between additive convolution semigroups and max-convolution semigroups through the law of large numbers for the free multiplicative convolution. Based on these relations, we give a formula related with the Belinschi–Nica semigroup and the max-Belinschi–Nica semigroup. Finally, we give several limit theorems for classical, free and Boolean extreme values.

Keywords: Belinschi–Nica semigroup; Bercovici–Pata bijection; extreme values; free multiplicative convolution; max-convolution

References

[1] Ando, T. (1989). Majorization, doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra Appl*. **118** 163–248. MR0995373 https://doi.org/10.1016/0024-3795(89)90580-6
[2] Arizmendi, O. and Hasebe, T. (2016). Classical scale mixtures of Boolean stable laws. *Trans*. *Amer. Math*. *Soc*. **368** 4873–4905. MR3456164 https://doi.org/10.1090/tran/6792
[3] Arizmendi, O., Hasebe, T. and Sakuma, N. (2013). On the law of free subordinators. *ALEA Lat*. *Am*. *J*. *Probab*. *Math*. *Stat*. **10** 271–291. MR3083927
[4] Barndorff-Nielsen, O.E. and Thorbjørnsen, S. (2002). Self-decomposability and Lévy processes in free probability. *Bernoulli* **8** 323–366. MR1913111
[5] Belinschi, S.T. (2003). The atoms of the free multiplicative convolution of two probability distributions. *Integral Equations Operator Theory* **46** 377–386. MR1997977 https://doi.org/10.1007/s00020-002-1145-4
[6] Belinschi, S.T. and Bercovici, H. (2004). Atoms and regularity for measures in a partially defined free convolution semigroup. *Math*. *Z*. **248** 665–674. MR2103535 https://doi.org/10.1007/s00209-004-0671-y
[7] Belinschi, S.T. and Nica, A. (2008). On a remarkable semigroup of homomorphisms with respect to free multiplicative convolution. *Indiana Univ*. *Math*. *J*. **57** 1679–1713. MR2440877 https://doi.org/10.1512/iumj.2008.57.3285
[8] Ben Arous, G. and Voiculescu, D.V. (2006). Free extreme values. *Ann*. *Probab*. **34** 2037–2059. MR2271490 https://doi.org/10.1214/009117906000000016
[9] Benaych-Georges, F. and Cabanal-Duvillard, T. (2010). A matrix interpolation between classical and free max operations. I. The univariate case. *J*. *Theoret*. *Probab*. **23** 447–465. MR2644869 https://doi.org/10.1007/s10959-009-0210-1
[10] Bercovici, H. and Pata, V. (1999). Stable laws and domains of attraction in free probability theory. *Ann*. *of Math*. (2) **149** 1023–1060. With an appendix by Philippe Biane. MR1709310 https://doi.org/10.2307/121080
[11] Bercovici, H. and Voiculescu, D. (1993). Free convolution of measures with unbounded support. *Indiana Univ*. *Math*. *J*. **42** 733–773. MR1254116 https://doi.org/10.1512/iumj.1993.42.42033
[12] Bercovici, H. and Voiculescu, D. (1995). Superconvergence to the central limit and failure of the Cramér theorem for free random variables. *Probab*. *Theory Related Fields* **103** 215–222. MR1355057 https://doi.org/10.1007/BF01204215

1350-7265

[13] Bercovici, H. and Voiculescu, D. (1998). Regularity questions for free convolution. In *Nonselfadjoint Operator Algebras*, *Operator Theory*, *and Related Topics*. *Oper. Theory Adv. Appl.* **104** 37–47. Basel: Birkhäuser. MR1639647

[14] Bertoin, J. (1999). Subordinators: Examples and applications. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1997). *Lecture Notes in Math.* **1717** 1–91. Berlin: Springer. MR1746300 https://doi.org/10.1007/978-3-540-48115-7_1

[15] Grela, J. and Nowak, M.A. On relations between extreme value statistics, extreme random matrices and Peak-Over-Threshold method. arXiv:1711.03459.

[16] Fisher, R.A. and Tippett, L.H.C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Philos. Soc.* **24** 180.

[17] Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Ann. Soc. Math. Polon.* **6** 93–116.

[18] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math.* (2) **44** 423–453. MR0008655 https://doi.org/10.2307/1968974

[19] Haagerup, U. and Larsen, F. (2000). Brown's spectral distribution measure for R-diagonal elements in finite von Neumann algebras. *J. Funct. Anal.* **176** 331–367. MR1784419 https://doi.org/10.1006/jfan.2000.3610

[20] Haagerup, U. and Möller, S. (2013). The law of large numbers for the free multiplicative convolution. In *Operator Algebra and Dynamics*. *Springer Proc. Math. Stat.* **58** 157–186. Heidelberg: Springer. MR3142036 https://doi.org/10.1007/978-3-642-39459-1_8

[21] Haagerup, U. and Schultz, H. (2007). Brown measures of unbounded operators affiliated with a finite von Neumann algebra. *Math. Scand.* **100** 209–263. MR2339369 https://doi.org/10.7146/math.scand.a-15023

[22] Haagerup, U. and Schultz, H. (2009). Invariant subspaces for operators in a general II_1-factor. *Publ. Math. Inst. Hautes Études Sci.* **109** 19–111. MR2511586 https://doi.org/10.1007/s10240-009-0018-7

[23] Kadison, R.V. (1951). Order properties of bounded self-adjoint operators. *Proc. Amer. Math. Soc.* **2** 505–510. MR0042064 https://doi.org/10.2307/2031784

[24] Lindsay, J.M. and Pata, V. (1997). Some weak laws of large numbers in noncommutative probability. *Math. Z.* **226** 533–543. MR1484709 https://doi.org/10.1007/PL00004356

[25] Maassen, H. (1992). Addition of freely independent random variables. *J. Funct. Anal.* **106** 409–438. MR1165862 https://doi.org/10.1016/0022-1236(92)90055-N

[26] Młotkowski, W., Sakuma, N. and Ueda, Y. (2020). Free self-decomposability and unimodality of the Fuss–Catalan distributions. *J. Stat. Phys.* **178** 1055–1075. MR4081219 https://doi.org/10.1007/s10955-020-02488-1

[27] Nica, A. and Speicher, R. (1996). On the multiplication of free N-tuples of noncommutative random variables. *Amer. J. Math.* **118** 799–837. MR1400060

[28] Nica, A. and Speicher, R. (2006). *Lectures on the Combinatorics of Free Probability*. *London Mathematical Society Lecture Note Series* **335**. Cambridge: Cambridge Univ. Press. MR2266879 https://doi.org/10.1017/CBO9780511735127

[29] Olson, M.P. (1971). The selfadjoint operators of a von Neumann algebra form a conditionally complete lattice. *Proc. Amer. Math. Soc.* **28** 537–544. MR0276788 https://doi.org/10.2307/2038007

[30] Resnick, S.I. (2008). *Extreme Values*, *Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2364939

[31] Sakuma, N. (2011). On free regular infinitely divisible distributions. *RIMS Kôkyûroku Bessatsu* **B27** 115–122.

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

[33] Speicher, R. and Woroudi, R. (1997). Boolean convolution. In *Free Probability Theory* (*Waterloo*, *ON*, 1995). *Fields Inst. Commun.* **12** 267–279. Providence, RI: Amer. Math. Soc. MR1426845

[34] Tucci, G.H. (2010). Limits laws for geometric means of free random variables. *Indiana Univ. Math. J.* **59** 1–13. MR2666470 https://doi.org/10.1512/iumj.2010.59.3775

[35] Ueda, Y. Limit theorems for classical, freely and Boolean max-infinitely divisible distributions. Available at arXiv:1907.11996.

[36] Vargas, J.G. and Voiculescu, D.V. (2020). Boolean extremes and dagum distributions. *Indiana Univ. Math. J.* To appear.

[37] Voiculescu, D. (1986). Addition of certain noncommuting random variables. *J. Funct. Anal.* **66** 323–346. MR0839105 https://doi.org/10.1016/0022-1236(86)90062-5

[38] Voiculescu, D. (1987). Multiplication of certain noncommuting random variables. *J. Operator Theory* **18** 223–235. MR0915507

[39] Voiculescu, D.V., Dykema, K.J. and Nica, A. (1992). *Free Random Variables. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups. CRM Monograph Series* **1**. Providence, RI: Amer. Math. Soc. MR1217253

Bernoulli **27**(1), 2021, 532–553
https://doi.org/10.3150/20-BEJ1248

Statistical estimation of ergodic Markov chain kernel over discrete state space

GEOFFREY WOLFER* and ARYEH KONTOROVICH†

Computer Science Department, Ben-Gurion University of the Negev, Beer-Sheva, Israel.
E-mail: *geoffrey@post.bgu.ac.il; †karyeh@cs.bgu.ac.il*

We investigate the statistical complexity of estimating the parameters of a discrete-state Markov chain kernel from a single long sequence of state observations. In the finite case, we characterize (modulo logarithmic factors) the minimax sample complexity of estimation with respect to the operator infinity norm, while in the countably infinite case, we analyze the problem with respect to a natural entry-wise norm derived from total variation. We show that in both cases, the sample complexity is governed by the mixing properties of the unknown chain, for which, in the finite-state case, there are known finite-sample estimators with fully empirical confidence intervals.

