

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Eight Number Two May 2022 ISSN: 1350-7265

CONTENTS

BELLEÇ, P.C. and ZHANG, C.-H.	713
De-biasing the lasso with degrees-of-freedom adjustment	
WANG, Y. and POLITIS, D.N.	744
Model-free bootstrap for a general class of stationary time series	
MAÏDA, M., NGUYEN, T.D., PHAM NGOC, T.M., RIVOIRARD, V. and TRAN, V.C.	771
Statistical deconvolution of the free Fokker-Planck equation at fixed time	
GAIGALL, D., GERSTENBERG, J. and TRINH, T.T.H.	803
Empirical process of concomitants for partly categorical data and applications in statistics	
JOURDAIN, B. and MARGHERITI, W.	830
Martingale Wasserstein inequality for probability measures in the convex order	
QIN, Q. and JONES, G.L.	859
Convergence rates of two-component MCMC samplers	
ALT, J. and KRÜGER, T.	886
Local elliptic law	
MICHEL, B. and NOUY, A.	910
Learning with tree tensor networks: Complexity estimates and model selection	
LU, J., TAN, Y. and XU, L.	937
Central limit theorem and self-normalized Cramér-type moderate deviation for Euler-Maruyama scheme	
PAINDAVEINE, D.	965
On the measure of anchored Gaussian simplices, with applications to multivariate medians	
BING, X., BUNEA, F. and WEGKAMP, M.	997
Inference in latent factor regression with clusterable features	
PADOAN, S.A. and STUPFLER, G.	1021
Joint inference on extreme expectiles for multivariate heavy-tailed distributions	
TANG, M., CAPE, J. and PRIEBE, C.E.	1049
Asymptotically efficient estimators for stochastic blockmodels: The naive MLE, the rank-constrained MLE, and the spectral estimator	
CHATTERJI, N.S., BARTLETT, P.L. and LONG, P.M.	1074
Oracle lower bounds for stochastic gradient sampling algorithms	
CHAKRABORTY, M. and GHOSAL, S.	1093
Rates and coverage for monotone densities using projection-posterior	
GOLDENSHLUGER, A. and LEPSKI, O.V.	1120
Minimax estimation of norms of a probability density: I. Lower bounds	

(continued)

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

BERNOULLI

*Official Journal of the Bernoulli Society for Mathematical Statistics
and Probability*

Volume Twenty Eight Number Two May 2022 ISSN: 1350-7265

CONTENTS

(continued)

GOLDENSHLUGER, A. and LEPSKI, O.V.	1155
Minimax estimation of norms of a probability density: II. Rate-optimal estimation procedures	
LAM-WEIL, J., CARPENTIER, A. and SRIPERUMBUDUR, B.K.	1179
Local minimax rates for closeness testing of discrete distributions	
ALISHAHI, K. and BARZEGAR, M.	1198
Paving property for real stable polynomials and strongly Rayleigh processes	
ZHANG, L., ZHOU, W. and WANG, H.	1224
Non-asymptotic properties of spectral decomposition of large Gram-type matrices and applications	
KURISU, D.	1250
Nonparametric regression for locally stationary random fields under stochastic sampling design	
BERGER, D. and LINDNER, A.	1276
A Cramér–Wold device for infinite divisibility of $\mathbb{Z}^d$ -valued distributions	
NAULET, Z.	1284
Adaptive Bayesian density estimation in sup-norm	
BEHME, A. and SIDERIS, A.	1309
Markov-modulated generalized Ornstein-Uhlenbeck processes and an application in risk theory	
FRANCESCHINI, C., GONÇALVES, P. and SAU, F.	1340
Symmetric inclusion process with slow boundary: Hydrodynamics and hydrostatics	
BELOMESTNY, D., IOSIPOI, L., PARIS, Q. and ZHIVOTOVSKIY, N.	1382
Empirical variance minimization with applications in variance reduction and optimal control	
KERSTING, G. and MINUESA, C.	1408
Defective Galton-Watson processes in a varying environment	
HILÁRIO, M. and UNGARETTI, D.	1432
A note on the phase transition for independent alignment percolation	
NAZAROV, A. and NIKITIN, Y.	1448
Spectral equivalence of Gaussian random functions: Operator approach	
HART, S. and RINOTT, Y.	1461
Posterior probabilities: Nonmonotonicity, asymptotic rates, log-concavity, and Turán’s inequality	
BARNES, C.	1491
Convergence of jump processes with stochastic intensity to Brownian motion with inert drift	

BERNOULLI

Volume 28 Number 2 May 2022 Pages 713–1518

ISI/BS

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope

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

Bernoulli Society for Mathematical Statistics and Probability

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

Meetings: <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

Detailed information about the members of the Executive Committee can be found on <http://www.bernoulli-society.org/who-is-who>

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

©2022 International Statistical Institute/Bernoulli Society

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the Publisher.

In 2022 Bernoulli consists of 4 issues published in February, May, August and November.



Bernoulli Society
for Mathematical Statistics
and Probability

De-biasing the lasso with degrees-of-freedom adjustment

PIERRE C. BELLEC^{*} and CUN-HUI ZHANG[†]

Department of Statistics, Hill Center, Busch Campus, Rutgers University, Piscataway, NJ 08854, USA.
E-mail: ^{*}pierre.bellec@rutgers.edu; [†]czhang@stat.rutgers.edu

This paper studies schemes to de-bias the Lasso in sparse linear regression with Gaussian design where the goal is to estimate and construct confidence intervals for a low-dimensional projection of the unknown coefficient vector in a preconceived direction $\mathbf{a}_0$. Our analysis reveals that previously analyzed propositions to de-bias the Lasso require a modification in order to enjoy nominal coverage and asymptotic efficiency in a full range of the level of sparsity. This modification takes the form of a degrees-of-freedom adjustment that accounts for the dimension of the model selected by the Lasso. The degrees-of-freedom adjustment (a) preserves the success of de-biasing methodologies in regimes where previous proposals were successful, and (b) repairs the nominal coverage and provides efficiency in regimes where previous proposals produce spurious inferences and provably fail to achieve the nominal coverage. Hence our theoretical and simulation results call for the implementation of this degrees-of-freedom adjustment in de-biasing methodologies.

Let s_0 denote the number of nonzero coefficients of the true coefficient vector and Σ the population Gram matrix. The unadjusted de-biasing scheme may fail to achieve the nominal coverage as soon as $s_0 \gg n^{2/3}$ if Σ is known. If Σ is unknown, the degrees-of-freedom adjustment grants efficiency for the contrast in a general direction $\mathbf{a}_0$ when

$$\frac{s_0 \log p}{n} + \min \left\{ \frac{s_\Omega \log p}{n}, \frac{\|\Sigma^{-1} \mathbf{a}_0\|_1 \sqrt{\log p}}{\|\Sigma^{-1/2} \mathbf{a}_0\|_2 \sqrt{n}} \right\} + \frac{\min(s_\Omega, s_0) \log p}{\sqrt{n}} \rightarrow 0$$

where $s_\Omega = \|\Sigma^{-1} \mathbf{a}_0\|_0$. The dependence in s_0 , s_Ω and $\|\Sigma^{-1} \mathbf{a}_0\|_1$ is optimal and closes a gap in previous upper and lower bounds. Our construction of the estimated score vector provides a novel methodology to handle dense directions $\mathbf{a}_0$.

Beyond the degrees-of-freedom adjustment, our proof techniques yield a sharp ℓ_∞ error bound for the Lasso which is of independent interest.

Keywords: Statistical inference; Lasso; semiparametric model; Fisher information; efficiency; confidence interval; p-value; regression; high-dimensional data

References

- [1] Bellec, P.C. and Zhang, C.-H. (2021). Second order Stein: Sure for sure and other applications in high-dimensional inference. *Ann. Statist.* To appear.
- [2] Bellec, P.C. and Zhang, C.-H. (2022). Supplement to “De-biasing the lasso with degrees-of-freedom adjustment.” <https://doi.org/10.3150/21-BEJ1348SUPP>
- [3] Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19** 521–547. [MR3037163 https://doi.org/10.3150/11-BEJ410](https://doi.org/10.3150/11-BEJ410)
- [4] Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Rev. Econ. Stud.* **81** 608–650. [MR3207983 https://doi.org/10.1093/restud/rdt044](https://doi.org/10.1093/restud/rdt044)
- [5] Belloni, A., Chernozhukov, V. and Wang, L. (2014). Pivotal estimation via square-root Lasso in nonparametric regression. *Ann. Statist.* **42** 757–788. [MR3210986 https://doi.org/10.1214/14-AOS1204](https://doi.org/10.1214/14-AOS1204)

- [6] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Series in the Mathematical Sciences. Baltimore, MD: Johns Hopkins Univ. Press. [MR1245941](#)
- [7] Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. [MR2533469](#) <https://doi.org/10.1214/08-AOS620>
- [8] Bühlmann, P. (2013). Statistical significance in high-dimensional linear models. *Bernoulli* **19** 1212–1242. [MR3102549](#) <https://doi.org/10.3150/12-BEJSP11>
- [9] Cai, T., Cai, T. and Guo, Z. (2019). Individualized treatment selection: An optimal hypothesis testing approach in high-dimensional models. ArXiv preprint. Available at [arXiv:1904.12891](#).
- [10] Cai, T.T. and Guo, Z. (2017). Confidence intervals for high-dimensional linear regression: Minimax rates and adaptivity. *Ann. Statist.* **45** 615–646. [MR3650395](#) <https://doi.org/10.1214/16-AOS1461>
- [11] Candès, E. and Tao, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n . *Ann. Statist.* **35** 2313–2351. [MR2382644](#) <https://doi.org/10.1214/009053606000001523>
- [12] Candès, E.J. and Tao, T. (2005). Decoding by linear programming. *IEEE Trans. Inf. Theory* **51** 4203–4215. [MR2243152](#) <https://doi.org/10.1109/TIT.2005.858979>
- [13] Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909. [MR3277152](#)
- [14] Javanmard, A. and Montanari, A. (2014). Hypothesis testing in high-dimensional regression under the Gaussian random design model: Asymptotic theory. *IEEE Trans. Inf. Theory* **60** 6522–6554. [MR3265038](#) <https://doi.org/10.1109/TIT.2014.2343629>
- [15] Javanmard, A. and Montanari, A. (2018). Debiasing the Lasso: Optimal sample size for Gaussian designs. *Ann. Statist.* **46** 2593–2622. [MR3851749](#) <https://doi.org/10.1214/17-AOS1630>
- [16] Lounici, K. (2008). Sup-norm convergence rate and sign concentration property of Lasso and Dantzig estimators. *Electron. J. Stat.* **2** 90–102. [MR2386087](#) <https://doi.org/10.1214/08-EJS177>
- [17] 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>
- [18] Schick, A. (1986). On asymptotically efficient estimation in semiparametric models. *Ann. Statist.* **14** 1139–1151. [MR0856811](#) <https://doi.org/10.1214/aos/1176350055>
- [19] Sun, T. and Zhang, C.-H. (2012). Scaled sparse linear regression. *Biometrika* **99** 879–898. [MR2999166](#) <https://doi.org/10.1093/biomet/ass043>
- [20] Tibshirani, R.J. and Taylor, J. (2012). Degrees of freedom in lasso problems. *Ann. Statist.* **40** 1198–1232. [MR2985948](#) <https://doi.org/10.1214/12-AOS1003>
- [21] Tropp, J.A. (2006). Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Trans. Inf. Theory* **52** 1030–1051. [MR2238069](#) <https://doi.org/10.1109/TIT.2005.864420>
- [22] van de Geer, S. (2016). *Estimation and Testing Under Sparsity*. *Lecture Notes in Math.* **2159**. Cham: Springer. [MR3526202](#) <https://doi.org/10.1007/978-3-319-32774-7>
- [23] van de Geer, S. (2017). On the efficiency of the de-biased lasso. arXiv preprint. Available at [arXiv:1708.07986](#).
- [24] van de Geer, S., Bühlmann, P., Ritov, Y. and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42** 1166–1202. [MR3224285](#) <https://doi.org/10.1214/14-AOS1221>
- [25] 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>
- [26] Wainwright, M.J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1 -constrained quadratic programming (Lasso). *IEEE Trans. Inf. Theory* **55** 2183–2202. [MR2729873](#) <https://doi.org/10.1109/TIT.2009.2016018>
- [27] Zhang, C.-H. (2005). Estimation of sums of random variables: Examples and information bounds. *Ann. Statist.* **33** 2022–2041. [MR2211078](#) <https://doi.org/10.1214/009053605000000390>
- [28] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. [MR2604701](#) <https://doi.org/10.1214/09-AOS729>
- [29] Zhang, C.-H. (2011). *Statistical inference for high-dimensional data*, Mathematisches Forschungsinstitut Oberwolfach: Very High Dimensional Semiparametric Models, Report 48, 28–31.

- [30] Zhang, C.-H. and Huang, J. (2008). The sparsity and bias of the LASSO selection in high-dimensional linear regression. *Ann. Statist.* **36** 1567–1594. [MR2435448](#) <https://doi.org/10.1214/07-AOS520>
- [31] Zhang, C.-H. and Zhang, S.S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 217–242. [MR3153940](#) <https://doi.org/10.1111/rssb.12026>
- [32] Zhang, C.-H. and Zhang, T. (2012). A general theory of concave regularization for high-dimensional sparse estimation problems. *Statist. Sci.* **27** 576–593. [MR3025135](#) <https://doi.org/10.1214/12-STS399>
- [33] Zhao, P. and Yu, B. (2006). On model selection consistency of Lasso. *J. Mach. Learn. Res.* **7** 2541–2563. [MR2274449](#)
- [34] Zhu, Y. and Bradic, J. (2018). Significance testing in non-sparse high-dimensional linear models. *Electron. J. Stat.* **12** 3312–3364. [MR3861831](#) <https://doi.org/10.1214/18-EJS1443>
- [35] Zhu, Y. and Bradic, J. (2018). Linear hypothesis testing in dense high-dimensional linear models. *J. Amer. Statist. Assoc.* **113** 1583–1600. [MR3902231](#) <https://doi.org/10.1080/01621459.2017.1356319>
- [36] Zhu, Y. and Bradic, J. (2018). Significance testing in non-sparse high-dimensional linear models. *Electron. J. Stat.* **12** 3312–3364. [MR3861831](#) <https://doi.org/10.1214/18-EJS1443>
- [37] Zou, H., Hastie, T. and Tibshirani, R. (2007). On the “degrees of freedom” of the lasso. *Ann. Statist.* **35** 2173–2192. [MR2363967](#) <https://doi.org/10.1214/009053607000000127>

Model-free bootstrap for a general class of stationary time series

YIREN WANG¹ and DIMITRIS N. POLITIS²

¹*Department of Mathematics, University of California, San Diego, La Jolla, CA, USA.*

E-mail: yiw518@ucsd.edu

²*Department of Mathematics and Halicioğlu Data Science Institute, University of California, San Diego, La Jolla, CA, USA.*

E-mail: dpolitis@ucsd.edu

A model-free bootstrap procedure for a general class of stationary time series is introduced. The theoretical framework is established, showing asymptotic validity of bootstrap confidence intervals for many statistics of interest. In addition, asymptotic validity of one-step ahead bootstrap prediction intervals is also demonstrated. Finite-sample experiments are conducted to empirically confirm the performance of the new method, and to compare with popular methods such as the block bootstrap and the autoregressive (AR)-sieve bootstrap.

Keywords: m-approximation; bootstrap validity; transformation function; confidence intervals

References

- [1] Alonso, A.M., Peña, D. and Romo, J. (2002). Forecasting time series with sieve bootstrap. *J. Statist. Plann. Inference* **100** 1–11. [MR1869807](#) [https://doi.org/10.1016/S0378-3758\(01\)00092-1](https://doi.org/10.1016/S0378-3758(01)00092-1)
- [2] Angus, J.E. (1994). The probability integral transform and related results. *SIAM Rev.* **36** 652–654. [MR1306928](#) <https://doi.org/10.1137/1036146>
- [3] Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U -processes of stationary mixing sequences. *J. Theoret. Probab.* **7** 47–71. [MR1256391](#) <https://doi.org/10.1007/BF02213360>
- [4] Berkes, I., Hörmann, S. and Schauer, J. (2009). Asymptotic results for the empirical process of stationary sequences. *Stochastic Process. Appl.* **119** 1298–1324. [MR2508575](#) <https://doi.org/10.1016/j.spa.2008.06.010>
- [5] Berkes, I., Hörmann, S. and Schauer, J. (2011). Split invariance principles for stationary processes. *Ann. Probab.* **39** 2441–2473. [MR2932673](#) <https://doi.org/10.1214/10-AOP603>
- [6] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*. New York, NY: Springer.
- [7] Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli* **3** 123–148. [MR1466304](#) <https://doi.org/10.2307/3318584>
- [8] Das, S. and Politis, D.N. (2020). Nonparametric Estimation of the Conditional Distribution at Regression Boundary Points. *Amer. Statist.* **74** 233–242. [MR4133499](#) <https://doi.org/10.1080/00031305.2018.1558109>
- [9] Das, S. and Politis, D.N. (2021). Predictive inference for locally stationary time series with an application to climate data. *J. Amer. Statist. Assoc.* **116** 919–934. [MR4270034](#) <https://doi.org/10.1080/01621459.2019.1708368>
- [10] Drmač, Z., Omladič, M. and Veselić, K. (1994). On the perturbation of the Cholesky factorization. *SIAM J. Matrix Anal. Appl.* **15** 1319–1332. [MR1293920](#) <https://doi.org/10.1137/S0895479893244717>
- [11] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* **7** 1–26. [MR0515681](#)
- [12] Giacomini, R., Politis, D.N. and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators. *Econometric Theory* **29** 567–589. [MR3064050](#) <https://doi.org/10.1017/S0266466612000655>
- [13] Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. [MR2662361](#) <https://doi.org/10.1214/09-AOS768>

- [14] Jentsch, C. and Politis, D.N. (2015). Covariance matrix estimation and linear process bootstrap for multivariate time series of possibly increasing dimension. *Ann. Statist.* **43** 1117–1140. [MR3346699](#) <https://doi.org/10.1214/14-AOS1301>
- [15] Kreiss, J.-P. and Paparoditis, E. (2020). *Bootstrap for Time Series: Theory and Applications*. New York: Springer.
- [16] Kreiss, J.-P., Paparoditis, E. and Politis, D.N. (2011). On the range of validity of the autoregressive sieve bootstrap. *Ann. Statist.* **39** 2103–2130. [MR2893863](#) <https://doi.org/10.1214/11-AOS900>
- [17] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241. [MR1015147](#) <https://doi.org/10.1214/aos/1176347265>
- [18] Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*. Springer Series in Statistics. New York: Springer. [MR2001447](#) <https://doi.org/10.1007/978-1-4757-3803-2>
- [19] Mase, S. (1977). Some theorems on normality-preserving transformations. *Sankhyā Ser. A* **39** 186–190. [MR0478435](#)
- [20] McElroy, T.S. and Politis, D.N. (2019). *Time Series: A First Course with Bootstrap Starter*. Boca Raton, FL: CRC Press/CRC Press.
- [21] McMurry, T.L. and Politis, D.N. (2010). Banded and tapered estimates for autocovariance matrices and the linear process bootstrap. *J. Time Series Anal.* **31** 471–482. [MR2732601](#) <https://doi.org/10.1111/j.1467-9892.2010.00679.x>
- [22] McMurry, T.L. and Politis, D.N. (2015). High-dimensional autocovariance matrices and optimal linear prediction. *Electron. J. Stat.* **9** 753–788. [MR3331856](#) <https://doi.org/10.1214/15-EJS1000>
- [23] Pan, L. and Politis, D.N. (2016). Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *J. Statist. Plann. Inference* **177** 1–27. [MR3511318](#) <https://doi.org/10.1016/j.jspi.2014.10.003>
- [24] Pan, L. and Politis, D.N. (2016). Bootstrap prediction intervals for Markov processes. *Comput. Statist. Data Anal.* **100** 467–494. [MR3505813](#) <https://doi.org/10.1016/j.csda.2015.05.010>
- [25] Patton, A., Politis, D.N. and White, H. (2009). Correction to “Automatic block-length selection for the dependent bootstrap” by D. Politis and H. White [MR2041534]. *Econometric Rev.* **28** 372–375. [MR2532278](#) <https://doi.org/10.1080/07474930802459016>
- [26] Politis, D.N. (2013). Model-free model-fitting and predictive distributions. *TEST* **22** 183–221. [MR3062250](#) <https://doi.org/10.1007/s11749-013-0317-7>
- [27] Politis, D.N. (2015). *Model-Free Prediction and Regression: A Transformation-Based Approach to Inference*. Frontiers in Probability and the Statistical Sciences. Cham: Springer. [MR3442999](#) <https://doi.org/10.1007/978-3-319-21347-7>
- [28] Politis, D.N. and Romano, J.P. (1992). A general resampling scheme for triangular arrays of α -mixing random variables with application to the problem of spectral density estimation. *Ann. Statist.* **20** 1985–2007. [MR1193322](#) <https://doi.org/10.1214/aos/1176348899>
- [29] Politis, D.N. and Romano, J.P. (1995). Bias-corrected nonparametric spectral estimation. *J. Time Series Anal.* **16** 67–103. [MR1323618](#) <https://doi.org/10.1111/j.1467-9892.1995.tb00223.x>
- [30] Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.* **23** 470–472. [MR0049525](#) <https://doi.org/10.1214/aoms/1177729394>
- [31] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** 43–47. [MR0074711](#) <https://doi.org/10.1073/pnas.42.1.43>
- [32] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. Stochastic Modeling. New York: CRC Press. [MR1280932](#)
- [33] Snelson, E., Ghahramani, Z. and Rasmussen, C.E. (2004). Warped Gaussian processes. In *Advances in Neural Information Processing Systems* 16 337–344. MIT Press.
- [34] Wang, Y. and Politis, D.N. (2022). Supplement to “Model-free bootstrap for a general class of stationary time series.” <https://doi.org/10.3150/21-BEJ1352SUPP>
- [35] Wu, W.B. (2005). Nonlinear system theory: Another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** 14150–14154. [MR2172215](#) <https://doi.org/10.1073/pnas.0506715102>
- [36] Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.* **33** 1934–1963. [MR2166566](#) <https://doi.org/10.1214/0090536050000000291>
- [37] Wu, W.B. and Pourahmadi, M. (2009). Banded sample autocovariance matrices of stationary processes. *Statist. Sinica* **19** 1755–1768. [MR2589209](#)

- [38] Zhang, Y., Liu, J., Pan, Z. and Politis, D.N. (2019). Estimating transformation function. *Electron. J. Stat.* **13** 3095–3119. [MR4010594](#) <https://doi.org/10.1214/19-ejs1603>

Statistical deconvolution of the free Fokker-Planck equation at fixed time

MYLÈNE MAÏDA¹, TIEN DAT NGUYEN^{2,*}, THANH MAI PHAM NGOC^{2,†},
VINCENT RIVOIRARD³ and VIET CHI TRAN⁴

¹Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France.

E-mail: mylene.maida@univ-lille.fr

²Laboratoire de Mathématiques d'Orsay, CNRS, Université Paris-Saclay, 91405, Orsay, France.

E-mail: ^{*}tien-dat.nguyen@math.u-psud.fr; [†]thanh.pham_ngoc@math.u-psud.fr

³CEREMADE, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, 75016 Paris, France.

E-mail: Vincent.Rivoirard@dauphine.fr

⁴LAMA, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77454 Marne-la-Vallée, France.

E-mail: chi.tran@univ-eiffel.fr

We are interested in reconstructing the initial condition of a non-linear partial differential equation (PDE), namely the Fokker-Planck equation, from the observation of a Dyson Brownian motion at a given time $t > 0$. The Fokker-Planck equation describes the evolution of electrostatic repulsive particle systems, and can be seen as the large particle limit of correctly renormalized Dyson Brownian motions. The solution of the Fokker-Planck equation can be written as the free convolution of the initial condition and the semi-circular distribution. We propose a non-parametric estimator for the initial condition obtained by performing the free deconvolution via the subordination functions method. This statistical estimator is original as it involves the resolution of a fixed point equation, and a classical deconvolution by a Cauchy distribution. This is due to the fact that, in free probability, the analogue of the Fourier transform is the R-transform, related to the Cauchy transform. In past literature, there has been a focus on the estimation of the initial conditions of linear PDEs such as the heat equation, but to the best of our knowledge, this is the first time that the problem is tackled for a non-linear PDE. The convergence of the estimator is proved and the integrated mean square error is computed, providing rates of convergence similar to the ones known for non-parametric deconvolution methods. Finally, a simulation study illustrates the good performances of our estimator.

Keywords: 35Q62; 65M32; 62G05; 46L53; 35R30; 60B20; 46L54

References

- [1] Anderson, G.W., Guionnet, A. and Zeitouni, O. (2010). *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics **118**. Cambridge: Cambridge Univ. Press. [MR2760897](#)
- [2] Arizmendi, O., Tarrago, P. and Vargas, C. (2020). Subordination methods for free deconvolution. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2565–2594. [MR4164848](#) <https://doi.org/10.1214/20-AIHP1050>
- [3] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. Springer Series in Statistics. New York: Springer. [MR2567175](#) <https://doi.org/10.1007/978-1-4419-0661-8>
- [4] Belinschi, S.T. and Bercovici, H. (2007). A new approach to subordination results in free probability. *J. Anal. Math.* **101** 357–365. [MR2346550](#) <https://doi.org/10.1007/s11854-007-0013-1>
- [5] Belomestny, D. and Goldenshluger, A. (2021). Density deconvolution under general assumptions on the distribution of measurement errors. *Ann. Statist.* **49** 615–649. [MR4255101](#) <https://doi.org/10.1214/20-aos1969>
- [6] Bertoin, J., Giraud, C. and Isozaki, Y. (2001). Statistics of a flux in Burgers turbulence with one-sided Brownian initial data. *Comm. Math. Phys.* **224** 551–564. [MR1870057](#) <https://doi.org/10.1007/PL00005589>
- [7] Biane, P. (1997). On the free convolution with a semi-circular distribution. *Indiana Univ. Math. J.* **46** 705–718. [MR1488333](#) <https://doi.org/10.1512/iumj.1997.46.1467>

- [8] Biane, P. and Speicher, R. (2001). Free diffusions, free entropy and free Fisher information. *Ann. Inst. Henri Poincaré Probab. Stat.* **37** 581–606. [MR1851716](#) [https://doi.org/10.1016/S0246-0203\(00\)01074-8](https://doi.org/10.1016/S0246-0203(00)01074-8)
- [9] Bourgain, J. (1994). Periodic nonlinear Schrödinger equation and invariant measures. *Comm. Math. Phys.* **166** 1–26. [MR1309539](#)
- [10] Burgers, J.M. (1974). *The Nonlinear Diffusion Equation*. Berlin: Springer.
- [11] Burq, N. and Tzvetkov, N. (2008). Random data Cauchy theory for supercritical wave equations. I. Local theory. *Invent. Math.* **173** 449–475. [MR2425133](#) <https://doi.org/10.1007/s00222-008-0124-z>
- [12] Burq, N. and Tzvetkov, N. (2008). Random data Cauchy theory for supercritical wave equations. II. A global existence result. *Invent. Math.* **173** 477–496. [MR2425134](#) <https://doi.org/10.1007/s00222-008-0123-0>
- [13] Butucea, C. and Tsybakov, A.B. (2007). Sharp optimality in density deconvolution with dominating bias. I. *Teor. Veroyatn. Primen.* **52** 111–128. [MR2354572](#) <https://doi.org/10.1137/S0040585X97982840>
- [14] Carrillo, J.A., McCann, R.J. and Villani, C. (2003). Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoam.* **19** 971–1018. [MR2053570](#) <https://doi.org/10.4171/RMI/376>
- [15] Comte, F., Cuenod, C.-A., Pensky, M. and Rozenholc, Y. (2017). Laplace deconvolution on the basis of time domain data and its application to dynamic contrast-enhanced imaging. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 69–94. [MR3597965](#) <https://doi.org/10.1111/rssb.12159>
- [16] Comte, F. and Lacour, C. (2013). Anisotropic adaptive kernel deconvolution. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 569–609. [MR3088382](#) <https://doi.org/10.1214/11-aihp470>
- [17] Constantin, P. and Wu, J. (1997). Statistical solutions of the Navier-Stokes equations on the phase space of vorticity and the inviscid limits. *J. Math. Phys.* **38** 3031–3045. [MR1449547](#) <https://doi.org/10.1063/1.532032>
- [18] Dallaporta, S. and Fevrier, M. (2019). Fluctuations of linear spectral statistics of deformed Wigner matrices. Submitted. [hal-02079313](#).
- [19] Dattner, I., Reiß, M. and Trabs, M. (2016). Adaptive quantile estimation in deconvolution with unknown error distribution. *Bernoulli* **22** 143–192. [MR3449779](#) <https://doi.org/10.3150/14-BEJ626>
- [20] Fan, J. and Koo, J.-Y. (2002). Wavelet deconvolution. *IEEE Trans. Inf. Theory* **48** 734–747. [MR1889978](#) <https://doi.org/10.1109/18.986021>
- [21] Flandoli, F. (2018). Weak vorticity formulation of 2D Euler equations with white noise initial condition. *Comm. Partial Differential Equations* **43** 1102–1149. [MR3910197](#) <https://doi.org/10.1080/03605302.2018.1467448>
- [22] Giraud, C. (2003). Some properties of Burgers turbulence with White noise initial conditions. In *Probabilistic Methods in Fluids* 161–178. River Edge, NJ: World Sci. Publ. [MR2083371](#) https://doi.org/10.1142/9789812703989_0011
- [23] Groetsch, C.W. (1984). *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. Research Notes in Mathematics* **105**. Boston, MA: Pitman (Advanced Publishing Program). [MR0742928](#)
- [24] Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Math. Acad. Sci. Paris* **342** 877–882. [MR2224640](#) <https://doi.org/10.1016/j.crma.2006.04.006>
- [25] Maïda, M., Nguyen, T.D., Pham Ngoc, T.M., Rivoirard, V. and Tran, V.C. (2022). Supplement to “Statistical deconvolution of the free Fokker-Planck equation at fixed time.” <https://doi.org/10.3150/21-BEJ1366SUPP>
- [26] Méléard, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995). Lecture Notes in Math.* **1627** 42–95. Berlin: Springer. [MR1431299](#) <https://doi.org/10.1007/BFb0093177>
- [27] Nguyen, T.D. (2021). Nonparametric free deconvolution and circular regression estimation Ph.D. thesis, Université Paris Saclay, Paris, France.
- [28] Pensky, M. (2017). Minimax theory of estimation of linear functionals of the deconvolution density with or without sparsity. *Ann. Statist.* **45** 1516–1541. [MR3670187](#) <https://doi.org/10.1214/16-AOS1498>
- [29] Pensky, M. and Sapatinas, T. (2009). Functional deconvolution in a periodic setting: Uniform case. *Ann. Statist.* **37** 73–104. [MR2488345](#) <https://doi.org/10.1214/07-AOS552>
- [30] Pensky, M. and Sapatinas, T. (2010). On convergence rates equivalency and sampling strategies in functional deconvolution models. *Ann. Statist.* **38** 1793–1844. [MR2662360](#) <https://doi.org/10.1214/09-AOS767>

- [31] Sznitman, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX—1989. Lecture Notes in Math.* **1464** 165–251. Berlin: Springer. [MR1108185](#) <https://doi.org/10.1007/BFb0085169>
- [32] Talay, D. and Vaillant, O. (2003). A stochastic particle method with random weights for the computation of statistical solutions of McKean-Vlasov equations. *Ann. Appl. Probab.* **13** 140–180. [MR1951996](#) <https://doi.org/10.1214/aoap/1042765665>
- [33] Tarrago, P. (2020). Spectral deconvolution of unitary invariant matrix models. [arXiv:2006.09356](#).
- [34] Tran, V.C. (2008). A wavelet particle approximation for McKean-Vlasov and 2D-Navier-Stokes statistical solutions. *Stochastic Process. Appl.* **118** 284–318. [MR2376903](#) <https://doi.org/10.1016/j.spa.2007.04.008>
- [35] Tzvetkov, N. (2019). Random data wave equations. In *Singular Random Dynamics. Lecture Notes in Math.* **2253** 221–313. Cham: Springer. [MR3971360](#)
- [36] Vergassola, M., Dubrulle, B., Frisch, U. and Noullez, A. (1994). Burgers' equation, Devil's staircases and the mass distribution for large-scale structures. *Astron. Astrophys.* **289** 325–356.
- [37] Vishik, M.J. and Fursikov, A.V. (1988). *Mathematical Problems of Statistical Hydromechanics. Mathematics and Its Applications (Soviet Series)* **9**. Dordrecht: Kluwer Academic. [MR3444271](#) <https://doi.org/10.1007/978-94-009-1423-0>
- [38] 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](https://doi.org/10.1016/0022-1236(86)90062-5)
- [39] Voiculescu, D. (1993). The analogues of entropy and of Fisher's information measure in free probability theory. I. *Comm. Math. Phys.* **155** 71–92. [MR1228526](#)

Empirical process of concomitants for partly categorical data and applications in statistics

DANIEL GAIGALL¹, JULIAN GERSTENBERG^{1,2,*} and THI THU HA TRINH¹

¹*Institute of Probability and Statistics, Leibniz University Hannover, Welfengarten 1, 30167 Hannover, Germany.*

²*Institute of Mathematics, Goethe University Frankfurt, Robert-Mayer-Straße 10, 60325 Frankfurt am Main, Germany.*

E-mail: *gerstenb@math.uni-frankfurt.de

On the basis of independent and identically distributed bivariate random vectors, where the components are categorical and continuous variables, respectively, the related concomitants, also called induced order statistic, are considered. The main theoretical result is a functional central limit theorem for the empirical process of the concomitants in a triangular array setting. A natural application is hypothesis testing. An independence test and a two-sample test are investigated in detail. The fairly general setting enables limit results under local alternatives and bootstrap samples. For the comparison with existing tests from the literature simulation studies are conducted. The empirical results obtained confirm the theoretical findings.