Keywords: discrete state space; ergodic Markov chain; minimax theory

References

[1] Anthony, M. and Bartlett, P.L. (1999). *Neural Network Learning*: *Theoretical Foundations*. Cambridge: Cambridge Univ. Press. MR1741038 https://doi.org/10.1017/CBO9780511624216
[2] Assouad, P. (1983). Deux remarques sur l'estimation. *C. R. Acad. Sci. Paris Sér. I Math*. **296** 1021–1024. MR0777600
[3] Berend, D. and Kontorovich, A. (2013). A sharp estimate of the binomial mean absolute deviation with applications. *Statist. Probab. Lett*. **83** 1254–1259. MR3041401 https://doi.org/10.1016/j.spl.2013.01.023
[4] Billingsley, P. (1961). Statistical methods in Markov chains. *Ann. Math. Stat*. **32** 12–40. MR0123420 https://doi.org/10.1214/aoms/1177705136
[5] Bubley, R. and Dyer, M. (1997). Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings* 38*th Annual Symposium on Foundations of Computer Science* 223–231. IEEE.
[6] Craig, B.A. and Sendi, P.P. (2002). Estimation of the transition matrix of a discrete-time Markov chain. *Health Econ*. **11** 33–42.
[7] Devroye, L. and Lugosi, G. (2001). *Combinatorial Methods in Density Estimation*. *Springer Series in Statistics*. New York: Springer. MR1843146 https://doi.org/10.1007/978-1-4613-0125-7
[8] Fill, J.A. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab*. **1** 62–87. MR1097464
[9] Han, Y., Jiao, J. and Weissman, T. (2015). Minimax estimation of discrete distributions under ℓ_1 loss. *IEEE Trans. Inf. Theory* **61** 6343–6354. MR3418968 https://doi.org/10.1109/TIT.2015.2478816
[10] Hao, Y., Orlitsky, A. and Pichapati, V. (2018). On learning Markov chains. In *Advances in Neural Information Processing Systems* 646–655.
[11] Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183 https://doi.org/10.1017/CBO9780511810817
[12] Hsu, D., Kontorovich, A., Levin, D.A., Peres, Y., Szepesvári, C. and Wolfer, G. (2019). Mixing time estimation in reversible Markov chains from a single sample path. *Ann. Appl. Probab*. **29** 2439–2480. MR3983341 https://doi.org/10.1214/18-AAP1457
[13] Kamath, S., Orlitsky, A., Pichapati, D. and Suresh, A.T. (2015). On learning distributions from their samples. In *Proceedings of the* 28*th Conference on Learning Theory*, *COLT* 2015, *Paris*, *France*, *July* 3–6, 2015 1066–1100. Available at http://jmlr.org/proceedings/papers/v40/Kamath15.html.

[14] Kearns, M.J., Mansour, Y., Ron, D., Rubinfeld, R., Schapire, R.E. and Sellie, L. (1994). On the learnability of discrete distributions. In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing* **1994** 273–282. Montréal, Québec, Canada. https://doi.org/10.1145/195058.195155
[15] Kemeny, J.G. and Snell, J.L. (1976). *Finite Markov Chains*. New York: Springer. Reprinting of the 1960 original, Undergraduate Texts in Mathematics. MR0410929
[16] Kontorovich, A. and Pinelis, I. (2019). Exact lower bounds for the agnostic probably-approximately-correct (PAC) machine learning model. *Ann. Statist.* **47** 2822–2854. MR3988774 https://doi.org/10.1214/18-AOS1766
[17] Kontorovich, A.L. (2007). Measure Concentration of Strongly Mixing Processes with Applications Ph.D. thesis, Carnegie Mellon University.
[18] Kontorovich, L. and Ramanan, K. (2008). Concentration inequalities for dependent random variables via the martingale method. *Ann. Probab.* **36** 2126–2158. MR2478678 https://doi.org/10.1214/07-AOP384
[19] Levin, D.A., Peres, Y. and Wilmer, E.L. (2009). *Markov Chains and Mixing Times*, 2nd ed. Providence: American Mathematical Soc.
[20] Markov, A.A. (1906). Extension of the law of large numbers to dependent quantities. *Izvestiia Fiz.-Matem. Obsch. Kazan Univ.* **15** 135–156.
[21] Orlitsky, A. and Suresh, A.T. (2015). Competitive distribution estimation: Why is good-Turing good. In *Advances in Neural Information Processing Systems* 28: *Annual Conference on Neural Information Processing Systems* 2143–2151. Available at http://papers.nips.cc/paper/5762-competitive-distribution-estimation-why-is-good-turing-good.
[22] Paulin, D. (2015). Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electron. J. Probab.* **20** 79. MR3383563 https://doi.org/10.1214/EJP.v20-4039
[23] Roberts, G.O. and Rosenthal, J.S. (1997). Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.* **2** 13–25. MR1448322 https://doi.org/10.1214/ECP.v2-981
[24] Silvester, J.R. (2000). Determinants of block matrices. *Math. Gaz.* **84** 460–467. Available at http://www.jstor.org/stable/3620776.
[25] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats. MR2724359 https://doi.org/10.1007/b13794
[26] Vidyasagar, M. (2014). An elementary derivation of the large deviation rate function for finite state Markov chains. *Asian J. Control* **16** 1–19. MR3160167 https://doi.org/10.1002/asjc.806
[27] Waggoner, B. (2015). ℓ_p testing and learning of discrete distributions. In *ITCS'15 – Proceedings of the 6th Innovations in Theoretical Computer Science* 347–356. New York: ACM. MR3419028 https://doi.org/10.1145/2688073.2688095
[28] Welton, N. and Ades, A.E. (2005). Estimation of Markov chain transition probabilities and rates from fully and partially observed data: Uncertainty propagation, evidence synthesis, and model calibration. *Med. Decis. Mak.* **25** 633–645.
[29] Wolfer, G. (2020). Mixing time estimation in ergodic Markov chains from a single trajectory with contraction methods. In *Proceedings of the 31st International Conference on Algorithmic Learning Theory* (A. Kontorovich and G. Neu, eds.). *Proceedings of Machine Learning Research* **117** 890–905. San Diego, CA, USA. Available at http://proceedings.mlr.press/v117/wolfer20a.html.
[30] Wolfer, G. and Kontorovich, A. (2019). Minimax learning of ergodic Markov chains. In *Proceedings of the 30th International Conference on Algorithmic Learning Theory*, *Proceedings of Machine Learning Research* **98** 904–930. Chicago, IL. Available at http://proceedings.mlr.press/v98/wolfer19a.html.
[31] Wolfer, G. and Kontorovich, A. (2019). Estimating the mixing time of ergodic Markov chains. In *Proceedings of the Thirty-Second Conference on Learning Theory*. *Proceedings of Machine Learning Research* **99** 3120–3159. Phoenix, USA. Available at http://proceedings.mlr.press/v99/wolfer19a.html.
[32] Yu, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam* 423–435. New York: Springer. MR1462963

Bernoulli **27**(1), 2021, 554–575
https://doi.org/10.3150/20-BEJ1249

Optimal covariance change point localization in high dimensions

DAREN WANG[1], YI YU[2] and ALESSANDRO RINALDO[3]

[1]*Department of Statistics, University of Chicago, 5747 S. Ellis Avenue, Jones 120A, Chicago, IL 60637, USA.*
E-mail: darenw@galton.uchicago.edu
[2]*Department of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom.*
E-mail: yi.yu.2@warwick.ac.uk
[3]*Department of Statistics and Data Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA.*
E-mail: arinaldo@cmu.edu

We study the problem of change point localization for covariance matrices in high dimensions. We assume that we observe a sequence of independent and centered p-dimensional sub-Gaussian random vectors whose covariance matrices are piecewise constant, and only change at unknown times. We are concerned with the localization task of estimating the positions of the change points. In our analysis, we allow for all the model parameters to change with the sample size n, including the dimension p, the minimal spacing between consecutive change points Δ, the maximal Orlicz-ψ_2 norm B of the sample points and the magnitude κ of the smallest distributional change, defined as the minimal operator norm of the difference between the covariance matrix at a change point and the covariance matrix at the previous time point.

We introduce two procedures, one based on the binary segmentation algorithm and the other on its popular extension known as wild binary segmentation, and demonstrate that, under suitable conditions, both procedures can consistently estimate the change points. In particular, our second algorithm, called Wild Binary Segmentation through Independent Projection (WBSIP), delivers a localization error of order $B^4\kappa^{-2}\log(n)$, which is shown to be minimax rate optimal, save, possibly, for the $\log(n)$ term. WBSIP requires the model parameters to satisfy the scaling $\Delta\kappa^2 \gtrsim pB^4\log^{1+\xi}(n)$, for any $\xi > 0$, which we demonstrate to be essentially necessary, in the sense that no algorithm can guarantee consistent localization if $\Delta\kappa^2 \lesssim pB^4$. This result reveals an interesting phase transition effect separating parameter combinations for which the localization task is feasible from the ones for which it is not.

Keywords: binary segmentation; change point detection; high-dimensional covariance testing; independent projection; minimax optimal; wild binary segmentation

References

[1] Anderson, T.W. (2003). *An Introduction to Multivariate Statistical Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR1990662

[2] Aston, J.A.D. and Kirch, C. (2014). Efficiency of change point tests in high dimensional settings. Preprint. Available at arXiv:1409.1771.

[3] Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087. MR2572452 https://doi.org/10.1214/09-AOS707

[4] Avanesov, V. and Buzun, N. (2018). Change-point detection in high-dimensional covariance structure. *Electron. J. Stat.* **12** 3254–3294. MR3861282 https://doi.org/10.1214/18-EJS1484

[5] Baranowski, R., Chen, Y. and Fryzlewicz, P. (2019). Narrowest-over-threshold detection of multiple change points and change-point-like features. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 649–672. MR3961502 https://doi.org/10.1111/rssb.12322

1350-7265

[6] Barigozzi, M., Cho, H. and Fryzlewicz, P. (2018). Simultaneous multiple change-point and factor analysis for high-dimensional time series. *J. Econometrics* **206** 187–225. MR3840788 https://doi.org/10.1016/j.jeconom.2018.05.003
[7] Berkes, I., Gabrys, R., Horváth, L. and Kokoszka, P. (2009). Detecting changes in the mean of functional observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 927–946. MR2750251 https://doi.org/10.1111/j.1467-9868.2009.00713.x
[8] Birke, M. and Dette, H. (2005). A note on testing the covariance matrix for large dimension. *Statist. Probab. Lett.* **74** 281–289. MR2189467 https://doi.org/10.1016/j.spl.2005.04.051
[9] Boysen, L., Kempe, A., Liebscher, V., Munk, A. and Wittich, O. (2009). Consistencies and rates of convergence of jump-penalized least squares estimators. *Ann. Statist.* **37** 157–183. MR2488348 https://doi.org/10.1214/07-AOS558
[10] Cai, T., Liu, W. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.* **108** 265–277. MR3174618 https://doi.org/10.1080/01621459.2012.758041
[11] Cai, T.T. and Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* **19** 2359–2388. MR3160557 https://doi.org/10.3150/12-BEJ455
[12] Cho, H. (2016). Change-point detection in panel data via double CUSUM statistic. *Electron. J. Stat.* **10** 2000–2038. MR3522667 https://doi.org/10.1214/16-EJS1155
[13] Cho, H. and Fryzlewicz, P. (2012). Multiscale and multilevel technique for consistent segmentation of nonstationary time series. *Statist. Sinica* **22** 207–229. MR2933173 https://doi.org/10.5705/ss.2009.280
[14] Cho, H. and Fryzlewicz, P. (2015). Multiple-change-point detection for high dimensional time series via sparsified binary segmentation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 475–507. MR3310536 https://doi.org/10.1111/rssb.12079
[15] Davies, P.L. and Kovac, A. (2001). Local extremes, runs, strings and multiresolution. *Ann. Statist.* **29** 1–65. MR1833958 https://doi.org/10.1214/aos/996986500
[16] Davis, R.A., Lee, T.C.M. and Rodriguez-Yam, G.A. (2006). Structural break estimation for nonstationary time series models. *J. Amer. Statist. Assoc.* **101** 223–239. MR2268041 https://doi.org/10.1198/016214505000000745
[17] Dette, H., Pan, G. and Yang, Q. (2018). Estimating a change point in a sequence of very high-dimensional covariance matrices. Preprint. Available at arXiv:1807.10797.
[18] Frick, K., Munk, A. and Sieling, H. (2014). Multiscale change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 495–580. MR3210728 https://doi.org/10.1111/rssb.12047
[19] Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.* **42** 2243–2281. MR3269979 https://doi.org/10.1214/14-AOS1245
[20] Gombay, E., Horváth, L. and Hušková, M. (1996). Estimators and tests for change in variances. *Statist. Decisions* **14** 145–159. MR1406521
[21] Harchaoui, Z. and Lévy-Leduc, C. (2010). Multiple change-point estimation with a total variation penalty. *J. Amer. Statist. Assoc.* **105** 1480–1493. MR2796565 https://doi.org/10.1198/jasa.2010.tm09181
[22] Horváth, L. and Hušková, M. (2012). Change-point detection in panel data. *J. Time Series Anal.* **33** 631–648. MR2944843 https://doi.org/10.1111/j.1467-9892.2012.00796.x
[23] Inclán, C. and Tiao, G.C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *J. Amer. Statist. Assoc.* **89** 913–923. MR1294735
[24] Jirak, M. (2015). Uniform change point tests in high dimension. *Ann. Statist.* **43** 2451–2483. MR3405600 https://doi.org/10.1214/15-AOS1347
[25] Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961 https://doi.org/10.1214/aos/1009210544
[26] Killick, R., Fearnhead, P. and Eckley, I.A. (2012). Optimal detection of changepoints with a linear computational cost. *J. Amer. Statist. Assoc.* **107** 1590–1598. MR3036418 https://doi.org/10.1080/01621459.2012.737745
[27] Lavielle, M. (1999). Detection of multiple changes in a sequence of dependent variables. *Stochastic Process. Appl.* **83** 79–102. MR1705601 https://doi.org/10.1016/S0304-4149(99)00023-X
[28] Lavielle, M. and Moulines, E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *J. Time Series Anal.* **21** 33–59. MR1766173 https://doi.org/10.1111/1467-9892.00172

[29] Li, H., Guo, Q. and Munk, A. (2019). Multiscale change-point segmentation: Beyond step functions. *Electron. J. Stat.* **13** 3254–3296. MR4010980 https://doi.org/10.1214/19-ejs1608
[30] Li, J. and Chen, S.X. (2012). Two sample tests for high-dimensional covariance matrices. *Ann. Statist.* **40** 908–940. MR2985938 https://doi.org/10.1214/12-AOS993
[31] Lin, K., Sharpnack, J.L., Rinaldo, A. and Tibshirani, R.J. (2017). A sharp error analysis for the fused lasso, with application to approximate changepoint screening. In *Advances in Neural Information Processing Systems* 30 (I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds.) 6887–6896.
[32] Olshen, A.B., Venkatraman, E., Lucito, R. and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5** 557–572.
[33] Page, E.S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115. MR0088850 https://doi.org/10.1093/biomet/41.1-2.100
[34] Picard, D. (1985). Testing and estimating change-points in time series. *Adv. in Appl. Probab.* **17** 841–867. MR0809433 https://doi.org/10.2307/1427090
[35] Qian, J. and Jia, J. (2012). On pattern recovery of the fused lasso. Technical Report. Available at arXiv:1211.5194.
[36] Rojas, C.R. and Wahlberg, B. (2014). On change point detection using the fused lasso method. Technical Report. Available at arXiv:1401.5408.
[37] Venkatraman, E.S. (1992). Consistency results in multiple change-point problems. Ph.D. thesis, Stanford Univ. MR2687536
[38] Vostrikova, L.J. (1981). Discovery of "discord" in multidimensional random processes. *Dokl. Akad. Nauk SSSR* **259** 270–274. MR0625215
[39] Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Stat.* **16** 117–186. MR0013275 https://doi.org/10.1214/aoms/1177731118
[40] Wang, D., Yu, Y. and Rinaldo, A. (2020). Univariate mean change point detection: Penalization, CUSUM and optimality. *Electron. J. Stat.* **14** 1917–1961. MR4091859 https://doi.org/10.1214/20-EJS1710
[41] Wang, D., Yu, Y. and Rinaldo, A. (2020). Supplement to "Optimal covariance change point localization in high dimensions." https://doi.org/10.3150/20-BEJ1249SUPP
[42] Wang, T. and Samworth, R.J. (2018). High dimensional change point estimation via sparse projection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 57–83. MR3744712 https://doi.org/10.1111/rssb.12243
[43] Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika* **82** 385–397. MR1354236 https://doi.org/10.1093/biomet/82.2.385
[44] Yao, Y.-C. and Au, S.T. (1989). Least-squares estimation of a step function. *Sankhyā Ser. A* **51** 370–381. MR1175613

Bernoulli **27**(1), 2021, 576–585
https://doi.org/10.3150/20-BEJ1250

Minimal $\mathcal{L}^p$-densities with prescribed marginals

PAOLO GUASONI[1], EBERHARD MAYERHOFER[2,*] and MINGCHUAN ZHAO[2,†]

[1]*Dublin City University, School of Mathematical Sciences, Glasnevin, Dublin, Ireland.*
E-mail: paolo.guasoni@dcu.ie
[2]*University of Limerick, Department of Mathematics and Statistics, Castletroy, Ireland.*
*E-mail: *eberhard.mayerhofer@ul.ie; †mingchuan.zhao@ul.ie*

We derive sharp lower bounds for $\mathcal{L}^p$-functions on the n-dimensional unit hypercube in terms of their p-ths marginal moments. Such bounds are the unique solutions of a system of constrained nonlinear integral equations depending on the marginals. For square-integrable functions, the bounds have an explicit expression in terms of the second marginals moments.