Keywords: Categorical variable; concomitant; induced order statistic; empirical process; independence test; two-sample test; bootstrap; triangular array; local alternatives

References

- [1] Bahadur, R.R. (1966). A note on quantiles in large samples. *Ann. Math. Stat.* **37** 577–580. [MR0189095 https://doi.org/10.1214/aoms/1177699450](https://doi.org/10.1214/aoms/1177699450)
- [2] Baringhaus, L. and Gaigall, D. (2019). On an asymptotic relative efficiency concept based on expected volumes of confidence regions. *Statistics* **53** 1396–1436. [MR4034868 https://doi.org/10.1080/02331888.2019.1683560](https://doi.org/10.1080/02331888.2019.1683560)
- [3] Bücher, A. and Kojadinovic, I. (2019). A note on conditional versus joint unconditional weak convergence in bootstrap consistency results. *J. Theoret. Probab.* **32** 1145–1165. [MR3979663 https://doi.org/10.1007/s10959-018-0823-3](https://doi.org/10.1007/s10959-018-0823-3)
- [4] Chatterjee, S. (2020). A new coefficient of correlation. *J. Amer. Statist. Assoc.* 1–21.
- [5] David, H.A. and Nagaraja, H.N. (1998). Concomitants of order statistics. *Handb. Statist.* **16**.
- [6] David, H.A. and Nagaraja, H.N. (2003). *Order Statistics*, 3rd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. [MR1994955 https://doi.org/10.1002/0471722162](https://doi.org/10.1002/0471722162)
- [7] Davydov, Y. and Egorov, V. (2001). Functional CLT and LIL for induced order statistics. In *Asymptotic Methods in Probability and Statistics with Applications (St. Petersburg, 1998)*. *Stat. Ind. Technol.* 333–349. Boston, MA: Birkhäuser. [MR1890337 https://doi.org/10.1007/s00025-007-0257-1](https://doi.org/10.1007/s00025-007-0257-1)
- [8] Gaenssler, P., Molnár, P. and Rost, D. (2007). On continuity and strict increase of the CDF for the sup-functional of a Gaussian process with applications to statistics. *Results Math.* **51** 51–60. [MR2375638 https://doi.org/10.1007/s00025-007-0257-1](https://doi.org/10.1007/s00025-007-0257-1)
- [9] Gaigall, D. (2019). Testing marginal homogeneity of a continuous bivariate distribution with possibly incomplete paired data. *Metrika*.
- [10] Genest, C., Nešlehová, J.G., Rémillard, B. and Murphy, O.A. (2019). Testing for independence in arbitrary distributions. *Biometrika* **106** 47–68. [MR3912383 https://doi.org/10.1093/biomet/asy059](https://doi.org/10.1093/biomet/asy059)
- [11] Genest, C. and Rémillard, B. (2004). Tests of independence and randomness based on the empirical copula process. *TEST* **13** 335–370. [MR2154005 https://doi.org/10.1007/BF02595777](https://doi.org/10.1007/BF02595777)

- [12] Gerstenberg, J. (2020). General erased-word processes: Product-type filtrations, ergodic laws and Martin boundaries. *Stochastic Process. Appl.* **130** 3540–3573. [MR4092414](#) <https://doi.org/10.1016/j.spa.2019.10.003>
- [13] Heller, R., Heller, Y., Kaufman, S., Brill, B. and Gorfine, M. (2016). Consistent distribution-free K -sample and independence tests for univariate random variables. *J. Mach. Learn. Res.* **17** 29. [MR3491123](#)
- [14] Kamps, U. (1995). *A Concept of Generalized Order Statistics*. Berlin: Springer.
- [15] 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>
- [16] Pitman, J. and Tran, N.M. (2015). Size-biased permutation of a finite sequence with independent and identically distributed terms. *Bernoulli* **21** 2484–2512. [MR3378475](#) <https://doi.org/10.3150/14-BEJ652>
- [17] Puri, M.L. and Sen, P.K. (1971). *Nonparametric Methods in Multivariate Analysis*. New York: Wiley. [MR0298844](#)
- [18] Robbins, M.W., Lund, R.B., Gallagher, C.M. and Lu, Q. (2011). Changepoints in the North Atlantic tropical cyclone record. *J. Amer. Statist. Assoc.* **106** 89–99. [MR2816704](#) <https://doi.org/10.1198/jasa.2011.ap10023>
- [19] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley. [MR0595165](#)
- [20] Shorack, G.R. and Wellner, J.A. (2009). *Empirical Processes with Applications to Statistics*. Philadelphia: SIAM.
- [21] Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. [MR2382665](#) <https://doi.org/10.1214/009053607000000505>
- [22] Ziegler, K. (1997). Functional central limit theorems for triangular arrays of function-indexed processes under uniformly integrable entropy conditions. *J. Multivariate Anal.* **62** 233–272. [MR1473875](#) <https://doi.org/10.1006/jmva.1997.1688>

Martingale Wasserstein inequality for probability measures in the convex order

BENJAMIN JOURDAIN^{*} and WILLIAM MARGHERITI[†]

CERMICS, Ecole des Ponts, INRIA, Marne-la-Vallée, France.

E-mail: ^{*}benjamin.jourdain@enpc.fr; [†]william.margheriti@enpc.fr

It was shown by the authors that two one-dimensional probability measures in the convex order admit a martingale coupling with respect to which the integral of $|x - y|$ is smaller than twice their $\mathcal{W}_1$ -distance (Wasserstein distance with index 1). We showed that replacing $|x - y|$ and $\mathcal{W}_1$ respectively with $|x - y|^\rho$ and $\mathcal{W}_\rho^\rho$ does not lead to a finite multiplicative constant. We show here that a finite constant is recovered when replacing $\mathcal{W}_\rho^\rho$ with the product of $\mathcal{W}_\rho$ times the centred ρ -th moment of the second marginal to the power $\rho - 1$. Then we study the generalisation of this new martingale Wasserstein inequality to higher dimension.

Keywords: Convex order; martingale optimal transport; Wasserstein distance; martingale couplings

References

- [1] Alfonsi, A., Corbetta, J. and Jourdain, B. (2019). Sampling of one-dimensional probability measures in the convex order and computation of robust option price bounds. *Int. J. Theor. Appl. Finance* **22** 1950002, 41. [MR3951948](#) <https://doi.org/10.1142/S021902491950002X>
- [2] Alfonsi, A., Corbetta, J. and Jourdain, B. (2020). Sampling of probability measures in the convex order by Wasserstein projection. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 1706–1729. [MR4116705](#) <https://doi.org/10.1214/19-AIHP1014>
- [3] Alfonsi, A. and Jourdain, B. (2014). A remark on the optimal transport between two probability measures sharing the same copula. *Statist. Probab. Lett.* **84** 131–134. [MR3131266](#) <https://doi.org/10.1016/j.spl.2013.09.035>
- [4] Aliprantis, C.D. and Border, K.C. (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd ed. Berlin: Springer. [MR2378491](#)
- [5] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. [MR2401600](#)
- [6] Backhoff-Veraguas, J. and Pammer, G. (2019). Stability of martingale optimal transport and weak optimal transport. ArXiv E-prints. Available at [arXiv:1904.04171](#).
- [7] Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17** 477–501. [MR3066985](#) <https://doi.org/10.1007/s00780-013-0205-8>
- [8] Beiglböck, M. and Juillet, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. [MR3456332](#) <https://doi.org/10.1214/14-AOP966>
- [9] Beiglböck, M., Lim, T., Oblój, J. and Oblój, J. (2019). Dual attainment for the martingale transport problem. *Bernoulli* **25** 1640–1658. [MR3961225](#) <https://doi.org/10.3150/17-BEJ1015>
- [10] Beiglböck, M., Nutz, M. and Touzi, N. (2017). Complete duality for martingale optimal transport on the line. *Ann. Probab.* **45** 3038–3074. [MR3706738](#) <https://doi.org/10.1214/16-AOP1131>
- [11] Brücknerhoff, M. and Juillet, N. (2021). Instability of martingale optimal transport in dimension $d \geq 2$. ArXiv E-prints. Available at [arXiv:2101.06964](#).
- [12] De March, H. and Touzi, N. (2019). Irreducible convex paving for decomposition of multidimensional martingale transport plans. *Ann. Probab.* **47** 1726–1774. [MR3945758](#) <https://doi.org/10.1214/18-AOP1295>

- [13] De March, H. (2018). Quasi-sure duality for multi-dimensional martingale optimal transport. ArXiv E-prints. Available at [arXiv:1805.01757](https://arxiv.org/abs/1805.01757).
- [14] De March, H. (2018). Local structure of multi-dimensional martingale optimal transport. ArXiv E-prints. Available at [arXiv:1805.09469](https://arxiv.org/abs/1805.09469).
- [15] De March, H. (2018). Entropic approximation for multi-dimensional martingale optimal transport. ArXiv E-prints. Available at [arXiv:1812.11104](https://arxiv.org/abs/1812.11104).
- [16] Galichon, A., Henry-Labordère, P. and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24** 312–336. [MR3161649 https://doi.org/10.1214/13-AAP925](https://doi.org/10.1214/13-AAP925)
- [17] Guo, G., and Oblój, J. (2019). Computational methods for martingale optimal transport problems. *Ann. Appl. Probab.* **29** 3311–3347. [MR4047982 https://doi.org/10.1214/19-AAP1481](https://doi.org/10.1214/19-AAP1481)
- [18] Henry-Labordère, P. (2019). (Martingale) Optimal transport and anomaly detection with neural networks: A primal-dual algorithm. ArXiv Preprint. Available at [arXiv:1904.04546](https://arxiv.org/abs/1904.04546).
- [19] Henry-Labordère, P., Tan, X. and Touzi, N. (2016). An explicit martingale version of the one-dimensional Brenier’s theorem with full marginals constraint. *Stochastic Process. Appl.* **126** 2800–2834. [MR3522302 https://doi.org/10.1016/j.spa.2016.03.003](https://doi.org/10.1016/j.spa.2016.03.003)
- [20] Henry-Labordère, P. and Touzi, N. (2016). An explicit martingale version of the one-dimensional Brenier theorem. *Finance Stoch.* **20** 635–668. [MR3519164 https://doi.org/10.1007/s00780-016-0299-x](https://doi.org/10.1007/s00780-016-0299-x)
- [21] Jourdain, B. and Margheriti, W. (2020). A new family of one dimensional martingale couplings. *Electron. J. Probab.* **25** Paper No. 136, 50. [MR4171389 https://doi.org/10.3390/mca25030049](https://doi.org/10.3390/mca25030049)
- [22] Juillet, N. (2016). Stability of the shadow projection and the left-curtain coupling. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1823–1843. [MR3573297 https://doi.org/10.1214/15-AIHP700](https://doi.org/10.1214/15-AIHP700)
- [23] Móricz, T.F. (2009). Sharp inequalities between centered moments. *JIPAM. J. Inequal. Pure Appl. Math.* **10** Article 99, 9. [MR2577869](https://doi.org/10.1214/15-AIHP700)
- [24] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36** 423–439. [MR0177430 https://doi.org/10.1214/aoms/1177700153](https://doi.org/10.1214/aoms/1177700153)
- [25] Villani, C. (2003). *Topics in Optimal Transportation. Graduate Studies in Mathematics* **58**. Providence, RI: Amer. Math. Soc. [MR1964483 https://doi.org/10.1090/gsm/058](https://doi.org/10.1090/gsm/058)
- [26] Villani, C. (2009). *Optimal Transport: Old and New. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. [MR2459454 https://doi.org/10.1007/978-3-540-71050-9](https://doi.org/10.1007/978-3-540-71050-9)
- [27] Wiesel, J. (2020). Continuity of the martingale optimal transport problem on the real line. ArXiv E-prints. Available at [arXiv:1905.04574](https://arxiv.org/abs/1905.04574).

Convergence rates of two-component MCMC samplers

QIAN QIN^{*} and GALIN L. JONES[†]

School of Statistics, University of Minnesota, Minneapolis, USA.

E-mail: ^{*}qqin@umn.edu; [†]galin@umn.edu

Component-wise MCMC algorithms, including Gibbs and conditional Metropolis-Hastings samplers, are commonly used for sampling from multivariate probability distributions. A long-standing question regarding Gibbs algorithms is whether a deterministic-scan (systematic-scan) sampler converges faster than its random-scan counterpart. We answer this question when the samplers involve two components by establishing an exact quantitative relationship between the L^2 convergence rates of the two samplers. The relationship shows that the deterministic-scan sampler converges faster. We also establish qualitative relations among the convergence rates of two-component Gibbs samplers and some conditional Metropolis-Hastings variants. For instance, it is shown that if some two-component conditional Metropolis-Hastings samplers are geometrically ergodic, then so are the associated Gibbs samplers.

Keywords: Deterministic-scan; geometric ergodicity; Gibbs; Metropolis-within-Gibbs; random-scan

References

- [1] Amit, Y. (1991). On rates of convergence of stochastic relaxation for Gaussian and non-Gaussian distributions. *J. Multivariate Anal.* **38** 82–99. [MR1128938](#) [https://doi.org/10.1016/0047-259X\(91\)90033-X](https://doi.org/10.1016/0047-259X(91)90033-X)
- [2] Amit, Y. (1996). Convergence properties of the Gibbs sampler for perturbations of Gaussians. *Ann. Statist.* **24** 122–140. [MR1389883](#) <https://doi.org/10.1214/aos/1033066202>
- [3] Amit, Y. and Grenander, U. (1991). Comparing sweep strategies for stochastic relaxation. *J. Multivariate Anal.* **37** 197–222. [MR1114473](#) [https://doi.org/10.1016/0047-259X\(91\)90080-L](https://doi.org/10.1016/0047-259X(91)90080-L)
- [4] Andrieu, C. (2016). On random- and systematic-scan samplers. *Biometrika* **103** 719–726. [MR3551794](#) <https://doi.org/10.1093/biomet/asw019>
- [5] Böttcher, A. and Spitkovsky, I.M. (2010). A gentle guide to the basics of two projections theory. *Linear Algebra Appl.* **432** 1412–1459. [MR2580440](#) <https://doi.org/10.1016/j.laa.2009.11.002>
- [6] Brooks, S., Gelman, A., Jones, G.L. and Meng, X.-L., eds. (2011). *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton, FL: CRC Press. [MR2742422](#) <https://doi.org/10.1201/b10905>
- [7] Chan, K.S. and Geyer, C.J. (1994). Discussion: Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1747–1758.
- [8] Dai, N. and Jones, G.L. (2017). Multivariate initial sequence estimators in Markov chain Monte Carlo. *J. Multivariate Anal.* **159** 184–199. [MR3668555](#) <https://doi.org/10.1016/j.jmva.2017.05.009>
- [9] Diaconis, P., Khare, K. and Saloff-Coste, L. (2008). Gibbs sampling, exponential families and orthogonal polynomials. *Statist. Sci.* **23** 151–178. With comments and a rejoinder by the authors. [MR2446500](#) <https://doi.org/10.1214/07-STS252>
- [10] Doss, H. and Hobert, J.P. (2010). Estimation of Bayes factors in a class of hierarchical random effects models using a geometrically ergodic MCMC algorithm. *J. Comput. Graph. Statist.* **19** 295–312. With supplementary material available online. [MR2675092](#) <https://doi.org/10.1198/jcgs.2010.09182>
- [11] Doss, C.R., Flegal, J.M., Jones, G.L. and Neath, R.C. (2014). Markov chain Monte Carlo estimation of quantiles. *Electron. J. Stat.* **8** 2448–2478. [MR3285872](#) <https://doi.org/10.1214/14-EJS957>

- [12] Duncan, J. and Taylor, P.J. (1976). Norm inequalities for C^* -algebras. *Proc. Roy. Soc. Edinburgh Sect. A* **75** 119–129. [MR0454647](#) <https://doi.org/10.1017/S0308210500017832>
- [13] Ekvall, K.O. and Jones, G.L. (2021). Convergence analysis of a collapsed Gibbs sampler for Bayesian vector autoregressions. *Electron. J. Stat.* **15** 691–721. [MR4202497](#) <https://doi.org/10.1214/21-EJS1800>
- [14] Fishman, G.S. (1996). Coordinate selection rules for Gibbs sampling. *Ann. Appl. Probab.* **6** 444–465. [MR1398053](#) <https://doi.org/10.1214/aoap/1034968139>
- [15] Flegal, J.M., Haran, M. and Jones, G.L. (2008). Markov chain Monte Carlo: Can we trust the third significant figure? *Statist. Sci.* **23** 250–260. [MR2516823](#) <https://doi.org/10.1214/08-STS257>
- [16] Flegal, J.M. and Jones, G.L. (2010). Batch means and spectral variance estimators in Markov chain Monte Carlo. *Ann. Statist.* **38** 1034–1070. [MR2604704](#) <https://doi.org/10.1214/09-AOS735>
- [17] Fort, G., Moulines, E., Roberts, G.O. and Rosenthal, J.S. (2003). On the geometric ergodicity of hybrid samplers. *J. Appl. Probab.* **40** 123–146. [MR1953771](#) <https://doi.org/10.1239/jap/1044476831>
- [18] Geyer, C.J. (1992). Practical Markov chain Monte Carlo (with discussion). *Statist. Sci.* **7** 473–511.
- [19] Greenwood, P.E., McKeague, I.W. and Wefelmeyer, W. (1998). Information bounds for Gibbs samplers. *Ann. Statist.* **26** 2128–2156. [MR1700224](#) <https://doi.org/10.1214/aos/1024691464>
- [20] Helmberg, G. (2014). *Introduction to Spectral Theory in Hilbert Space*. Elsevier. [MR0243367](#)
- [21] Herbei, R. and McKeague, I.W. (2009). Hybrid samplers for ill-posed inverse problems. *Scand. J. Stat.* **36** 839–853. [MR2573311](#) <https://doi.org/10.1111/j.1467-9469.2009.00649.x>
- [22] Hobert, J.P. (2011). The data augmentation algorithm: Theory and methodology. In *Handbook of Markov Chain Monte Carlo* (S. Brooks, A. Gelman, G. Jones and X.-L. Meng, eds.). Chapman & Hall/CRC Handb. Mod. Stat. Methods 253–293. Boca Raton, FL: CRC Press. [MR2858452](#)
- [23] Hobert, J.P. and Geyer, C.J. (1998). Geometric ergodicity of Gibbs and block Gibbs samplers for a hierarchical random effects model. *J. Multivariate Anal.* **67** 414–430. [MR1659196](#) <https://doi.org/10.1006/jmva.1998.1778>
- [24] Hobert, J.P., Jones, G.L., Presnell, B. and Rosenthal, J.S. (2002). On the applicability of regenerative simulation in Markov chain Monte Carlo. *Biometrika* **89** 731–743. [MR1946508](#) <https://doi.org/10.1093/biomet/89.4.731>
- [25] Johnson, A.A. and Jones, G.L. (2008). Comment: Gibbs sampling, exponential families, and orthogonal polynomials. *Statist. Sci.* **23** 183–186.
- [26] Johnson, A.A. and Jones, G.L. (2015). Geometric ergodicity of random scan Gibbs samplers for hierarchical one-way random effects models. *J. Multivariate Anal.* **140** 325–342. [MR3372572](#) <https://doi.org/10.1016/j.jmva.2015.06.002>
- [27] Johnson, A.A., Jones, G.L. and Neath, R.C. (2013). Component-wise Markov chain Monte Carlo: Uniform and geometric ergodicity under mixing and composition. *Statist. Sci.* **28** 360–375. [MR3135537](#) <https://doi.org/10.1214/13-STS423>
- [28] Jones, G.L. (2004). On the Markov chain central limit theorem. *Probab. Surv.* **1** 299–320. [MR2068475](#) <https://doi.org/10.1214/154957804100000051>
- [29] Jones, G.L. and Hobert, J.P. (2001). Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statist. Sci.* **16** 312–334. [MR1888447](#) <https://doi.org/10.1214/ss/1015346317>
- [30] Jones, G.L. and Hobert, J.P. (2004). Sufficient burn-in for Gibbs samplers for a hierarchical random effects model. *Ann. Statist.* **32** 784–817. [MR2060178](#) <https://doi.org/10.1214/0090536040000000184>
- [31] Jones, G.L., Roberts, G.O. and Rosenthal, J.S. (2014). Convergence of conditional Metropolis-Hastings samplers. *Adv. in Appl. Probab.* **46** 422–445. [MR3215540](#) <https://doi.org/10.1239/aap/1401369701>
- [32] Jones, G.L., Haran, M., Caffo, B.S. and Neath, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **101** 1537–1547. [MR2279478](#) <https://doi.org/10.1198/016214506000000492>
- [33] Kayalar, S. and Weinert, H.L. (1988). Error bounds for the method of alternating projections. *Math. Control Signals Systems* **1** 43–59. [MR0923275](#) <https://doi.org/10.1007/BF02551235>
- [34] Khare, K. and Hobert, J.P. (2013). Geometric ergodicity of the Bayesian lasso. *Electron. J. Stat.* **7** 2150–2163. [MR3104915](#) <https://doi.org/10.1214/13-EJS841>
- [35] Levine, R.A. and Casella, G. (2008). Comment: On random scan Gibbs samplers. *Statist. Sci.* **23** 192–195.
- [36] Liu, J.S. (1996). Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Stat. Comput.* **6** 113–119.

- [37] Liu, J.S., Wong, W.H. and Kong, A. (1994). Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes. *Biometrika* **81** 27–40. [MR1279653](#) <https://doi.org/10.1093/biomet/81.1.27>
- [38] Liu, J.S., Wong, W.H. and Kong, A. (1995). Covariance structure and convergence rate of the Gibbs sampler with various scans. *J. Roy. Statist. Soc. Ser. B* **57** 157–169. [MR1325382](#)
- [39] Marchev, D. and Hobert, J.P. (2004). Geometric ergodicity of van Dyk and Meng’s algorithm for the multivariate Student’s t model. *J. Amer. Statist. Assoc.* **99** 228–238. [MR2054301](#) <https://doi.org/10.1198/016214504000000223>
- [40] Mengersen, K.L. and Tweedie, R.L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.* **24** 101–121. [MR1389882](#) <https://doi.org/10.1214/aos/1033066201>
- [41] Robert, C.P. (1995). Convergence control methods for Markov chain Monte Carlo algorithms. *Statist. Sci.* **10** 231–253. [MR1390517](#)
- [42] Roberts, G.O. and Polson, N.G. (1994). On the geometric convergence of the Gibbs sampler. *J. Roy. Statist. Soc. Ser. B* **56** 377–384. [MR1281941](#)
- [43] 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>
- [44] Roberts, G.O. and Rosenthal, J.S. (1998). Two convergence properties of hybrid samplers. *Ann. Appl. Probab.* **8** 397–407. [MR1624941](#) <https://doi.org/10.1214/aoap/1028903533>
- [45] Roberts, G.O. and Rosenthal, J.S. (2001). Markov chains and de-initializing processes. *Scand. J. Stat.* **28** 489–504. [MR1858413](#) <https://doi.org/10.1111/1467-9469.00250>
- [46] Roberts, G.O. and Rosenthal, J.S. (2016). Surprising convergence properties of some simple Gibbs samplers under various scans. *International Journal of Statistics and Probability* **5** 51–60.
- [47] Roberts, G.O. and Sahu, S.K. (1997). Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler. *J. Roy. Statist. Soc. Ser. B* **59** 291–317. [MR1440584](#) <https://doi.org/10.1111/1467-9868.00070>
- [48] Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika* **83** 95–110. [MR1399158](#) <https://doi.org/10.1093/biomet/83.1.95>
- [49] Roberts, G.O. and Tweedie, R.L. (2001). Geometric L^2 and L^1 convergence are equivalent for reversible Markov chains. *J. Appl. Probab.* **38** 37–41. [MR1915532](#) <https://doi.org/10.1239/jap/1085496589>
- [50] Robertson, N., Flegal, J.M., Vats, D. and Jones, G.L. (2021). Assessing and visualizing simultaneous simulation error. *J. Comput. Graph. Statist.* **30** 324–334. [MR4270507](#) <https://doi.org/10.1080/10618600.2020.1824871>
- [51] Rosenthal, J.S. and Rosenthal, P. (2015). Spectral bounds for certain two-factor non-reversible MCMC algorithms. *Electron. Commun. Probab.* **20** no. 91, 10. [MR3434208](#) <https://doi.org/10.1214/ECP.v20-4528>
- [52] Roy, V. (2012). Convergence rates for MCMC algorithms for a robust Bayesian binary regression model. *Electron. J. Stat.* **6** 2463–2485. [MR3020272](#) <https://doi.org/10.1214/12-EJS756>
- [53] Rudolf, D. and Ullrich, M. (2013). Positivity of hit-and-run and related algorithms. *Electron. Commun. Probab.* **18** no. 49, 8. [MR3078012](#) <https://doi.org/10.1214/ECP.v18-2507>
- [54] Smith, R.L. and Tierney, L. (1996). Exact transition probabilities for the independence Metropolis sampler. Preprint.
- [55] Tan, A. and Hobert, J.P. (2009). Block Gibbs sampling for Bayesian random effects models with improper priors: Convergence and regeneration. *J. Comput. Graph. Statist.* **18** 861–878. [MR2598033](#) <https://doi.org/10.1198/jcgs.2009.08153>
- [56] Tan, A., Jones, G.L. and Hobert, J.P. (2013). On the geometric ergodicity of two-variable Gibbs samplers. In *Advances in Modern Statistical Theory and Applications: A Festschrift in Honor of Morris L. Eaton*. *Inst. Math. Stat. (IMS) Collect.* **10** 25–42. Beachwood, OH: IMS. [MR3586937](#)
- [57] Tanner, M.A. and Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82** 528–550. With discussion and with a reply by the authors. [MR0898357](#)
- [58] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1762. With discussion and a rejoinder by the author. [MR1329166](#) <https://doi.org/10.1214/aos/1176325750>
- [59] Tierney, L. (1998). A note on Metropolis-Hastings kernels for general state spaces. *Ann. Appl. Probab.* **8** 1–9. [MR1620401](#) <https://doi.org/10.1214/aoap/1027961031>

- [60] van Dyk, D.A. and Meng, X.-L. (2001). The art of data augmentation. *J. Comput. Graph. Statist.* **10** 1–111. With discussions, and a rejoinder by the authors. [MR1936358](#) <https://doi.org/10.1198/10618600152418584>
- [61] Vats, D., Flegal, J.M. and Jones, G.L. (2018). Strong consistency of multivariate spectral variance estimators in Markov chain Monte Carlo. *Bernoulli* **24** 1860–1909. [MR3757517](#) <https://doi.org/10.3150/16-BEJ914>
- [62] Vats, D., Flegal, J.M. and Jones, G.L. (2019). Multivariate output analysis for Markov chain Monte Carlo. *Biometrika* **106** 321–337. [MR3949306](#) <https://doi.org/10.1093/biomet/asz002>
- [63] Vats, D., Robertson, N., Flegal, J.M. and Jones, G.L. (2020). Analyzing Markov chain Monte Carlo output. *Wiley Interdiscip. Rev.: Comput. Stat.* **12** e1501, 12. [MR4114459](#) <https://doi.org/10.1002/wics.1501>
- [64] Vidav, I. (1977). The norm of the sum of two projections. *Proc. Amer. Math. Soc.* **65** 297–298. [MR0442703](#) <https://doi.org/10.2307/2041910>
- [65] Wang, X. and Roy, V. (2018). Convergence analysis of the block Gibbs sampler for Bayesian probit linear mixed models with improper priors. *Electron. J. Stat.* **12** 4412–4439. [MR3892344](#) <https://doi.org/10.1214/18-ejs1506>
- [66] Wang, X. and Roy, V. (2018). Geometric ergodicity of Pólya-gamma Gibbs sampler for Bayesian logistic regression with a flat prior. *Electron. J. Stat.* **12** 3295–3311. [MR3861283](#) <https://doi.org/10.1214/18-EJS1481>

Local elliptic law

JOHANNES ALT^{1,2} and TORBEN KRÜGER³

¹*Section of Mathematics, University of Geneva, 24, rue du Général Dufour, Case postale 64, 1211 Genève 4, Switzerland.*

E-mail: johannes.alt@unige.ch

²*Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1110, United States of America.*

E-mail: johannes.alt@unige.ch

³*Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 København, Denmark.*

E-mail: tk@math.ku.dk

The empirical eigenvalue distribution of the elliptic random matrix ensemble tends to the uniform measure on an ellipse in the complex plane as its dimension tends to infinity. We show this convergence on all mesoscopic scales slightly above the typical eigenvalue spacing in the bulk spectrum with an optimal convergence rate. As a corollary we obtain complete delocalisation for the corresponding eigenvectors in any basis.

Keywords: Local law; eigenvector delocalisation; elliptic ensemble; matrix Dyson equation

References

- [1] Aggarwal, A., Lopatto, P. and Yau, H.-T. (2021). GOE statistics for Lévy matrices. *J. Eur. Math. Soc. (JEMS)* **23** 3707–3800. [MR4310816](#) <https://doi.org/10.4171/jems/1089>
- [2] Ajanki, O., Erdős, L. and Krüger, T. (2017). Singularities of solutions to quadratic vector equations on the complex upper half-plane. *Comm. Pure Appl. Math.* **70** 1672–1705. [MR3684307](#) <https://doi.org/10.1002/cpa.21639>
- [3] Ajanki, O.H., Erdős, L. and Krüger, T. (2017). Universality for general Wigner-type matrices. *Probab. Theory Related Fields* **169** 667–727. [MR3719056](#) <https://doi.org/10.1007/s00440-016-0740-2>
- [4] Ajanki, O.H., Erdős, L. and Krüger, T. (2019). Stability of the matrix Dyson equation and random matrices with correlations. *Probab. Theory Related Fields* **173** 293–373. [MR3916109](#) <https://doi.org/10.1007/s00440-018-0835-z>
- [5] Alt, J., Erdős, L. and Krüger, T. (2018). Local inhomogeneous circular law. *Ann. Appl. Probab.* **28** 148–203. [MR3770875](#) <https://doi.org/10.1214/17-AAP1302>
- [6] Alt, J., Erdős, L. and Krüger, T. (2021). Spectral radius of random matrices with independent entries. *Probab. Math. Phys.* **2** no. 2, 1–60. <https://doi.org/10.2140/pmp.2021.2.1>
- [7] Alt, J., Erdős, L., Krüger, T. and Nemish, Y. (2019). Location of the spectrum of Kronecker random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 661–696. [MR3949949](#) <https://doi.org/10.1214/18-aihp894>
- [8] Alt, J. and Krüger, T. (2021). Inhomogeneous circular law for correlated matrices. *J. Funct. Anal.* **281** Paper No. 109120, 73. [MR4271784](#) <https://doi.org/10.1016/j.jfa.2021.109120>
- [9] Anderson, G.W. (2015). A local limit law for the empirical spectral distribution of the anticommutator of independent Wigner matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 809–841. [MR3365962](#) <https://doi.org/10.1214/14-AIHP602>
- [10] Bai, Z.D. (1997). Circular law. *Ann. Probab.* **25** 494–529. [MR1428519](#) <https://doi.org/10.1214/aop/1024404298>
- [11] Bao, Z., Erdős, L. and Schnelli, K. (2019). Local single ring theorem on optimal scale. *Ann. Probab.* **47** 1270–1334. [MR3945747](#) <https://doi.org/10.1214/18-AOP1284>
- [12] Bauerschmidt, R., Huang, J. and Yau, H.-T. (2019). Local Kesten-McKay law for random regular graphs. *Comm. Math. Phys.* **369** 523–636. [MR3962004](#) <https://doi.org/10.1007/s00220-019-03345-3>

- [13] Bauerschmidt, R., Knowles, A. and Yau, H.-T. (2017). Local semicircle law for random regular graphs. *Comm. Pure Appl. Math.* **70** 1898–1960. [MR3688032](#) <https://doi.org/10.1002/cpa.21709>
- [14] Bender, M. (2010). Edge scaling limits for a family of non-Hermitian random matrix ensembles. *Probab. Theory Related Fields* **147** 241–271. [MR2594353](#) <https://doi.org/10.1007/s00440-009-0207-9>
- [15] Bloemendal, A., Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2014). Isotropic local laws for sample covariance and generalized Wigner matrices. *Electron. J. Probab.* **19** no. 33, 53. [MR3183577](#) <https://doi.org/10.1214/ejp.v19-3054>
- [16] Bordenave, C. and Guionnet, A. (2013). Localization and delocalization of eigenvectors for heavy-tailed random matrices. *Probab. Theory Related Fields* **157** 885–953. [MR3129806](#) <https://doi.org/10.1007/s00440-012-0473-9>
- [17] Bourgade, P., Erdős, L. and Yau, H.-T. (2014). Universality of general β -ensembles. *Duke Math. J.* **163** 1127–1190. [MR3192527](#) <https://doi.org/10.1215/00127094-2649752>
- [18] Bourgade, P., Yau, H.-T. and Yin, J. (2014). Local circular law for random matrices. *Probab. Theory Related Fields* **159** 545–595. [MR3230002](#) <https://doi.org/10.1007/s00440-013-0514-z>
- [19] Bourgade, P., Yau, H.-T. and Yin, J. (2020). Random band matrices in the delocalized phase I: Quantum unique ergodicity and universality. *Comm. Pure Appl. Math.* **73** 1526–1596. [MR4156609](#) <https://doi.org/10.1002/cpa.21895>
- [20] Cipolloni, G., Erdős, L. and Schröder, D. (2021). Edge universality for non-Hermitian random matrices. *Probab. Theory Related Fields* **179** 1–28. [MR4221653](#) <https://doi.org/10.1007/s00440-020-01003-7>
- [21] Crisanti, A. and Sompolinsky, H. (1987). Dynamics of spin systems with randomly asymmetric bonds: Langevin dynamics and a spherical model. *Phys. Rev. A* (3) **36** 4922–4939. [MR0918489](#) <https://doi.org/10.1103/PhysRevA.36.4922>
- [22] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Delocalization and diffusion profile for random band matrices. *Comm. Math. Phys.* **323** 367–416. [MR3085669](#) <https://doi.org/10.1007/s00220-013-1773-3>
- [23] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. [MR3098073](#) <https://doi.org/10.1214/11-AOP734>
- [24] Erdős, L., Krüger, T. and Nemish, Y. (2019). Scattering in quantum dots via noncommutative rational functions. Preprint, available at: [arXiv:1911.05112](#).
- [25] Erdős, L., Krüger, T. and Nemish, Y. (2020). Local laws for polynomials of Wigner matrices. *J. Funct. Anal.* **278** 108507, 59. [MR4078529](#) <https://doi.org/10.1016/j.jfa.2020.108507>
- [26] Erdős, L., Krüger, T. and Renfrew, D. (2019). Randomly coupled differential equations with elliptic correlations. Preprint, available at [arXiv:1908.05178](#).
- [27] Erdős, L., Krüger, T. and Schröder, D. (2019). Random matrices with slow correlation decay. *Forum Math. Sigma* **7** Paper No. e8, 89. [MR3941370](#) <https://doi.org/10.1017/fms.2019.2>
- [28] Erdős, L., Schlein, B. and Yau, H.-T. (2009). Local semicircle law and complete delocalization for Wigner random matrices. *Comm. Math. Phys.* **287** 641–655. [MR2481753](#) <https://doi.org/10.1007/s00220-008-0636-9>
- [29] Erdős, L. and Yau, H.-T. (2017). *A Dynamical Approach to Random Matrix Theory*. *Courant Lecture Notes in Mathematics* **28**. New York: Courant Institute of Mathematical Sciences; Providence, RI: Amer. Math. Soc. [MR3699468](#)
- [30] Erdős, L., Yau, H.-T. and Yin, J. (2011). Universality for generalized Wigner matrices with Bernoulli distribution. *J. Comb.* **2** 15–81. [MR2847916](#) <https://doi.org/10.4310/JOC.2011.v2.n1.a2>
- [31] Fyodorov, Y.V. and Khoruzhenko, B.A. (2016). Nonlinear analogue of the May-Wigner instability transition. *Proc. Natl. Acad. Sci. USA* **113** 6827–6832. [MR3521630](#) <https://doi.org/10.1073/pnas.1601136113>
- [32] Girko, V.L. (1985). Circular law. *Theory Probab. Appl.* **29** 694–706. <https://doi.org/10.1137/1129095>
- [33] Girko, V.L. (1985). The elliptic law. *Teor. Veroyatn. Primen.* **30** 640–651. [MR0816278](#)
- [34] Girko, V.L. (1997). Strong elliptic law. *Random Oper. Stoch. Equ.* **5** 269–306. [MR1483014](#) <https://doi.org/10.1515/rose.1997.5.3.269>
- [35] Götze, F., Naumov, A. and Tikhomirov, A. (2015). On minimal singular values of random matrices with correlated entries. *Random Matrices Theory Appl.* **4** 1550006, 30. [MR3356884](#) <https://doi.org/10.1142/S2010326315500069>