Keywords: Banach spaces; integral equations; multivariate distributions; sharp estimates

References

[1] Breeden, D.T. and Litzenberger, R.H. (1978). Prices of state-contingent claims implicit in option prices. *J. Bus.* 621–651.
[2] Carr, P. and Madan, D. (2001). Towards a theory of volatility trading. In *Option Pricing*, *Interest Rates and Risk Management*. *Handb. Math. Finance* 458–476. Cambridge: Cambridge Univ. Press. MR1848559 https://doi.org/10.1017/CBO9780511569708.013
[3] Green, R.C. and Jarrow, R.A. (1987). Spanning and completeness in markets with contingent claims. *J. Econom. Theory* **41** 202–210. MR0875990 https://doi.org/10.1016/0022-0531(87)90014-7
[4] Guasoni, P. and Mayerhofer, E. (2020). Technical note – options portfolio selection. *Oper. Res.* **68** 733–740. MR4105244 https://doi.org/10.1287/opre.2019.1925
[5] Megginson, R.E. (2012). *An Introduction to Banach Space Theory* **183**. Berlin: Springer.
[6] Nachman, D.C. (1988). Spanning and completeness with options. *Rev. Financ. Stud.* **1** 311–328.
[7] Nelsen, R.B. (2007). *An Introduction to Copulas*. Berlin: Springer.
[8] Rudin, W. (2006). *Real and Complex Analysis*. New York: McGraw-Hill.
[9] Shapiro, H.S. (2006). *Topics in Approximation Theory* **187**. Berlin: Springer.

Bernoulli **27**(1), 2021, 586–614
https://doi.org/10.3150/20-BEJ1251

Hanson–Wright inequality in Hilbert spaces with application to K-means clustering for non-Euclidean data

XIAOHUI CHEN[1,2,*] and YUN YANG[1,†]

[1]*Department of Statistics, University of Illinois at Urbana-Champaign, 725 S. Wright Street, Champaign, IL 61820, USA.*
E-mail: [*]*xhchen@illinois.edu;* [†]*yy84@illinois.edu*
[2]*Institute for Data, Systems, and Society, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, USA.*
E-mail: xiaohui@mit.edu

We derive a dimension-free Hanson–Wright inequality for quadratic forms of independent sub-gaussian random variables in a separable Hilbert space. Our inequality is an infinite-dimensional generalization of the classical Hanson–Wright inequality for finite-dimensional Euclidean random vectors. We illustrate an application to the generalized K-means clustering problem for non-Euclidean data. Specifically, we establish the exponential rate of convergence for a semidefinite relaxation of the generalized K-means, which together with a simple rounding algorithm imply the exact recovery of the true clustering structure.

Keywords: Hanson–Wright inequality; Hilbert space; K-means; semidefinite relaxation

References

[1] Adamczak, R. (2015). A note on the Hanson–Wright inequality for random vectors with dependencies. *Electron. Commun. Probab.* **20** Art. ID 72. MR3407216 https://doi.org/10.1214/ECP.v20-3829

[2] Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404. MR0051437 https://doi.org/10.2307/1990404

[3] Barthe, F. and Milman, E. (2013). Transference principles for log-Sobolev and spectral-gap with applications to conservative spin systems. *Comm. Math. Phys.* **323** 575–625. MR3096532 https://doi.org/10.1007/s00220-013-1782-2

[4] Bunea, F., Giraud, C., Royer, M. and Verzelen, N. (2016). PECOK: A convex optimization approach to variable clustering. Available at arXiv:1606.05100.

[5] Chen, X. (2018). Gaussian and bootstrap approximations for high-dimensional U-statistics and their applications. *Ann. Statist.* **46** 642–678. MR3782380 https://doi.org/10.1214/17-AOS1563

[6] Fei, Y. and Chen, Y. (2018). Hidden integrality of SDP relaxation for sub-Gaussian mixture models. Available at arXiv:1803.06510.

[7] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis: Theory and Practice. Springer Series in Statistics*. New York: Springer. MR2229687

[8] Filippone, M., Camastra, F., Masulli, F. and Rovetta, S. (2008). A survey of kernel and spectral methods for clustering. *Pattern Recognit.* **41** 176–190.

[9] Foucart, S. and Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing. Applied and Numerical Harmonic Analysis*. New York: Birkhäuser/Springer. MR3100033 https://doi.org/10.1007/978-0-8176-4948-7

[10] Giraud, C. and Verzelen, N. (2018). Partial recovery bounds for clustering with the relaxed K means. Available at arXiv:1807.07547.

[11] Giuliano Antonini, R. (1997). Sub-Gaussian random variables in Hilbert spaces. *Rend. Semin. Mat. Univ. Padova* **98** 89–99. MR1492970

[12] Hanson, D.L. and Wright, F.T. (1971). A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Stat.* **42** 1079–1083. MR0279864 https://doi.org/10.1214/aoms/1177693335

[13] Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators. Wiley Series in Probability and Statistics*. Chichester: Wiley. MR3379106 https://doi.org/10.1002/9781118762547

[14] Hsu, D., Kakade, S.M. and Zhang, T. (2012). A tail inequality for quadratic forms of subgaussian random vectors. *Electron. Commun. Probab.* **17** Art. ID 52. MR2994877 https://doi.org/10.1214/ECP.v17-2079

[15] Ieva, F., Paganoni, A.M., Pigoli, D. and Vitelli, V. (2013). Multivariate functional clustering for the morphological analysis of electrocardiograph curves. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **62** 401–418. MR3060623 https://doi.org/10.1111/j.1467-9876.2012.01062.x

[16] Laurent, B. and Massart, P. (2000). Adaptive estimation of a quadratic functional by model selection. *Ann. Statist.* **28** 1302–1338. MR1805785 https://doi.org/10.1214/aos/1015957395

[17] Li, X., Li, Y., Ling, S., Strohmer, T. and Wei, K. (2020). When do birds of a feather flock together? k-means, proximity, and conic programming. *Math. Program.* **179** 295–341. MR4050142 https://doi.org/10.1007/s10107-018-1333-x

[18] Lloyd, S.P. (1982). Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28** 129–137. MR0651807 https://doi.org/10.1109/TIT.1982.1056489

[19] Peng, J. and Wei, Y. (2007). Approximating K-means-type clustering via semidefinite programming. *SIAM J. Optim.* **18** 186–205. MR2299680 https://doi.org/10.1137/050641983

[20] Rasmussen, C.E. (2004). Gaussian processes in machine learning. In *Advanced Lectures on Machine Learning* 63–71. Berlin: Springer.

[21] Royer, M. (2017). Adaptive clustering through semidefinite programming. In *Advances in Neural Information Processing Systems* 30 (I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds.) 1795–1803. Red Hook, NY: Curran Associates.

[22] Rudelson, M. and Vershynin, R. (2013). Hanson–Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.* **18** Art. ID 82. MR3125258 https://doi.org/10.1214/ECP.v18-2865

[23] Schölkopf, B. and Smola, A. (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, 1st ed. Cambridge, MA: MIT Press.

[24] Song, L., Smola, A., Gretton, A. and Borgwardt, K. (2007). A dependence maximization view of clustering. In *Proceedings of the* 24*th International Conference on Machine Learning*.

[25] Tarpey, T. and Kinateder, K.K.J. (2003). Clustering functional data. *J. Classification* **20** 93–114. MR1983123 https://doi.org/10.1007/s00357-003-0007-3

[26] van der Vaart, A.W. and van Zanten, J.H. (2008). Reproducing kernel Hilbert spaces of Gaussian priors. In *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh. Inst. Math. Stat.* (*IMS*) *Collect.* **3** 200–222. Beachwood, OH: IMS. MR2459226 https://doi.org/10.1214/074921708000000156

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

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

[29] Vu, V. and Wang, K. (2015). Random weighted projections, random quadratic forms and random eigenvectors. *Random Structures Algorithms* **47** 792–821. MR3418916 https://doi.org/10.1002/rsa.20561

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

[31] Wright, F.T. (1973). A bound on tail probabilities for quadratic forms in independent random variables whose distributions are not necessarily symmetric. *Ann. Probab.* **1** 1068–1070. MR0353419 https://doi.org/10.1214/aop/1176996815

Bernoulli **27**(1), 2021, 615–636
https://doi.org/10.3150/20-BEJ1252

An invariance principle for biased voter model interfaces

RONGFENG SUN[1], JAN M. SWART[2] and JINJIONG YU[3]

[1]*Department of Mathematics, National University of Singapore, 10 Lower Kent Ridge Road, 119076 Singapore.*
E-mail: matsr@nus.edu.sg
[2]*The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod vodárenskou věží 4, 18200 Praha 8, Czech Republic.*
E-mail: swart@utia.cas.cz
[3]*NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China.*
E-mail: jinjiongyu@nyu.edu

We consider one-dimensional biased voter models, where 1's replace 0's at a faster rate than the other way round, started in a Heaviside initial state describing the interface between two infinite populations of 0's and 1's. In the limit of weak bias, for a diffusively rescaled process, we consider a measure-valued process describing the local fraction of type 1 sites as a function of time. Under a finite second moment condition on the rates, we show that in the diffusive scaling limit there is a drifted Brownian path with the property that all but a vanishingly small fraction of the sites on the left (resp. right) of this path are of type 0 (resp. 1). This extends known results for unbiased voter models. Our proofs depend crucially on recent results about interface tightness for biased voter models.

Keywords: biased voter model; branching and coalescing random walks; interface tightness; invariance principle

References

[1] Aldous, D. (1978). Stopping times and tightness. *Ann. Probab.* **6** 335–340. MR0474446 https://doi.org/10.1214/aop/1176995579
[2] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. *Applications of Mathematics* (*New York*): *Stochastic Modelling and Applied Probability* **51**. New York: Springer. MR1978607
[3] Athreya, S.R. and Sun, R. (2011). One-dimensional voter model interface revisited. *Electron. Commun. Probab.* **16** 792–800. MR2868600 https://doi.org/10.1214/ECP.v16-1688
[4] Belhaouari, S., Mountford, T., Sun, R. and Valle, G. (2006). Convergence results and sharp estimates for the voter model interfaces. *Electron. J. Probab.* **11** 768–801. MR2242663 https://doi.org/10.1214/EJP.v11-349
[5] Belhaouari, S., Mountford, T. and Valle, G. (2007). Tightness for the interfaces of one-dimensional voter models. *Proc. Lond. Math. Soc.* (3) **94** 421–442. MR2308233 https://doi.org/10.1112/plms/pdl016
[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1700749 https://doi.org/10.1002/9780470316962
[7] Cox, J.T. and Durrett, R. (1995). Hybrid zones and voter model interfaces. *Bernoulli* **1** 343–370. MR1369166 https://doi.org/10.2307/3318488
[8] Dawson, D.A. (1993). Measure-valued Markov processes. In *École d'Été de Probabilités de Saint-Flour XXI*—1991. *Lecture Notes in Math.* **1541** 1–260. Berlin: Springer. MR1242575 https://doi.org/10.1007/BFb0084190
[9] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes*: *Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658
[10] Fontes, L.R.G., Isopi, M., Newman, C.M. and Ravishankar, K. (2004). The Brownian web: Characterization and convergence. *Ann. Probab.* **32** 2857–2883. MR2094432 https://doi.org/10.1214/009117904000000568

1350-7265

[11] Newman, C.M., Ravishankar, K. and Sun, R. (2005). Convergence of coalescing nonsimple random walks to the Brownian web. *Electron*. *J*. *Probab.* **10** 21–60. MR2120239 https://doi.org/10.1214/EJP.v10-235
[12] Sun, R. and Swart, J.M. (2008). The Brownian net. *Ann*. *Probab.* **36** 1153–1208. MR2408586 https://doi.org/10.1214/07-AOP357
[13] Sun, R., Swart, J.M. and Yu, J. (2019). Equilibrium interfaces of biased voter models. *Ann*. *Appl*. *Probab.* **29** 2556–2593. MR3984257 https://doi.org/10.1214/19-AAP1461

Bernoulli **27**(1), 2021, 637–672
https://doi.org/10.3150/20-BEJ1253

Bayesian graph selection consistency under model misspecification

YABO NIU[*], DEBDEEP PATI[†] and BANI K. MALLICK[‡]

Department of Statistics, Texas A&M University, College Station, TX, USA.
E-mail: [*]*ybniu@stat.tamu.edu;* [†]*debdeep@stat.tamu.edu;* [‡]*bmallick@stat.tamu.edu*

Gaussian graphical models are a popular tool to learn the dependence structure in the form of a graph among variables of interest. Bayesian methods have gained in popularity in the last two decades due to their ability to simultaneously learn the covariance and the graph. There is a wide variety of model-based methods to learn the underlying graph assuming various forms of the graphical structure. Although for scalability of the Markov chain Monte Carlo algorithms, decomposability is commonly imposed on the graph space, its possible implication on the posterior distribution of the graph is not clear. *An open problem* in Bayesian decomposable structure learning is whether the posterior distribution is able to select a meaningful decomposable graph that is "close" to the true non-decomposable graph, when the dimension of the variables increases with the sample size. In this article, we explore specific conditions on the true precision matrix and the graph, which results in an affirmative answer to this question with a commonly used hyper-inverse Wishart prior on the covariance matrix and a suitable complexity prior on the graph space. In absence of structural sparsity assumptions, our strong selection consistency holds in a high-dimensional setting where $p = O(n^{\alpha})$ for $\alpha < 1/3$. We show when the true graph is non-decomposable, the posterior distribution concentrates on a set of graphs that are *minimal triangulations* of the true graph.

Keywords: decomposable graph; Gaussian graphical model; graph selection consistency; hyper-inverse Wishart distribution; minimal triangulation; model misspecification; partial correlation

References

[1] Armstrong, H., Carter, C.K., Wong, K.F.K. and Kohn, R. (2009). Bayesian covariance matrix estimation using a mixture of decomposable graphical models. *Stat. Comput.* **19** 303–316. MR2516221 https://doi.org/10.1007/s11222-008-9093-8

[2] Atay-Kayis, A. and Massam, H. (2005). A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. *Biometrika* **92** 317–335. MR2201362 https://doi.org/10.1093/biomet/92.2.317

[3] Banerjee, S. and Ghosal, S. (2014). Posterior convergence rates for estimating large precision matrices using graphical models. *Electron. J. Stat.* **8** 2111–2137. MR3273620 https://doi.org/10.1214/14-EJS945

[4] Banerjee, S. and Ghosal, S. (2015). Bayesian structure learning in graphical models. *J. Multivariate Anal.* **136** 147–162. MR3321485 https://doi.org/10.1016/j.jmva.2015.01.015

[5] Ben-David, E., Li, T., Massam, H. and Rajaratnam, B. (2011). High dimensional Bayesian inference for Gaussian directed acyclic graph models. Preprint. Available at arXiv:1109.4371.

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

[7] Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **106** 672–684. MR2847949 https://doi.org/10.1198/jasa.2011.tm10560

[8] Cao, X., Khare, K. and Ghosh, M. (2019). Posterior graph selection and estimation consistency for high-dimensional Bayesian DAG models. *Ann. Statist.* **47** 319–348. MR3909935 https://doi.org/10.1214/18-AOS1689

[9] Carvalho, C.M., Massam, H. and West, M. (2007). Simulation of hyper-inverse Wishart distributions in graphical models. *Biometrika* **94** 647–659. MR2410014 https://doi.org/10.1093/biomet/asm056

[10] Carvalho, C.M. and Scott, J.G. (2009). Objective Bayesian model selection in Gaussian graphical models. *Biometrika* **96** 497–512. MR2538753 https://doi.org/10.1093/biomet/asp017

[11] Dawid, A.P. and Lauritzen, S.L. (1993). Hyper-Markov laws in the statistical analysis of decomposable graphical models. *Ann. Statist.* **21** 1272–1317. MR1241267 https://doi.org/10.1214/aos/1176349260

[12] Dellaportas, P., Giudici, P. and Roberts, G. (2003). Bayesian inference for nondecomposable graphical Gaussian models. *Sankhyā* **65** 43–55. MR2016776

[13] Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281. MR0520238

[14] Dobra, A., Hans, C., Jones, B., Nevins, J.R., Yao, G. and West, M. (2004). Sparse graphical models for exploring gene expression data. *J. Multivariate Anal.* **90** 196–212. MR2064941 https://doi.org/10.1016/j.jmva.2004.02.009

[15] Donnet, S. and Marin, J.-M. (2012). An empirical Bayes procedure for the selection of Gaussian graphical models. *Stat. Comput.* **22** 1113–1123. MR2950089 https://doi.org/10.1007/s11222-011-9285-5

[16] Drton, M. and Perlman, M.D. (2007). Multiple testing and error control in Gaussian graphical model selection. *Statist. Sci.* **22** 430–449. MR2416818 https://doi.org/10.1214/088342307000000113

[17] El Karoui, N. (2008). Operator norm consistent estimation of large-dimensional sparse covariance matrices. *Ann. Statist.* **36** 2717–2756. MR2485011 https://doi.org/10.1214/07-AOS559

[18] Fitch, A.M., Jones, M.B. and Massam, H. (2014). The performance of covariance selection methods that consider decomposable models only. *Bayesian Anal.* **9** 659–684. MR3256059 https://doi.org/10.1214/14-BA874

[19] Giudici, P. (1996). Learning in graphical Gaussian models. In *Bayesian Statistics*, 5 (*Alicante*, 1994). *Oxford Sci. Publ.* 621–628. New York: Oxford Univ. Press. MR1425431

[20] Giudici, P. and Green, P.J. (1999). Decomposable graphical Gaussian model determination. *Biometrika* **86** 785–801. MR1741977 https://doi.org/10.1093/biomet/86.4.785

[21] Green, P.J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika* **82** 711–732. MR1380810 https://doi.org/10.1093/biomet/82.4.711

[22] Heggernes, P. (2006). Minimal triangulations of graphs: A survey. *Discrete Math.* **306** 297–317. MR2204109 https://doi.org/10.1016/j.disc.2005.12.003

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

[24] Jones, B., Carvalho, C., Dobra, A., Hans, C., Carter, C. and West, M. (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statist. Sci.* **20** 388–400. MR2210226 https://doi.org/10.1214/088342305000000304

[25] Khare, K., Rajaratnam, B. and Saha, A. (2018). Bayesian inference for Gaussian graphical models beyond decomposable graphs. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 727–747. MR3849341 https://doi.org/10.1111/rssb.12276

[26] Lam, C. and Fan, J. (2009). Sparsistency and rates of convergence in large covariance matrix estimation. *Ann. Statist.* **37** 4254–4278. MR2572459 https://doi.org/10.1214/09-AOS720

[27] Lauritzen, S.L. (1996). *Graphical Models*. *Oxford Statistical Science Series* **17**. New York: The Clarendon Press. MR1419991

[28] Lee, K., Lee, J. and Lin, L. (2019). Minimax posterior convergence rates and model selection consistency in high-dimensional DAG models based on sparse Cholesky factors. *Ann. Statist.* **47** 3413–3437. MR4025747 https://doi.org/10.1214/18-AOS1783

[29] Letac, G. and Massam, H. (2007). Wishart distributions for decomposable graphs. *Ann. Statist.* **35** 1278–1323. MR2341706 https://doi.org/10.1214/009053606000001235

[30] Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. MR2278363 https://doi.org/10.1214/009053606000000281

[31] Moghaddam, B., Khan, E., Murphy, K.P. and Marlin, B.M. (2009). Accelerating Bayesian structural inference for non-decomposable Gaussian graphical models. In *Advances in Neural Information Processing Systems* 1285–1293.