- [36] Götze, F., Naumov, A. and Tikhomirov, A. (2020). Local laws for non-Hermitian random matrices and their products. *Random Matrices Theory Appl.* **9** 2150004, 53. [MR4133077](#) <https://doi.org/10.1142/S2010326321500040>
- [37] He, Y., Knowles, A. and Rosenthal, R. (2018). Isotropic self-consistent equations for mean-field random matrices. *Probab. Theory Related Fields* **171** 203–249. [MR3800833](#) <https://doi.org/10.1007/s00440-017-0776-y>
- [38] Helton, J.W., Rashidi Far, R. and Speicher, R. (2007). Operator-valued semicircular elements: Solving a quadratic matrix equation with positivity constraints. *Int. Math. Res. Not. IMRN* **22** Art. ID rnm086, 15. [MR2376207](#) <https://doi.org/10.1093/imrn/rnm086>
- [39] Johansson, K. (2007). From Gumbel to Tracy-Widom. *Probab. Theory Related Fields* **138** 75–112. [MR2288065](#) <https://doi.org/10.1007/s00440-006-0012-7>
- [40] Knowles, A. and Yin, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1750. [MR3103909](#) <https://doi.org/10.1002/cpa.21450>
- [41] Ledoux, M. (2008). Complex Hermite polynomials: From the semi-circular law to the circular law. *Commun. Stoch. Anal.* **2** 27–32. [MR2446909](#) <https://doi.org/10.31390/cosa.2.1.03>
- [42] Lee, J.O., Schnell, K., Stetler, B. and Yau, H.-T. (2016). Bulk universality for deformed Wigner matrices. *Ann. Probab.* **44** 2349–2425. [MR3502606](#) <https://doi.org/10.1214/15-AOP1023>
- [43] Luh, K. and O’Rourke, S. (2020). Eigenvector delocalization for non-Hermitian random matrices and applications. *Random Structures Algorithms* **57** 169–210. [MR4120597](#) <https://doi.org/10.1002/rsa.20917>
- [44] Lytova, A. and Tikhomirov, K. (2020). On delocalization of eigenvectors of random non-Hermitian matrices. *Probab. Theory Related Fields* **177** 465–524. [MR4095020](#) <https://doi.org/10.1007/s00440-019-00956-8>
- [45] Martí, D., Brunel, N. and Ostojic, S. (2018). Correlations between synapses in pairs of neurons slow down dynamics in randomly connected neural networks. *Phys. Rev. E* **97** 062314. <https://doi.org/10.1103/PhysRevE.97.062314>
- [46] Mehl, B. and Chalker, J.T. (2000). Statistical properties of eigenvectors in non-Hermitian Gaussian random matrix ensembles. *J. Math. Phys.* **41** 3233–3256. [MR1755501](#) <https://doi.org/10.1063/1.533302>
- [47] Naumov, A.A. (2013). The elliptic law for random matrices. *Vestnik Moskov. Univ. Ser. XV Vychisl. Mat. Kibernet.* **1** 31–38, 48. [MR3114382](#)
- [48] Nemish, Y. (2017). Local law for the product of independent non-Hermitian random matrices with independent entries. *Electron. J. Probab.* **22** Paper No. 22, 35. [MR3622892](#) <https://doi.org/10.1214/17-EJP38>
- [49] Nguyen, H.H. and O’Rourke, S. (2015). The elliptic law. *Int. Math. Res. Not. IMRN* **17** 7620–7689. [MR3403996](#) <https://doi.org/10.1093/imrn/rnu174>
- [50] O’Rourke, S. and Renfrew, D. (2014). Low rank perturbations of large elliptic random matrices. *Electron. J. Probab.* **19** no. 43, 65. [MR3210544](#) <https://doi.org/10.1214/EJP.v19-3057>
- [51] O’Rourke, S. and Renfrew, D. (2016). Central limit theorem for linear eigenvalue statistics of elliptic random matrices. *J. Theoret. Probab.* **29** 1121–1191. [MR3540493](#) <https://doi.org/10.1007/s10959-015-0609-9>
- [52] Rudelson, M. and Vershynin, R. (2015). Delocalization of eigenvectors of random matrices with independent entries. *Duke Math. J.* **164** 2507–2538. [MR3405592](#) <https://doi.org/10.1215/00127094-3129809>
- [53] Rudelson, M. and Vershynin, R. (2016). No-gaps delocalization for general random matrices. *Geom. Funct. Anal.* **26** 1716–1776. [MR3579707](#) <https://doi.org/10.1007/s00039-016-0389-0>
- [54] Song, S., Sjöström, P.J., Reigl, M., Nelson, S. and Chklovskii, D.B. (2005). Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**. <https://doi.org/10.1371/journal.pbio.0030068>
- [55] Tao, T. and Vu, V. (2010). Random matrices: Universality of ESDs and the circular law. *Ann. Probab.* **38** 2023–2065. With an appendix by Manjunath Krishnapur. [MR2722794](#) <https://doi.org/10.1214/10-AOP534>
- [56] 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>
- [57] Wang, Y., Markram, H., Goodman, P.H., Berger, T.K., Ma, J. and Goldman-Rakic, P.S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat. Neurosci.* **9** 534.
- [58] Wigner, E.P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. of Math.* (2) **62** 548–564. [MR0077805](#) <https://doi.org/10.2307/1970079>
- [59] Xi, H., Yang, F. and Yin, J. (2017). Local circular law for the product of a deterministic matrix with a random matrix. *Electron. J. Probab.* **22** Paper No. 60, 77. [MR3683369](#) <https://doi.org/10.1214/17-EJP76>
- [60] Yin, J. (2014). The local circular law III: General case. *Probab. Theory Related Fields* **160** 679–732. [MR3278919](#) <https://doi.org/10.1007/s00440-013-0539-3>

Learning with tree tensor networks: Complexity estimates and model selection

BERTRAND MICHEL^{*} and ANTHONY NOUY[†]

Centrale Nantes, Laboratoire de Mathématiques Jean Leray, CNRS UMR 6629, France.

E-mail: ^{*}bertrand.michel@ec-nantes.fr; [†]anthony.nouy@ec-nantes.fr

In this paper, we propose and analyze a model selection method for tree tensor networks in an empirical risk minimization framework and analyze its performance over a wide range of smoothness classes. Tree tensor networks, or tree-based tensor formats, are prominent model classes for the approximation of high-dimensional functions in numerical analysis and data science. They correspond to sum-product neural networks with a sparse connectivity associated with a dimension partition tree T , widths given by a tuple r of tensor ranks, and multilinear activation functions (or units). The approximation power of these model classes has been proved to be optimal (or near to optimal) for classical smoothness classes. However, in an empirical risk minimization framework with a limited number of observations, the dimension tree T and ranks r should be selected carefully to balance estimation and approximation errors. In this paper, we propose a complexity-based model selection strategy à la Barron, Birgé, Massart. Given a family of model classes associated with different trees, ranks, tensor product feature spaces and sparsity patterns for sparse tensor networks, a model is selected by minimizing a penalized empirical risk, with a penalty depending on the complexity of the model class. After deriving bounds of the metric entropy of tree tensor networks with bounded parameters, we deduce a form of the penalty from bounds on suprema of empirical processes. This choice of penalty yields a risk bound for the predictor associated with the selected model. In a least-squares setting, after deriving fast rates of convergence of the risk, we show that the proposed strategy is (near to) minimax adaptive to a wide range of smoothness classes including Sobolev or Besov spaces (with isotropic, anisotropic or mixed dominating smoothness) and analytic functions. We discuss the role of sparsity of the tensor network for obtaining optimal performance in several regimes. In practice, the amplitude of the penalty is calibrated with a slope heuristics method. Numerical experiments in a least-squares regression setting illustrate the performance of the strategy for the approximation of multivariate functions and univariate functions identified with tensors by tensorization (quantization).

Keywords: Tensor networks; statistical learning; metric entropy; model selection; minimax adaptive

References

- [1] Ali, M. and Nouy, A. (2020). Approximation theory of tree tensor networks: Tensorized univariate functions. Part I. Preprint. Available at [arxiv:2007.00118](https://arxiv.org/abs/2007.00118).
- [2] Ali, M. and Nouy, A. (2020). Approximation theory of tree tensor networks: Tensorized univariate functions. Part II. Preprint. Available at [arxiv:2007.00128](https://arxiv.org/abs/2007.00128).
- [3] Ali, M. and Nouy, A. (2021). Approximation theory of tree tensor networks: Tensorized multivariate functions. Preprint. Available at [arXiv:2101.11932](https://arxiv.org/abs/2101.11932).
- [4] Arlot, S. (2019). Minimal penalties and the slope heuristics: A survey. *J. SFS* **160** 1–106. [MR4021408](#)
- [5] Bachmayr, M., Nouy, A. and Schneider, R. (2021). Approximation power of tree tensor networks for compositional functions. Preprint.
- [6] Bachmayr, M., Schneider, R. and Uschmajew, A. (2016). Tensor networks and hierarchical tensors for the solution of high-dimensional partial differential equations. *Found. Comput. Math.* **16** 1423–1472. [MR3579714](#) <https://doi.org/10.1007/s10208-016-9317-9>
- [7] Baudry, J.-P., Maugis, C. and Michel, B. (2012). Slope heuristics: Overview and implementation. *Stat. Comput.* **22** 455–470. [MR2865029](#) <https://doi.org/10.1007/s11222-011-9236-1>

- [8] Belitser, E. (1998). Efficient estimation of analytic density under random censorship. *Bernoulli* **4** 519–543. [MR1679796 https://doi.org/10.2307/3318664](https://doi.org/10.2307/3318664)
- [9] Birgé, L. and Massart, P. (2007). Minimal penalties for Gaussian model selection. *Probab. Theory Related Fields* **138** 33–73. [MR2288064 https://doi.org/10.1007/s00440-006-0011-8](https://doi.org/10.1007/s00440-006-0011-8)
- [10] Bousquet, O., Boucheron, S. and Lugosi, G. (2004). Introduction to statistical learning theory. In *Advanced Lectures on Machine Learning* 169–207. Berlin: Springer.
- [11] Cichocki, A., Lee, N., Oseledets, I., Phan, A.-H., Zhao, Q. and Mandic, D. (2016). Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions. *Found. Trends Mach. Learn.* **9** 249–429.
- [12] Cichocki, A., Phan, A.-H., Zhao, Q., Lee, N., Oseledets, I., Sugiyama, M. and Mandic, D. (2017). Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives. *Found. Trends Mach. Learn.* **9** 431–673.
- [13] Cohen, N., Sharir, O. and Shashua, A. (2016). On the expressive power of deep learning: A tensor analysis. In *Conference on Learning Theory* 698–728.
- [14] DeVore, R.A. and Popov, V.A. (1988). Interpolation of Besov spaces. *Trans. Amer. Math. Soc.* **305** 397–414. [MR0920166 https://doi.org/10.2307/2001060](https://doi.org/10.2307/2001060)
- [15] Donoho, D.L. and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. [MR1635414 https://doi.org/10.1214/aos/1024691081](https://doi.org/10.1214/aos/1024691081)
- [16] Donoho, D.L., Johnstone, I.M., Kerkycharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24** 508–539. [MR1394974 https://doi.org/10.1214/aos/1032894451](https://doi.org/10.1214/aos/1032894451)
- [17] Dũng, D., Temlyakov, V. and Ullrich, T. (2018). *Hyperbolic Cross Approximation. Advanced Courses in Mathematics. CRM Barcelona*. Cham: Birkhäuser/Springer. [MR3887571 https://doi.org/10.1007/978-3-319-92240-9](https://doi.org/10.1007/978-3-319-92240-9)
- [18] Falcó, A., Hackbusch, W. and Nouy, A. (2021). Tree-based tensor formats. *SeMA J.* **78** 159–173. [MR4270843 https://doi.org/10.1007/s40324-018-0177-x](https://doi.org/10.1007/s40324-018-0177-x)
- [19] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge Series in Statistical and Probabilistic Mathematics* **40**. New York: Cambridge Univ. Press. [MR3588285 https://doi.org/10.1017/CBO9781107337862](https://doi.org/10.1017/CBO9781107337862)
- [20] Grelier, E., Nouy, A. and Chevreuil, M. (2018). Learning with tree-based tensor formats. Preprint. Available at [arXiv:1811.04455](https://arxiv.org/abs/1811.04455).
- [21] Grelier, E., Nouy, A. and Lebrun, R. (2019). Learning high-dimensional probability distributions using tree tensor networks. Preprint. Available at [arXiv:1912.07913](https://arxiv.org/abs/1912.07913).
- [22] Gribonval, R., Kutyniok, G., Nielsen, M. and Voigtlaender, F. (2019). Approximation spaces of deep neural networks. Preprint. Available at [arXiv:1905.01208](https://arxiv.org/abs/1905.01208).
- [23] Griebel, M. and Harbrecht, H. (2019). Analysis of tensor approximation schemes for continuous functions. Preprint. Available at [arXiv:1903.04234](https://arxiv.org/abs/1903.04234).
- [24] Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. (2006). *A Distribution-Free Theory of Nonparametric Regression. Springer Series in Statistics*. New York: Springer. [MR1920390 https://doi.org/10.1007/b97848](https://doi.org/10.1007/b97848)
- [25] Hackbusch, W. (2012). *Tensor Spaces and Numerical Tensor Calculus. Springer Series in Computational Mathematics* **42**. Heidelberg: Springer. [MR3236394 https://doi.org/10.1007/978-3-642-28027-6](https://doi.org/10.1007/978-3-642-28027-6)
- [26] Hansen, M. and Sickel, W. (2010). Best m -term approximation and tensor product of Sobolev and Besov spaces—The case of non-compact embeddings. *East J. Approx.* **16** 345–388. [MR2808578](https://doi.org/10.1007/s00365-012-9161-3)
- [27] Hansen, M. and Sickel, W. (2012). Best m -term approximation and Sobolev-Besov spaces of dominating mixed smoothness—The case of compact embeddings. *Constr. Approx.* **36** 1–51. [MR2926304 https://doi.org/10.1007/s00365-012-9161-3](https://doi.org/10.1007/s00365-012-9161-3)
- [28] Kazeev, V., Oseledets, I., Rakhuba, M. and Schwab, C. (2017). QTT-finite-element approximation for multiscale problems I: Model problems in one dimension. *Adv. Comput. Math.* **43** 411–442. [MR3627473 https://doi.org/10.1007/s10444-016-9491-y](https://doi.org/10.1007/s10444-016-9491-y)
- [29] Kazeev, V. and Schwab, C. (2015). Approximation of singularities by quantized-tensor fem. *PAMM* **15** 743–746.
- [30] Khoromskij, B. (2012). Tensors-structured numerical methods in scientific computing: Survey on recent advances. *Chemom. Intell. Lab. Syst.* **110** 1–19.

- [31] Khrulkov, V., Novikov, A. and Oseledets, I. (2018). Expressive power of recurrent neural networks. In *International Conference on Learning Representations*.
- [32] Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems: Ecole d'Eté de Probabilités de Saint-Flour XXXVIII-2008. Lecture Notes in Math.* **2033**. Heidelberg: Springer. [MR2829871](#) <https://doi.org/10.1007/978-3-642-22147-7>
- [33] Leisner, C. (2003). Nonlinear wavelet approximation in anisotropic Besov spaces. *Indiana Univ. Math. J.* **52** 437–455. [MR1976084](#) <https://doi.org/10.1512/iumj.2003.52.2224>
- [34] Massart, P. (2007). *Concentration Inequalities and Model Selection. Lecture Notes in Math.* **1896**. Berlin: Springer. [MR2319879](#)
- [35] Michel, B. and Nouy, A. (2022). Supplement to “Learning with tree tensor networks: Complexity estimates and model selection.” <https://doi.org/10.3150/21-BEJ1371SUPP>
- [36] Neumann, M.H. (2000). Multivariate wavelet thresholding in anisotropic function spaces. *Statist. Sinica* **10** 399–431. [MR1769750](#)
- [37] Nouy, A. (2017). Low-rank methods for high-dimensional approximation and model order reduction. In *Model Reduction and Approximation. Comput. Sci. Eng.* **15** 171–226. Philadelphia, PA: SIAM. [MR3672148](#) <https://doi.org/10.1137/1.9781611974829.ch4>
- [38] Recht, B., Fazel, M. and Parrilo, P.A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **52** 471–501. [MR2680543](#) <https://doi.org/10.1137/070697835>
- [39] Schneider, R. and Uschmajew, A. (2014). Approximation rates for the hierarchical tensor format in periodic Sobolev spaces. *J. Complexity* **30** 56–71. [MR3166521](#) <https://doi.org/10.1016/j.jco.2013.10.001>
- [40] Signoretto, M., De Lathauwer, L. and Suykens, J.A. (2010). Nuclear norms for tensors and their use for convex multilinear estimation. *Linear Algebra Appl.* **43**. To appear.
- [41] Suzuki, T. (2018). Adaptivity of deep relu network for learning in besov and mixed smooth besov spaces: Optimal rate and curse of dimensionality. Preprint. Available at [arXiv:1810.08033](#).
- [42] Vapnik, V.N. (2013). *The Nature of Statistical Learning Theory*. New York: Springer. [MR1367965](#) <https://doi.org/10.1007/978-1-4757-2440-0>
- [43] Yuan, M. and Zhang, C.-H. (2016). On tensor completion via nuclear norm minimization. *Found. Comput. Math.* **16** 1031–1068. [MR3529132](#) <https://doi.org/10.1007/s10208-015-9269-5>

Central limit theorem and self-normalized Cramér-type moderate deviation for Euler-Maruyama scheme

JIANYA LU^{1,2,*}, YUZHEN TAN³ and LIHU XU^{1,2,†}

¹Department of Mathematics, Faculty of Science and Technology, University of Macau, Macau, China.

E-mail: ^{*}jianya.lu@connect.um.edu.mo; [†]lihuxu@um.edu.mo

²UM Zhuhai Research Institute, Hengqin New District, Zhuhai, China

³Zhongtai Securities Institute for Financial Studies, Shandong University, 250100, Jinan, China.

E-mail: tanyuzhensdu@gmail.com

We consider a stochastic differential equation and its Euler-Maruyama (EM) scheme, under some appropriate conditions, they both admit a unique invariant measure, denoted by π and π_η respectively (η is the step size of the EM scheme). We construct an empirical measure Π_η of the EM scheme as a statistic of π_η , and use Stein's method developed in Fang, Shao and Xu (*Probab. Theory Related Fields* **174** (2019) 945–979) to prove a central limit theorem of Π_η . The proof of the self-normalized Cramér-type moderate deviation (SNCMD) is based on a standard decomposition on Markov chain, splitting $\eta^{-1/2}(\Pi_\eta(\cdot) - \pi(\cdot))$ into a martingale difference series sum $\mathcal{H}_\eta$ and a negligible remainder $\mathcal{R}_\eta$. We handle $\mathcal{H}_\eta$ by the time-change technique for martingale, while prove that $\mathcal{R}_\eta$ is exponentially negligible by concentration inequalities, which have their independent interest. Moreover, we show that SNCMD holds for $x = o(\eta^{-1/6})$, which has the same order as that of the classical result in Shao (*J. Theoret. Probab.* **12** (1999) 385–398), Jing, Shao and Wang (*Ann. Probab.* **31** (2003) 2167–2215).

Keywords: Stochastic differential equation; Euler-Maruyama scheme; central limit theorem; self-normalized Cramér-type moderate deviation; Stein's method

References

- [1] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge: Cambridge Univ. Press. MR2512800 <https://doi.org/10.1017/CBO9780511809781>
- [2] Bao, J., Huang, X. and Yuan, C. (2019). New regularity of Kolmogorov equation and application on approximation of semi-linear SPDEs with Hölder continuous drifts. *Commun. Pure Appl. Anal.* **18** 341–360. MR3845569 <https://doi.org/10.3934/cpaa.2019018>
- [3] Bao, J. and Shao, J. (2018). Weak convergence of path-dependent SDEs with irregular coefficients. Preprint [arXiv:1809.03088](https://arxiv.org/abs/1809.03088).
- [4] Bao, J. and Yuan, C. (2013). Convergence rate of EM scheme for SDDEs. *Proc. Amer. Math. Soc.* **141** 3231–3243. MR3068976 <https://doi.org/10.1090/S0002-9939-2013-11886-1>
- [5] Barlow, M.T., Jacka, S.D. and Yor, M. (1986). Inequalities for a pair of processes stopped at a random time. *Proc. Lond. Math. Soc.* (3) **52** 142–172. MR0812449 <https://doi.org/10.1112/plms/s3-52.1.142>
- [6] Chatterji, N., Diakonikolas, J., Jordan, M.I. and Bartlett, P. (2020). Langevin Monte Carlo without smoothness. In *International Conference on Artificial Intelligence and Statistics* 1716–1726. PMLR.
- [7] Chen, L.H.Y., Fang, X. and Shao, Q.-M. (2013). From Stein identities to moderate deviations. *Ann. Probab.* **41** 262–293. MR3059199 <https://doi.org/10.1214/12-AOP746>
- [8] Chen, X., Shao, Q.-M., Wu, W.B. and Xu, L. (2016). Self-normalized Cramér-type moderate deviations under dependence. *Ann. Statist.* **44** 1593–1617. MR3519934 <https://doi.org/10.1214/15-AOS1429>

- [9] 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>
- [10] Dedecker, J. and Gouëzel, S. (2015). Subgaussian concentration inequalities for geometrically ergodic Markov chains. *Electron. Commun. Probab.* **20** no. 64, 12. [MR3407208](#) <https://doi.org/10.1214/ECP.v20-3966>
- [11] Del Moral, P., Hu, S. and Wu, L. (2015). Moderate deviations for interacting processes. *Statist. Sinica* **25** 921–951. [MR3409731](#)
- [12] Dupuis, P. and Johnson, D. (2017). Moderate deviations-based importance sampling for stochastic recursive equations. *Adv. in Appl. Probab.* **49** 981–1010. [MR3732184](#) <https://doi.org/10.1017/apr.2017.31>
- [13] 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>
- [14] Durmus, A. and Moulines, É. (2019). High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli* **25** 2854–2882. [MR4003567](#) <https://doi.org/10.3150/18-BEJ1073>
- [15] Fan, X. (2020). Cramér type moderate deviations for self-normalized ψ -mixing sequences. *J. Math. Anal. Appl.* **486** 123902, 17. [MR4060085](#) <https://doi.org/10.1016/j.jmaa.2020.123902>
- [16] Fan, X., Grama, I., Liu, Q. and Shao, Q.-M. (2019). Self-normalized Cramér type moderate deviations for martingales. *Bernoulli* **25** 2793–2823. [MR4003565](#) <https://doi.org/10.3150/18-BEJ1071>
- [17] Fan, X., Grama, I., Liu, Q. and Shao, Q.-M. (2020). Self-normalized Cramér type moderate deviations for stationary sequences and applications. *Stochastic Process. Appl.* **130** 5124–5148. [MR4108484](#) <https://doi.org/10.1016/j.spa.2020.03.001>
- [18] Fang, X., Luo, L. and Shao, Q.-M. (2020). A refined Cramér-type moderate deviation for sums of local statistics. *Bernoulli* **26** 2319–2352. [MR4091111](#) <https://doi.org/10.3150/20-BEJ1195>
- [19] Fang, X., Shao, Q.-M. and Xu, L. (2019). Multivariate approximations in Wasserstein distance by Stein’s method and Bismut’s formula. *Probab. Theory Related Fields* **174** 945–979. [MR3980309](#) <https://doi.org/10.1007/s00440-018-0874-5>
- [20] Jing, B.-Y., Shao, Q.-M. and Wang, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann. Probab.* **31** 2167–2215. [MR2016616](#) <https://doi.org/10.1214/aop/1068646382>
- [21] Jing, B.-Y., Wang, Q. and Zhou, W. (2015). Cramér-type moderate deviation for studentized compound Poisson sum. *J. Theoret. Probab.* **28** 1556–1570. [MR3422942](#) <https://doi.org/10.1007/s10959-014-0542-3>
- [22] Krylov, N.V. and Priola, E. (2010). Elliptic and parabolic second-order PDEs with growing coefficients. *Comm. Partial Differential Equations* **35** 1–22. [MR2748616](#) <https://doi.org/10.1080/03605300903424700>
- [23] Lin, Z. and Bai, Z. (2011). *Probability Inequalities*. Berlin: Springer.
- [24] Lu, J., Tan, Y. and Xu, L. (2022). Supplement to “Central limit theorem and self-normalized Cramér-type moderate deviation for Euler-Maruyama scheme.” <https://doi.org/10.3150/21-BEJ1372SUPP>
- [25] Majka, M.B., Mijatović, A. and Szpruch, Ł. (2020). Nonasymptotic bounds for sampling algorithms without log-concavity. *Ann. Appl. Probab.* **30** 1534–1581. [MR4132634](#) <https://doi.org/10.1214/19-AAP1535>
- [26] Mao, X. (2008). *Stochastic Differential Equations and Applications*, 2nd ed. Chichester: Horwood Publishing Limited. [MR2380366](#) <https://doi.org/10.1533/9780857099402>
- [27] McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628. [MR0358933](#) <https://doi.org/10.1214/aop/1176996608>
- [28] 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>
- [29] 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>
- [30] Nyquist, P. (2017). Moderate deviation principles for importance sampling estimators of risk measures. *J. Appl. Probab.* **54** 490–506. [MR3668479](#) <https://doi.org/10.1017/jpr.2017.13>
- [31] Partington, J.R. (2004). *Linear Operators and Linear Systems: An Analytical Approach to Control Theory*. London Mathematical Society Student Texts **60**. Cambridge: Cambridge Univ. Press. [MR2158502](#) <https://doi.org/10.1017/CBO9780511616693>
- [32] Revuz, D. and Yor, M. (2013). *Continuous martingales and Brownian motion* **293**. Heidelberg: Springer.
- [33] Roberts, G.O. and Rosenthal, J.S. (2004). General state space Markov chains and MCMC algorithms. *Probab. Surv.* **1** 20–71. [MR2095565](#) <https://doi.org/10.1214/154957804100000024>

- [34] 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>
- [35] Shao, J. (2018). Invariant measures and Euler-Maruyama’s approximations of state-dependent regime-switching diffusions. *SIAM J. Control Optim.* **56** 3215–3238. [MR3853602](#) <https://doi.org/10.1137/18M116678X>
- [36] Shao, Q.-M. (1999). A Cramér type large deviation result for Student’s t -statistic. *J. Theoret. Probab.* **12** 385–398. [MR1684750](#) <https://doi.org/10.1023/A:1021626127372>
- [37] Shao, Q.-M., Zhang, M. and Zhang, Z.-S. (2021). Cramér-type moderate deviation theorems for nonnormal approximation. *Ann. Appl. Probab.* **31** 247–283. [MR4254480](#) <https://doi.org/10.1214/20-aap1589>
- [38] Shao, Q.-M. and Zhou, W.-X. (2016). Cramér type moderate deviation theorems for self-normalized processes. *Bernoulli* **22** 2029–2079. [MR3498022](#) <https://doi.org/10.3150/15-BEJ719>
- [39] 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](#)
- [40] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1762. [MR1329166](#) <https://doi.org/10.1214/aos/1176325750>

On the measure of anchored Gaussian simplices, with applications to multivariate medians

DAVY PAINDAVEINE^{1,2}

¹*ECARES and Département de Mathématique, Université libre de Bruxelles, Avenue F.D. Roosevelt, 50, ECARES, CP114/04, B-1050, Brussels, Belgium.*

E-mail: dpaindav@ulb.ac.be

²*Toulouse School of Economics, Université Toulouse Capitole, 1, Esplanade de l'Université, 31080 Toulouse Cedex 06, France*

We consider anchored Gaussian ℓ -simplices in the d -dimensional Euclidean space, that is, simplices with one fixed vertex $y \in \mathbb{R}^d$ and the remaining vertices $X_1, \dots, X_\ell$ randomly sampled from the d -variate standard normal distribution. We determine the distribution of the measure of such simplices for any d , any ℓ , and any anchor point y , which is of interest, e.g., when studying the asymptotic behaviour of U-statistics based on such simplex measures. We provide two proofs of the results. The first one is short but is not self-contained as it crucially relies on a technical result for non-central Wishart distributions. The second one is a simple and self-contained proof, that also provides some geometric insight on the results. Quite nicely, variations on this second argument reveal intriguing distributional identities on products of central and non-central chi-square distributions with Beta-distributed non-centrality parameters. We independently establish these distributional identities by making use of Mellin transforms. Beyond the aforementioned use to study the asymptotic behaviour of some U-statistics, our results do find natural applications in the context of robust location estimation, as we illustrate by considering a class of simplex-based multivariate medians that contains the celebrated spatial median and Oja median as special cases. Throughout, our results are confirmed by numerical experiments.

Keywords: Distributional identities; Mellin transforms; Multivariate medians; Oja median; Random simplices; Spatial median; Stochastic geometry

References

- [1] Abramowitz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*, 9th ed. New York: Dover Publications. [MR0208797](#)
- [2] Anderson, W.J. (1986). On certain random simplices in $\mathbb{R}^n$. *J. Multivariate Anal.* **19** 265–272. [MR0853057](#) [https://doi.org/10.1016/0047-259X\(86\)90031-X](https://doi.org/10.1016/0047-259X(86)90031-X)
- [3] Arcones, M.A. (1995). Asymptotic normality of multivariate trimmed means. *Statist. Probab. Lett.* **25** 43–53. [MR1364817](#) [https://doi.org/10.1016/0167-7152\(94\)00204-L](https://doi.org/10.1016/0167-7152(94)00204-L)
- [4] Arcones, M.A. (1998). Asymptotic theory for M -estimators over a convex kernel. *Econometric Theory* **14** 387–422. [MR1650029](#) <https://doi.org/10.1017/S0266466698144018>
- [5] Baryshnikov, Y.M. and Vitale, R.A. (1994). Regular simplices and Gaussian samples. *Discrete Comput. Geom.* **11** 141–147. [MR1254086](#) <https://doi.org/10.1007/BF02574000>
- [6] Bilodeau, M. and Brenner, D. (1999). *Theory of Multivariate Statistics*. Springer Texts in Statistics. New York: Springer. [MR1705291](#)
- [7] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge: Cambridge Univ. Press. [MR2061575](#) <https://doi.org/10.1017/CBO9780511804441>
- [8] Brown, B.M. (1983). Statistical uses of the spatial median. *J. Roy. Statist. Soc. Ser. B* **45** 25–30. [MR0701072](#)
- [9] Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.* **91** 862–872. [MR1395753](#) <https://doi.org/10.2307/2291681>
- [10] Dahel, S. and Giri, N. (1994). Some distributions related to a noncentral Wishart distribution. *Comm. Statist. Theory Methods* **23** 229–237. [MR1261743](#) <https://doi.org/10.1080/03610929408831249>

- [11] Dürre, A. and Paindaveine, D. (2021). Affine-equivariant inference for multivariate location under L_p loss functionals. Submitted.
- [12] Epstein, B. (1948). Some applications of the Mellin transform in statistics. *Ann. Math. Stat.* **19** 370–379. [MR0029128 https://doi.org/10.1214/aoms/1177730201](https://doi.org/10.1214/aoms/1177730201)
- [13] Fischer, D., Mosler, K., Möttönen, J., Nordhausen, K., Pokotylo, O. and Vogel, D. (2020). Computing the Oja median in R: The package OjaNP. *J. Stat. Softw.* **92**.
- [14] Grote, J., Kabluchko, Z. and Thäle, C. (2019). Limit theorems for random simplices in high dimensions. *ALEA Lat. Am. J. Probab. Math. Stat.* **16** 141–177. [MR3903027 https://doi.org/10.30757/alea.v16-06](https://doi.org/10.30757/alea.v16-06)
- [15] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics: The Approach Based on Influence Functions*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: Wiley. [MR0829458](https://doi.org/10.1002/9781118133265)
- [16] Hettmansperger, T.P., Möttönen, J. and Oja, H. (1997). Affine-invariant multivariate one-sample signed-rank tests. *J. Amer. Statist. Assoc.* **92** 1591–1600. [MR1615268 https://doi.org/10.2307/2965430](https://doi.org/10.2307/2965430)
- [17] Hettmansperger, T.P., Nyblom, J. and Oja, H. (1994). Affine invariant multivariate one-sample sign tests. *J. Roy. Statist. Soc. Ser. B* **56** 221–234. [MR1257809](https://doi.org/10.2307/2346499)
- [18] Hettmansperger, T.P. and Randles, R.H. (2002). A practical affine equivariant multivariate median. *Biometrika* **89** 851–860. [MR1946515 https://doi.org/10.1093/biomet/89.4.851](https://doi.org/10.1093/biomet/89.4.851)
- [19] Hsu, D. and Sabato, S. (2016). Loss minimization and parameter estimation with heavy tails. *J. Mach. Learn. Res.* **17** Paper No. 18, 40. [MR3491112](https://doi.org/10.26434/chemrxiv-2016-08-00000)
- [20] Kabluchko, Z. and Zaporozhets, D. (2020). Absorption probabilities for Gaussian polytopes and regular spherical simplices. *Adv. in Appl. Probab.* **52** 588–616. [MR4123647 https://doi.org/10.1017/apr.2020.7](https://doi.org/10.1017/apr.2020.7)
- [21] León, C.A. and Massé, J.-C. (1993). La médiane simpliciale d'Oja: Existence, unicité et stabilité. *Canad. J. Statist.* **21** 397–408. [MR1254286 https://doi.org/10.2307/3315703](https://doi.org/10.2307/3315703)
- [22] Lopuhaä, H.P. and Rousseeuw, P.J. (1991). Breakdown points of affine equivariant estimators of multivariate location and covariance matrices. *Ann. Statist.* **19** 229–248. [MR1091847 https://doi.org/10.1214/aos/1176347978](https://doi.org/10.1214/aos/1176347978)
- [23] Lugosi, G. and Mendelson, S. (2019). Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.* **19** 1145–1190. [MR4017683 https://doi.org/10.1007/s10208-019-09427-x](https://doi.org/10.1007/s10208-019-09427-x)
- [24] Magyar, A. and Tyler, D.E. (2011). The asymptotic efficiency of the spatial median for elliptically symmetric distributions. *Sankhya B* **73** 165–192. [MR2924190 https://doi.org/10.1007/s13571-011-0032-x](https://doi.org/10.1007/s13571-011-0032-x)
- [25] Mathai, A.M. (1982). On a conjecture in geometric probability regarding asymptotic normality of a random simplex. *Ann. Probab.* **10** 247–251. [MR0637392](https://doi.org/10.2307/2286732)
- [26] Mathai, A.M. (1999). Random p -content of a p -parallelotope in Euclidean n -space. *Adv. in Appl. Probab.* **31** 343–354. [MR1724556 https://doi.org/10.1239/aap/1029955138](https://doi.org/10.1239/aap/1029955138)
- [27] Miles, R.E. (1971). Isotropic random simplices. *Adv. in Appl. Probab.* **3** 353–382. [MR0309164 https://doi.org/10.2307/1426176](https://doi.org/10.2307/1426176)
- [28] Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** 2308–2335. [MR3378468 https://doi.org/10.3150/14-BEJ645](https://doi.org/10.3150/14-BEJ645)
- [29] Möttönen, J., Hettmansperger, T.P., Oja, H. and Tienari, J. (1998). On the efficiency of affine invariant multivariate rank tests. *J. Multivariate Anal.* **66** 118–132. [MR1648529 https://doi.org/10.1006/jmva.1998.1740](https://doi.org/10.1006/jmva.1998.1740)
- [30] Möttönen, J., Nordhausen, K. and Oja, H. (2010). Asymptotic theory of the spatial median. In *Nonparametrics and Robustness in Modern Statistical Inference and Time Series Analysis: A Festschrift in Honor of Professor Jana Jurečková*. Inst. Math. Stat. (IMS) Collect. **7** 182–193. Beachwood, OH: IMS. [MR2808378](https://doi.org/10.1002/9781118133265.ch10)
- [31] Möttönen, J. and Oja, H. (1995). Multivariate spatial sign and rank methods. *J. Nonparametr. Stat.* **5** 201–213. [MR1346895 https://doi.org/10.1080/10485259508832643](https://doi.org/10.1080/10485259508832643)
- [32] Muirhead, R.J. (2005). *Aspects of Multivariate Statistical Theory*. Hoboken, NJ: Wiley. [MR0652932](https://doi.org/10.1002/9781118133265.ch10)
- [33] Niinimaa, A. and Oja, H. (1995). On the influence functions of certain bivariate medians. *J. Roy. Statist. Soc. Ser. B* **57** 565–574. [MR1341324](https://doi.org/10.2307/2346499)
- [34] Niinimaa, A., Oja, H. and Tableman, M. (1990). The finite-sample breakdown point of the Oja bivariate median and of the corresponding half-samples version. *Statist. Probab. Lett.* **10** 325–328. [MR1069910 https://doi.org/10.1016/0167-7152\(90\)90050-H](https://doi.org/10.1016/0167-7152(90)90050-H)
- [35] Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statist. Probab. Lett.* **1** 327–332. [MR0721446 https://doi.org/10.1016/0167-7152\(83\)90054-8](https://doi.org/10.1016/0167-7152(83)90054-8)

- [36] Oja, H. (1999). Affine invariant multivariate sign and rank tests and corresponding estimates: A review. *Scand. J. Stat.* **26** 319–343. [MR1712063](#) <https://doi.org/10.1111/1467-9469.00152>
- [37] Oja, H. (2010). *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks. Lecture Notes in Statistics* **199**. New York: Springer. [MR2598854](#) <https://doi.org/10.1007/978-1-4419-0468-3>
- [38] Oja, H. and Niinimaa, A. (1985). Asymptotic properties of the generalized median in the case of multivariate normality. *J. Roy. Statist. Soc. Ser. B* **47** 372–377. [MR0816103](#)
- [39] Ollila, E., Oja, H. and Hettmansperger, T.P. (2002). Estimates of regression coefficients based on the sign covariance matrix. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 447–466. [MR1924300](#) <https://doi.org/10.1111/1467-9868.00344>
- [40] Paouris, G., Pivovarov, P. and Valettas, P. (2017). Gaussian convex bodies: A non-asymptotic approach. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **457** 286–316. [MR3723587](#) <https://doi.org/10.1007/s10958-019-04256-3>
- [41] Paouris, G., Pivovarov, P. and Zinn, J. (2014). A central limit theorem for projections of the cube. *Probab. Theory Related Fields* **159** 701–719. [MR3230006](#) <https://doi.org/10.1007/s00440-013-0518-8>
- [42] Pivovarov, P. (2010). On determinants and the volume of random polytopes in isotropic convex bodies. *Geom. Dedicata* **149** 45–58. [MR2737677](#) <https://doi.org/10.1007/s10711-010-9462-2>
- [43] Ruben, H. (1979). The volume of an isotropic random parallelotope. *J. Appl. Probab.* **16** 84–94. [MR0520939](#) <https://doi.org/10.1017/s0021900200046222>
- [44] Ruben, H. and Miles, R.E. (1980). A canonical decomposition of the probability measure of sets of isotropic random points in $\mathbb{R}^n$. *J. Multivariate Anal.* **10** 1–18. [MR0569792](#) [https://doi.org/10.1016/0047-259X\(80\)90077-9](https://doi.org/10.1016/0047-259X(80)90077-9)
- [45] Serfling, R. (2010). Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardisation. *J. Nonparametr. Stat.* **22** 915–936. [MR2738875](#) <https://doi.org/10.1080/10485250903431710>
- [46] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics. Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. [MR0595165](#)
- [47] Shen, G. (2008). Asymptotics of Oja median estimate. *Statist. Probab. Lett.* **78** 2137–2141. [MR2458023](#) <https://doi.org/10.1016/j.spl.2008.02.004>
- [48] 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>
- [49] Vitale, R.A. (2007). Multivariate medians and measure-symmetrization. In *Complex Datasets and Inverse Problems. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **54** 260–267. Beachwood, OH: IMS. [MR2459194](#) <https://doi.org/10.1214/074921707000000201>
- [50] Wells, W.T., Anderson, R.L. and Cell, J.W. (1962). The distribution of the product of two central or noncentral chi-square variates. *Ann. Math. Stat.* **33** 1016–1020. [MR0142170](#) <https://doi.org/10.1214/aoms/1177704469>
- [51] Zhou, W. and Serfling, R. (2008). Multivariate spatial U-quantiles: A Bahadur-Kiefer representation, a Theil-Sen estimator for multiple regression, and a robust dispersion estimator. *J. Statist. Plann. Inference* **138** 1660–1678. [MR2400473](#) <https://doi.org/10.1016/j.jspi.2007.05.043>

Inference in latent factor regression with clusterable features

XIN BING^{1,*}, FLORENTINA BUNEA^{1,†} and MARTEN WEGKAMP^{1,2}

¹Department of Statistics and Data Science, Cornell University, Ithaca, New York, USA.