[32] Niu, Y., Pati, D. and Mallick, B.K. (2020). Supplement to "Bayesian graph selection consistency under model misspecification." https://doi.org/10.3150/20-BEJ1253SUPP

[33] Rajaratnam, B., Massam, H. and Carvalho, C.M. (2008). Flexible covariance estimation in graphical Gaussian models. *Ann. Statist.* **36** 2818–2849. MR2485014 https://doi.org/10.1214/08-AOS619

[34] Raskutti, G., Yu, B., Wainwright, M.J. and Ravikumar, P.K. (2009). Model selection in Gaussian graphical models: High-dimensional consistency of lregularized MLE. In *Advances in Neural Information Processing Systems* 1329–1336.

[35] Rose, D.J., Tarjan, R.E. and Lueker, G.S. (1976). Algorithmic aspects of vertex elimination on graphs. *SIAM J. Comput.* **5** 266–283. MR0408312 https://doi.org/10.1137/0205021

[36] Roverato, A. (2000). Cholesky decomposition of a hyper inverse Wishart matrix. *Biometrika* **87** 99–112. MR1766831 https://doi.org/10.1093/biomet/87.1.99

[37] Roverato, A. (2002). Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. *Scand. J. Stat.* **29** 391–411. MR1925566 https://doi.org/10.1111/1467-9469.00297

[38] Scott, J.G. and Carvalho, C.M. (2008). Feature-inclusion stochastic search for Gaussian graphical models. *J. Comput. Graph. Statist.* **17** 790–808. MR2649067 https://doi.org/10.1198/106186008X382683

[39] Spokoiny, V. (2013). Bernstein–von Mises theorem for growing parameter dimension. Preprint. Available at arXiv:1302.3430.

[40] Uhler, C., Lenkoski, A. and Richards, D. (2018). Exact formulas for the normalizing constants of Wishart distributions for graphical models. *Ann. Statist.* **46** 90–118. MR3766947 https://doi.org/10.1214/17-AOS1543

[41] Wang, H. and Carvalho, C.M. (2010). Simulation of hyper-inverse Wishart distributions for non-decomposable graphs. *Electron. J. Stat.* **4** 1470–1475. MR2741209 https://doi.org/10.1214/10-EJS591

[42] Yuan, M. and Lin, Y. (2007). Model selection and estimation in the Gaussian graphical model. *Biometrika* **94** 19–35. MR2367824 https://doi.org/10.1093/biomet/asm018

Bernoulli **27**(1), 2021, 673–701
https://doi.org/10.3150/20-BEJ1254

Flexible integrated functional depths

STANISLAV NAGY[1], SAMI HELANDER[2,*], GERMAIN VAN BEVER[3], LAURI VIITASAARI[4] and PAULIINA ILMONEN[2,†]

[1]*Charles University, Prague, Czech Republic.*
E-mail: nagy@karlin.mff.cuni.cz
[2]*Aalto University School of Science, Espoo, Finland.*
*E-mail: *sami.helander@aalto.fi; †pauliina.ilmonen@aalto.fi*
[3]*University of Namur, Namur, Belgium.*
E-mail: germain.vanbever@unamur.be
[4]*Aalto University School of Business, Espoo, Finland.*
E-mail: lauri.viitasaari@aalto.fi

This paper develops a new class of functional depths. A generic member of this class is coined Jth order kth moment integrated depth. It is based on the distribution of the cross-sectional halfspace depth of a function in the marginal evaluations (in time) of the random process. Asymptotic properties of the proposed depths are provided: we show that they are uniformly consistent and satisfy an inequality related to the law of the iterated logarithm. Moreover, limiting distributions are derived under mild regularity assumptions. The versatility displayed by the new class of depths makes them particularly amenable for capturing important features of functional distributions. This is illustrated in supervised learning, where we show that the corresponding maximum depth classifiers outperform classical competitors.

Keywords: asymptotics; data depth; functional data analysis; integrated depths; supervised classification

References

[1] Biau, G., Bunea, F. and Wegkamp, M.H. (2005). Functional classification in Hilbert spaces. *IEEE Trans. Inf. Theory* **51** 2163–2172. MR2235289 https://doi.org/10.1109/TIT.2005.847705

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

[3] Borggaard, C. and Thodberg, H.H. (1992). Optimal minimal neural interpretation of spectra. *Anal. Chem.* **64** 545–551.

[4] Cérou, F. and Guyader, A. (2006). Nearest neighbor classification in infinite dimension. *ESAIM Probab. Stat.* **10** 340–355. MR2247925 https://doi.org/10.1051/ps:2006014

[5] Chakraborty, A. and Chaudhuri, P. (2014). The spatial distribution in infinite dimensional spaces and related quantiles and depths. *Ann. Statist.* **42** 1203–1231. MR3224286 https://doi.org/10.1214/14-AOS1226

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

[7] Claeskens, G., Hubert, M., Slaets, L. and Vakili, K. (2014). Multivariate functional halfspace depth. *J. Amer. Statist. Assoc.* **109** 411–423. MR3180573 https://doi.org/10.1080/01621459.2013.856795

[8] Cuevas, A., Febrero, M. and Fraiman, R. (2007). Robust estimation and classification for functional data via projection-based depth notions. *Comput. Statist.* **22** 481–496. MR2336349 https://doi.org/10.1007/s00180-007-0053-0

[9] Cuevas, A. and Fraiman, R. (2009). On depth measures and dual statistics. A methodology for dealing with general data. *J. Multivariate Anal.* **100** 753–766. MR2478196 https://doi.org/10.1016/j.jmva.2008.08.002

[10] Delaigle, A. and Hall, P. (2012). Achieving near perfect classification for functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 267–286. MR2899863 https://doi.org/10.1111/j.1467-9868.2011.01003.x

[11] Donoho, D.L. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Ann. Statist.* **20** 1803–1827. MR1193313 https://doi.org/10.1214/aos/1176348890
[12] 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
[13] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358 https://doi.org/10.1017/CBO9780511755347
[14] Ferraty, F. and Vieu, P. (2003). Curves discrimination: A nonparametric functional approach **44** 161–173. MR2020144 https://doi.org/10.1016/S0167-9473(03)00032-X
[15] Ferraty, F. and Vieu, P. (2006). *Nonparametric Functional Data Analysis*: *Theory and Practice*. *Springer Series in Statistics*. New York: Springer. MR2229687
[16] Fraiman, R. and Muniz, G. (2001). Trimmed means for functional data. *TEST* **10** 419–440. MR1881149 https://doi.org/10.1007/BF02595706
[17] Ghosh, A.K. and Chaudhuri, P. (2005). On maximum depth and related classifiers. *Scand. J. Stat.* **32** 327–350. MR2188677 https://doi.org/10.1111/j.1467-9469.2005.00423.x
[18] Gijbels, I. and Nagy, S. (2015). Consistency of non-integrated depths for functional data. *J. Multivariate Anal.* **140** 259–282. MR3372567 https://doi.org/10.1016/j.jmva.2015.05.012
[19] Grinblat, L.Š. (1976). A limit theorem for measurable random processes and its applications. *Proc. Amer. Math. Soc.* **61** 371–376. MR0423450 https://doi.org/10.2307/2041344
[20] Hall, P., Poskitt, D.S. and Presnell, B. (2001). A functional data-analytic approach to signal discrimination. *Technometrics* **43** 1–9. MR1847775 https://doi.org/10.1198/00401700152404273
[21] Hallin, M., Paindaveine, D. and Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: From L_1 optimization to halfspace depth. *Ann. Statist.* **38** 635–669. MR2604670 https://doi.org/10.1214/09-AOS723
[22] Kuelbs, J. and Dudley, R.M. (1980). Log log laws for empirical measures. *Ann. Probab.* **8** 405–418. MR0573282
[23] Kuelbs, J. and Zinn, J. (2013). Concerns with functional depth. *ALEA Lat. Am. J. Probab. Math. Stat.* **10** 831–855. MR3125749
[24] Li, B., Van Bever, G., Oja, H., Sabolova, R. and Critchley, F. (2019). Functional independent component analysis: An extension of fourth order blind identification. Technical Report, Univ. Namur.
[25] Li, J., Cuesta-Albertos, J.A. and Liu, R.Y. (2012). DD-classifier: Nonparametric classification procedure based on DD-plot. *J. Amer. Statist. Assoc.* **107** 737–753. MR2980081 https://doi.org/10.1080/01621459.2012.688462
[26] Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414. MR1041400 https://doi.org/10.1214/aos/1176347507
[27] Liu, R.Y., Parelius, J.M. and Singh, K. (1999). Multivariate analysis by data depth: Descriptive statistics, graphics and inference. *Ann. Statist.* **27** 783–858. MR1724033 https://doi.org/10.1214/aos/1018031260
[28] Liu, X., Mosler, K. and Mozharovskyi, P. (2019). Fast computation of Tukey trimmed regions and median in dimension $p > 2$. *J. Comput. Graph. Statist.* **28** 682–697. MR4007750 https://doi.org/10.1080/10618600.2018.1546595
[29] López-Pintado, S. and Romo, J. (2009). On the concept of depth for functional data. *J. Amer. Statist. Assoc.* **104** 718–734. MR2541590 https://doi.org/10.1198/jasa.2009.0108
[30] López-Pintado, S. and Romo, J. (2011). A half-region depth for functional data. *Comput. Statist. Data Anal.* **55** 1679–1695. MR2748671 https://doi.org/10.1016/j.csda.2010.10.024
[31] Magnano, L., Boland, J.W. and Hyndman, R.J. (2008). Generation of synthetic sequences of half-hourly temperature. *Environmetrics* **19** 818–835. MR2654607 https://doi.org/10.1002/env.905
[32] Massart, P. (1986). Rates of convergence in the central limit theorem for empirical processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** 381–423. MR0871904
[33] Massé, J.-C. (2004). Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean. *Bernoulli* **10** 397–419. MR2061438 https://doi.org/10.3150/bj/1089206404
[34] Mosler, K. (2013). Depth statistics. In *Robustness and Complex Data Structures* (C. Becker, R. Fried and S. Kuhnt, eds.) 17–34. Heidelberg: Springer. MR3135871 https://doi.org/10.1007/978-3-642-35494-6_2

[35] Mosler, K. and Polyakova, Y. (2016). General notions of depth for functional data. Preprint. Available at arXiv:1208.1981.
[36] Nagy, S. (2017). Integrated depth for measurable functions and sets. *Statist. Probab. Lett.* **123** 165–170. MR3598634 https://doi.org/10.1016/j.spl.2016.12.012
[37] Nagy, S. and Ferraty, F. (2019). Data depth for measurable noisy random functions. *J. Multivariate Anal.* **170** 95–114. MR3913030 https://doi.org/10.1016/j.jmva.2018.11.003
[38] Nagy, S., Gijbels, I. and Hlubinka, D. (2017). Depth-based recognition of shape outlying functions. *J. Comput. Graph. Statist.* **26** 883–893. MR3765352 https://doi.org/10.1080/10618600.2017.1336445
[39] Nagy, S., Gijbels, I., Omelka, M. and Hlubinka, D. (2016). Integrated depth for functional data: Statistical properties and consistency. *ESAIM Probab. Stat.* **20** 95–130. MR3528619 https://doi.org/10.1051/ps/2016005
[40] Nagy, S., Helander, S., Van Bever, G., Viitasaari, L. and Ilmonen, P. (2020). Supplement to "Flexible integrated functional depths." https://doi.org/10.3150/20-BEJ1254SUPP
[41] Narisetty, N.N. and Nair, V.N. (2016). Extremal depth for functional data and applications. *J. Amer. Statist. Assoc.* **111** 1705–1714. MR3601729 https://doi.org/10.1080/01621459.2015.1110033
[42] Pokotylo, O., Mozharovskyi, P. and Dyckerhoff, R. (2019). Depth and depth-based classification with R package ddalpha. *J. Stat. Softw.* **91** 1–46.
[43] Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993
[44] Sguera, C., Galeano, P. and Lillo, R. (2014). Spatial depth-based classification for functional data. *TEST* **23** 725–750. MR3274472 https://doi.org/10.1007/s11749-014-0379-1
[45] Shao, W. and Zuo, Y. (2020). Computing the halfspace depth with multiple try algorithm and simulated annealing algorithm. *Comput. Statist.* **35** 203–226.
[46] Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians* (*Vancouver, B. C.*, 1974), *Vol.* 2 523–531. MR0426989
[47] 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
[48] Zuo, Y. and Serfling, R. (2000). General notions of statistical depth function. *Ann. Statist.* **28** 461–482. MR1790005 https://doi.org/10.1214/aos/1016218226

Bernoulli **27**(1), 2021, 702–726
https://doi.org/10.3150/20-BEJ1255

A class of models for Bayesian predictive inference

PATRIZIA BERTI[1], EMANUELA DREASSI[2], LUCA PRATELLI[3] and PIETRO RIGO[4]

[1]*Dipartimento di Matematica Pura ed Applicata "G. Vitali", Università di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy.*
E-mail: patrizia.berti@unimore.it
[2]*Dipartimento di Statistica, Informatica, Applicazioni Università di Firenze, viale Morgagni 59, 50134 Firenze, Italy.*
E-mail: emanuela.dreassi@unifi.it
[3]*Accademia Navale di Livorno, viale Italia 72, 57100 Livorno, Italy.*
E-mail: pratel@mail.dm.unipi.it
[4]*Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, via delle Belle Arti 41, 40126 Bologna, Italy.*
E-mail: pietro.rigo@unibo.it

In a Bayesian framework, to make predictions on a sequence $X_1, X_2, \dots$ of random observations, the inferrer needs to assign the predictive distributions $\sigma_n(\cdot) = P(X_{n+1} \in \cdot \mid X_1, \dots, X_n)$. In this paper, we propose to assign σ_n directly, without passing through the usual prior/posterior scheme. One main advantage is that no prior probability has to be assessed. The data sequence (X_n) is assumed to be conditionally identically distributed (c.i.d.) in the sense of (*Ann. Probab.* **32** (2004) 2029–2052). To realize this programme, a class Σ of predictive distributions is introduced and investigated. Such a Σ is rich enough to model various real situations and (X_n) is actually c.i.d. if σ_n belongs to Σ. Furthermore, when a new observation X_{n+1} becomes available, σ_{n+1} can be obtained by a simple recursive update of σ_n. If μ is the a.s. weak limit of σ_n, conditions for μ to be a.s. discrete are provided as well.

Keywords: Bayesian nonparametrics; conditional identity in distribution; exchangeability; predictive distribution; random probability measure; sequential predictions; strategy

References

[1] Airoldi, E.M., Costa, T., Bassetti, F., Leisen, F. and Guindani, M. (2014). Generalized species sampling priors with latent beta reinforcements. *J. Amer. Statist. Assoc.* **109** 1466–1480. MR3293604 https://doi.org/10.1080/01621459.2014.950735

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

[3] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P. (2009). Rate of convergence of predictive distributions for dependent data. *Bernoulli* **15** 1351–1367. MR2597596 https://doi.org/10.3150/09-BEJ191

[4] Berti, P., Dreassi, E., Pratelli, L. and Rigo, P. (2020). Asymptotics of certain conditionally identically distributed sequences. *Statist. Probab. Lett.* To appear.

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

[6] Berti, P., Pratelli, L. and Rigo, P. (2012). Limit theorems for empirical processes based on dependent data. *Electron. J. Probab.* **17** no. 9, 18. MR2878788 https://doi.org/10.1214/EJP.v17-1765

[7] Berti, P., Pratelli, L. and Rigo, P. (2013). Exchangeable sequences driven by an absolutely continuous random measure. *Ann. Probab.* **41** 2090–2102. MR3098068 https://doi.org/10.1214/12-AOP786