E-mail: *xb43@cornell.edu; †fb238@cornell.edu

²Department of Mathematics, Cornell University, Ithaca, New York, USA.

E-mail: mhw73@cornell.edu

Regression models, in which the observed features $X \in \mathbb{R}^p$ and the response $Y \in \mathbb{R}$ depend, jointly, on a lower dimensional, unobserved, latent vector $Z \in \mathbb{R}^K$, with $K \ll p$, are popular in a large array of applications, and mainly used for predicting a response from correlated features. In contrast, methodology and theory for inference on the regression coefficient $\beta \in \mathbb{R}^K$ relating Y to Z are scarce, since typically the un-observable factor Z is hard to interpret. Furthermore, the determination of the asymptotic variance of an estimator of β is a long-standing problem, with solutions known only in a few particular cases.

To address some of these outstanding questions, we develop inferential tools for β in a class of factor regression models in which the observed features are signed mixtures of the latent factors. The model specifications are both practically desirable, in a large array of applications, render interpretability to the components of Z , and are sufficient for parameter identifiability.

Without assuming that the number of latent factors K or the structure of the mixture is known in advance, we construct computationally efficient estimators of β , along with estimators of other important model parameters. We benchmark the rate of convergence of β by first establishing its ℓ_2 -norm minimax lower bound, and show that our proposed estimator $\hat{\beta}$ is minimax-rate adaptive. Our main contribution is the provision of a unified analysis of the component-wise Gaussian asymptotic distribution of $\hat{\beta}$ and, especially, the derivation of a closed form expression of its asymptotic variance, together with consistent variance estimators. The resulting inferential tools can be used when both K and p are independent of the sample size n , and also when both, or either, p and K vary with n , while allowing for $p > n$. This complements the only asymptotic normality results obtained for a particular case of the model under consideration, in the regime $K = O(1)$ and $p \rightarrow \infty$, but without a variance estimate.

As an application, we provide, within our model specifications, a statistical platform for inference in regression on latent cluster centers, thereby increasing the scope of our theoretical results.

We benchmark the newly developed methodology on a recently collected data set for the study of the effectiveness of a new SIV vaccine. Our analysis enables the determination of the top latent antibody-centric mechanisms associated with the vaccine response.

Keywords: High dimensional regression; latent factor model; identification; uniform inference; minimax estimation; pure variables; post clustering inference/regression; adaptive estimation

References

- [1] Ackerman, M.E., Das, J., Pittala, S., Broge, T., Linde, C., Suscovich, T.J., Brown, E.P., Bradley, T., Nataraajan, H., Lin, S., Sassic, J.K., O’Keefe, S., Mehta, N., Goodman, D., Sips, M., Weiner, J.A., Tomaras, G.D., Haynes, B.F., Lauffenburger, D.A., Bailey-Kellogg, C., Roederer, M. and Alter, G. (2018). Route of immunization defines multiple mechanisms of vaccine-mediated protection against SIV. *Nat. Med.* **24** 1590–1598. <https://doi.org/10.1038/s41591-018-0161-0>
- [2] Ahn, S.C. and Horenstein, A.R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81** 1203–1227. [MR3064065 https://doi.org/10.3982/ECTA8968](https://doi.org/10.3982/ECTA8968)
- [3] Anderson, T.W. and Amemiya, Y. (1988). The asymptotic normal distribution of estimators in factor analysis under general conditions. *Ann. Statist.* **16** 759–771. [MR0947576 https://doi.org/10.1214/aos/1176350834](https://doi.org/10.1214/aos/1176350834)

- [4] Anderson, T.W. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, Vol. V 111–150. Berkeley and Los Angeles, CA: Univ. California Press. [MR0084943](#)
- [5] Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* **71** 135–171. [MR1956857](#) <https://doi.org/10.1111/1468-0262.00392>
- [6] Bai, J. and Li, K. (2012). Statistical analysis of factor models of high dimension. *Ann. Statist.* **40** 436–465. [MR3014313](#) <https://doi.org/10.1214/11-AOS966>
- [7] 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>
- [8] Bai, J. and Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica* **74** 1133–1150. [MR2238213](#) <https://doi.org/10.1111/j.1468-0262.2006.00696.x>
- [9] Bai, J. and Ng, S. (2008). Forecasting economic time series using targeted predictors. *J. Econometrics* **146** 304–317. [MR2465175](#) <https://doi.org/10.1016/j.jeconom.2008.08.010>
- [10] Bair, E., Hastie, T., Paul, D. and Tibshirani, R. (2006). Prediction by supervised principal components. *J. Amer. Statist. Assoc.* **101** 119–137. [MR2252436](#) <https://doi.org/10.1198/016214505000000628>
- [11] Belloni, A., Rosenbaum, M. and Tsybakov, A.B. (2017). Linear and conic programming estimators in high dimensional errors-in-variables models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 939–956. [MR3641415](#) <https://doi.org/10.1111/rssb.12196>
- [12] Bien, J., Bunea, F. and Xiao, L. (2016). Convex banding of the covariance matrix. *J. Amer. Statist. Assoc.* **111** 834–845. [MR3538709](#) <https://doi.org/10.1080/01621459.2015.1058265>
- [13] Bing, X., Bunea, F., Ning, Y. and Wegkamp, M. (2020). Adaptive estimation in structured factor models with applications to overlapping clustering. *Ann. Statist.* **48** 2055–2081. [MR4134786](#) <https://doi.org/10.1214/19-AOS1877>
- [14] Bing, X., Bunea, F. and Wegkamp, M. (2020). Detecting approximate replicate components of a high-dimensional random vector with latent structure. [arXiv:2010.02288](#).
- [15] Bing, X., Bunea, F. and Wegkamp, M. (2022). Supplement to “Inference in latent factor regression with clusterable features.” <https://doi.org/10.3150/21-BEJ1374SUPP>
- [16] Blei, D.M. and McAuliffe, J.D. (2007). Supervised topic models. In *Proceedings of the 20th International Conference on Neural Information Processing Systems. NIPS'07* 121–128. USA: Curran Associates Inc.
- [17] Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *J. Econometrics* **132** 169–194. [MR2271395](#) <https://doi.org/10.1016/j.jeconom.2005.01.027>
- [18] Bollen, K.A. (1989). *Structural Equations with Latent Variables*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: Wiley. [MR0996025](#) <https://doi.org/10.1002/9781118619179>
- [19] Bournazos, S. and Ravetch, J.V. (2017). Fcγ receptor function and the design of vaccination strategies. *Immunity* **47** 224–233. <https://doi.org/10.1016/j.immuni.2017.07.009>
- [20] Bunea, F., She, Y. and Wegkamp, M.H. (2011). Optimal selection of reduced rank estimators of high-dimensional matrices. *Ann. Statist.* **39** 1282–1309. [MR2816355](#) <https://doi.org/10.1214/11-AOS876>
- [21] Bunea, F., Strimas-Mackey, S. and Wegkamp, M. (2020). Interpolating Predictors in High-Dimensional Factor Regression. [arXiv:2002.02525](#).
- [22] Bunea, F. and Xiao, L. (2015). On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA. *Bernoulli* **21** 1200–1230. [MR3338661](#) <https://doi.org/10.3150/14-BEJ602>
- [23] Chamberlain, G. and Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* **51** 1281–1304. [MR0736050](#) <https://doi.org/10.2307/1912275>
- [24] Chandrasekaran, V., Sanghavi, S., Parrilo, P.A. and Willsky, A.S. (2009). Sparse and low-rank matrix decompositions. *IFAC Proc. Vol.* **42** 1493–1498. 15th IFAC Symposium on System Identification. <https://doi.org/10.3182/20090706-3-FR-2004.00249>
- [25] Connor, G. and Korajczyk, R. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *J. Financ. Econ.* **15** 373–394.
- [26] Fan, J., Liao, Y. and Mincheva, M. (2011). High-dimensional covariance matrix estimation in approximate factor models. *Ann. Statist.* **39** 3320–3356. [MR3012410](#) <https://doi.org/10.1214/11-AOS944>

- [27] Fan, J., Liao, Y. and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 603–680. [MR3091653](#) <https://doi.org/10.1111/rssb.12016>
- [28] Fan, J., Xue, L. and Yao, J. (2017). Sufficient forecasting using factor models. *J. Econometrics* **201** 292–306. [MR3717565](#) <https://doi.org/10.1016/j.jeconom.2017.08.009>
- [29] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *Rev. Econ. Stat.* **82** 540–554.
- [30] Forni, M. and Reichlin, L. (1996). Dynamic common factors in large cross-sections. *Empir. Econ.* **21** 27–42.
- [31] Giraud, C. (2015). *Introduction to High-Dimensional Statistics. Monographs on Statistics and Applied Probability* **139**. Boca Raton, FL: CRC Press. [MR3307991](#)
- [32] Hahn, P.R., Carvalho, C.M. and Mukherjee, S. (2013). Partial factor modeling: Predictor-dependent shrinkage for linear regression. *J. Amer. Statist. Assoc.* **108** 999–1008. [MR3174679](#) <https://doi.org/10.1080/01621459.2013.779843>
- [33] Jöreskog, K.G. (1970). A general method for analysis of covariance structures. *Biometrika* **57** 239–251. [MR0269024](#) <https://doi.org/10.2307/2334833>
- [34] Jöreskog, K.G. (1970). A general method for estimating a linear structural equation system. *ETS Research Bulletin Series* **1970** i–41. <https://doi.org/10.1002/j.2333-8504.1970.tb00783.x>
- [35] Kelly, B. and Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *J. Econometrics* **186** 294–316. [MR3343788](#) <https://doi.org/10.1016/j.jeconom.2015.02.011>
- [36] Koopmans, T.C. and Reiersøl, O. (1950). The identification of structural characteristics. *Ann. Math. Stat.* **21** 165–181. [MR0039967](#) <https://doi.org/10.1214/aoms/1177729837>
- [37] Lawley, D.N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proc. R. Soc. Edinb.* **60** 64–82. [MR0002754](#)
- [38] Lawley, D.N. and Maxwell, A.E. (1971). *Factor Analysis as a Statistical Method*, 2nd ed. New York: American Elsevier Publishing Co., Inc. [MR0343471](#)
- [39] McDonald, R.P. (1999). *Test Theory: A Unified Treatment*. London: Taylor & Francis.
- [40] Nimmerjahn, F. and Ravetch, J.V. (2007). Fcγ receptors as regulators of immunity. *Advances in Immunology* **96** 179–204. Academic Press. [https://doi.org/10.1016/S0065-2776\(07\)96005-8](https://doi.org/10.1016/S0065-2776(07)96005-8)
- [41] 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>
- [42] Reiß, M. and Wahl, M. (2016). Non-asymptotic upper bounds for the reconstruction error of PCA.
- [43] Stock, J.H. and Watson, M.W. (2002). Forecasting using principal components from a large number of predictors. *J. Amer. Statist. Assoc.* **97** 1167–1179. [MR1951271](#) <https://doi.org/10.1198/016214502388618960>
- [44] Stock, J.H. and Watson, M.W. (2002). Macroeconomic forecasting using diffusion indexes. *J. Bus. Econom. Statist.* **20** 147–162. [MR1963257](#) <https://doi.org/10.1198/073500102317351921>
- [45] Thurstone, L.L. (1931). Multiple factor analysis. *Psychol. Rev.* **38** 406–427.
- [46] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. New York: Springer. [MR2724359](#) <https://doi.org/10.1007/b13794>
- [47] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. [MR2963170](#)
- [48] 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>
- [49] Wegkamp, M. and Zhao, Y. (2016). Adaptive estimation of the copula correlation matrix for semiparametric elliptical copulas. *Bernoulli* **22** 1184–1226. [MR3449812](#) <https://doi.org/10.3150/14-BEJ690>
- [50] Yalcin, I. and Amemiya, Y. (2001). Nonlinear factor analysis as a statistical method. *Statist. Sci.* **16** 275–294. [MR1874155](#) <https://doi.org/10.1214/ss/1009213729>

Joint inference on extreme expectiles for multivariate heavy-tailed distributions

SIMONE A. PADOAN¹ and GILLES STUPFLER²

¹Department of Decision Sciences, Bocconi University, via Roentgen 1, 20136 Milano, Italy.

E-mail: simone.padoan@unibocconi.it

²Univ Rennes, Ensai, CNRS, CREST - UMR 9194, Rennes, France.

E-mail: gilles.stupfler@ensai.fr

Expectiles induce a law-invariant, coherent and elicitable risk measure that has received substantial attention in actuarial and financial risk management contexts. A number of recent papers have focused on the behaviour and estimation of extreme expectile-based risk measures and their potential for risk assessment was highlighted in financial and actuarial real data applications. Joint inference of several extreme expectiles has however been left untouched; in fact, even the inference about a marginal extreme expectile turns out to be a difficult problem in finite samples, even though an accurate idea of estimation uncertainty is crucial for the construction of confidence intervals in applications to risk management. We investigate the joint estimation of extreme marginal expectiles of a random vector with heavy-tailed marginal distributions, in a general extremal dependence model. We use these results to derive corrected confidence regions for extreme expectiles, as well as a test for the equality of tail expectiles. The methods are showcased in a finite-sample simulation study and on real financial data.

Keywords: Expectiles; extremal dependence; heavy tails; joint convergence; joint inference; tail copula; testing

References

- [1] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Math. Finance* **9** 203–228. [MR1850791](#) <https://doi.org/10.1111/1467-9965.00068>
- [2] Bellini, F. and Di Bernardino, E. (2017). Risk management with expectiles. *Eur. J. Finance* **23** 487–506.
- [3] Bellini, F., Klar, B., Müller, A. and Rosazza Gianin, E. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.* **54** 41–48. [MR3145849](#) <https://doi.org/10.1016/j.insmatheco.2013.10.015>
- [4] Beranger, B. and Padoan, S.A. (2015). Extreme dependence models. In *Extreme Value Modeling and Risk Analysis: Methods and Applications* (D. Dey and J. Yan, eds.) CRC Press/CRC.
- [5] Cai, J.-J., de Haan, L. and Zhou, C. (2013). Bias correction in extreme value statistics with index around zero. *Extremes* **16** 173–201. [MR3057195](#) <https://doi.org/10.1007/s10687-012-0158-x>
- [6] Cai, J.-J., Einmahl, J.H.J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: The mean when a related variable is extreme. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 417–442. [MR3310533](#) <https://doi.org/10.1111/rssb.12069>
- [7] Daouia, A., Girard, S. and Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 263–292. [MR3763692](#) <https://doi.org/10.1111/rssb.12254>
- [8] Daouia, A., Girard, S. and Stupfler, G. (2019). Extreme M-quantiles as risk measures: From L^1 to L^p optimization. *Bernoulli* **25** 264–309. [MR3892320](#) <https://doi.org/10.3150/17-bej987>
- [9] Daouia, A., Girard, S. and Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli* **26** 531–556. [MR4036043](#) <https://doi.org/10.3150/19-BEJ1137>
- [10] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Series in Operations Research and Financial Engineering. New York: Springer. [MR2234156](#) <https://doi.org/10.1007/0-387-34471-3>
- [11] de Haan, L., Mercadier, C. and Zhou, C. (2016). Adapting extreme value statistics to financial time series: Dealing with bias and serial dependence. *Finance Stoch.* **20** 321–354. [MR3479324](#) <https://doi.org/10.1007/s00780-015-0287-6>

- [12] Drees, H. (1998). Optimal rates of convergence for estimates of the extreme value index. *Ann. Statist.* **26** 434–448. [MR1608148](#) <https://doi.org/10.1214/aos/1030563992>
- [13] Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli* **9** 617–657. [MR1996273](#) <https://doi.org/10.3150/bj/1066223272>
- [14] Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function. *J. Multivariate Anal.* **64** 25–47. [MR1619974](#) <https://doi.org/10.1006/jmva.1997.1708>
- [15] Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 505–562. [MR3506792](#) <https://doi.org/10.1111/rssb.12154>
- [16] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. [MR1458613](#) <https://doi.org/10.1007/978-3-642-33483-2>
- [17] Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. [MR2847988](#) <https://doi.org/10.1198/jasa.2011.r10138>
- [18] Goegebeur, Y., Beirlant, J. and de Wet, T. (2010). Kernel estimators for the second order parameter in extreme value statistics. *J. Statist. Plann. Inference* **140** 2632–2652. [MR2644084](#) <https://doi.org/10.1016/j.jspi.2010.03.029>
- [19] Gomes, M.I., Pestana, D. and Caeiro, F. (2009). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statist. Probab. Lett.* **79** 295–303. [MR2493012](#) <https://doi.org/10.1016/j.spl.2008.08.016>
- [20] Hall, P. and Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation. *Ann. Statist.* **13** 331–341. [MR0773171](#) <https://doi.org/10.1214/aos/1176346596>
- [21] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174. [MR0378204](#)
- [22] Hoga, Y. (2018). Detecting tail risk differences in multivariate time series. *J. Time Series Anal.* **39** 665–689. [MR3849521](#) <https://doi.org/10.1111/jtsa.12292>
- [23] Hurlin, C., Laurent, S., Quaevlieg, R. and Smeeke, S. (2017). Risk measure inference. *J. Bus. Econom. Statist.* **35** 499–512. [MR3716038](#) <https://doi.org/10.1080/07350015.2015.1127815>
- [24] Joe, H. (2014). *Dependence Modeling with Copulas*. Boca Raton, USA: CRC Press/CRC Press.
- [25] Knight, K. (1999). Epi-convergence in distribution and stochastic equi-semicontinuity. Technical Report Univ. Toronto.
- [26] Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50. [MR0474644](#) <https://doi.org/10.2307/1913643>
- [27] Mainik, G., Mitov, G. and Rüschendorf, L. (2015). Portfolio optimization for heavy-tailed assets: Extreme risk index vs. Markowitz. *J. Empir. Finance* **32** 115–134.
- [28] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847. [MR0906565](#) <https://doi.org/10.2307/1911031>
- [29] Padoan, S.A. and Stupfler, G. (2020). Extreme expectile estimation for heavy-tailed time series. [arXiv:2004.04078v2](#).
- [30] Padoan, S.A. and Stupfler, G. (2022). Supplement to “Joint inference on extreme expectiles for multivariate heavy-tailed distributions.” <https://doi.org/10.3150/21-BEJ1375SUPP>
- [31] R Core Team (2020). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- [32] Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scand. J. Stat.* **33** 307–335. [MR2279645](#) <https://doi.org/10.1111/j.1467-9469.2005.00483.x>
- [33] Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression. *Comput. Statist. Data Anal.* **56** 755–767. [MR2888723](#) <https://doi.org/10.1016/j.csda.2010.11.015>
- [34] Straetmans, S.T.M., Verschoor, W.F.C. and Wolff, C.C.P. (2008). Extreme US stock market fluctuations in the wake of 9/11. *J. Appl. Econometrics* **23** 17–42. [MR2427314](#) <https://doi.org/10.1002/jae.973>
- [35] Stupfler, G. (2019). On a relationship between randomly and non-randomly thresholded empirical average excesses for heavy tails. *Extremes* **22** 749–769. [MR4031856](#) <https://doi.org/10.1007/s10687-019-00351-5>
- [36] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73** 812–815. [MR0521329](#)

- [37] Ziegel, J.F. (2016). Coherence and elicibility. *Math. Finance* **26** 901–918. [MR3551510](#) <https://doi.org/10.1111/mafi.12080>

Asymptotically efficient estimators for stochastic blockmodels: The naive MLE, the rank-constrained MLE, and the spectral estimator

MINH TANG¹, JOSHUA CAPE² and CAREY E. PRIEBE³

¹*Department of Statistics, North Carolina State University, Raleigh, NC, USA .
E-mail: mtang8@ncsu.edu*

²*Department of Statistics, University of Pittsburgh, Pittsburgh, PA, USA .
E-mail: joshua.cape@pitt.edu*

³*Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, USA .
E-mail: cep@jhu.edu*

We establish asymptotic normality results for estimation of the block probability matrix $\mathbf{B}$ in stochastic block-model graphs using spectral embedding when the average degrees grows at the rate of $\omega(\sqrt{n})$ in n , the number of vertices. As a corollary, we show that when $\mathbf{B}$ is of full-rank, estimates of $\mathbf{B}$ obtained from spectral embedding are asymptotically efficient. When $\mathbf{B}$ is singular the estimates obtained from spectral embedding can have smaller mean square error than those obtained from maximizing the log-likelihood under no rank assumption, and furthermore, can be almost as efficient as the true MLE that assumes the rank of $\mathbf{B}$ is known. Our results indicate, in the context of stochastic blockmodel graphs, that spectral embedding is not just computationally tractable, but that the resulting estimates are also admissible, even when compared to the purportedly optimal but computationally intractable maximum likelihood estimation under no rank assumption.

Keywords: Asymptotic efficiency; random dot product graph; stochastic blockmodels; asymptotic normality; spectral embedding

References

- [1] Abbe, E. (2017). Community detection and stochastic block models: Recent developments. *J. Mach. Learn. Res.* **18** Paper No. 177, 86. [MR3827065](#)
- [2] Abbe, E., Bandeira, A.S. and Hall, G. (2016). Exact recovery in the stochastic block model. *IEEE Trans. Inf. Theory* **62** 471–487. [MR3447993](#) <https://doi.org/10.1109/TIT.2015.2490670>
- [3] Abbe, E., Fan, J., Wang, K. and Zhong, Y. (2020). Entrywise eigenvector analysis of random matrices with low expected rank. *Ann. Statist.* **48** 1452–1474. [MR4124330](#) <https://doi.org/10.1214/19-AOS1854>
- [4] Airoldi, E.M., Blei, D.M., Fienberg, S.E. and Xing, E.P. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.
- [5] Athreya, A., Priebe, C.E., Tang, M., Lyzinski, V., Marchette, D.J. and Sussman, D.L. (2016). A limit theorem for scaled eigenvectors of random dot product graphs. *Sankhya A* **78** 1–18. [MR3494576](#) <https://doi.org/10.1007/s13171-015-0071-x>
- [6] Banerjee, D. and Ma, Z. (2017). Optimal hypothesis testing for stochastic blockmodels with growing degrees. Preprint. Available at <https://arxiv.org/abs/1705.05305>.
- [7] Bickel, P., Choi, D., Chang, X. and Zhang, H. (2013). Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. *Ann. Statist.* **41** 1922–1943. [MR3127853](#) <https://doi.org/10.1214/13-AOS1124>

- [8] Bickel, P.J. and Chen, A. (2009). A nonparametric view of network models and Newman-Girvan and other modularities. *Proc. Natl. Acad. Sci. USA* **106** 21068–73.
- [9] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins Series in the Mathematical Sciences. Baltimore, MD: Johns Hopkins Univ. Press. [MR1245941](#)
- [10] Bickel, P.J. and Sarkar, P. (2016). Hypothesis testing for automated community detection in networks. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 253–273. [MR3453655](#) <https://doi.org/10.1111/rssb.12117>
- [11] Cape, J., Tang, M. and Priebe, C.E. (2019). The two-to-infinity norm and singular subspace geometry with applications to high-dimensional statistics. *Ann. Statist.* **47** 2405–2439. [MR3988761](#) <https://doi.org/10.1214/18-AOS1752>
- [12] Celisse, A., Daudin, J.-J. and Pierre, L. (2012). Consistency of maximum-likelihood and variational estimators in the stochastic block model. *Electron. J. Stat.* **6** 1847–1899. [MR2988467](#) <https://doi.org/10.1214/12-EJS729>
- [13] Chatterjee, S. (2015). Matrix estimation by universal singular value thresholding. *Ann. Statist.* **43** 177–214. [MR3285604](#) <https://doi.org/10.1214/14-AOS1272>
- [14] Choi, D.S., Wolfe, P.J. and Airolidi, E.M. (2012). Stochastic blockmodels with a growing number of classes. *Biometrika* **99** 273–284. [MR2931253](#) <https://doi.org/10.1093/biomet/asr053>
- [15] Coja-Oghlan, A. (2010). Graph partitioning via adaptive spectral techniques. *Combin. Probab. Comput.* **19** 227–284. [MR2593622](#) <https://doi.org/10.1017/S0963548309990514>
- [16] DasGupta, A. and Johnstone, I.M. (2014). Asymptotic risk and Bayes risk of thresholding and superefficient estimates and optimal thresholding. In *Contemporary Developments in Statistical Theory*. Springer Proc. Math. Stat. **68** 41–67. Cham: Springer. [MR3149915](#) https://doi.org/10.1007/978-3-319-02651-0_4
- [17] Daudin, J.-J., Picard, F. and Robin, S. (2008). A mixture model for random graphs. *Stat. Comput.* **18** 173–183. [MR2390817](#) <https://doi.org/10.1007/s11222-007-9046-7>
- [18] Eldridge, J., Belkin, M. and Wang, Y. (2018). Unperturbed: Spectral analysis beyond Davis-Kahan. In *Proceedings of Algorithmic Learning Theory* 321–358.
- [19] Erdős, L., Knowles, A., Yau, H.-T. and Yin, J. (2013). Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.* **41** 2279–2375. [MR3098073](#) <https://doi.org/10.1214/11-AOP734>
- [20] Fan, J., Fan, Y., Han, X. and Lv, J. (2021). Asymptotic theory of eigenvectors for large random matrices. *J. Amer. Statist. Assoc.*
- [21] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. [MR1946581](#) <https://doi.org/10.1198/016214501753382273>
- [22] Fishkind, D.E., Sussman, D.L., Tang, M., Vogelstein, J.T. and Priebe, C.E. (2013). Consistent adjacency-spectral partitioning for the stochastic block model when the model parameters are unknown. *SIAM J. Matrix Anal. Appl.* **34** 23–39. [MR3032990](#) <https://doi.org/10.1137/120875600>
- [23] Fortunato, S. (2010). Community detection in graphs. *Phys. Rep.* **486** 75–174. [MR2580414](#) <https://doi.org/10.1016/j.physrep.2009.11.002>
- [24] Füredi, Z. and Komlós, J. (1981). The eigenvalues of random symmetric matrices. *Combinatorica* **1** 233–241. [MR0637828](#) <https://doi.org/10.1007/BF02579329>
- [25] Gao, C., Lu, Y. and Zhou, H.H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43** 2624–2652. [MR3405606](#) <https://doi.org/10.1214/15-AOS1354>
- [26] Hajek, B., Wu, Y. and Xu, J. (2016). Achieving exact cluster recovery threshold via semidefinite programming. *IEEE Trans. Inf. Theory* **62** 2788–2797. [MR3493879](#) <https://doi.org/10.1109/TIT.2016.2546280>
- [27] Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of Statistics 175–194. [MR0400513](#)
- [28] Hoff, P.D., Raftery, A.E. and Handcock, M.S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. [MR1951262](#) <https://doi.org/10.1198/016214502388618906>
- [29] Holland, P.W., Laskey, K.B. and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. [MR0718088](#) [https://doi.org/10.1016/0378-8733\(83\)90021-7](https://doi.org/10.1016/0378-8733(83)90021-7)
- [30] Joseph, A. and Yu, B. (2016). Impact of regularization on spectral clustering. *Ann. Statist.* **44** 1765–1791. [MR3519940](#) <https://doi.org/10.1214/16-AOS1447>

- [31] Karrer, B. and Newman, M.E.J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* (3) **83** 016107, 10. [MR2788206](https://doi.org/10.1103/PhysRevE.83.016107) <https://doi.org/10.1103/PhysRevE.83.016107>
- [32] Klopp, O., Tsybakov, A.B. and Verzelen, N. (2017). Oracle inequalities for network models and sparse graphon estimation. *Ann. Statist.* **45** 316–354. [MR3611494](https://doi.org/10.1214/16-AOS1454) <https://doi.org/10.1214/16-AOS1454>
- [33] Leeb, H. and Pötscher, B.M. (2008). Sparse estimators and the oracle property, or the return of Hodges’ estimator. *J. Econometrics* **142** 201–211. [MR2394290](https://doi.org/10.1016/j.jeconom.2007.05.017) <https://doi.org/10.1016/j.jeconom.2007.05.017>
- [34] Lei, J. (2016). A goodness-of-fit test for stochastic block models. *Ann. Statist.* **44** 401–424. [MR3449773](https://doi.org/10.1214/15-AOS1370) <https://doi.org/10.1214/15-AOS1370>
- [35] 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) <https://doi.org/10.1214/14-AOS1274>
- [36] Lovász, L. (2012). *Large Networks and Graph Limits*. American Mathematical Society Colloquium Publications **60**. Providence, RI: Amer. Math. Soc. [MR3012035](https://doi.org/10.1090/coll/060) <https://doi.org/10.1090/coll/060>
- [37] Lyzinski, V., Sussman, D.L., Tang, M., Athreya, A. and Priebe, C.E. (2014). Perfect clustering for stochastic blockmodel graphs via adjacency spectral embedding. *Electron. J. Stat.* **8** 2905–2922. [MR3299126](https://doi.org/10.1214/14-EJS978) <https://doi.org/10.1214/14-EJS978>
- [38] Lyzinski, V., Tang, M., Athreya, A., Park, Y. and Priebe, C.E. (2017). Community detection and classification in hierarchical stochastic blockmodels. *IEEE Trans. Netw. Sci. Eng.* **4** 13–26. [MR3625952](https://doi.org/10.1109/TNSE.2016.2634322) <https://doi.org/10.1109/TNSE.2016.2634322>
- [39] Magnus, J.R. and Neudecker, H. (1979). The commutation matrix: Some properties and applications. *Ann. Statist.* **7** 381–394. [MR0520247](https://doi.org/10.1214/aos/1176344977)
- [40] Magnus, J.R. and Neudecker, H. (1980). The elimination matrix: Some lemmas and applications. *SIAM J. Algebr. Discrete Methods* **1** 422–449. [MR0593853](https://doi.org/10.1137/0601049) <https://doi.org/10.1137/0601049>
- [41] Mao, X., Sarkar, P. and Chakrabarti, D. (2021). Estimating mixed memberships with sharp eigenvector deviations. *J. Amer. Statist. Assoc.*
- [42] McSherry, F. (2001). Spectral partitioning of random graphs. In *42nd IEEE Symposium on Foundations of Computer Science (Las Vegas, NV, 2001)* 529–537. Los Alamitos, CA: IEEE Computer Soc. [MR1948742](https://doi.org/10.1109/SFCS.2001.1055542)
- [43] Newman, M.E.J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Phys. Rev. E* **69** 1–15.
- [44] Oliveira, R.I. (2009). Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges. [http://arxiv.org/abs/0911.0600](https://arxiv.org/abs/0911.0600).
- [45] Pons, P. and Latapy, M. (2005). Computing communities in large networks using random walks. In *Proceedings of the 20th International Conference on Computer and Information Sciences* 284–293.
- [46] 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) <https://doi.org/10.1214/11-AOS887>
- [47] Rosvall, M. and Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* **105**.
- [48] Rubin-Delanchy, P., Cape, J., Tang, M. and Priebe, C.E. (2017). A statistical interpretation of spectral embedding: The generalised random dot product graph. Preprint. Available at <https://arxiv.org/abs/1709.05506>.
- [49] Sarkar, P. and Bickel, P.J. (2015). Role of normalization in spectral clustering for stochastic blockmodels. *Ann. Statist.* **43** 962–990. [MR3346694](https://doi.org/10.1214/14-AOS1285) <https://doi.org/10.1214/14-AOS1285>
- [50] Snijders, T.A.B. and Nowicki, K. (1997). Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classification* **14** 75–100. [MR1449742](https://doi.org/10.1007/s003579900004) <https://doi.org/10.1007/s003579900004>
- [51] Sussman, D.L., Tang, M., Fishkind, D.E. and Priebe, C.E. (2012). A consistent adjacency spectral embedding for stochastic blockmodel graphs. *J. Amer. Statist. Assoc.* **107** 1119–1128. [MR3010899](https://doi.org/10.1080/01621459.2012.699795) <https://doi.org/10.1080/01621459.2012.699795>
- [52] Tang, M., Cape, J. and Priebe, C.E. (2022). Supplement to “Asymptotically efficient estimators for stochastic blockmodels: The naive MLE, the rank-constrained MLE, and the spectral estimator.” <https://doi.org/10.3150/21-BEJ1376SUPP>
- [53] Tang, M. and Priebe, C.E. (2018). Limit theorems for eigenvectors of the normalized Laplacian for random graphs. *Ann. Statist.* **46** 2360–2415. [MR3845021](https://doi.org/10.1214/17-AOS1623) <https://doi.org/10.1214/17-AOS1623>
- [54] 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) <https://doi.org/10.1017/CBO9780511802256>

- [55] von Luxburg, U. (2007). A tutorial on spectral clustering. *Stat. Comput.* **17** 395–416. [MR2409803](#) <https://doi.org/10.1007/s11222-007-9033-z>
- [56] Vu, T., Chunikhina, E. and Raich, R. (2021). Perturbation expansions and error bounds for the truncated singular value decomposition. *Linear Algebra Appl.* **627** 94–139. [MR4275028](#) <https://doi.org/10.1016/j.laa.2021.05.020>
- [57] Wang, Y.X.R. and Bickel, P.J. (2017). Likelihood-based model selection for stochastic block models. *Ann. Statist.* **45** 500–528. [MR3650391](#) <https://doi.org/10.1214/16-AOS1457>
- [58] Wolfe, D.A. and Olhede, S.C. (2013). Nonparametric graphon estimation. Preprint. Available at <http://arxiv.org/abs/1309.5936>.
- [59] Xu, J. (2018). Rates of convergence of spectral methods for graphon estimation. In *Proceedings of the 35th International Conference on Machine Learning* 5433–5442.
- [60] Young, S.J. and Scheinerman, E.R. (2007). Random dot product graph models for social networks. In *Algorithms and Models for the Web-Graph. Lecture Notes in Computer Science* **4863** 138–149. Berlin: Springer. [MR2504912](#) https://doi.org/10.1007/978-3-540-77004-6_11

Oracle lower bounds for stochastic gradient sampling algorithms

NILADRI S. CHATTERJI¹, PETER L. BARTLETT² and PHILIP M. LONG³

¹*Department of Computer Science, Stanford University, 353 Jane Stanford Way, Stanford, CA 94305, USA.*

E-mail: niladri@cs.stanford.edu

²*University of California, Berkeley & Google, 367 Evans Hall #3860, Berkeley, CA 94720-3860, USA.*

E-mail: peter@berkeley.edu

³*Google, 1600 Amphitheatre Parkway, Mountain View, CA, 94043, USA.*

E-mail: plong@google.com

We consider the problem of sampling from a strongly log-concave density in $\mathbb{R}^d$, and prove an information theoretic lower bound on the number of stochastic gradient queries of the log density needed. Several popular sampling algorithms (including many Markov chain Monte Carlo methods) operate by using stochastic gradients of the log density to generate a sample; our results establish an information theoretic limit for all these algorithms.