1350-7265

[8] Berti, P., Regazzini, E. and Rigo, P. (1997). Well-calibrated, coherent forecasting systems. *Teor. Veroyatn. Primen.* **42** 144–168. MR1453335 https://doi.org/10.1137/S0040585X97975988
[9] Berti, P. and Rigo, P. (2007). 0–1 laws for regular conditional distributions. *Ann. Probab.* **35** 649–662. MR2308591 https://doi.org/10.1214/009117906000000845
[10] Cassese, A., Zhu, W., Guindani, M. and Vannucci, M. (2019). A Bayesian nonparametric spiked process prior for dynamic model selection. *Bayesian Anal.* **14** 553–572. MR3934097 https://doi.org/10.1214/18-BA1116
[11] Cifarelli, D.M. and Regazzini, E. (1996). De Finetti's contribution to probability and statistics. *Statist. Sci.* **11** 253–282. MR1445983 https://doi.org/10.1214/ss/1032280303
[12] de Finetti, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036
[13] Diaconis, P. and Freedman, D.A. (1990). Cauchy's equation and de Finetti's theorem. *Scand. J. Stat.* **17** 235–249. MR1092946
[14] Diaconis, P. and Ylvisaker, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281. MR0520238
[15] Dubins, L.E. and Savage, L.J. (1965). *How to Gamble If You Must. Inequalities for Stochastic Processes.* New York: McGraw-Hill. MR0236983
[16] Fortini, S., Ladelli, L. and Regazzini, E. (2000). Exchangeability, predictive distributions and parametric models. *Sankhyā Ser. A* **62** 86–109. MR1769738
[17] Fortini, S. and Petrone, S. (2012). Predictive construction of priors in Bayesian nonparametrics. *Braz. J. Probab. Stat.* **26** 423–449. MR2949087 https://doi.org/10.1214/11-BJPS176
[18] Fortini, S., Petrone, S. and Sporysheva, P. (2018). On a notion of partially conditionally identically distributed sequences. *Stochastic Process. Appl.* **128** 819–846. MR3758339 https://doi.org/10.1016/j.spa.2017.06.008
[19] Hahn, P.R., Martin, R. and Walker, S.G. (2018). On recursive Bayesian predictive distributions. *J. Amer. Statist. Assoc.* **113** 1085–1093. MR3862341 https://doi.org/10.1080/01621459.2017.1304219
[20] Hill, B.M. (1993). Parametric models for A_n: Splitting processes and mixtures. *J. Roy. Statist. Soc. Ser. B* **55** 423–433. MR1224406
[21] Kallenberg, O. (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16** 508–534. MR0929061
[22] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[23] Krasker, W.S. and Pratt, J.W. (1986). Discussion: On the consistency of Bayes estimates. *Ann. Statist.* **14** 55–58.
[24] Lee, J., Quintana, F.A., Müller, P. and Trippa, L. (2013). Defining predictive probability functions for species sampling models. *Statist. Sci.* **28** 209–222. MR3112406 https://doi.org/10.1214/12-sts407
[25] Martin, R. and Tokdar, S.T. (2011). Semiparametric inference in mixture models with predictive recursion marginal likelihood. *Biometrika* **98** 567–582. MR2836407 https://doi.org/10.1093/biomet/asr030
[26] Newton, M.A. and Zhang, Y. (1999). A recursive algorithm for nonparametric analysis with missing data. *Biometrika* **86** 15–26. MR1688068 https://doi.org/10.1093/biomet/86.1.15
[27] Pitman, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory. Institute of Mathematical Statistics Lecture Notes – Monograph Series* **30** 245–267. Hayward, CA: IMS. MR1481784 https://doi.org/10.1214/lnms/1215453576
[28] Pitman, J. (2006). *Combinatorial Stochastic Processes: Lectures from the* 32*nd Summer School on Probability Theory Held in Saint-Flour, July* 7–24, 2002. *Lecture Notes in Math.* **1875**. Berlin: Springer. With a foreword by Jean Picard. MR2245368
[29] Tokdar, S.T., Martin, R. and Ghosh, J.K. (2009). Consistency of a recursive estimate of mixing distributions. *Ann. Statist.* **37** 2502–2522. MR2543700 https://doi.org/10.1214/08-AOS639
[30] Walker, S. and Muliere, P. (1997). A characterisation of Polya tree distributions. *Statist. Probab. Lett.* **31** 163–168. MR1440632 https://doi.org/10.1016/S0167-7152(96)00028-4

Bernoulli Forthcoming Papers

CARPENTIER, A. and VERZELEN, N.
Optimal sparsity testing in linear regression model

GENTIL, I. and ZUGMEYER, S.
A family of Beckner inequalities under various curvature-dimension conditions

NA, S. and KOLAR, M.
High-dimensional index volatility models via Stein's identity

KRAMPE, J., KREISS, J.-P. and PAPARODITIS, E.
Bootstrap based inference for sparse high-dimensional time series models

DÜMBGEN, L., SAMWORTH, R.J. and WELLNER, J.A.
Bounding distributional errors via density ratios

GRANDITS, P.
Asymptotics of the hitting probability for a small sphere and a two dimensional Brownian motion with discontinuous anisotropic drift

WANG, C., YANG, S. and ZHANG, T.
Reflected Brownian motion with singular drift

BRIAND, P., GEISS, C., GEISS, S. and LABART, C.
Donsker-type theorem for BSDEs: Rate of convergence

ZHAO, P., WANG, L. and SHAO, J.
Sufficient dimension reduction and instrument search for data with nonignorable nonresponse

DJEHICHE, B., MAZHAR, O. and ROJAS, C.R.
Finite impulse response models: A non-asymptotic analysis of the least squares estimator

KUPPER, M. and ZAPATA, J.M.
Large deviations built on max-stability

JAKUBOWSKI, A., RODIONOV, I. and SOJA-KUKIEŁA, N.
Directional phantom distribution functions for stationary random fields

LI, Y. and XU, Y.
On fluctuations of global and mesoscopic linear statistics of generalized Wigner matrices

BOURGUIN, S., DIEZ, C.-P. and TUDOR, C.A.
Limiting behavior of large correlated Wishart matrices with chaotic entries

CHAKRABORTY, A. and PANARETOS, V.M.
Functional registration and local variations: Identifiability, rank, and tuning

RÖCKNER, M. and ZHANG, X.
Well-posedness of distribution dependent SDEs with singular drifts

BOGERD, K., CASTRO, R.M., VAN DER HOFSTAD, R. and VERZELEN, N.
Detecting a planted community in an inhomogeneous random graph

Continues

Bernoulli Forthcoming Papers—*Continued*

ZHU, C. and SHAO, X.
Interpoint distance based two sample tests in high dimension

YANO, K., KANEKO, R. and KOMAKI, F.
Minimax predictive density for sparse count data

DREES, H. and NEBLUNG, S.
Asymptotics for sliding blocks estimators of rare events

FAN, A. and KARAGULYAN, D.
On μ-Dvoretzky random covering of the circle

MIKOSCH, T. and RODIONOV, I.
Precise large deviations for dependent subexponential variables

ARMSTRONG, T.B.
Adaptation bounds for confidence bands under self-similarity

BASRAK, B. and PLANINIĆ, H.
Compound Poisson approximation for regularly varying fields with application to sequence alignment

JENTSCH, C. and KULIK, R.
Bootstrapping Hill estimator and tail array sums for regularly varying time series

LAW, M. and RITOV, Y.
Inference without compatibility: Using exponential weighting for inference on a parameter of a linear model

KRAAIJ, R.C.
Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton–Jacobi equation

BETZ, V., MÜHLBAUER, J., SCHÄFER, H. and ZEINDLER, D.
Precise asymptotics of longest cycles in random permutations without macroscopic cycles

CATOR, E. and DON, H.
Explicit bounds for critical infection rates and expected extinction times of the contact process on finite random graphs

CARDONA-TOBÓN, N. and PALAU, S.
Yaglom's limit for critical Galton–Watson processes in varying environment: A probabilistic approach

GARINO, V., NOURDIN, I., NUALART, D. and SALAMAT, M.
Limit theorems for integral functionals of Hermite-driven processes

BRUNEL, V.-E., KLUSOWSKI, J.M. and YANG, D.
Estimation of convex supports from noisy measurements

QIAO, W.
Asymptotic confidence regions for density ridges

Continues

Bernoulli Forthcoming Papers—*Continued*

ABI JABER, E.
Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels

AGAPIOU, S., DASHTI, M. and HELIN, T.
Rates of contraction of posterior distributions based on p-exponential priors

KREISS, A.
Correlation bounds, mixing and m-dependence under random time-varying network distances with an application to Cox-processes

KREBS, J.
On the law of the iterated logarithm and strong invariance principles in stochastic geometry

FONTES, L.R., GOMES, P.A. and SANCHIS, R.
Contact process under heavy-tailed renewals on finite graphs

KOVCHEGOV, Y. and ZALIAPIN, I.
Invariance and attraction properties of Galton–Watson trees

BOBBIA, B., DOMBRY, C. and VARRON, D.
The coupling method in extreme value theory

JAKUBOWSKI, J. and WIŚNIEWOLSKI, M.
A convolution formula for the local time of an Itô diffusion reflecting at 0 and a generalized Stroock–Williams equation

FAN, W.-T.L.
Stochastic PDEs on graphs as scaling limits of discrete interacting systems

MA, Y.-A., CHATTERJI, N.S., CHENG, X., FLAMMARION, N., BARTLETT, P.L. and JORDAN, M.I.
Is there an analog of Nesterov acceleration for gradient-based MCMC?

ZHOU, X., FOUCART, C. and LI, P.-S.
Time-changed spectrally positive Lévy processes started from infinity

GUHA, A., HO, N. and NGUYEN, X.
On posterior contraction of parameters and interpretability in Bayesian mixture modeling

GNEDIN, A. and SEKSENBAYEV, A.
Asymptotics and renewal approximation in the online selection of increasing subsequence

HUANG, D. and TROPP, J.A.
From Poincaré inequalities to nonlinear matrix concentration

Continues

Bernoulli Forthcoming Papers—*Continued*

BREHMER, J. and GNEITING, T.
Scoring interval forecasts: Equal-tailed, shortest, and modal interval

CÉNAC, P., CHAUVIN, B., NOÛS, C., PACCAUT, F. and POUYANNE, N.
Variable Length Memory Chains: Characterization of stationary probability measures