We show that for every algorithm, there exists a well-conditioned strongly log-concave target density for which the distribution of points generated by the algorithm would be at least ε away from the target in total variation distance if the number of gradient queries is less than $\Omega(\sigma^2 d / \varepsilon^2)$, where $\sigma^2 d$ is the variance of the stochastic gradient. Our lower bound follows by combining the ideas of Le Cam deficiency routinely used in the comparison of statistical experiments along with standard information theoretic tools used in lower bounding Bayes risk functions. To the best of our knowledge our results provide the first nontrivial dimension-dependent lower bound for this problem.

Keywords: Sampling lower bounds; information theoretic lower bounds; Markov chain Monte Carlo; stochastic gradient Monte Carlo

References

- [1] Agarwal, A., Bartlett, P.L., Ravikumar, P. and Wainwright, M.J. (2012). Information-theoretic lower bounds on the oracle complexity of stochastic convex optimization. *IEEE Trans. Inf. Theory* **58** 3235–3249. [MR2952543](#) <https://doi.org/10.1109/TIT.2011.2182178>
- [2] Bélisle, C.J.P., Romeijn, H.E. and Smith, R.L. (1993). Hit-and-run algorithms for generating multivariate distributions. *Math. Oper. Res.* **18** 255–266. [MR1250117](#) <https://doi.org/10.1287/moor.18.2.255>
- [3] Blackwell, D. (1953). Equivalent comparisons of experiments. *Ann. Math. Stat.* **24** 265–272. [MR0056251](#) <https://doi.org/10.1214/aoms/1177729032>
- [4] Bou-Rabee, N. and Hairer, M. (2013). Nonasymptotic mixing of the MALA algorithm. *IMA J. Numer. Anal.* **33** 80–110. [MR3020951](#) <https://doi.org/10.1093/imanum/drs003>
- [5] Brooks, S., Gelman, A., Jones, G.L. and Meng, X.-L., eds. (2011). *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton, FL: CRC Press. [MR2742422](#) <https://doi.org/10.1201/b10905>
- [6] Bubeck, S., Eldan, R. and Lehec, J. (2018). Sampling from a log-concave distribution with projected Langevin Monte Carlo. *Discrete Comput. Geom.* **59** 757–783. [MR3802303](#) <https://doi.org/10.1007/s00454-018-9992-1>
- [7] Chatterji, N.S., Bartlett, P.L. and Long, P.M. (2020). Oracle lower bounds for stochastic gradient sampling algorithms. Preprint. Available at [arXiv:2002.00291](https://arxiv.org/abs/2002.00291).
- [8] Cheng, X. and Bartlett, P. (2018). Convergence of Langevin MCMC in KL-divergence. In *Proc. ALT'18* 186–211. Lanzarote, Canary Islands: PMLR.

- [9] Cheng, X., Chatterji, N., Bartlett, P. and Jordan, M. (2018). Underdamped Langevin MCMC: A non-asymptotic analysis. In *Proc. COLT'18* 300–323. Stockholm: PMLR.
- [10] 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>
- [11] 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>
- [12] Durmus, A., Majewski, S. and Miasojedow, B. (2019). Analysis of Langevin Monte Carlo via convex optimization. *J. Mach. Learn. Res.* **20** Paper No. 73, 46 pp. [MR3960927](#)
- [13] 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>
- [14] Dwivedi, R., Chen, Y., Wainwright, M.J. and Yu, B. (2019). Log-concave sampling: Metropolis-Hastings algorithms are fast. *J. Mach. Learn. Res.* **20** Paper No. 183, 42 pp. [MR4048994](#)
- [15] Dyer, M., Frieze, A. and Kannan, R. (1991). A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.* **38** 1–17. [MR1095916](#) <https://doi.org/10.1145/102782.102783>
- [16] Eberle, A., Guillin, A. and Zimmer, R. (2019). Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47** 1982–2010. [MR3980913](#) <https://doi.org/10.1214/18-AOP1299>
- [17] Ermak, D. (1975). A computer simulation of charged particles in solution. *J. Chem. Phys.* **62** 4189–4196.
- [18] Bohnenblust, H.F., Shapley, L. and Sherman, S. (1949). *Reconnaissance in Game Theory*. RAND Corporation Research Memorandum Series.
- [19] Ge, R., Lee, H. and Lu, J. (2020). Estimating normalizing constants for log-concave distributions: Algorithms and lower bounds. In *Proc. STOC'20* 579–586. New York: ACM. [MR4141783](#) <https://doi.org/10.1145/3357713.3384289>
- [20] Grenander, U. and Miller, M.I. (1994). Representations of knowledge in complex systems. *J. Roy. Statist. Soc. Ser. B* **56** 549–603. [MR1293234](#)
- [21] Johndrow, J., Pillai, N. and Smith, A. (2020). No free lunch for approximate MCMC. Preprint. Available at [arXiv:2010.12514](#).
- [22] Kallenberg, O. (2006). *Foundations of Modern Probability. Probability Theory and Stochastic Modelling* **99**. Cham: Springer. [MR4226142](#) <https://doi.org/10.1007/978-3-030-61871-1>
- [23] Kannan, R., Lovász, L. and Simonovits, M. (1995). Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* **13** 541–559. [MR1318794](#) <https://doi.org/10.1007/BF02574061>
- [24] Le Cam, L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Stat.* **35** 1419–1455. [MR0207093](#) <https://doi.org/10.1214/aoms/1177700372>
- [25] Le Cam, L. and Yang, G.L. (2012). *Asymptotics in Statistics: Some Basic Concepts*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR1784901](#) <https://doi.org/10.1007/978-1-4612-1166-2>
- [26] Lovász, L. (1999). Hit-and-run mixes fast. *Math. Program.* **86** 443–461. [MR1733749](#) <https://doi.org/10.1007/s101070050099>
- [27] Lovász, L. and Simonovits, M. (1990). The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *31st Annual Symposium on Foundations of Computer Science, Vol. I, II (St. Louis, MO, 1990)* 346–354. Los Alamitos, CA: IEEE Comput. Soc. Press. [MR1150706](#) <https://doi.org/10.1109/FSCS.1990.89553>
- [28] Lovász, L. and Simonovits, M. (1993). Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms* **4** 359–412. [MR1238906](#) <https://doi.org/10.1002/rsa.3240040402>
- [29] Lovász, L. and Vempala, S. (2006). Hit-and-run from a corner. *SIAM J. Comput.* **35** 985–1005. [MR2203735](#) <https://doi.org/10.1137/S009753970544727X>
- [30] Lovász, L. and Vempala, S. (2007). The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms* **30** 307–358. [MR2309621](#) <https://doi.org/10.1002/rsa.20135>
- [31] Mangoubi, O. and Smith, A. (2017). Rapid mixing of Hamiltonian Monte Carlo on strongly log-concave distributions. Preprint. Available at [arXiv:1708.07114](#).
- [32] Mangoubi, O. and Vishnoi, N. (2018). Dimensionally tight bounds for second-order Hamiltonian Monte Carlo. In *Proc. NeurIPS'18* 6030–6040. Montréal: PMLR.

- [33] Mengersen, K.L. and Tweedie, R.L. (1996). Rates of convergence of the Hastings and Metropolis algorithms. *Ann. Statist.* **24** 101–121. [MR1389882](#) <https://doi.org/10.1214/aos/1033066201>
- [34] Neal, R.M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC Handb. Mod. Stat. Methods 113–162. Boca Raton, FL: CRC Press. [MR2858447](#)
- [35] Nemirovsky, A. and Yudin, D. (1985). Problem complexity and method efficiency in optimization. *SIAM Rev.* **27** 264–265.
- [36] Parisi, G. (1981). Correlation functions and computer simulations. *Nuclear Phys. B* **180** 378–384. [MR0615176](#) [https://doi.org/10.1016/0550-3213\(81\)90056-0](https://doi.org/10.1016/0550-3213(81)90056-0)
- [37] Pereyra, M. (2016). Proximal Markov chain Monte Carlo algorithms. *Stat. Comput.* **26** 745–760. [MR3515019](#) <https://doi.org/10.1007/s11222-015-9567-4>
- [38] Rademacher, L. and Vempala, S. (2008). Dispersion of mass and the complexity of randomized geometric algorithms. *Adv. Math.* **219** 1037–1069. [MR2442061](#) <https://doi.org/10.1016/j.aim.2008.06.004>
- [39] Raginsky, M. and Rakhlin, A. (2011). Information-based complexity, feedback and dynamics in convex programming. *IEEE Trans. Inf. Theory* **57** 7036–7056. [MR2882278](#) <https://doi.org/10.1109/TIT.2011.2154375>
- [40] Ramdas, A. and Singh, A. (2013). Optimal rates for stochastic convex optimization under Tsybakov noise condition. In *Proc. ICML'13* 365–373. Atlanta: JMLR.org.
- [41] Robert, C.P. and Casella, G. (2013). *Monte Carlo Statistical Methods*. Springer Texts in Statistics. New York: Springer. [MR1707311](#) <https://doi.org/10.1007/978-1-4757-3071-5>
- [42] Roberts, G.O. and Stramer, O. (2002). Langevin diffusions and Metropolis-Hastings algorithms. *Methodol. Comput. Appl. Probab.* **4** 337–357. [MR2002247](#) <https://doi.org/10.1023/A:1023562417138>
- [43] 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>
- [44] Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika* **83** 95–110. [MR1399158](#) <https://doi.org/10.1093/biomet/83.1.95>
- [45] Smith, R.L. (1984). Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Oper. Res.* **32** 1296–1308. [MR0775260](#) <https://doi.org/10.1287/opre.32.6.1296>
- [46] Talwar, K. (2019). Computational separations between sampling and optimization. In *Proc. NeurIPS'19* 14997–15007. Vancouver: PMLR.
- [47] Torgersen, E. (1991). *Comparison of Statistical Experiments*. *Encyclopedia of Mathematics and Its Applications* **36**. Cambridge: Cambridge Univ. Press. [MR1104437](#) <https://doi.org/10.1017/CBO9780511666353>
- [48] 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>
- [49] Welling, M. and Teh, Y. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proc. ICML'11* 681–688. Bellevue: Omnipress.
- [50] Xifara, T., Sherlock, C., Livingstone, S., Byrne, S. and Girolami, M. (2014). Langevin diffusions and the Metropolis-adjusted Langevin algorithm. *Statist. Probab. Lett.* **91** 14–19. [MR3208109](#) <https://doi.org/10.1016/j.spl.2014.04.002>

Rates and coverage for monotone densities using projection-posterior

MOUMITA CHAKRABORTY* and SUBHASHIS GHOSAL†

Department of Statistics, North Carolina State University, 2311 Stinson Drive, Raleigh, North Carolina 27695-8203, USA.

E-mail: *mchakra@ncsu.edu; †sghosal@ncsu.edu

We consider Bayesian inference for a monotone density on the unit interval and study the resulting asymptotic properties. We consider a “projection-posterior” approach, where we construct a prior on density functions through random histograms without imposing the monotonicity constraint, but induce a random distribution by projecting a sample from the posterior on the space of monotone functions. The approach allows us to retain posterior conjugacy, allowing explicit expressions extremely useful for studying asymptotic properties. We show that the projection-posterior contracts at the optimal $n^{-1/3}$ -rate. We then construct a consistent test based on the posterior distribution for testing the hypothesis of monotonicity. Finally, we obtain the limiting coverage of a projection-posterior credible interval for the value of the function at an interior point. Interestingly, the limiting coverage turns out to be higher than the nominal credibility level, the opposite of the undercoverage phenomenon observed in a smoothness regime. Moreover, we show that a recalibration method using a lower credibility level gives an intended limiting coverage. We also discuss extensions of the obtained results for densities on the half-line. We conduct a simulation study to demonstrate the accuracy of the asymptotic results in finite samples.

Keywords: Monotone density; contraction rate; Bayesian test for monotonicity; credible interval; coverage

References

- [1] Banerjee, M. (2000). *Likelihood Ratio Inference in Regular and Non-regular Problems*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)—University of Washington. [MR2701606](#)
- [2] Banerjee, M. (2007). Likelihood based inference for monotone response models. *Ann. Statist.* **35** 931–956. [MR2341693](#) <https://doi.org/10.1214/009053606000001578>
- [3] Banerjee, M. and Wellner, J.A. (2001). Likelihood ratio tests for monotone functions. *Ann. Statist.* **29** 1699–1731. [MR1891743](#) <https://doi.org/10.1214/aos/1015345959>
- [4] Barlow, R.E., Bartholomew, D.J., Bremner, J.M. and Brunk, H.D. (1972). *Statistical Inference Under Order Restrictions. The Theory and Application of Isotonic Regression*. Wiley Series in Probability and Mathematical Statistics. London–Sydney: Wiley. [MR0326887](#)
- [5] Barlow, R.E. and Brunk, H.D. (1972). The isotonic regression problem and its dual. *J. Amer. Statist. Assoc.* **67** 140–147. [MR0314205](#)
- [6] Bhaumik, P. and Ghosal, S. (2015). Bayesian two-step estimation in differential equation models. *Electron. J. Stat.* **9** 3124–3154. [MR3453972](#) <https://doi.org/10.1214/15-EJS1099>
- [7] Bhaumik, P. and Ghosal, S. (2017). Efficient Bayesian estimation and uncertainty quantification in ordinary differential equation models. *Bernoulli* **23** 3537–3570. [MR3654815](#) <https://doi.org/10.3150/16-BEJ856>
- [8] Bhaumik, P. and Ghosal, S. (2017). Bayesian inference for higher-order ordinary differential equation models. *J. Multivariate Anal.* **157** 103–114. [MR3641739](#) <https://doi.org/10.1016/j.jmva.2017.03.003>
- [9] Bhaumik, P., Shi, W. and Ghosal, S. (2021). Bayesian generalized regression in partial differential equation models. *Bernoulli*. To appear.
- [10] Bowman, A.W., Jones, M.C. and Gijbels, I. (1998). Testing monotonicity of regression. *J. Comput. Graph. Statist.* **7** 489–500.

- [11] Brunk, H.D. (1970). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)* 177–197. London: Cambridge Univ. Press. [MR0277070](#)
- [12] Brunner, L.J. (1992). Bayesian nonparametric methods for data from a unimodal density. *Statist. Probab. Lett.* **14** 195–199. [MR1173617](#) [https://doi.org/10.1016/0167-7152\(92\)90021-V](https://doi.org/10.1016/0167-7152(92)90021-V)
- [13] Brunner, L.J. and Lo, A.Y. (1989). Bayes methods for a symmetric unimodal density and its mode. *Ann. Statist.* **17** 1550–1566. [MR1026299](#) <https://doi.org/10.1214/aos/1176347381>
- [14] 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>
- [15] Chakraborty, M. and Ghosal, S. (2021). Convergence rates for Bayesian estimation and testing in monotone regression. *Electron. J. Stat.* **15** 3478–3503. [MR4280172](#) <https://doi.org/10.1214/21-ejs1861>
- [16] Chakraborty, M. and Ghosal, S. (2021). Coverage of credible intervals in nonparametric monotone regression. *Ann. Statist.* **49** 1011–1028. [MR4255117](#) <https://doi.org/10.1214/20-aos1989>
- [17] Chernoff, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.* **16** 31–41. [MR0172382](#) <https://doi.org/10.1007/BF02868560>
- [18] Cox, D.D. (1993). An analysis of Bayesian inference for nonparametric regression. *Ann. Statist.* **21** 903–923. [MR1232525](#) <https://doi.org/10.1214/aos/1176349157>
- [19] de Boor, C. (2001). *A Practical Guide to Splines*, revised ed. *Applied Mathematical Sciences* **27**. New York: Springer. [MR1900298](#)
- [20] De Leeuw, J., Kurt, H. and Mair, P. (2009). Isotone optimization in R: Pool-adjacent-violators algorithm (PAVA) and active set methods. *J. Stat. Softw.* **32**. <https://doi.org/10.18637/jss.v032.i05>
- [21] Dümbgen, L. (2003). Optimal confidence bands for shape-restricted curves. *Bernoulli* **9** 423–449. [MR1997491](#) <https://doi.org/10.3150/bj/1065444812>
- [22] Dümbgen, L. and Johns, R.B. (2004). Confidence bands for isotonic median curves using sign tests. *J. Comput. Graph. Statist.* **13** 519–533. [MR2063998](#) <https://doi.org/10.1198/1061860043506>
- [23] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. [MR1790007](#) <https://doi.org/10.1214/aos/1016218228>
- [24] Ghosal, S., Sen, A. and van der Vaart, A.W. (2000). Testing monotonicity of regression. *Ann. Statist.* **28** 1054–1082. [MR1810919](#) <https://doi.org/10.1214/aos/1015956707>
- [25] Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference. Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. [MR3587782](#) <https://doi.org/10.1017/9781139029834>
- [26] Gijbels, I., Hall, P., Jones, M.C. and Koch, I. (2000). Tests for monotonicity of a regression mean with guaranteed level. *Biometrika* **87** 663–673. [MR1789816](#) <https://doi.org/10.1093/biomet/87.3.663>
- [27] Grenander, U. (1957). On the theory of mortality measurement. II. *Skand. Aktuarietidskr.* **39** 125–153. [MR0093415](#) <https://doi.org/10.1080/03461238.1956.10414944>
- [28] Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields* **81** 79–109. [MR0981568](#) <https://doi.org/10.1007/BF00343738>
- [29] Groeneboom, P. and Jongbloed, G. (2014). *Nonparametric Estimation Under Shape Constraints: Estimators, Algorithms and Asymptotics. Cambridge Series in Statistical and Probabilistic Mathematics* **38**. New York: Cambridge Univ. Press. [MR3445293](#) <https://doi.org/10.1017/CBO9781139020893>
- [30] Groeneboom, P. and Wellner, J.A. (2001). Computing Chernoff’s distribution. *J. Comput. Graph. Statist.* **10** 388–400. [MR1939706](#) <https://doi.org/10.1198/10618600152627997>
- [31] Hall, P. and Heckman, N.E. (2000). Testing for monotonicity of a regression mean by calibrating for linear functions. *Ann. Statist.* **28** 20–39. [MR1762902](#) <https://doi.org/10.1214/aos/1016120363>
- [32] Huang, J. and Wellner, J.A. (1995). Estimation of a monotone density or monotone hazard under random censoring. *Scand. J. Stat.* **22** 3–33. [MR1334065](#)
- [33] Huang, Y. and Zhang, C.-H. (1994). Estimating a monotone density from censored observations. *Ann. Statist.* **22** 1256–1274. [MR1311975](#) <https://doi.org/10.1214/aos/1176325628>
- [34] Lin, L. and Dunson, D.B. (2014). Bayesian monotone regression using Gaussian process projection. *Biometrika* **101** 303–317. [MR3215349](#) <https://doi.org/10.1093/biomet/ast063>
- [35] Prakasa Rao, B.L.S. (1969). Estimation of a unimodal density. *Sankhyā Ser. A* **31** 23–36. [MR0267677](#)

- [36] Salomond, J.-B. (2014). Concentration rate and consistency of the posterior distribution for selected priors under monotonicity constraints. *Electron. J. Stat.* **8** 1380–1404. [MR3263126](#) <https://doi.org/10.1214/14-EJS929>
- [37] Salomond, J.-B. (2014). Adaptive Bayes test for monotonicity. In *The Contribution of Young Researchers to Bayesian Statistics. Springer Proc. Math. Stat.* **63** 29–33. Cham: Springer. [MR3133254](#) https://doi.org/10.1007/978-3-319-02084-6_7
- [38] Salomond, J.-B. (2018). Testing un-separated hypotheses by estimating a distance. *Bayesian Anal.* **13** 461–484. [MR3780431](#) <https://doi.org/10.1214/17-BA1059>
- [39] Schmidt-Hieber, J., Munk, A. and Dümbgen, L. (2013). Multiscale methods for shape constraints in deconvolution: Confidence statements for qualitative features. *Ann. Statist.* **41** 1299–1328. [MR3113812](#) <https://doi.org/10.1214/13-AOS1089>
- [40] Schwartz, L. (1965). On Bayes procedures. *Z. Wahrsch. Verw. Gebiete* **4** 10–26. [MR0184378](#) <https://doi.org/10.1007/BF00535479>
- [41] Sen, B., Banerjee, M. and Woodroffe, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *Ann. Statist.* **38** 1953–1977. [MR2676880](#) <https://doi.org/10.1214/09-AOS777>
- [42] Shen, W. and Ghosal, S. (2015). Adaptive Bayesian procedures using random series priors. *Scand. J. Stat.* **42** 1194–1213. [MR3426318](#) <https://doi.org/10.1111/sjos.12159>
- [43] van der Vaart, A.W. (2000). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. [MR1652247](#) <https://doi.org/10.1017/CBO9780511802256>
- [44] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Process with Applications to Statistics*. New York: Springer-Verlag.
- [45] Wu, Y. and Ghosal, S. (2008). Kullback–Leibler property of kernel mixture priors in Bayesian density estimation. *Electron. J. Stat.* **2** 298–331. [MR2399197](#) <https://doi.org/10.1214/07-EJS130>

Minimax estimation of norms of a probability density: I. Lower bounds

ALEXANDER GOLDENSHLUGER¹ and OLEG V. LEPSKI²

¹*Department of Statistics, University of Haifa, Haifa, Israel.*

E-mail: goldensh@stat.haifa.ac.il

²*Institut de Mathématique de Marseille, Aix-Marseille Université, Marseille, France.*

E-mail: oleg.lepski@univ-amu.fr

The paper deals with the problem of nonparametric estimating the $\mathbb{L}_p$ -norm, $p \in (1, \infty)$, of a probability density on $\mathbb{R}^d$, $d \geq 1$ from independent observations. The unknown density is assumed to belong to a ball in the anisotropic Nikolskii's space. We adopt the minimax approach, and derive lower bounds on the minimax risk. In particular, we demonstrate that accuracy of estimation procedures essentially depends on whether p is integer or not. Moreover, we develop a general technique for derivation of lower bounds on the minimax risk in the problems of estimating nonlinear functionals. The proposed technique is applicable for a broad class of nonlinear functionals, and it is used for derivation of the lower bounds in the $\mathbb{L}_p$ -norm estimation.

Keywords: Estimation of nonlinear functionals; minimax estimation; minimax risk; anisotropic Nikolskii's class; best approximation

References

- [1] Bickel, P.J. and Ritov, Y. (1988). Estimating integrated squared density derivatives: Sharp best order of convergence estimates. *Sankhyā Ser. A* **50** 381–393. [MR1065550](#)
- [2] Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11–29. [MR1331653](#) <https://doi.org/10.1214/aos/1176324452>
- [3] 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>
- [4] Cai, T.T. and Low, M.G. (2004). Minimax estimation of linear functionals over nonconvex parameter spaces. *Ann. Statist.* **32** 552–576. [MR2060169](#) <https://doi.org/10.1214/009053604000000094>
- [5] Cai, T.T. and Low, M.G. (2005). Nonquadratic estimators of a quadratic functional. *Ann. Statist.* **33** 2930–2956. [MR2253108](#) <https://doi.org/10.1214/009053605000000147>
- [6] Cai, T.T. and Low, M.G. (2011). Testing composite hypotheses, Hermite polynomials and optimal estimation of a nonsmooth functional. *Ann. Statist.* **39** 1012–1041. [MR2816346](#) <https://doi.org/10.1214/10-AOS849>
- [7] Donoho, D.L. and Liu, R.C. (1991). Geometrizing rates of convergence. II, III. *Ann. Statist.* **19** 633–667, 668–701. [MR1105839](#) <https://doi.org/10.1214/aos/1176348114>
- [8] Donoho, D.L. and Nussbaum, M. (1990). Minimax quadratic estimation of a quadratic functional. *J. Complexity* **6** 290–323. [MR1081043](#) [https://doi.org/10.1016/0885-064X\(90\)90025-9](https://doi.org/10.1016/0885-064X(90)90025-9)
- [9] Giné, E. and Nickl, R. (2008). A simple adaptive estimator of the integrated square of a density. *Bernoulli* **14** 47–61. [MR2401653](#) <https://doi.org/10.3150/07-BEJ110>
- [10] Goldenshluger, A. and Lepski, O. (2014). On adaptive minimax density estimation on $\mathbb{R}^d$. *Probab. Theory Related Fields* **159** 479–543. [MR3230001](#) <https://doi.org/10.1007/s00440-013-0512-1>
- [11] Goldenshluger, A. and Lepski, O. (2020). Minimax estimation of norms of a probability density: II. Rate-optimal estimation procedures. Manuscript.
- [12] Goldenshluger, A. and Lepski, O.V. (2022). Supplement to “Minimax estimation of norms of a probability density: I. Lower bounds.” <https://doi.org/10.3150/21-BEJ1380SUPP>

- [13] Han, Y., Jiao, J. and Mukherjee, R. (2020). On estimation of L_r -norms in Gaussian white noise models. *Probab. Theory Related Fields* **177** 1243–1294. [MR4126939](#) <https://doi.org/10.1007/s00440-020-00982-x>
- [14] Han, Y., Jiao, J., Weissman, T. and Wu, Y. (2020). Optimal rates of entropy estimation over Lipschitz balls. *Ann. Statist.* **48** 3228–3250. [MR4185807](#) <https://doi.org/10.1214/19-AOS1927>
- [15] Ibragimov, I.A. and Khas'minskiĭ, R.Z. (1986). An estimate for the value of a linear functional of the density of a distribution. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* **153** 45–59, 174, 178. [MR0859537](#) <https://doi.org/10.1007/BF01095383>
- [16] Ibragimov, I.A., Nemirovskii, A.S. and Khas'minskiĭ, R.Z. (1986). Some problems of nonparametric estimation in Gaussian white noise. *Teor. Veroyatnost. i Primenen.* **31** 451–466. [MR0866866](#)
- [17] Juditsky, A. and Nemirovski, A. (2020). *Statistical Inference via Convex Optimization. Princeton Series in Applied Mathematics*. Princeton, NJ: Princeton Univ. Press. [MR4264425](#)
- [18] Kerkycharian, G. and Picard, D. (1996). Estimating nonquadratic functionals of a density using Haar wavelets. *Ann. Statist.* **24** 485–507. [MR1394973](#) <https://doi.org/10.1214/aos/1032894450>
- [19] Kozachenko, L.F. and Leonenko, N.N. (1987). Sample estimate of the entropy of a random vector. *Probl. Inform. Transm.* **23** 95–101.
- [20] Laurent, B. (1996). Efficient estimation of integral functionals of a density. *Ann. Statist.* **24** 659–681. [MR1394981](#) <https://doi.org/10.1214/aos/1032894458>
- [21] Laurent, B. (1997). Estimation of integral functionals of a density and its derivatives. *Bernoulli* **3** 181–211. [MR1466306](#) <https://doi.org/10.2307/3318586>
- [22] Leonenko, N., Pronzato, L. and Savani, V. (2008). A class of Rényi information estimators for multidimensional densities. *Ann. Statist.* **36** 2153–2182. [MR2458183](#) <https://doi.org/10.1214/07-AOS539>
- [23] Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the L_r norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. [MR1670867](#) <https://doi.org/10.1007/s004409970006>
- [24] Natanson, I.P. (1964). *Constructive Function Theory. Vol. I. Uniform Approximation*. New York: Frederick Ungar Publishing Co. The Russian original appeared in 1949 in one volume *Constructive theory of functions* (Russian), GITTL, Moscow, 1949. [MR0196340](#)
- [25] Natanson, I.P. (1965). *Constructive Function Theory. Vol. II. Approximation in Mean*. New York: Frederick Ungar Publishing Co. The Russian original appeared in 1949 in one volume *Constructive theory of functions* (Russian), GITTL, Moscow, 1949.
- [26] Nemirovski, A. (2000). Topics in non-parametric statistics. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1998). Lecture Notes in Math.* **1738** 85–277. Berlin: Springer. [MR1775640](#)
- [27] Nemirovskii, A.S. (1990). Necessary conditions for efficient estimation of functionals of a nonparametric signal observed in white noise. *Teor. Veroyatnost. i Primenen.* **35** 83–91. [MR1050056](#) <https://doi.org/10.1137/1135009>
- [28] Nikol'skii, S.M. (1977). *Priblizhenie Funktsii Mnogikh Peremennykh i Teoremy Vlozheniya*, 2nd ed. Moscow: Nauka.
- [29] Rényi, A. (1961). On measures of entropy and information. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 547–561. Berkeley, CA: Univ. California Press. [MR0132570](#)
- [30] Suetin, P.K. (1979). *Classical Orthogonal Polynomials*. 2nd rev. ed. Moscow: “Nauka”.
- [31] Tchetgen, E., Li, L., Robins, J. and van der Vaart, A. (2008). Minimax estimation of the integral of a power of a density. *Statist. Probab. Lett.* **78** 3307–3311. [MR2479495](#) <https://doi.org/10.1016/j.spl.2008.07.001>
- [32] Timan, A.F. (1963). *Theory of Approximation of Functions of a Real Variable. A Pergamon Press Book*. New York: The Macmillan Co. [MR0192238](#)
- [33] Tsallis, C. (1988). Possible generalization of Boltzmann–Gibbs statistics. *J. Stat. Phys.* **52** 479–487. [MR0968597](#) <https://doi.org/10.1007/BF01016429>
- [34] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. New York: Springer. [MR2724359](#) <https://doi.org/10.1007/b13794>
- [35] Wu, Y. and Yang, P. (2016). Minimax rates of entropy estimation on large alphabets via best polynomial approximation. *IEEE Trans. Inform. Theory* **62** 3702–3720. [MR3506758](#) <https://doi.org/10.1109/TIT.2016.2548468>

Minimax estimation of norms of a probability density: II. Rate-optimal estimation procedures

ALEXANDER GOLDENSHLUGER¹ and OLEG V. LEPSKI²

¹*Department of Statistics, University of Haifa, Haifa, Israel.*

E-mail: goldensh@stat.haifa.ac.il

²*Institut de Mathématique de Marseille, Aix-Marseille Université, Marseille, France.*

E-mail: oleg.lepski@univ-amu.fr

In this paper we develop rate-optimal estimation procedures in the problem of estimating the $\mathbb{L}_p$ -norm, $p \in (1, \infty)$ of a probability density from independent observations. The density is assumed to be defined on $\mathbb{R}^d$, $d \geq 1$ and to belong to a ball in the anisotropic Nikolskii space. We adopt the minimax approach and construct rate-optimal estimators in the case of integer $p \geq 2$. We demonstrate that, depending on the parameters of the Nikolskii class and the norm index p , the minimax rates of convergence may vary from inconsistency to the parametric $\sqrt{n}$ -estimation. The results in this paper complement the minimax lower bounds derived in the companion paper (Goldenshluger and Lepski (2020)).

Keywords: Density estimation; minimax risk; $\mathbb{L}_p$ -norm; U -statistics; anisotropic Nikol'skii class

References

- [1] Bickel, P.J. and Ritov, Y. (1988). Estimating integrated squared density derivatives: Sharp best order of convergence estimates. *Sankhyā Ser. A* **50** 381–393. [MR1065550](#)
- [2] Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11–29. [MR1331653](#) <https://doi.org/10.1214/aos/1176324452>
- [3] Cai, T.T. and Low, M.G. (2005). Nonquadratic estimators of a quadratic functional. *Ann. Statist.* **33** 2930–2956. [MR2253108](#) <https://doi.org/10.1214/009053605000000147>
- [4] de Guzmán, M. (1975). *Differentiation of Integrals in $\mathbb{R}^n$* . *Lecture Notes in Mathematics* **481**. Berlin: Springer. [MR0457661](#)
- [5] Folland, G.B. (1999). *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. *Pure and Applied Mathematics (New York)*. New York: Wiley. [MR1681462](#)
- [6] Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39** 1608–1632. [MR2850214](#) <https://doi.org/10.1214/11-AOS883>
- [7] Goldenshluger, A. and Lepski, O. (2014). On adaptive minimax density estimation on $\mathbb{R}^d$. *Probab. Theory Related Fields* **159** 479–543. [MR3230001](#) <https://doi.org/10.1007/s00440-013-0512-1>
- [8] Goldenshluger, A. and Lepski, O.V. (2020). Minimax estimation of norms of a probability density: I. Lower bounds. Manuscript.
- [9] Han, Y., Jiao, J. and Mukherjee, R. (2020). On estimation of L_r -norms in Gaussian white noise models. *Probab. Theory Related Fields* **177** 1243–1294. [MR4126939](#) <https://doi.org/10.1007/s00440-020-00982-x>
- [10] Houdré, C. and Reynaud-Bouret, P. (2003). Exponential inequalities, with constants, for U -statistics of order two. In *Stochastic Inequalities and Applications. Progress in Probability* **56** 55–69. Basel: Birkhäuser. [MR2073426](#)
- [11] Kerkycharian, G. and Picard, D. (1996). Estimating nonquadratic functionals of a density using Haar wavelets. *Ann. Statist.* **24** 485–507. [MR1394973](#) <https://doi.org/10.1214/aos/1032894450>
- [12] Laurent, B. (1997). Estimation of integral functionals of a density and its derivatives. *Bernoulli* **3** 181–211. [MR1466306](#) <https://doi.org/10.2307/3318586>

- [13] Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the L_r norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. [MR1670867](#) <https://doi.org/10.1007/s004409970006>
- [14] Lepski, O.V. and Spokoiny, V.G. (1999). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. *Bernoulli* **5** 333–358. [MR1681702](#) <https://doi.org/10.2307/3318439>
- [15] Lepski, O.V. and Willer, T. (2019). Oracle inequalities and adaptive estimation in the convolution structure density model. *Ann. Statist.* **47** 233–287. [MR3909933](#) <https://doi.org/10.1214/18-AOS1687>
- [16] Lepskiĭ, O.V. (1991). Asymptotically minimax adaptive estimation. I. Upper bounds. Optimally adaptive estimates. *Theory Probab. Appl.* **36** 682–697.
- [17] Lepskiĭ, O.V. (1992). Asymptotically minimax adaptive estimation. II. Schemes without optimal adaptation. Adaptive estimates. *Teor. Veroyatn. Primen.* **37** 468–481. [MR1214353](#) <https://doi.org/10.1137/1137095>
- [18] Major, P. (2013). *On the Estimation of Multiple Random Integrals and U-Statistics. Lecture Notes in Math.* **2079**. Heidelberg: Springer. [MR3087566](#) <https://doi.org/10.1007/978-3-642-37617-7>
- [19] Nikol'skii, S.M. (1977). *Priblizhenie Funktsii Mnogikh Peremennykh i Teoremy Vlozheniya* (in Russian). [Approximation of functions of several variables and embedding theorems.] 2nd ed., revised and supplemented. Nauka, Moscow.
- [20] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics. Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. [MR0595165](#)
- [21] Tchetgen, E., Li, L., Robins, J. and van der Vaart, A. (2008). Minimax estimation of the integral of a power of a density. *Statist. Probab. Lett.* **78** 3307–3311. [MR2479495](#) <https://doi.org/10.1016/j.spl.2008.07.001>

Local minimax rates for closeness testing of discrete distributions

JOSEPH LAM-WEIL^{1,*}, ALEXANDRA CARPENTIER^{1,†} and
BHARATH K. SRIPERUMBUDUR²

¹Magdeburg University, Universitätsplatz 2, 39106 Magdeburg, Germany.

E-mail: ^{*}joseph.lam@ovgu.de; [†]alexandra.carpentier@ovgu.de

²Pennsylvania State University, 201 Old Main, Pennsylvania 16802, USA.

E-mail: bks18@psu.edu

We consider the closeness testing problem for discrete distributions. The goal is to distinguish whether two samples are drawn from the same unspecified distribution, or whether their respective distributions are separated in L_1 -norm. In this paper, we focus on adapting the rate to the shape of the underlying distributions, i.e. we consider a *local minimax setting*. We provide, to the best of our knowledge, the first local minimax rate for the separation distance up to logarithmic factors, together with a test that achieves it. In view of the rate, closeness testing turns out to be substantially harder than the related one-sample testing problem over a wide range of cases.

Keywords: Local minimax optimality; closeness testing; two-sample; instance optimal; discrete distributions; hypothesis testing; composite-composite testing

References

- [1] Acharya, J., Das, H., Jafarpour, A., Orlitsky, A., Pan, S. and Suresh, A. (2012). Competitive classification and closeness testing. In *Conference on Learning Theory* 22–1.
- [2] Balakrishnan, S. and Wasserman, L. (2018). Hypothesis testing for high-dimensional multinomials: A selective review. *Ann. Appl. Stat.* **12** 727–749. [MR3834283](#) <https://doi.org/10.1214/18-AOAS1155SF>
- [3] Balakrishnan, S. and Wasserman, L. (2019). Hypothesis testing for densities and high-dimensional multinomials: Sharp local minimax rates. *Ann. Statist.* **47** 1893–1927. [MR3953439](#) <https://doi.org/10.1214/18-AOS1729>
- [4] Batu, T., Fortnow, L., Rubinfeld, R., Smith, W.D. and White, P. (2000). Testing that distributions are close. In *41st Annual Symposium on Foundations of Computer Science (Redondo Beach, CA, 2000)* 259–269. Los Alamitos, CA: IEEE Comput. Soc. Press. [MR1931824](#) <https://doi.org/10.1109/SFCS.2000.892113>
- [5] Bhattacharya, B. and Valiant, G. (2015). Testing closeness with unequal sized samples. In *Advances in Neural Information Processing Systems* 2611–2619.
- [6] Chan, S.-O., Diakonikolas, I., Valiant, G. and Valiant, P. (2014). Optimal algorithms for testing closeness of discrete distributions. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms* 1193–1203. New York: ACM. [MR3376448](#) <https://doi.org/10.1137/1.9781611973402.88>
- [7] Diakonikolas, I., Gouleakis, T., Peebles, J. and Price, E. (2018). Sample-optimal identity testing with high probability. In *45th International Colloquium on Automata, Languages, and Programming. LIPIcs. Leibniz Int. Proc. Inform.* **107** Art. No. 41, 14. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. [MR3829972](#)
- [8] Diakonikolas, I. and Kane, D.M. (2016). A new approach for testing properties of discrete distributions. In *57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016* 685–694. Los Alamitos, CA: IEEE Computer Soc. [MR3631031](#)
- [9] Diakonikolas, I., Kane, D.M. and Nikishkin, V. (2015). Optimal algorithms and lower bounds for testing closeness of structured distributions. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science—FOCS 2015* 1183–1202. Los Alamitos, CA: IEEE Computer Soc. [MR3473364](#) <https://doi.org/10.1109/FOCS.2015.76>

- [10] Diakonikolas, I., Kane, D.M. and Nikishkin, V. (2017). Near-optimal closeness testing of discrete histogram distributions. In *44th International Colloquium on Automata, Languages, and Programming. LIPIcs. Leibniz Int. Proc. Inform.* **80** Art. No. 8, 15. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. [MR3685748](#)
- [11] Goldreich, O., Goldwasser, S. and Ron, D. (1998). Property testing and its connection to learning and approximation. *J. ACM* **45** 653–750. [MR1675099](#) <https://doi.org/10.1145/285055.285060>
- [12] Ingster, Y. and Suslina, I.A. (2012). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models* **169**. Springer Science & Business Media. [MR1991446](#) <https://doi.org/10.1007/978-0-387-21580-8>
- [13] Kim, I., Balakrishnan, S. and Wasserman, L. (2020). Robust multivariate nonparametric tests via projection averaging. *Ann. Statist.* **48** 3417–3441. [MR4185814](#) <https://doi.org/10.1214/19-AOS1936>
- [14] Lam-Weil, J., Carpentier, A. and Sriperumbudur, B.K. (2022). Supplement to “Local minimax rates for closeness testing of discrete distributions.” <https://doi.org/10.3150/21-BEJ1382SUPP>
- [15] Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. [MR2135927](#)
- [16] Neyman, J. and Pearson, E.S. (1933). IX. On the problem of the most efficient tests of statistical hypotheses. *Philos. Trans. R. Soc. Lond. A* **231** 289–337.
- [17] Paninski, L. (2008). A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Trans. Inf. Theory* **54** 4750–4755. [MR2591136](#) <https://doi.org/10.1109/TIT.2008.928987>
- [18] Rubinfeld, R. and Sudan, M. (1996). Robust characterizations of polynomials with applications to program testing. *SIAM J. Comput.* **25** 252–271. [MR1379300](#) <https://doi.org/10.1137/S0097539793255151>
- [19] Valiant, G. and Valiant, P. (2017). An automatic inequality prover and instance optimal identity testing. *SIAM J. Comput.* **46** 429–455. [MR3614697](#) <https://doi.org/10.1137/151002526>
- [20] Valiant, P. (2011). Testing symmetric properties of distributions. *SIAM J. Comput.* **40** 1927–1968. [MR2863200](#) <https://doi.org/10.1137/080734066>

Paving property for real stable polynomials and strongly Rayleigh processes

KASRA ALISHAHI^{*} and MILAD BARZEGAR[†]

Department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran.

E-mail: ^{*}alishahi@sharif.edu; [†]milad.barzegar@sharif.edu

One of the equivalent formulations of the Kadison–Singer problem which was resolved in 2013 by Marcus, Spielman and Srivastava, is the “paving conjecture”. In this paper, we first extend this result to real stable polynomials. We prove that for every multi-affine real stable polynomial satisfying a simple condition, it is possible to partition its set of variables to a small number of subsets such that the “restriction” of the polynomial to each subset has small roots. Then, we derive a probabilistic interpretation of this result. We show that there exists a partition of the underlying space of every strongly Rayleigh process into a small number of sets such that the restriction of the point process to each set has “almost independent” points. This result implies that the dependence structure of strongly Rayleigh processes is constrained—a phenomenon that is to be expected in negatively dependent measures. To prove this result, we introduce and study the notion of kernel polynomial for strongly Rayleigh processes. This notion is a natural generalization of the kernel of determinantal processes. We also derive an entropy bound for strongly Rayleigh processes in terms of the roots of the kernel polynomial which is interesting on its own.

Keywords: Paving property; real stable polynomials; strongly Rayleigh processes

References

- [1] Alishahi, K. and Barzegar, M. (2022). Supplement to “Paving property for real stable polynomials and strongly Rayleigh processes.” <https://doi.org/10.3150/21-BEJ1383SUP>
- [2] Anari, N. and Gharan, S.O. (2014). The Kadison–Singer Problem for strongly rayleigh measures and applications to asymmetric TSP. ArXiv E-prints. Available at <https://arxiv.org/abs/1412.1143>.
- [3] Anari, N., Gharan, S.O. and Rezaei, A. (2016). Monte Carlo Markov chain algorithms for sampling strongly Rayleigh distributions and determinantal point processes. In *Conference on Learning Theory* 103–115.
- [4] Anari, N. and Oveis Gharan, S. (2017). A generalization of permanent inequalities and applications in counting and optimization. In *STOC’17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing* 384–396. New York: ACM. [MR3678196 https://doi.org/10.1145/3055399.3055469](https://doi.org/10.1145/3055399.3055469)
- [5] Anari, N., Oveis Gharan, S. and Vintzilez, C. (2018). Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. In *59th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2018* 35–46. Los Alamitos, CA: IEEE Computer Soc. [MR3899575 https://doi.org/10.1109/FOCS.2018.00013](https://doi.org/10.1109/FOCS.2018.00013)
- [6] Anderson, J. (1979). Extensions, restrictions, and representations of states on C^* -algebras. *Trans. Amer. Math. Soc.* **249** 303–329. [MR0525675 https://doi.org/10.2307/1998793](https://doi.org/10.2307/1998793)
- [7] Borcea, J. and Brändén, P. (2008). Applications of stable polynomials to mixed determinants: Johnson’s conjectures, unimodality, and symmetrized Fischer products. *Duke Math. J.* **143** 205–223. [MR2420507 https://doi.org/10.1215/00127094-2008-018](https://doi.org/10.1215/00127094-2008-018)
- [8] Borcea, J. and Brändén, P. (2010). Multivariate Pólya–Schur classification problems in the Weyl algebra. *Proc. Lond. Math. Soc.* (3) **101** 73–104. [MR2661242 https://doi.org/10.1112/plms/pdp049](https://doi.org/10.1112/plms/pdp049)
- [9] Borcea, J., Brändén, P. and Liggett, T.M. (2009). Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* **22** 521–567. [MR2476782 https://doi.org/10.1090/S0894-0347-08-00618-8](https://doi.org/10.1090/S0894-0347-08-00618-8)
- [10] Bownik, M. (2018). The Kadison–Singer problem. In *Frames and Harmonic Analysis. Contemp. Math.* **706** 63–92. Amer. Math. Soc., Providence, RI. [MR3796632 https://doi.org/10.1090/conm/706/14218](https://doi.org/10.1090/conm/706/14218)

- [11] Brändén, P. (2007). Polynomials with the half-plane property and matroid theory. *Adv. Math.* **216** 302–320. [MR2353258 https://doi.org/10.1016/j.aim.2007.05.011](https://doi.org/10.1016/j.aim.2007.05.011)
- [12] Brändén, P. (2018). Hyperbolic polynomials and the Kadison–Singer problem. ArXiv E-prints. Available at <https://arxiv.org/abs/1809.03255>.
- [13] Brändén, P. and Jonasson, J. (2012). Negative dependence in sampling. *Scand. J. Stat.* **39** 830–838. [MR3000852 https://doi.org/10.1111/j.1467-9469.2011.00766.x](https://doi.org/10.1111/j.1467-9469.2011.00766.x)
- [14] Choe, Y.-B., Oxley, J.G., Sokal, A.D. and Wagner, D.G. (2004). Homogeneous multivariate polynomials with the half-plane property. *Adv. Pure Appl. Math.* **32** 88–187.
- [15] Ghosh, S., Liggett, T.M. and Pemantle, R. (2017). Multivariate CLT follows from strong Rayleigh property. In *2017 Proceedings of the Fourteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)* 139–147. Philadelphia, PA: SIAM. [MR3630942 https://doi.org/10.1137/1.9781611974775.14](https://doi.org/10.1137/1.9781611974775.14)
- [16] Gurvits, L. (2004). Combinatorics hidden in hyperbolic polynomials and related topics. ArXiv E-prints. Available at <https://arxiv.org/abs/math/0402088>.
- [17] Gurvits, L. (2008). Van der Waerden/Schrijver–Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: One theorem for all. *Electron. J. Combin.* **15** Research Paper 66. [MR2411443](https://doi.org/10.1137/1.9781611974775.14)
- [18] Hough, J.B., Krishnapur, M., Peres, Y. and Virág, B. (2009). *Zeros of Gaussian Analytic Functions and Determinantal Point Processes. University Lecture Series* **51**. Providence, RI: Amer. Math. Soc. [MR2552864 https://doi.org/10.1090/ulect/051](https://doi.org/10.1090/ulect/051)
- [19] Kadison, R.V. and Singer, I.M. (1959). Extensions of pure states. *Amer. J. Math.* **81** 383–400. [MR0123922 https://doi.org/10.2307/2372748](https://doi.org/10.2307/2372748)
- [20] Marcus, A.W., Spielman, D.A. and Srivastava, N. (2015). Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. *Ann. of Math. (2)* **182** 327–350. [MR3374963 https://doi.org/10.4007/annals.2015.182.1.8](https://doi.org/10.4007/annals.2015.182.1.8)
- [21] Marcus, A.W., Spielman, D.A. and Srivastava, N. (2015). Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)* **182** 307–325. [MR3374962 https://doi.org/10.4007/annals.2015.182.1.7](https://doi.org/10.4007/annals.2015.182.1.7)
- [22] Marshall, A.W., Olkin, I. and Arnold, B.C. (2011). *Inequalities: Theory of Majorization and Its Applications*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR2759813 https://doi.org/10.1007/978-0-387-68276-1](https://doi.org/10.1007/978-0-387-68276-1)
- [23] Michelen, M. and Sahasrabudhe, J. (2019). Central limit theorems and the geometry of polynomials. ArXiv Preprint. Available at <https://arxiv.org/abs/1908.09020>.
- [24] Oveis Gharan, S., Saberi, A. and Singh, M. (2011). A randomized rounding approach to the traveling salesman problem. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science—FOCS 2011* 550–559. Los Alamitos, CA: IEEE Computer Soc. [MR2932731 https://doi.org/10.1109/FOCS.2011.80](https://doi.org/10.1109/FOCS.2011.80)
- [25] Pemantle, R. (2000). Towards a theory of negative dependence. *J. Math. Phys.* **41** 1371–1390. [MR1757964 https://doi.org/10.1063/1.533200](https://doi.org/10.1063/1.533200)
- [26] Pemantle, R. (2012). Hyperbolicity and stable polynomials in combinatorics and probability. In *Current Developments in Mathematics*, 2011 57–123. Somerville, MA: Int. Press. [MR3098077](https://doi.org/10.1090/S0963548313000345)
- [27] Pemantle, R. and Peres, Y. (2014). Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combin. Probab. Comput.* **23** 140–160. [MR3197973 https://doi.org/10.1017/S0963548313000345](https://doi.org/10.1017/S0963548313000345)
- [28] Rahman, Q.I. and Schmeisser, G. (2002). *Analytic Theory of Polynomials. London Mathematical Society Monographs. New Series* **26**. Oxford: The Clarendon Press. [MR1954841](https://doi.org/10.1017/S0025571802000000)
- [29] Ravichandran, M. and Leake, J. (2020). Mixed determinants and the Kadison–Singer problem. *Math. Ann.* **377** 511–541. [MR4099640 https://doi.org/10.1007/s00208-020-01986-7](https://doi.org/10.1007/s00208-020-01986-7)
- [30] Straszak, D. and Vishnoi, N.K. (2017). Real stable polynomials and matroids: Optimization and counting. In *STOC’17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing* 370–383. New York: ACM. [MR3678195 https://doi.org/10.1145/3055399.3055457](https://doi.org/10.1145/3055399.3055457)
- [31] Tao, T. (2013). Real stable polynomials and the Kadison–Singer problem. Available at <https://terrytao.wordpress.com/2013/11/04/real-stable-polynomials-and-the-kadison-singer-problem/>.
- [32] Wagner, D.G. (2011). Multivariate stable polynomials: Theory and applications. *Bull. Amer. Math. Soc. (N.S.)* **48** 53–84. [MR2738906 https://doi.org/10.1090/S0273-0979-2010-01321-5](https://doi.org/10.1090/S0273-0979-2010-01321-5)

Non-asymptotic properties of spectral decomposition of large Gram-type matrices and applications

LYUOU ZHANG¹, WEN ZHOU^{2,*} and HAONAN WANG^{2,†}

¹*School of Statistics and Management, Shanghai University of Finance of Economics, 777 Guoding Road, Shanghai, 200433, P.R. China.*

E-mail: zhanglvou@mail.shufe.edu.cn

²*Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA.*

*E-mail: *riczw@stat.colostate.edu; †wanghn@stat.colostate.edu*

Gram-type matrices and their spectral decomposition are of central importance for numerous problems in statistics, applied mathematics, physics, and machine learning. In this paper, we carefully study the non-asymptotic properties of spectral decomposition of large Gram-type matrices when data are not necessarily independent. Specifically, we derive the exponential tail bounds for the deviation between eigenvectors of the right Gram matrix to their population counterparts as well as the Berry-Esseen type bound for these deviations. We also obtain the non-asymptotic tail bound of the ratio between eigenvalues of the left Gram matrix, namely the sample covariance matrix, and their population counterparts regardless of the size of the data matrix. The documented non-asymptotic properties are further demonstrated in a suite of applications, including the non-asymptotic characterization of the estimated number of latent factors in factor models and relate machine learning problems, the estimation and forecasting of high-dimensional time series, the spectral properties of large sample covariance matrix such as perturbation bounds and inference on the spectral projectors, and low-rank matrix denoising using dependent data.

Keywords: Approximate factor model; Gram-type matrices; high-dimensional time series; non-asymptotic analysis; principal component analysis; spectral decomposition

References

- [1] Adamczak, R., Litvak, A., Pajor, A. and Tomczak-Jaegermann, N. (2010). Quantitative estimates of the convergence of the empirical covariance matrix in log-concave ensembles. *J. Amer. Math. Soc.* **23** 535–561.
- [2] Ahn, S.C. and Horenstein, A.R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81** 1203–1227.
- [3] Ahn, S.C., Lee, Y.H. and Schmidt, P. (2001). GMM estimation of linear panel data models with time-varying individual effects. *J. Econometrics* **101** 219–255.
- [4] Anderson, T.W. (1962). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.
- [5] Anderson, T.W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Stat.* **34** 122–148.
- [6] Anderson, T.W. and Rubin, H. (1956). Statistical inference in factor analysis. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 5: Contributions to Econometrics, Industrial Research, and Psychometry* 111–150. Univ. California Press.
- [7] Athreya, K.B. and Lahiri, S.N. (2006). *Measure Theory and Probability Theory*. New York: Springer.
- [8] Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* **71** 135–171.
- [9] Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica* **77** 1229–1279.
- [10] Bai, J. and Li, K. (2014). Theory and methods of panel data models with interactive effects. *Ann. Statist.* **42** 142–170.
- [11] Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* **70** 191–221.

- [12] Bai, J. and Ng, S. (2008). Large dimensional factor analysis. *Found Trends Econom.* **3** 89–163.
- [13] Bai, J. and Ng, S. (2013). Principal components estimation and identification of static factors. *J. Econometrics* **176** 18–29.
- [14] 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.
- [15] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*. New York: Springer.
- [16] Bai, Z. and Yin, Y. (1993). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *Ann. Probab.* **21** 1275–1294.
- [17] Bai, Z., Yin, Y. and Krishnaiah, P.R. (1986). On limiting spectral distribution of product of two random matrices when the underlying distribution is isotropic. *J. Multivariate Anal.* **19** 189–200.
- [18] Bai, Z., Yin, Y. and Krishnaiah, P.R. (1988). On the limiting empirical distribution function of the eigenvalues of a multivariate F matrix. *Theory Probab. Appl.* **32** 490–500.
- [19] Baik, J., Arous, G.B. and Pécché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697.
- [20] Bartholomew, D.J., Knott, M. and Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*. New York: Wiley.
- [21] Bialecki, B. and Fairweather, G. (1995). Matrix decomposition algorithms in orthogonal spline collocation for separable elliptic boundary value problems. *SIAM J. Sci. Comput.* **16** 330–347.
- [22] Bickel, P.J. and Levina, E. (2008). Covariance regularization by thresholding. *Ann. Statist.* **36** 2577–2604.
- [23] Bickel, P.J. and Levina, E. (2008). Regularized estimation of large covariance matrices. *Ann. Statist.* **36** 199–227.
- [24] Bien, J., Bunea, F. and Xiao, L. (2016). Convex banding of the covariance matrix. *J. Amer. Statist. Assoc.* **111** 834–845. <https://doi.org/10.1080/01621459.2015.1058265>
- [25] Bobkov, S.G., Chistyakov, G. and Götze, F. (2018). Berry–Esseen bounds for typical weighted sums. *Electron. J. Probab.* **23**.
- [26] Bobkov, S.G. and Chistyakov, G.P. (2015). On concentration functions of random variables. *J. Theoret. Probab.* **28** 976–988.
- [27] Brockwell, P.J., Davis, R.A. and Fienberg, S.E. (1991). *Time Series: Theory and Methods: Theory and Methods*. New York: Springer.
- [28] Bühlmann, P. and Künsch, H.R. (1999). Block length selection in the bootstrap for time series. *Comput. Statist. Data Anal.* **31** 295–310.
- [29] Bunea, F. and Xiao, L. (2015). On the sample covariance matrix estimator of reduced effective rank population matrices, with applications to fPCA. *Bernoulli* **21** 1200–1230.
- [30] Cai, T., Han, X. and Pan, G. (2020). Limiting laws for divergent spiked eigenvalues and largest nonspiked eigenvalue of sample covariance matrices. *Ann. Statist.* **48** 1255–1280.
- [31] 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.
- [32] Cai, T. and Zhang, A. (2016). Minimax rate-optimal estimation of high-dimensional covariance matrices with incomplete data. *J. Multivariate Anal.* **150** 55–74.
- [33] Callaert, H. and Janssen, P. (1978). The Berry–Esseen theorem for U -statistics. *Ann. Statist.* **6** 417–421.
- [34] Chamberlain, G. and Rothschild, M. (1983). Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica* **51** 1281–1304.
- [35] Chan, Y.-K. and Wierman, J. (1977). On the Berry–Esseen theorem for U -statistics. *Ann. Probab.* **5** 136–139.
- [36] Chen, L., Wang, W. and Wu, W.B. (2021). Dynamic semiparametric factor model with structural breaks. *J. Bus. Econom. Statist.* **39** 757–771. [MR4272933 https://doi.org/10.1080/07350015.2020.1730857](https://doi.org/10.1080/07350015.2020.1730857)
- [37] Chen, L.H. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028.
- [38] Chen, L.H. and Shao, Q.-M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** 581–599.
- [39] Chen, X., Womersley, R.S. and Ye, J.J. (2011). Minimizing the condition number of a Gram matrix. *SIAM J. Optim.* **21** 127–148.

- [40] Choi, Y., Taylor, J. and Tibshirani, R. (2017). Selecting the number of principal components: Estimation of the true rank of a noisy matrix. *Ann. Statist.* **45** 2590–2617.
- [41] Davis, C. and Kahan, W.M. (1970). The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* **7** 1–46.
- [42] De Almeida, M.C., Asada, E.N. and Garcia, A.V. (2008). On the use of gram matrix in observability analysis. *IEEE Trans. Power Syst.* **23** 249–251.
- [43] De Almeida, M.C., Asada, E.N. and Garcia, A.V. (2008). Power system observability analysis based on gram matrix and minimum norm solution. *IEEE Trans. Power Syst.* **23** 1611–1618.
- [44] Donath, W.E. and Hoffman, A.J. (1973). Lower bounds for the partitioning of graphs. *IBM J. Res. Develop.* **17** 420–425.
- [45] Drineas, P. and Mahoney, M.W. (2005). On the Nystrom method for approximating a Gram matrix for improved kernel-based learning. *J. Mach. Learn. Res.* **6** 2153–2175.
- [46] Fan, J., Ke, Y., Sun, Q. and Zhou, W.-X. (2019). FarmTest: Factor-adjusted robust multiple testing with approximate false discovery control. *J. Amer. Statist. Assoc.* **114** 1880–1893.
- [47] Fan, J., Liao, Y. and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *J. Roy. Statist. Soc. Ser. B* **75** 603–680.
- [48] Fan, J., Liao, Y. and Wang, W. (2016). Projected principal component analysis in factor models. *Ann. Statist.* **44** 219–254.
- [49] Fan, J., Sun, Q., Zhou, W. and Zhu, Z. (2018). Principal component analysis for big data. Wiley StatsRef: Statistics Reference Online. to appear. <https://doi.org/10.1002/9781118445112.stat08122>
- [50] Fan, J., Wang, W. and Zhong, Y. (2018). An ℓ^∞ eigenvector perturbation bound and its application to robust covariance estimation. *J. Mach. Learn. Res.* **18** 1–42.
- [51] Florescu, L. and Perkins, W. (2016). Spectral thresholds in the bipartite stochastic block model. *Proc. Mach. Learn. Res.* **49** 943–959.
- [52] Goldstein, L. and Shao, Q.-M. (2009). Berry-Esseen bounds for projections of coordinate symmetric random vectors. *Electron. Commun. Probab.* **14** 474–485.
- [53] Hall, P., Horowitz, J.L. and Jing, B. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82** 561–574.
- [54] He, W., Zhang, H., Zhang, L. and Shen, H. (2015). Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **8** 3050–3061.
- [55] Hörmann, S. (2009). Berry-Esseen bounds for econometric time series. *ALEA Lat. Am. J. Probab. Math. Stat.* **6** 377–397.
- [56] Horst, P. (1965). *Factor Analysis of Data Matrices*. New York: Holt, Rinehart and Winston.
- [57] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **24** 417.
- [58] James, G. and Murphy, G. (1979). The determinant of the gram matrix for a Specht module. *J. Algebra* **59** 222–235.
- [59] James, W. and Stein, C. (1961). Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the Univ. California.
- [60] Ji, H., Huang, S., Shen, Z. and Xu, Y. (2011). Robust video restoration by joint sparse and low rank matrix approximation. *SIAM J. Imaging Sci.* **4** 1122–1142.
- [61] Jirak, M. (2016). Berry–Esseen theorems under weak dependence. *Ann. Probab.* **44** 2024–2063.
- [62] Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327.
- [63] Johnstone, I.M. and Lu, A.Y. (2009). On consistency and sparsity for principal components analysis in high dimensions. *J. Amer. Statist. Assoc.* **104** 682–693.
- [64] Johnstone, I.M. and Paul, D. (2018). PCA in high dimensions: An orientation. *Proc. IEEE* **106** 1277–1292.
- [65] Jolliffe, I. (2002). *Principal Component Analysis*, 2nd ed. ed. New York: Springer.
- [66] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38.
- [67] Koltchinskii, V. and Lounici, K. (2016). Asymptotics and concentration bounds for bilinear forms of spectral projectors of sample covariance. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1976–2013.

- [68] Koltchinskii, V. and Lounici, K. (2017). Normal approximation and concentration of spectral projectors of sample covariance. *Ann. Statist.* **45** 121–157.
- [69] Lam, C. and Yao, Q. (2012). Factor modeling for high-dimensional time series: Inference for the number of factors. *Ann. Statist.* **40** 694–726.
- [70] Lawley, D.N. and Maxwell, A.E. (1962). Factor analysis as a statistical method. *J. R. Stat. Soc., Ser. D* **12** 209–229.
- [71] Mandel, J. (1982). Use of the singular value decomposition in regression analysis. *Amer. Statist.* **36** 15–24.
- [72] Merlevède, F., Peligrad, M. and Rio, E. (2011). A Bernstein type inequality and moderate deviations for weakly dependent sequences. *Probab. Theory Related Fields* **151** 435–474.
- [73] Moon, H.R. and Weidner, M. (2017). Dynamic linear panel regression models with interactive fixed effects. *Econometric Theory* **33** 158–195.
- [74] Naumov, A., Spokoiny, V. and Ulyanov, V. (2019). Bootstrap confidence sets for spectral projectors of sample covariance. *Probab. Theory Related Fields* **174** 1091–1132.
- [75] Ng, A.Y., Jordan, M.I. and Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems* 849–856.
- [76] Onatski, A. (2012). Asymptotics of the principal components estimator of large factor models with weakly influential factors. *J. Econometrics* **168** 244–258.
- [77] Pal, P. and Vaidyanathan, P.P. (2014). A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising. *IEEE Signal Process. Lett.* **21** 737–741.
- [78] Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *Philos. Mag.* **2** 559–572.
- [79] Ramona, M., Richard, G. and David, B. (2012). Multiclass feature selection with kernel Gram-matrix-based criteria. *IEEE Trans. Neural Netw. Learn. Syst.* **23** 1611–1623.
- [80] Rummel, R.J. (1988). *Applied Factor Analysis*. Evanston, IL: Northwestern Univ. Press.
- [81] Schölkopf, B., Shawe-Taylor, J., Smola, A.J. and Williamson, R.C. (1999). Generalization bounds via eigenvalues of the Gram matrix Technical Report 99-035 NeuroCOLT.
- [82] Shawe-Taylor, J., Williams, C., Cristianini, N. and Kandola, J. (2002). On the eigenspectrum of the Gram matrix and its relationship to the operator eigenspectrum. In *International Conference on Algorithmic Learning Theory* 23–40. Springer.
- [83] Shawe-Taylor, J., Williams, C.K., Cristianini, N. and Kandola, J. (2005). On the eigenspectrum of the Gram matrix and the generalization error of kernel-PCA. *IEEE Trans. Inf. Theory* **51** 2510–2522.
- [84] Stark, C. (2014). Self-consistent tomography of the state-measurement Gram matrix. *Phys. Rev. A* **89** 052109.
- [85] Stein, C. (1956). Some problems in multivariate analysis, Part I Technical report, Stanford Univ.
- [86] Stock, J.H. and Watson, M.W. (2002). Forecasting using principal components from a large number of predictors. *J. Amer. Statist. Assoc.* **97** 1167–1179.
- [87] Wachter, K.W. (1978). The strong limits of random matrix spectra for sample matrices of independent elements. *Ann. Probab.* **6** 1–18.
- [88] Wang, W. and Fan, J. (2017). Asymptotics of empirical eigenstructure for high dimensional spiked covariance. *Ann. Statist.* **45** 1342–1374.
- [89] Wedin, P.Å. (1972). Perturbation bounds in connection with singular value decomposition. *BIT* **12** 99–111.
- [90] Yu, Y., Wang, T. and Samworth, R.J. (2014). A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika* **102** 315–323.
- [91] Zhang, A., Cai, T.T. and Wu, Y. (2019). Heteroskedastic PCA: Algorithm, optimality, and applications. arXiv preprint, [arXiv:1810.08316](https://arxiv.org/abs/1810.08316).
- [92] Zhang, H., He, W., Zhang, L., Shen, H. and Yuan, Q. (2013). Hyperspectral image restoration using low-rank matrix recovery. *IEEE Trans. Geosci. Remote Sens.* **52** 4729–4743.
- [93] Zhang, L., Zhou, W. and Wang, H. (2022). Supplement to “Non-Asymptotic Properties of Spectral Decomposition of Large Gram-Type Matrices and Applications.” <https://doi.org/10.3150/21-BEJ1384SUPP>
- [94] Zhang, X. and Cheng, G. (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli* **24** 2640–2675.
- [95] Zou, H., Hastie, T. and Tibshirani, R. (2006). Sparse principal component analysis. *J. Comput. Graph. Statist.* **15** 265–286.
- [96] Zou, H. and Xue, L. (2018). A selective overview of sparse principal component analysis. *Proc. IEEE* **106** 1311–1320.

Nonparametric regression for locally stationary random fields under stochastic sampling design

DAISUKE KURISU

Department of Industrial Engineering and Economics, Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo, 152-8552, Japan.

E-mail: kurisu.d.aa@m.titech.ac.jp

In this study, we develop an asymptotic theory of nonparametric regression for locally stationary random fields (LSRFs) $\{X_{s,A_n} : s \in R_n\}$ in $\mathbb{R}^P$ observed at irregularly spaced locations in $R_n = [0, A_n]^d \subset \mathbb{R}^d$. We first derive the uniform convergence rate of general kernel estimators, followed by the asymptotic normality of an estimator for the mean function of the model. Moreover, we consider additive models to avoid the curse of dimensionality arising from the dependence of the convergence rate of estimators on the number of covariates. Subsequently, we derive the uniform convergence rate and joint asymptotic normality of the estimators for additive functions. We also introduce *approximately m_n -dependent* RFs to provide examples of LSRFs. We find that these RFs include a wide class of Lévy-driven moving average RFs.

Keywords: Nonparametric regression; locally stationary random field; irregularly spaced data; additive model; Lévy-driven moving average random field

References

- [1] Altmeyer, R. and Reiss, M. (2021). Nonparametric estimation for linear SPDEs from local measurements. *Ann. Appl. Probab.* **31** 1–38. [MR4254472](#) <https://doi.org/10.1214/20-aap1581>
- [2] Bandyopadhyay, S., Lahiri, S.N. and Nordman, D.J. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *Ann. Statist.* **43** 519–545. [MR3316189](#) <https://doi.org/10.1214/14-AOS1291>
- [3] Bernstein, S.N. (1926). Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes. *Ann. of Math.* **21** 1–59.
- [4] Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields. *Ann. Probab.* **10** 1047–1050. [MR0672305](#)
- [5] Bradley, R.C. (1989). A caution on mixing conditions for random fields. *Statist. Probab. Lett.* **8** 489–491. [MR1040812](#) [https://doi.org/10.1016/0167-7152\(89\)90032-1](https://doi.org/10.1016/0167-7152(89)90032-1)
- [6] Brockwell, P.J. (2001). Lévy-driven CARMA processes. *Ann. Inst. Statist. Math.* **53** 113–124. Nonlinear non-Gaussian models and related filtering methods (Tokyo, 2000). [MR1820952](#) <https://doi.org/10.1023/A:1017972605872>
- [7] Brockwell, P.J. and Matsuda, Y. (2017). Continuous auto-regressive moving average random fields on $\mathbb{R}^n$. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 833–857. [MR3641410](#) <https://doi.org/10.1111/rssb.12197>
- [8] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. [MR1429916](#) <https://doi.org/10.1214/aos/1034276620>
- [9] Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli* **25** 1013–1044. [MR3920364](#) <https://doi.org/10.3150/17-bej1011>
- [10] Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying ARCH processes. *Ann. Statist.* **34** 1075–1114. [MR2278352](#) <https://doi.org/10.1214/009053606000000227>
- [11] Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence: With Examples and Applications. Lecture Notes in Statistics* **190**. New York: Springer. [MR2338725](#)
- [12] Doukhan, P. (1994). *Mixing: Properties and Examples. Lecture Notes in Statistics* **85**. New York: Springer. [MR1312160](#) <https://doi.org/10.1007/978-1-4612-2642-0>

- [13] Fryzlewicz, P., Sapatinas, T. and Subba Rao, S. (2008). Normalized least-squares estimation in time-varying ARCH models. *Ann. Statist.* **36** 742–786. [MR2396814](#) <https://doi.org/10.1214/07-AOS510>
- [14] Fuglstad, G.-A., Simpson, D., Lindgren, F. and Rue, H. (2015). Does non-stationary spatial data always require non-stationary random fields? *Spat. Stat.* **14** 505–531. [MR3431054](#) <https://doi.org/10.1016/j.spasta.2015.10.001>
- [15] Gelfand, A.E., Kim, H.-J., Sirmans, C.F. and Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. *J. Amer. Statist. Assoc.* **98** 387–396. [MR1995715](#) <https://doi.org/10.1198/0162145030000170>
- [16] Hafner, C.M. and Linton, O. (2010). Efficient estimation of a multivariate multiplicative volatility model. *J. Econometrics* **159** 55–73. [MR2720844](#) <https://doi.org/10.1016/j.jeconom.2010.04.007>
- [17] Hallin, M., Lu, Z. and Tran, L.T. (2004). Local linear spatial regression. *Ann. Statist.* **32** 2469–2500. [MR2153992](#) <https://doi.org/10.1214/009053604000000850>
- [18] Hallin, M., Lu, Z. and Yu, K. (2009). Local linear spatial quantile regression. *Bernoulli* **15** 659–686. [MR2555194](#) <https://doi.org/10.3150/08-BEJ168>
- [19] Hansen, B.E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* **24** 726–748. [MR2409261](#) <https://doi.org/10.1017/S0266466608080304>
- [20] Horowitz, J.L. and Lee, S. (2012). Uniform confidence bands for functions estimated nonparametrically with instrumental variables. *J. Econometrics* **168** 175–188. [MR2923762](#) <https://doi.org/10.1016/j.jeconom.2011.12.001>
- [21] Jenish, N. (2012). Nonparametric spatial regression under near-epoch dependence. *J. Econometrics* **167** 224–239. [MR2885448](#) <https://doi.org/10.1016/j.jeconom.2011.11.008>
- [22] Karcher, W., Roth, S., Spodarev, E. and Walk, C. (2019). An inverse problem for infinitely divisible moving average random fields. *Stat. Inference Stoch. Process.* **22** 263–306. [MR3959289](#) <https://doi.org/10.1007/s11203-018-9188-6>
- [23] Kato, K. and Kurisu, D. (2020). Bootstrap confidence bands for spectral estimation of Lévy densities under high-frequency observations. *Stochastic Process. Appl.* **130** 1159–1205. [MR4058270](#) <https://doi.org/10.1016/j.spa.2019.04.012>
- [24] Koo, B. and Linton, O. (2012). Estimation of semiparametric locally stationary diffusion models. *J. Econometrics* **170** 210–233. [MR2955950](#) <https://doi.org/10.1016/j.jeconom.2012.05.003>
- [25] Kristensen, D. (2009). Uniform convergence rates of kernel estimators with heterogeneous dependent data. *Econometric Theory* **25** 1433–1445. [MR2540506](#) <https://doi.org/10.1017/S0266466609090744>
- [26] Kristensen, D. (2011). Stationary approximations of time-inhomogenous Markov chains with applications. Mimeo.
- [27] Kurisu, D. (2019). On nonparametric inference for spatial regression models under domain expanding and infill asymptotics. *Statist. Probab. Lett.* **154** 108543. [MR3980502](#) <https://doi.org/10.1016/j.spl.2019.06.019>
- [28] Kurisu, D. (2021). Nonparametric regression for locally stationary functional time series. [arXiv:2105.07613](#).
- [29] Kurisu, D. (2021). On the estimation of locally stationary functional time series. [arXiv:2105.11873](#).
- [30] Kurisu, D. (2022). Supplement to “Nonparametric regression for locally stationary random fields under stochastic sampling design.” <https://doi.org/10.3150/21-BEJ1385SUPP>
- [31] Kurisu, D., Kato, K. and Shao, X. (2021). Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. [arXiv:2103.10720](#).
- [32] Lahiri, S.N. (2003). Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **65** 356–388. [MR2028905](#)
- [33] Lahiri, S.N. and Zhu, J. (2006). Resampling methods for spatial regression models under a class of stochastic designs. *Ann. Statist.* **34** 1774–1813. [MR2283717](#) <https://doi.org/10.1214/0090536060000000551>
- [34] Lu, Z. and Tjøstheim, D. (2014). Nonparametric estimation of probability density functions for irregularly observed spatial data. *J. Amer. Statist. Assoc.* **109** 1546–1564. [MR3293609](#) <https://doi.org/10.1080/01621459.2014.947376>
- [35] Mammen, E., Linton, O. and Nielsen, J. (1999). The existence and asymptotic properties of a backfitting projection algorithm under weak conditions. *Ann. Statist.* **27** 1443–1490. [MR1742496](#) <https://doi.org/10.1214/aos/1017939137>
- [36] Marquardt, T. and Stelzer, R. (2007). Multivariate CARMA processes. *Stochastic Process. Appl.* **117** 96–120. [MR2287105](#) <https://doi.org/10.1016/j.spa.2006.05.014>

- [37] Masry, E. (1996). Multivariate local polynomial regression for time series: Uniform strong consistency and rates. *J. Time Series Anal.* **17** 571–599. [MR1424907](#) <https://doi.org/10.1111/j.1467-9892.1996.tb00294.x>
- [38] Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on $\mathbb{R}^d$. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 191–217. [MR2655530](#) <https://doi.org/10.1111/j.1467-9868.2008.00685.x>
- [39] Matsuda, Y. and Yajima, Y. (2018). Locally stationary spatio-temporal processes. *Jpn. J. Stat. Data Sci.* **1** 41–57. [MR4010751](#) <https://doi.org/10.1007/s42081-018-0003-9>
- [40] Matsuda, Y. and Yuan, X. (2020). Multivariate CARMA random fields. DSSR Discussion Papers No. 113.
- [41] Rajput, B.S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probab. Theory Related Fields* **82** 451–487. [MR1001524](#) <https://doi.org/10.1007/BF00339998>
- [42] Robinson, P.M. (2011). Asymptotic theory for nonparametric regression with spatial data. *J. Econometrics* **165** 5–19. [MR2842798](#) <https://doi.org/10.1016/j.jeconom.2011.05.002>
- [43] Schlemm, E. and Stelzer, R. (2012). Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli* **18** 46–63. [MR2888698](#) <https://doi.org/10.3150/10-BEJ329>
- [44] Steel, M.F.J. and Fuentes, M. (2010). Non-Gaussian and nonparametric models for continuous spatial data. In *Handbook of Spatial Statistics* (A.E. Gelfand et al., eds.). *Chapman & Hall/CRC Handb. Mod. Stat. Methods* 149–167. Boca Raton, FL: CRC Press. [MR2730943](#) <https://doi.org/10.1201/9781420072884-c11>
- [45] Truquet, L. (2017). Parameter stability and semiparametric inference in time varying auto-regressive conditional heteroscedasticity models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1391–1414. [MR3731668](#) <https://doi.org/10.1111/rssb.12221>
- [46] Truquet, L. (2019). Local stationarity and time-inhomogeneous Markov chains. *Ann. Statist.* **47** 2023–2050. [MR3953443](#) <https://doi.org/10.1214/18-AOS1739>
- [47] Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. [MR3097614](#) <https://doi.org/10.1214/12-AOS1043>
- [48] Yu, B. (1994). Rates of convergence for empirical processes of stationary mixing sequences. *Ann. Probab.* **22** 94–116. [MR1258867](#)
- [49] Zhang, T. and Wu, W.B. (2015). Time-varying nonlinear regression models: Nonparametric estimation and model selection. *Ann. Statist.* **43** 741–768. [MR3319142](#) <https://doi.org/10.1214/14-AOS1299>
- [50] Zhao, Z. and Wu, W.B. (2008). Confidence bands in nonparametric time series regression. *Ann. Statist.* **36** 1854–1878. [MR2435458](#) <https://doi.org/10.1214/07-AOS533>
- [51] Zhou, Z. (2014). Nonparametric specification for non-stationary time series regression. *Bernoulli* **20** 78–108. [MR3160574](#) <https://doi.org/10.3150/12-BEJ477>
- [52] Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. [MR2541444](#) <https://doi.org/10.1214/08-AOS636>

A Cramér–Wold device for infinite divisibility of $\mathbb{Z}^d$ -valued distributions

DAVID BERGER¹ and ALEXANDER LINDNER²

¹*Institut für Mathematische Stochastik, Technische Universität Dresden, D-01062 Dresden, Germany.
 E-mail: david.berger2@tu-dresden.de*

²*Institute of Mathematical Finance, Ulm University, D-89081 Ulm, Germany.
 E-mail: alexander.lindner@uni-ulm.de*

We show that a Cramér–Wold device holds for infinite divisibility of $\mathbb{Z}^d$ -valued distributions, i.e. that the distribution of a $\mathbb{Z}^d$ -valued random vector X is infinitely divisible if and only if the distribution of $a^T X$ is infinitely divisible for all $a \in \mathbb{R}^d$, and that this in turn is equivalent to infinite divisibility of the distribution of $a^T X$ for all $a \in \mathbb{N}_0^d$. A key tool for proving this is a Lévy–Khintchine type representation with a signed Lévy measure for the characteristic function of a $\mathbb{Z}^d$ -valued distribution, provided the characteristic function is zero-free.

Keywords: Cramér–Wold device; infinitely divisible distribution; quasi-infinitely divisible distribution; signed Lévy measure

References

- [1] Alexeev, I.A. and Khartov, A.A. (2021). Spectral representations of characteristic functions of discrete probability laws. Preprint. Available at [arXiv:2101.06038](https://arxiv.org/abs/2101.06038).
- [2] Basrak, B., Davis, R.A. and Mikosch, T. (2002). A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12** 908–920. [MR1925445 https://doi.org/10.1214/aoap/1031863174](https://doi.org/10.1214/aoap/1031863174)
- [3] Berger, D. (2019). On quasi-infinitely divisible distributions with a point mass. *Math. Nachr.* **292** 1674–1684. [MR3994295 https://doi.org/10.1002/mana.201800073](https://doi.org/10.1002/mana.201800073)
- [4] Calderón, A., Spitzer, F. and Widom, H. (1959). Inversion of Toeplitz matrices. *Illinois J. Math.* **3** 490–498. [MR0121652 https://doi.org/10.1215/ijm/1255631652](https://doi.org/10.1215/ijm/1255631652)
- [5] Chhaiba, H., Demni, N. and Mouayn, Z. (2016). Analysis of generalized negative binomial distributions attached to hyperbolic Landau levels. *J. Math. Phys.* **57** 072103, 14 pp. [MR3522606 https://doi.org/10.1063/1.4958724](https://doi.org/10.1063/1.4958724)
- [6] Cramér, H. and Wold, H. (1936). Some Theorems on Distribution Functions. *J. Lond. Math. Soc.* **11** 290–294. [MR1574927 https://doi.org/10.1112/jlms/s1-11.4.290](https://doi.org/10.1112/jlms/s1-11.4.290)
- [7] Cuppens, R. (1975). *Decomposition of Multivariate Probabilities. Probability and Mathematical Statistics*, Vol. 29. New York: Academic Press. [MR0517412 https://doi.org/10.1080/00036817508839142](https://doi.org/10.1080/00036817508839142)
- [8] Demni, N. and Mouayn, Z. (2015). Analysis of generalized Poisson distributions associated with higher Landau levels. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **18** 1550028, 13 pp. [MR3447229 https://doi.org/10.1142/S0219025715500289](https://doi.org/10.1142/S0219025715500289)
- [9] Dwass, M. and Teicher, H. (1957). On infinitely divisible random vectors. *Ann. Math. Stat.* **28** 461–470. [MR0091550 https://doi.org/10.1214/aoms/1177706974](https://doi.org/10.1214/aoms/1177706974)
- [10] Grafakos, L. (2014). *Classical Fourier Analysis*, 3rd ed. *Graduate Texts in Mathematics* **249**. New York: Springer. [MR3243734 https://doi.org/10.1007/978-1-4939-1194-3](https://doi.org/10.1007/978-1-4939-1194-3)
- [11] Hult, H. and Lindskog, F. (2006). On Kesten’s counterexample to the Cramér–Wold device for regular variation. *Bernoulli* **12** 133–142. [MR2202325 https://doi.org/10.1080/10236190600572325](https://doi.org/10.1080/10236190600572325)
- [12] Ibragimov, I.A. (1972). On a problem of C. R. Rao on I. D. laws. *Sankhyā Ser. A* **34** 447–448. [MR0331464 https://doi.org/10.2307/3293464](https://doi.org/10.2307/3293464)
- [13] Khartov, A.A. (2019). Compactness criteria for quasi-infinitely divisible distributions on the integers. *Statist. Probab. Lett.* **153** 1–6. [MR3957474 https://doi.org/10.1016/j.spl.2019.05.008](https://doi.org/10.1016/j.spl.2019.05.008)

- [14] Lindner, A., Pan, L. and Sato, K. (2018). On quasi-infinitely divisible distributions. *Trans. Amer. Math. Soc.* **370** 8483–8520. [MR3864385](#) <https://doi.org/10.1090/tran/7249>
- [15] Linnik, Y.V. (1964). *Decomposition of Probability Distributions*. Edinburgh: Oliver and Boyd Ltd. [MR0189085](#)
- [16] Nakamura, T. (2015). A complete Riemann zeta distribution and the Riemann hypothesis. *Bernoulli* **21** 604–617. [MR3322332](#) <https://doi.org/10.3150/13-BEJ581>
- [17] Passeggeri, R. (2020). Spectral representations of quasi-infinitely divisible processes. *Stochastic Process. Appl.* **130** 1735–1791. [MR4058288](#) <https://doi.org/10.1016/j.spa.2019.05.014>
- [18] Rudin, W. (1962). *Fourier Analysis on Groups*. Interscience Tracts in Pure and Applied Mathematics, No. 12. New York: Wiley. [MR0152834](#)
- [19] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic models with infinite variance*. *Stochastic Modeling*. New York: CRC Press. [MR1280932](#)
- [20] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. [MR1739520](#)
- [21] Zhang, H., Liu, Y. and Li, B. (2014). Notes on discrete compound Poisson model with applications to risk theory. *Insurance Math. Econom.* **59** 325–336. [MR3283233](#) <https://doi.org/10.1016/j.insmatheco.2014.09.012>

Adaptive Bayesian density estimation in sup-norm

ZACHARIE NAULET

Laboratoire de mathématiques d'Orsay, CNRS, Université Paris-Saclay, 91405, Orsay, France.
E-mail: zacharie.naulet@universite-paris-saclay.fr

We investigate the problem of deriving adaptive posterior rates of contraction on $\mathbb{L}_\infty$ balls in density estimation. Although it is known that log-density priors can achieve optimal rates when the true density is sufficiently smooth, adaptive rates were still to be proven. Here we establish that the so-called *spike-and-slab* prior can achieve adaptive and optimal posterior contraction rates. Along the way, we prove a generic $\mathbb{L}_\infty$ contraction result for log-density priors with independent wavelet coefficients. Interestingly, our approach is different from previous works on $\mathbb{L}_\infty$ contraction and is reminiscent of the classical test-based approach used in Bayesian nonparametrics. Moreover, we require no lower bound on the smoothness of the true density, albeit the rates are deteriorated by an extra $\log(n)$ factor in the case of low smoothness.

Keywords: Bayesian density estimation; supremum norm; adaptation

References

- [1] Barron, A., Schervish, M.J. and Wasserman, L. (1999). The consistency of posterior distributions in non-parametric problems. *Ann. Statist.* **27** 536–561. [MR1714718](#) <https://doi.org/10.1214/aos/1018031206>
- [2] 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>
- [3] Brown, L.D. and Zhang, C.-H. (1998). Asymptotic nonequivalence of nonparametric experiments when the smoothness index is $1/2$. *Ann. Statist.* **26** 279–287. [MR1611772](#) <https://doi.org/10.1214/aos/1030563986>
- [4] Castillo, I. (2014). On Bayesian supremum norm contraction rates. *Ann. Statist.* **42** 2058–2091. [MR3262477](#) <https://doi.org/10.1214/14-AOS1253>
- [5] Castillo, I. (2017). Pólya tree posterior distributions on densities. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 2074–2102. [MR3729648](#) <https://doi.org/10.1214/16-AIHP784>
- [6] Castillo, I. and Mismar, R. (2019). Spike and Slab Pólya tree posterior distributions. Preprint. Available at [arXiv:1911.12106](https://arxiv.org/abs/1911.12106).
- [7] Cohen, A., Daubechies, I. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.* **1** 54–81. [MR1256527](#) <https://doi.org/10.1006/acha.1993.1005>
- [8] Donoho, D.L., Johnstone, I.M., Kerkycharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24** 508–539. [MR1394974](#) <https://doi.org/10.1214/aos/1032894451>
- [9] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. [MR1790007](#) <https://doi.org/10.1214/aos/1016218228>
- [10] Ghosal, S. and van der Vaart, A. (2007). Posterior convergence rates of Dirichlet mixtures at smooth densities. *Ann. Statist.* **35** 697–723. [MR2336864](#) <https://doi.org/10.1214/009053606000001271>
- [11] Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist.* **35** 192–223. [MR2332274](#) <https://doi.org/10.1214/009053606000001172>
- [12] Giné, E. and Nickl, R. (2011). Rates of contraction for posterior distributions in L^r -metrics, $1 \leq r \leq \infty$. *Ann. Statist.* **39** 2883–2911. [MR3012395](#) <https://doi.org/10.1214/11-AOS924>
- [13] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge Series in Statistical and Probabilistic Mathematics **40**. New York: Cambridge Univ. Press. [MR3588285](#) <https://doi.org/10.1017/CBO9781107337862>

- [14] Goldenshluger, A. and Lepski, O. (2014). On adaptive minimax density estimation on R^d . *Probab. Theory Related Fields* **159** 479–543. [MR3230001](#) <https://doi.org/10.1007/s00440-013-0512-1>
- [15] Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, A. (2000). *Wavelets, Approximation, and Statistical Applications. Lecture Notes in Statistics* **129**. New York: Springer. [MR1618204](#) <https://doi.org/10.1007/978-1-4612-2222-4>
- [16] Hasminskii, R.Z. (1978). A lower bound for risks of nonparametric density estimates in the uniform metric. *Theory Probab. Appl.* **23** 794–798.
- [17] Hoffmann, M., Rousseau, J. and Schmidt-Hieber, J. (2015). On adaptive posterior concentration rates. *Ann. Statist.* **43** 2259–2295. [MR3396985](#) <https://doi.org/10.1214/15-AOS1341>
- [18] Ibragimov, I.A. and Hasminskii, R.Z. (1980). On estimate of the density function. *Zap. Nauchn. Sem. POMI* **98** 61–85.
- [19] Lijoi, A., Prünster, I. and Walker, S.G. (2005). On consistency of nonparametric normal mixtures for Bayesian density estimation. *J. Amer. Statist. Assoc.* **100** 1292–1296. [MR2236442](#) <https://doi.org/10.1198/016214505000000358>
- [20] Mitchell, T.J. and Beauchamp, J.J. (1988). Bayesian variable selection in linear regression. *J. Amer. Statist. Assoc.* **83** 1023–1032. [MR0997578](#)
- [21] Naulet, Z. (2022). Supplement to “Adaptive Bayesian density estimation in sup-norm.” <https://doi.org/10.3150/21-BEJ1387SUPP>
- [22] Pollard, D. (1990). *Empirical Processes: Theory and Applications. NSF-CBMS Regional Conference Series in Probability and Statistics* **2**. Hayward, CA: IMS. [MR1089429](#)
- [23] Rivoirard, V. and Rousseau, J. (2012). Bernstein–von Mises theorem for linear functionals of the density. *Ann. Statist.* **40** 1489–1523. [MR3015033](#) <https://doi.org/10.1214/12-AOS1004>
- [24] Schwartz, L. (1965). On Bayes procedures. *Z. Wahrsch. Verw. Gebiete* **4** 10–26. [MR0184378](#) <https://doi.org/10.1007/BF00535479>
- [25] Scricciolo, C. (2014). Adaptive Bayesian density estimation in L^p -metrics with Pitman–Yor or normalized inverse-Gaussian process kernel mixtures. *Bayesian Anal.* **9** 475–520. [MR3217004](#) <https://doi.org/10.1214/14-BA863>
- [26] Shen, X. and Wasserman, L. (2001). Rates of convergence of posterior distributions. *Ann. Statist.* **29** 687–714. [MR1865337](#) <https://doi.org/10.1214/aos/1009210686>
- [27] Walker, S.G., Lijoi, A. and Prünster, I. (2007). On rates of convergence for posterior distributions in infinite-dimensional models. *Ann. Statist.* **35** 738–746. [MR2336866](#) <https://doi.org/10.1214/009053606000001361>
- [28] Yoo, W.W. and Ghosal, S. (2016). Supremum norm posterior contraction and credible sets for nonparametric multivariate regression. *Ann. Statist.* **44** 1069–1102. [MR3485954](#) <https://doi.org/10.1214/15-AOS1398>
- [29] Yoo, W.W., Rousseau, J. and Rivoirard, V. (2017). Adaptive Supremum Norm Posterior Contraction: Spike-and-Slab Priors and Anisotropic Besov Spaces. Preprint. Available at [arXiv:1708.01909](https://arxiv.org/abs/1708.01909).

Markov-modulated generalized Ornstein-Uhlenbeck processes and an application in risk theory

ANITA BEHME^{*} and APOSTOLOS SIDERIS[†]

Technische Universität Dresden, Institute of Mathematical Stochastics, 01062 Dresden, Germany.

E-mail: ^{*}anita.behme@tu-dresden.de; [†]apostolos.sideris@tu-dresden.de

By embedding a Markov-modulated random recurrence equation in continuous time, we derive the Markov-modulated generalized Ornstein-Uhlenbeck process. This process turns out to be the unique solution of a stochastic differential equation driven by a bivariate Markov-additive process. We present this stochastic differential equation as well as its solution explicitly in terms of the driving Markov-additive process. Moreover, we give necessary and sufficient conditions for strict stationarity of the Markov-modulated generalized Ornstein-Uhlenbeck process, and prove that its stationary distribution is given by the distribution of a certain exponential functional of Markov-additive processes. Finally, we propose a Markov-modulated risk model with investment that generalizes Paulsen's risk process to a Markov-switching environment, and derive a formula for the ruin probability in this risk model.

Keywords: Exponential functional; generalized Ornstein-Uhlenbeck process; Lévy process; Markov additive process; Markov-modulated random recurrence equation; Markov-switching model; risk theory; ruin probability; stationary process

References

- [1] Alsmeyer, G. and Buckmann, F. (2017). Stability of perpetuities in Markovian environment. *J. Difference Equ. Appl.* **23** 699–740. [MR3678532](#) <https://doi.org/10.1080/10236198.2016.1271878>
- [2] Asmussen, S. (2003). *Applied Probability and Queues*, 2nd ed. *Stochastic Modelling and Applied Probability* **51**. New York: Springer. [MR1978607](#)
- [3] Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 167–241. [MR1841412](#) <https://doi.org/10.1111/1467-9868.00282>
- [4] Behme, A., Lindner, A. and Maller, R. (2011). Stationary solutions of the stochastic differential equation $dV_t = V_t - dU_t + dL_t$ with Lévy noise. *Stochastic Process. Appl.* **121** 91–108. [MR2739007](#) <https://doi.org/10.1016/j.spa.2010.09.003>
- [5] Behme, A., Lindner, A., Reker, J. and Rivero, V. (2021). Continuity properties and the support of killed exponential functionals. *Stochastic Process. Appl.* **140** 115–146. [MR4276495](#) <https://doi.org/10.1016/j.spa.2021.06.002>
- [6] Behme, A. and Sideris, A. (2020). Exponential functionals of Markov additive processes. *Electron. J. Probab.* **25** Paper No. 37, 25. [MR4089787](#) <https://doi.org/10.1214/20-ejp441>
- [7] Breiman, L. (1968). *Probability*. Reading, Mass.-London-Don Mills, Ont.: Addison-Wesley Company. [MR0229267](#)
- [8] Çinlar, E. (1972). Markov additive processes. I. *Z. Wahrsch. Verw. Gebiete* **24** 85–93. <https://doi.org/10.1007/BF00532536>
- [9] Çinlar, E. (1972). Markov additive processes. II. *Z. Wahrsch. Verw. Gebiete* **24** 95–121. <https://doi.org/10.1007/BF00532537>
- [10] Dereich, S., Döring, L. and Kyprianou, A.E. (2017). Real self-similar processes started from the origin. *Ann. Probab.* **45** 1952–2003. [MR3650419](#) <https://doi.org/10.1214/16-AOP1105>

- [11] Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **322** 549–560.
- [12] Goldie, C.M. and Maller, R.A. (2000). Stability of perpetuities. *Ann. Probab.* **28** 1195–1218. [MR1797309 https://doi.org/10.1214/aop/1019160331](https://doi.org/10.1214/aop/1019160331)
- [13] Grigelionis, B. (1978). Markov additive processes. *Litovsk. Mat. Sb.* **18** 43–47, 213. [MR0518317](https://doi.org/10.1016/0304-4149(89)90077-X)
- [14] de Haan, L. and Karandikar, R.L. (1989). Embedding a stochastic difference equation into a continuous-time process. *Stochastic Process. Appl.* **32** 225–235. [MR1014451 https://doi.org/10.1016/0304-4149\(89\)90077-X](https://doi.org/10.1016/0304-4149(89)90077-X)
- [15] Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57** 357–384. [MR0996941 https://doi.org/10.2307/1912559](https://doi.org/10.2307/1912559)
- [16] Hamilton, J.D. and Raj, B. (2002). *Advances in Markov-Switching Models: Applications in Business Cycle Research and Finance*. Heidelberg: Physica-Verlag.
- [17] Huang, G., Jansen, H.M., Mandjes, M., Spreij, P. and De Turck, K. (2016). Markov-modulated Ornstein-Uhlenbeck processes. *Adv. in Appl. Probab.* **48** 235–254. [MR3473576](https://doi.org/10.1016/j.spa.2017.04.010)
- [18] Jacod, J. (1985). Grossissement de filtration et processus d’Ornstein-Uhlenbeck généralisé. In *Grossissements de Filtrations: Exemples et Applications* (Jeulin, Th. and Yor, M., eds.) *Lecture Notes in Math.* **1118**. Berlin: Springer.
- [19] Janssen, J. (1980/81). Some transient results on the $M/SM/1$ special semi-Markov model in risk and queueing theories. *Astin Bull.* **11** 41–51. [MR0580670 https://doi.org/10.1017/S0515036100006607](https://doi.org/10.1017/S0515036100006607)
- [20] 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](https://doi.org/10.1016/j.spa.2017.04.010)
- [21] Klüppelberg, C., Lindner, A. and Maller, R. (2004). A continuous-time GARCH process driven by a Lévy process: Stationarity and second-order behaviour. *J. Appl. Probab.* **41** 601–622. [MR2074811 https://doi.org/10.1017/s0021900200020428](https://doi.org/10.1017/s0021900200020428)
- [22] Kötter, M. and Bäuerle, N. (2007). The Markov-modulated risk model with investment. *Oper. Res. Proc.* **2006** 575–580.
- [23] Lindner, A. and Maller, R. (2005). Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. *Stochastic Process. Appl.* **115** 1701–1722. [MR2165340 https://doi.org/10.1016/j.spa.2005.05.004](https://doi.org/10.1016/j.spa.2005.05.004)
- [24] Lindskog, F. and Pal Majumder, A. (2020). Exact long time behavior of some regime switching stochastic processes. *Bernoulli* **26** 2572–2604. [MR4140522 https://doi.org/10.3150/20-BEJ1196](https://doi.org/10.3150/20-BEJ1196)
- [25] Maller, R., Müller, G. and Szimayer, A. (2009). Ornstein-Uhlenbeck processes and extensions. In *Handbook of Financial Time Series* (T. Andersen, R.A. Davis, J.P. Kreiß and T. Mikosch, eds.) 421–437. Springer Berlin.
- [26] Ornstein, L. and Uhlenbeck, G. (1930). On the theory of the Brownian motion. *Phys. Rev.* **36** 823–84.
- [27] Paulsen, J. (1993). Risk theory in a stochastic economic environment. *Stochastic Process. Appl.* **46** 327–361. [MR1226415 https://doi.org/10.1016/0304-4149\(93\)90010-2](https://doi.org/10.1016/0304-4149(93)90010-2)
- [28] Paulsen, J. and Gjessing, H.K. (1997). Ruin theory with stochastic return on investments. *Adv. in Appl. Probab.* **29** 965–985. [MR1484776 https://doi.org/10.2307/1427849](https://doi.org/10.2307/1427849)
- [29] Protter, P.E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **21**. Berlin: Springer. [MR2020294](https://doi.org/10.1016/j.cam.2016.09.010)
- [30] Ramsden, L. and Papaioannou, A.D. (2017). Asymptotic results for a Markov-modulated risk process with stochastic investment. *J. Comput. Appl. Math.* **313** 38–53. [MR3573225 https://doi.org/10.1016/j.cam.2016.09.010](https://doi.org/10.1016/j.cam.2016.09.010)
- [31] Reinhard, J.M. (1984). On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment. *Astin Bull.* **14** 23–43.
- [32] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. [MR1739520](https://doi.org/10.1080/03610918.2015.1132321)
- [33] Skorokhod, A.V. (1982). Operator stochastic differential equations and stochastic semigroups. *Uspekhi Mat. Nauk* **37** 157–183. [MR0683278](https://doi.org/10.1016/0304-4149(82)90050-3)
- [34] Wolfe, S.J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$. *Stochastic Process. Appl.* **12** 301–312. [MR0656279 https://doi.org/10.1016/0304-4149\(82\)90050-3](https://doi.org/10.1016/0304-4149(82)90050-3)
- [35] Zhang, Z. and Wang, W. (2017). The stationary distribution of Ornstein-Uhlenbeck process with a two-state Markov switching. *Comm. Statist. Simulation Comput.* **46** 4783–4794. [MR3672577 https://doi.org/10.1080/03610918.2015.1132321](https://doi.org/10.1080/03610918.2015.1132321)

- [36] Zhao, K. (2012). Lévy driven Markov-modulated Ornstein-Uhlenbeck processes - Applications to Alberta electricity market Ph.D. thesis Univ. Calgary, Alberta.

Symmetric inclusion process with slow boundary: Hydrodynamics and hydrostatics

CHIARA FRANCESCHINI^{1,*}, PATRÍCIA GONÇALVES^{1,†} and FEDERICO SAU²

¹Center for Mathematical Analysis, Geometry and Dynamical Systems, IST, Universidade de Lisboa, 1049-001 Lisboa, Portugal.

E-mail: *chiara.franceschini@tecnico.ulisboa.pt; †pgoncalves@tecnico.ulisboa.pt

²IST Austria, Am Campus 1, 3400, Klosterneuburg, Austria.

E-mail: federico.sau@ist.ac.at

We study the hydrodynamic and hydrostatic limits of the one-dimensional open symmetric inclusion process with slow boundary. Depending on the value of the parameter tuning the interaction rate of the bulk of the system with the boundary, we obtain a linear heat equation with either Dirichlet, Robin or Neumann boundary conditions as hydrodynamic equation. In our approach, we combine duality and first-second class particle techniques to reduce the scaling limit of the inclusion process to the limiting behavior of a single, non-interacting, particle.

Keywords: Symmetric inclusion process; slow boundary; hydrodynamic limit; hydrostatic limit; non-equilibrium steady state; duality for Markov processes; second class particles

References

- [1] Adams, R.A. and Fournier, J.J.F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics (Amsterdam)* **140**. Amsterdam: Elsevier/Academic Press. [MR2424078](#)
- [2] Arendt, W. and Warma, M. (2003). Dirichlet and Neumann boundary conditions: What is in between? *J. Evol. Equ.* **3** 119–135. Dedicated to Philippe Bérnilan. [MR1977430](#) <https://doi.org/10.1007/s000280300005>
- [3] Arendt, W. and Warma, M. (2003). The Laplacian with Robin boundary conditions on arbitrary domains. *Potential Anal.* **19** 341–363. [MR1988110](#) <https://doi.org/10.1023/A:1024181608863>
- [4] Baldasso, R., Menezes, O., Neumann, A. and Souza, R.R. (2017). Exclusion process with slow boundary. *J. Stat. Phys.* **167** 1112–1142. [MR3647054](#) <https://doi.org/10.1007/s10955-017-1763-5>
- [5] Bernardin, C., Gonçalves, P., Jara, M. and Scotta, S. (2020). Equilibrium fluctuations for diffusive symmetric exclusion with long jumps and infinitely extended reservoirs. [arXiv:2002.12841](#).
- [6] Bernardin, C., Gonçalves, P. and Jiménez-Oviedo, B. (2021). A microscopic model for a one parameter class of fractional Laplacians with Dirichlet boundary conditions. *Arch. Ration. Mech. Anal.* **239** 1–48. [MR4198713](#) <https://doi.org/10.1007/s00205-020-01549-9>
- [7] Bianchi, A., Dommers, S. and Giardinà, C. (2017). Metastability in the reversible inclusion process. *Electron. J. Probab.* **22** Paper No. 70, 34. [MR3698739](#) <https://doi.org/10.1214/17-EJP98>
- [8] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. [MR1700749](#) <https://doi.org/10.1002/9780470316962>
- [9] Bogachev, V.I. (2007). *Measure Theory. Vol. I, II*. Berlin: Springer. [MR2267655](#) <https://doi.org/10.1007/978-3-540-34514-5>
- [10] Borodin, A.N. and Salminen, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Basel: Birkhäuser. [MR1912205](#) <https://doi.org/10.1007/978-3-0348-8163-0>
- [11] Carinci, G., Franceschini, C., Giardinà, C., Groenevelt, W. and Redig, F. (2019). Orthogonal dualities of Markov processes and unitary symmetries. *SIGMA Symmetry Integrability Geom. Methods Appl.* **15** Paper No. 053, 27. [MR3980548](#) <https://doi.org/10.3842/SIGMA.2019.053>

- [12] Carinci, G., Giardinà, C., Giberti, C. and Redig, F. (2013). Duality for stochastic models of transport. *J. Stat. Phys.* **152** 657–697. [MR3092998](#) <https://doi.org/10.1007/s10955-013-0786-9>
- [13] De Masi, A. and Presutti, E. (1991). *Mathematical Methods for Hydrodynamic Limits. Lecture Notes in Math.* **1501**. Berlin: Springer. [MR1175626](#) <https://doi.org/10.1007/BFb0086457>
- [14] Derrida, B., Evans, M.R., Hakim, V. and Pasquier, V. (1993). Exact solution of a 1D asymmetric exclusion model using a matrix formulation. *J. Phys. A* **26** 1493–1517. [MR1219679](#)
- [15] Donnelly, P. and Kurtz, T.G. (1996). A countable representation of the Fleming-Viot measure-valued diffusion. *Ann. Probab.* **24** 698–742. [MR1404525](#) <https://doi.org/10.1214/aop/1039639359>
- [16] Erignoux, C. (2018). Hydrodynamic limit of boundary driven exclusion processes with nonreversible boundary dynamics. *J. Stat. Phys.* **172** 1327–1357. [MR3856946](#) <https://doi.org/10.1007/s10955-018-2102-1>
- [17] Erignoux, C., Gonçalves, P. and Nahum, G. (2020). Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part I, the critical regime and beyond. *J. Stat. Phys.* **181** 1433–1469. [MR4163507](#) <https://doi.org/10.1007/s10955-020-02633-w>
- [18] Erignoux, C., Gonçalves, P. and Nahum, G. (2020). Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part II, below the critical regime. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** 791–823. [MR4153070](#) <https://doi.org/10.30757/alea.v17-31>
- [19] Evans, L.C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Providence, RI: Amer. Math. Soc. [MR2597943](#) <https://doi.org/10.1090/gsm/019>
- [20] Eyink, G., Lebowitz, J.L. and Spohn, H. (1990). Hydrodynamics of stationary nonequilibrium states for some stochastic lattice gas models. *Comm. Math. Phys.* **132** 253–283. [MR1069212](#)
- [21] Eyink, G., Lebowitz, J.L. and Spohn, H. (1991). Lattice gas models in contact with stochastic reservoirs: Local equilibrium and relaxation to the steady state. *Comm. Math. Phys.* **140** 119–131. [MR1124262](#)
- [22] Farfan, J., Landim, C. and Mourragui, M. (2011). Hydrostatics and dynamical large deviations of boundary driven gradient symmetric exclusion processes. *Stochastic Process. Appl.* **121** 725–758. [MR2770905](#) <https://doi.org/10.1016/j.spa.2010.11.014>
- [23] Floreani, S., Redig, F. and Sau, F. (2021). Orthogonal polynomial duality of boundary driven particle systems and non-equilibrium correlations. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at [arXiv:2007.08272](#). <https://doi.org/10.1214/21-AIHP1163>
- [24] Franco, T., Gonçalves, P. and Neumann, A. (2013). Phase transition in equilibrium fluctuations of symmetric slowed exclusion. *Stochastic Process. Appl.* **123** 4156–4185. [MR3096351](#) <https://doi.org/10.1016/j.spa.2013.06.016>
- [25] Franco, T., Gonçalves, P. and Neumann, A. (2017). Equilibrium fluctuations for the slow boundary exclusion process. In *From Particle Systems to Partial Differential Equations. Springer Proc. Math. Stat.* **209** 177–197. Cham: Springer. [MR3746752](#) https://doi.org/10.1007/978-3-319-66839-0_9
- [26] Frómeta, S., Misturini, R. and Neumann, A. (2021). The Boundary Driven Zero-Range Process. In *From Particle Systems to Partial Differential Equations. ICPS 2019, ICPS 2018, ICPS 2017* (C. Bernardin, F. Golse, P. Gonçalves, V. Ricci and A.J. Soares, eds.). *Springer Proceedings in Mathematics & Statistics* **352**. Cham: Springer. https://doi.org/10.1007/978-3-030-69784-6_12
- [27] Giardinà, C., Kurchan, J. and Redig, F. (2007). Duality and exact correlations for a model of heat conduction. *J. Math. Phys.* **48** 033301, 15. [MR2314497](#) <https://doi.org/10.1063/1.2711373>
- [28] Giardinà, C., Kurchan, J., Redig, F. and Vafayi, K. (2009). Duality and hidden symmetries in interacting particle systems. *J. Stat. Phys.* **135** 25–55. [MR2505724](#) <https://doi.org/10.1007/s10955-009-9716-2>
- [29] Gonçalves, P. (2019). Hydrodynamics for symmetric exclusion in contact with reservoirs. In *Stochastic Dynamics Out of Equilibrium. Springer Proc. Math. Stat.* **282** 137–205. Cham: Springer. [MR3986064](#) https://doi.org/10.1007/978-3-030-15096-9_4
- [30] Gonçalves, P., Jara, M., Menezes, O. and Neumann, A. (2020). Non-equilibrium and stationary fluctuations for the SSEP with slow boundary. *Stochastic Process. Appl.* **130** 4326–4357. [MR4102268](#) <https://doi.org/10.1016/j.spa.2019.12.006>
- [31] Holley, R.A. and Stroock, D.W. (1978). Generalized Ornstein-Uhlenbeck processes and infinite particle branching Brownian motions. *Publ. Res. Inst. Math. Sci.* **14** 741–788. [MR0527199](#) <https://doi.org/10.2977/prims/1195188837>

- [32] Jara, M.D. and Landim, C. (2008). Quenched non-equilibrium central limit theorem for a tagged particle in the exclusion process with bond disorder. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 341–361. [MR2446327 https://doi.org/10.1214/07-AIHP112](https://doi.org/10.1214/07-AIHP112)
- [33] Jatuviriyapornchai, W., Chleboun, P. and Grosskinsky, S. (2020). Structure of the condensed phase in the inclusion process. *J. Stat. Phys.* **178** 682–710. [MR4059956 https://doi.org/10.1007/s10955-019-02451-9](https://doi.org/10.1007/s10955-019-02451-9)
- [34] Kallianpur, G. and Xiong, J. (1995). *Stochastic Differential Equations in Infinite-Dimensional Spaces. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **26**. Hayward, CA: IMS. Expanded version of the lectures delivered as part of the 1993 Barrett Lectures at the University of Tennessee, Knoxville, TN, March 25–27, 1993, With a foreword by Balram S. Rajput and Jan Rosinski. [MR1465436](https://doi.org/10.1007/978-3-662-03752-2)
- [35] Kipnis, C. and Landim, C. (1999). *Scaling Limits of Interacting Particle Systems. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **320**. Berlin: Springer. [MR1707314 https://doi.org/10.1007/978-3-662-03752-2](https://doi.org/10.1007/978-3-662-03752-2)
- [36] Kipnis, C., Marchioro, C. and Presutti, E. (1982). Heat flow in an exactly solvable model. *J. Stat. Phys.* **27** 65–74. [MR0656869 https://doi.org/10.1007/BF01011740](https://doi.org/10.1007/BF01011740)
- [37] Landim, C., Milanés, A. and Olla, S. (2008). Stationary and nonequilibrium fluctuations in boundary driven exclusion processes. *Markov Process. Related Fields* **14** 165–184. [MR2437527](https://doi.org/10.1007/s10955-019-02451-9)
- [38] Landim, C. and Tsunoda, K. (2018). Hydrostatics and dynamical large deviations for a reaction-diffusion model. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 51–74. [MR3765880 https://doi.org/10.1214/16-AIHP794](https://doi.org/10.1214/16-AIHP794)
- [39] Liggett, T.M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **324**. Berlin: Springer. [MR1717346 https://doi.org/10.1007/978-3-662-03990-8](https://doi.org/10.1007/978-3-662-03990-8)
- [40] Liggett, T.M. (2005). *Interacting Particle Systems. Classics in Mathematics*. Berlin: Springer. Reprint of the 1985 original. [MR2108619 https://doi.org/10.1007/b138374](https://doi.org/10.1007/b138374)
- [41] Mitoma, I. (1983). Tightness of probabilities on $C([0, 1]; \mathcal{S}')$ and $\mathcal{D}([0, 1]; \mathcal{S}')$. *Ann. Probab.* **11** 989–999. [MR0714961](https://doi.org/10.1007/BF01011740)
- [42] Netrusov, Yu. and Safarov, Yu. (2005). Weyl asymptotic formula for the Laplacian on domains with rough boundaries. *Comm. Math. Phys.* **253** 481–509. [MR2140257 https://doi.org/10.1007/s00220-004-1158-8](https://doi.org/10.1007/s00220-004-1158-8)
- [43] Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. *Universitext*. Berlin: Springer. [MR2001996 https://doi.org/10.1007/978-3-642-14394-6](https://doi.org/10.1007/978-3-642-14394-6)
- [44] Pérez-Abreu, V. and Tudor, C. (1992). Regularity and convergence of stochastic convolutions in duals of nuclear Fréchet spaces. *J. Multivariate Anal.* **43** 185–199. [MR1193611 https://doi.org/10.1016/0047-259X\(92\)90033-C](https://doi.org/10.1016/0047-259X(92)90033-C)
- [45] Redig, F. and Sau, F. (2018). Factorized duality, stationary product measures and generating functions. *J. Stat. Phys.* **172** 980–1008. [MR3830295 https://doi.org/10.1007/s10955-018-2090-1](https://doi.org/10.1007/s10955-018-2090-1)
- [46] Stone, C. (1963). Limit theorems for random walks, birth and death processes, and diffusion processes. *Illinois J. Math.* **7** 638–660. [MR0158440](https://doi.org/10.1007/BF01011740)
- [47] Tsunoda, K. (2020). Hydrostatic limit for exclusion process with slow boundary revisited. In *Stochastic Analysis on Large Scale Interacting Systems. RIMS Kôkyûroku Bessatsu* **B79** 149–162. Kyoto: Res. Inst. Math. Sci. (RIMS).
- [48] Vafayi, K. and Duong, M.H. (2014). Weakly nonequilibrium properties of a symmetric inclusion process with open boundaries. *Phys. Rev. E* **90** 052143.

Empirical variance minimization with applications in variance reduction and optimal control

DENIS BELOMESTNY^{1,2,*}, LEONID IOSIPOI^{2,†}, QUENTIN PARIS^{2,‡} and NIKITA ZHIVOTOVSKIY³

¹*Faculty of Mathematics, Duisburg-Essen University, Essen, Germany.*

E-mail: *denis.belomestny@uni-due.de

²*Faculty of Computer Science, HSE University, Moscow, Russia.*

E-mail: †liosipoi@hse.ru; ‡qparis@hse.ru

³*Department of Mathematics, ETH, Zürich, Switzerland.*

E-mail: nikita.zhivotovskii@math.ethz.ch

We study the problem of empirical minimization for variance-type functionals over functional classes. Sharp non-asymptotic bounds for the excess variance are derived under mild conditions. In particular, it is shown that under some restrictions imposed on the functional class fast convergence rates can be achieved including the optimal non-parametric rates for expressive classes in the non-Donsker regime under some additional assumptions. Our main applications include variance reduction and optimal control.

Keywords: Empirical variance minimization; variance reduction; control variates; optimal control

References

- [1] Assaraf, R. and Caffarel, M. (1999). Zero-variance principle for Monte Carlo algorithms. *Phys. Rev. Lett.* **83** 4682–4685.
- [2] Barron, A., Birgé, L. and Massart, P. (1999). Risk bounds for model selection via penalization. *Probab. Theory Related Fields* **113** 301–413. [MR1679028](https://doi.org/10.1007/s0044000050210) <https://doi.org/10.1007/s0044000050210>
- [3] Bartlett, P.L., Bousquet, O. and Mendelson, S. (2005). Local Rademacher complexities. *Ann. Statist.* **33** 1497–1537. [MR2166554](https://doi.org/10.1214/009053605000000282) <https://doi.org/10.1214/009053605000000282>
- [4] Bartlett, P.L. and Mendelson, S. (2006). Empirical minimization. *Probab. Theory Related Fields* **135** 311–334. [MR2240689](https://doi.org/10.1007/s00440-005-0462-3) <https://doi.org/10.1007/s00440-005-0462-3>
- [5] Belomestny, D., Iosipoi, L., Moulines, E., Naumov, A. and Samsonov, S. (2020). Variance reduction for Markov chains with application to MCMC. *Stat. Comput.* **30** 973–997. [MR4108687](https://doi.org/10.1007/s11222-020-09931-z) <https://doi.org/10.1007/s11222-020-09931-z>
- [6] Belomestny, D., Iosipoi, L., Moulines, E., Naumov, A. and Samsonov, S. (2021). Variance reduction for dependent sequences with applications to stochastic gradient MCMC. *SIAM/ASA J. Uncertain. Quantificat.* **9** 507–535. [MR4252078](https://doi.org/10.1137/19M1301199) <https://doi.org/10.1137/19M1301199>
- [7] Belomestny, D., Iosipoi, L., Paris, Q. and Zhivotovskiy, N. (2022). Supplement to “Empirical Variance Minimization with Applications in Variance Reduction and Optimal Control.” <https://doi.org/10.3150/21-BEJ1392SUPP>
- [8] Belomestny, D. and Schoenmakers, J. (2018). *Advanced Simulation-Based Methods for Optimal Stopping and Control: With Applications in Finance*. London: Palgrave Macmillan. [MR3752620](https://doi.org/10.1007/978-1-4939-9822-2)
- [9] Belomestny, D.V., Iosipoi, L.S. and Zhivotovskiy, N.K. (2018). Variance reduction in Monte Carlo estimators via empirical variance minimization. *Dokl. Math.* **98** 494–497.
- [10] Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probab. Theory Related Fields* **97** 113–150. [MR1240719](https://doi.org/10.1007/BF01199316) <https://doi.org/10.1007/BF01199316>

- [11] 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>
- [12] Cléménçon, S. (2014). A statistical view of clustering performance through the theory of U -processes. *J. Multivariate Anal.* **124** 42–56. [MR3147310](#) <https://doi.org/10.1016/j.jmva.2013.10.001>
- [13] Cléménçon, S., Colin, I. and Bellet, A. (2016). Scaling-up empirical risk minimization: Optimization of incomplete U -statistics. *J. Mach. Learn. Res.* **17** Paper No. 76, 36. [MR3517099](#)
- [14] Cléménçon, S., Lugosi, G. and Vayatis, N. (2008). Ranking and empirical minimization of U -statistics. *Ann. Statist.* **36** 844–874. [MR2396817](#) <https://doi.org/10.1214/009052607000000910>
- [15] Cléménçon, S.J. (2011). On U -processes and clustering performance. In *Advances in Neural Information Processing Systems* 24 37–45. Curran Associates.
- [16] Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition. Applications of Mathematics (New York)* **31**. New York: Springer. [MR1383093](#) <https://doi.org/10.1007/978-1-4612-0711-5>
- [17] Glasserman, P. (2013). *Monte Carlo Methods in Financial Engineering* **53**. Springer Science & Business Media. [MR1999614](#)
- [18] Halmos, P.R. (1946). The theory of unbiased estimation. *Ann. Math. Stat.* **17** 34–43. [MR0015746](#) <https://doi.org/10.1214/aoms/1177731020>
- [19] 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>
- [20] Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Stat.* **19** 293–325. [MR0026294](#) <https://doi.org/10.1214/aoms/1177730196>
- [21] Hoeffding, W. (1961). The strong law of large numbers for U -statistics. *Institute of Statistics Mimeo Series* **302**.
- [22] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. [MR0144363](#)
- [23] Joly, E. and Lugosi, G. (2016). Robust estimation of U -statistics. *Stochastic Process. Appl.* **126** 3760–3773. [MR3565476](#) <https://doi.org/10.1016/j.spa.2016.04.021>
- [24] Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *Ann. Statist.* **34** 2593–2656. [MR2329442](#) <https://doi.org/10.1214/009053606000001019>
- [25] Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems. Lecture Notes in Math.* **2033**. Heidelberg: Springer. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, École d’Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. [MR2829871](#) <https://doi.org/10.1007/978-3-642-22147-7>
- [26] Korolyuk, V.S. and Borovskich, Yu.V. (2013). *Theory of U -Statistics* **273**. Springer Science & Business Media. [MR1472486](#) <https://doi.org/10.1007/978-94-017-3515-5>
- [27] Kur, G., Dagan, Y. and Rakhlin, A. (2019). Optimality of Maximum Likelihood for Log-Concave Density Estimation and Bounded Convex Regression. arXiv preprint, [arXiv:1903.05315](#).
- [28] Laforgue, P., Cléménçon, S. and Bertail, P. (2019). On medians of (randomized) pairwise means. In *Proceedings of the 36th International Conference on Machine Learning* **97** 1272–1281. PMLR.
- [29] Lecué, G. and Mendelson, S. (2012). General nonexact oracle inequalities for classes with a subexponential envelope. *Ann. Statist.* **40** 832–860. [MR2933668](#) <https://doi.org/10.1214/11-AOS965>
- [30] Massart, P. (2007). *Concentration Inequalities and Model Selection. Lecture Notes in Math.* **1896**. Berlin: Springer. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard. [MR2319879](#)
- [31] Maurer, A. and Pontil, M. (2019). Uniform concentration and symmetrization for weak interactions. In *Proceedings of the Thirty-Second Conference on Learning Theory* **99** 2372–2387. PMLR.
- [32] Mira, A., Solgi, R. and Imparato, D. (2013). Zero variance Markov chain Monte Carlo for Bayesian estimators. *Stat. Comput.* **23** 653–662. [MR3094805](#) <https://doi.org/10.1007/s11222-012-9344-6>
- [33] Nickl, R. and Pötscher, B.M. (2007). Bracketing metric entropy rates and empirical central limit theorems for function classes of Besov- and Sobolev-type. *J. Theoret. Probab.* **20** 177–199. [MR2324525](#) <https://doi.org/10.1007/s10959-007-0058-1>
- [34] Oates, C.J., Girolami, M. and Chopin, N. (2017). Control functionals for Monte Carlo integration. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 695–718. [MR3641403](#) <https://doi.org/10.1111/rssb.12185>

- [35] Rakhlin, A., Sridharan, K. and Tsybakov, A.B. (2017). Empirical entropy, minimax regret and minimax risk. *Bernoulli* **23** 789–824. [MR3606751](#) <https://doi.org/10.3150/14-BEJ679>
- [36] Robert, C.P. and Casella, G. (1999). *Monte Carlo Statistical Methods*. *Springer Texts in Statistics*. New York: Springer. [MR1707311](#) <https://doi.org/10.1007/978-1-4757-3071-5>
- [37] Rogers, L.C.G. (2007). Pathwise stochastic optimal control. *SIAM J. Control Optim.* **46** 1116–1132. [MR2338441](#) <https://doi.org/10.1137/050642885>
- [38] Rubinstein, R.Y. and Kroese, D.P. (2017). *Simulation and the Monte Carlo Method*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. Third edition [of MR0624270]. [MR3617204](#)
- [39] van de Geer, S. (2000). *Empirical Processes in M-Estimation*. Cambridge.
- [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>
- [41] 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>
- [42] Wong, W.H. and Shen, X. (1995). Probability inequalities for likelihood ratios and convergence rates of sieve MLEs. *Ann. Statist.* **23** 339–362. [MR1332570](#) <https://doi.org/10.1214/aos/1176324524>
- [43] Yang, Y. and Barron, A. (1999). Information-theoretic determination of minimax rates of convergence. *Ann. Statist.* **27** 1564–1599. [MR1742500](#) <https://doi.org/10.1214/aos/1017939142>
- [44] Zhivotovskiy, N. (2017). Optimal learning via local entropies and sample compression. In *Proceedings of the 2017 Conference on Learning Theory* **65** 2023–2065. PMLR.

Defective Galton-Watson processes in a varying environment

GÖTZ KERSTING¹ and CARMEN MINUESA^{2,3}

¹*Institut für Mathematik, Goethe Universität, Robert-Mayer-Strasse, 6-10, 60325 Frankfurt am Main, Germany.
 E-mail: kersting@math.uni-frankfurt.de*

²*Department of Mathematics, Autonomous University of Madrid, Ciudad Universitaria de Cantoblanco, s/n, 28049 Madrid, Spain.*

E-mail: carmen.minuesa@uam.es,  0000-0002-8858-3145

³*Department of Mathematics, University of Extremadura, Avda. de Elvas s/n, 06006 Badajoz, Spain.
 E-mail: cminuesa@unex.es*

We study an extension of the so-called defective Galton-Watson processes obtained by allowing the offspring distribution to change over the generations. Thus, in these processes, the individuals reproduce independently of the others and in accordance to some possibly defective offspring distribution depending on the generation. Moreover, the defect $1 - f_n(1)$ of the offspring distribution at generation n represents the probability that the process hits an absorbing state Δ at that generation. We focus on the asymptotic behaviour of these processes. We establish the almost sure convergence of the process to a random variable with values in $\mathbb{N}_0 \cup \{\Delta\}$ and we provide two characterisations of the duality extinction-absorption at Δ . We also state some results on the absorption time and the properties of the process conditioned upon its non-absorption, some of which require us to introduce the notion of defective branching trees in varying environment.

Keywords: Branching process; varying environment; defective distribution; absorption; family tree

References

- [1] Agresti, A. (1975). On the extinction times of varying and random environment branching processes. *J. Appl. Probab.* **12** 39–46. [MR0365733](#) <https://doi.org/10.2307/3212405>
- [2] Braunstains, P. and Hautphenne, S. (2019). Extinction in lower Hessenberg branching processes with countably many types. *Ann. Appl. Probab.* **29** 2782–2818. [MR4019875](#) <https://doi.org/10.1214/19-AAP1464>
- [3] D’Souza, J.C. and Biggins, J.D. (1992). The supercritical Galton–Watson process in varying environments. *Stochastic Process. Appl.* **42** 39–47. [MR1172506](#) [https://doi.org/10.1016/0304-4149\(92\)90025-L](https://doi.org/10.1016/0304-4149(92)90025-L)
- [4] Foster, J.H. and Goettge, R.T. (1976). The rates of growth of the Galton–Watson process in varying environment. *J. Appl. Probab.* **13** 144–147. [MR0394910](#) <https://doi.org/10.2307/3212675>
- [5] Fujimagari, T. (1980). On the extinction time distribution of a branching process in varying environments. *Adv. in Appl. Probab.* **12** 350–366. [MR0569432](#) <https://doi.org/10.2307/1426601>
- [6] Geiger, J. (1999). Elementary new proofs of classical limit theorems for Galton–Watson processes. *J. Appl. Probab.* **36** 301–309. [MR1724856](#) <https://doi.org/10.1239/jap/1032374454>
- [7] Goettge, R.T. (1976). Limit theorems for the supercritical Galton–Watson process in varying environments. *Math. Biosci.* **28** 171–190. [MR0436358](#) [https://doi.org/10.1016/0025-5564\(76\)90100-0](https://doi.org/10.1016/0025-5564(76)90100-0)
- [8] Hsu, P.L. and Robbins, H. (1947). Complete convergence and the law of large numbers. *Proc. Natl. Acad. Sci. USA* **33** 25–31. [MR0019852](#) <https://doi.org/10.1073/pnas.33.2.25>
- [9] Hu, Y., Hu, W. and Yin, Y. (2011). The extinction of a branching process in a varying or random environment. In *Nonlinear Mathematics for Uncertainty and Its Applications. Advances in Intelligent and Soft Computing* (S. Li, X. Wang, Y. Okazaki, J. Kawabe, T. Murofushi and L. Guan, eds.) **100** 309–315. Springer, Berlin.
- [10] Ispány, M. (2016). Some asymptotic results for strongly critical branching processes with immigration in varying environment. In *Branching Processes and Their Applications* (I. del Puerto, M. González, C. Gutiérrez

- rez, R. Martínez, C. Minuesa, M. Molina, M. Mota and A. Ramos, eds.). *Lect. Notes Stat.* **219** 77–95. Springer Cham. [MR3587996](#)
- [11] Iwasa, Y., Michor, F. and Nowak, M.A. (2004). Evolutionary dynamics of invasion and escape. *J. Theoret. Biol.* **226** 205–214. [MR2069303](#) <https://doi.org/10.1016/j.jtbi.2003.08.014>
 - [12] Jagers, P. (1974). Galton–Watson processes in varying environments. *J. Appl. Probab.* **11** 174–178. [MR0368197](#) <https://doi.org/10.2307/3212594>
 - [13] Karlin, S. and Tavaré, S. (1982). Linear birth and death processes with killing. *J. Appl. Probab.* **19** 477–487. [MR0664833](#) <https://doi.org/10.1017/s0021900200036986>
 - [14] Keiding, N. and Nielsen, J.E. (1975). Branching processes with varying and random geometric offspring distributions. *J. Appl. Probab.* **12** 135–141. [MR0368198](#) <https://doi.org/10.2307/3212415>
 - [15] Kersting, G. (2020). A unifying approach to branching processes in a varying environment. *J. Appl. Probab.* **57** 196–220. [MR4094390](#) <https://doi.org/10.1017/jpr.2019.84>
 - [16] Kersting, G. and Vatutin, V. (2017). *Discrete Time Branching Processes in Random Environment*. ISTE Ltd and John Wiley and Sons, Inc.
 - [17] Lindvall, T. (1974). Almost sure convergence of branching processes in varying and random environments. *Ann. Probab.* **2** 344–346. [MR0378130](#) <https://doi.org/10.1214/aop/1176996717>
 - [18] MacPhee, I.M. and Schuh, H.-J. (1983). A Galton–Watson branching process in varying environments with essentially constant offspring means and two rates of growth. *Aust. J. Stat.* **25** 329–338. [MR0725212](#)
 - [19] Robertson, A. (1978). The time of detection of recessive visible genes in small populations. *Genetics Research* **31** 255–264.
 - [20] Sagitov, S. and Minuesa, C. (2017). Defective Galton–Watson processes. *Stoch. Models* **33** 451–472. [MR3765109](#) <https://doi.org/10.1080/15326349.2017.1349614>
 - [21] Yu, J. and Pei, J. (2009). Extinction of branching processes in varying environments. *Statist. Probab. Lett.* **79** 1872–1877. [MR2560486](#) <https://doi.org/10.1016/j.spl.2009.05.023>

A note on the phase transition for independent alignment percolation

MARCELO HILÁRIO¹ and DANIEL UNGARETTI^{1,2,*}

¹ICEx, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.

E-mail: mhilario@mat.ufmg.com

²IME, Universidade de São Paulo, São Paulo, Brazil.

E-mail: *danielungaretti@gmail.com

We study the independent alignment percolation model on $\mathbb{Z}^d$ introduced by Beaton, Grimmett and Holmes. It is a model for random intersecting line segments defined as follows. First the sites of $\mathbb{Z}^d$ are independently declared occupied with probability p and vacant otherwise. Conditional on the configuration of occupied vertices, consider the set of all line segments that are parallel to the coordinate axis, whose extremes are occupied vertices and that do not traverse any other occupied vertex. Declare independently the segments on this set open with probability λ and closed otherwise. All the edges that lie on open segments are also declared open giving rise to a bond percolation model in $\mathbb{Z}^d$. We show that for any $d \geq 2$ and $p \in (0, 1]$ the critical value for λ satisfies $\lambda_c(p) < 1$ completing the proof that the phase transition is non-trivial over the whole interval $(0, 1]$. We also show that the critical curve $p \mapsto \lambda_c(p)$ is continuous at $p = 1$.

Keywords: Percolation; renormalization; phase transition

References

- [1] Baldasso, R. and Teixeira, A. (2018). How can a clairvoyant particle escape the exclusion process? *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 2177–2202. [MR3865670](#) <https://doi.org/10.1214/17-AIHP868>
- [2] Beaton, N.R., Grimmett, G.R. and Holmes, M. (2021). Alignment percolation. *Math. Phys. Anal. Geom.* **24** Paper No. 3. [MR4219398](#) <https://doi.org/10.1007/s11040-021-09373-7>
- [3] Blondel, O., Hilário, M.R. and Teixeira, A. (2020). Random walks on dynamical random environments with nonuniform mixing. *Ann. Probab.* **48** 2014–2051. [MR4124532](#) <https://doi.org/10.1214/19-AOP1414>
- [4] Burton, R.M. and Keane, M. (1989). Density and uniqueness in percolation. *Comm. Math. Phys.* **121** 501–505. [MR0990777](#)
- [5] Grimmett, G. (1999). *Percolation. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **321**. Berlin: Springer. [MR1707339](#) <https://doi.org/10.1007/978-3-662-03981-6>
- [6] Grimmett, G. (2006). *The Random-Cluster Model. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **333**. Berlin: Springer. [MR2243761](#) <https://doi.org/10.1007/978-3-540-32891-9>
- [7] Hilário, M.R. and Sidoravicius, V. (2019). Bernoulli line percolation. *Stochastic Process. Appl.* **129** 5037–5072. [MR4025699](#) <https://doi.org/10.1016/j.spa.2019.01.002>
- [8] Liggett, T.M. (2005). *Interacting Particle Systems. Classics in Mathematics*. Berlin: Springer. Reprint of the 1985 original. [MR2108619](#) <https://doi.org/10.1007/b138374>
- [9] Sznitman, A.-S. (2010). Vacant set of random interlacements and percolation. *Ann. of Math. (2)* **171** 2039–2087. [MR2680403](#) <https://doi.org/10.4007/annals.2010.171.2039>
- [10] Tykesson, J. and Windisch, D. (2012). Percolation in the vacant set of Poisson cylinders. *Probab. Theory Related Fields* **154** 165–191. [MR2981421](#) <https://doi.org/10.1007/s00440-011-0366-3>

Spectral equivalence of Gaussian random functions: Operator approach

ALEXANDER NAZAROV^{1,2} and YAKOV NIKITIN^{2,3,†}

¹*St.Petersburg Department of Steklov Mathematical Institute of Russian Academy of Science, St.Petersburg, Russia.*

E-mail: al.il.nazarov@gmail.com

²*Mathematics and Mechanics Faculty, St.Petersburg State University, St.Petersburg, Russia*

³*National Research University — Higher School of Economics, St.Petersburg, Russia.*

E-mail: y.nikitin@spbu.ru

We introduce a new approach to the spectral equivalence of Gaussian processes and fields, based on the methods of operator theory in Hilbert space. Besides several new results including identities in law of quadratic norms for integrated and multiply integrated Gaussian random functions we give an application to goodness-of-fit testing.

Keywords: Gaussian random functions; identity in law; spectral equivalence; tensor product; Brownian sheet

References

- [1] Ai, X. and Li, W.V. (2014). Karhunen-Loeve expansions for the m -th order detrended Brownian motion. *Sci. China Math.* **57** 2043–2052. [MR3247532](#) <https://doi.org/10.1007/s11425-014-4873-4>
- [2] Ai, X., Li, W.V. and Liu, G. (2012). Karhunen-Loeve expansions for the detrended Brownian motion. *Statist. Probab. Lett.* **82** 1235–1241. [MR2929770](#) <https://doi.org/10.1016/j.spl.2012.03.007>
- [3] Ash, R.B. and Gardner, M.F. (1975). *Topics in Stochastic Processes. Probability and Mathematical Statistics, Vol. 27*. New York-London: Academic Press [Harcourt Brace Jovanovich, Publishers]. [MR0448463](#)
- [4] Bahadur, R.R. (1960). Stochastic comparison of tests. *Ann. Math. Stat.* **31** 276–295. [MR0116413](#) <https://doi.org/10.1214/aoms/1177705894>
- [5] Beghin, L., Nikitin, Ya. and Orsingher, E. (2003). Exact small ball constants for some Gaussian processes under L^2 -norm. *J. Math. Sci.* **128** 2493–2502.
- [6] Birman, M.Sh. and Solomjak, M.Z. (2010). *Spectral Theory of Selfadjoint Operators in Hilbert Space*, 2nd ed., revised and extended. Lan', St.Petersburg (Russian); English transl. of the 1st ed.: *Mathematics and Its Applications (Soviet Series)* **5**. Dordrecht: Kluwer.
- [7] Cairoli, R. and Walsh, J.B. (1975). Stochastic integrals in the plane. *Acta Math.* **134** 111–183. [MR0420845](#) <https://doi.org/10.1007/BF02392100>
- [8] Deheuvels, P. (2006). Karhunen-Loève expansions of mean-centered Wiener processes. In *High Dimensional Probability* (E. Giné, V. Koltchinskii, W. Li, J. Zinn, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **51** 62–76. Beachwood, OH: IMS. [MR2387761](#) <https://doi.org/10.1214/074921706000000761>
- [9] Deheuvels, P., Peccati, G. and Yor, M. (2006). On quadratic functionals of the Brownian sheet and related processes. *Stochastic Process. Appl.* **116** 493–538. [MR2199561](#) <https://doi.org/10.1016/j.spa.2005.10.004>
- [10] Donati-Martin, C. and Yor, M. (1991). Fubini's theorem for double Wiener integrals and the variance of the Brownian path. *Ann. Inst. Henri Poincaré Probab. Stat.* **27** 181–200. [MR1118933](#)
- [11] Dozzi, M. (1989). *Stochastic Processes with a Multidimensional Parameter*. *Pitman Research Notes in Mathematics Series* **194**. New York: Longman Scientific & Technical, Harlow; copublished in the United States with Wiley. [MR0991563](#)
- [12] Durbin, J. (1970). Asymptotic distributions of some statistics based on the bivariate sample distribution function. In *Nonparametric Techniques in Statistical Inference* (*Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969*) 435–451. London: Cambridge Univ. Press. [MR0278472](#)

- [13] Gosselin, P. and Wurzbacher, T. (1997). An Itô type isometry for loops in $\mathbf{R}^d$ via the Brownian bridge. In *Séminaire de Probabilités, XXXI* (J. Azéma, M. Émery and M. Yor, eds.). *Lecture Notes in Math.* **1655** 225–231. Berlin: Springer. [MR1478731](#) <https://doi.org/10.1007/BFb0119307>
- [14] Jeulin, T. and Yor, M. (1979). Inégalité de Hardy, semimartingales, et faux-amis. In *Séminaire de Probabilités, XIII* (Univ. Strasbourg, Strasbourg, 1977/78). *Lecture Notes in Math.* **721** 332–359. Berlin: Springer. [MR0544805](#)
- [15] Karol', A., Nazarov, A. and Nikitin, Y. (2008). Small ball probabilities for Gaussian random fields and tensor products of compact operators. *Trans. Amer. Math. Soc.* **360** 1443–1474. [MR2357702](#) <https://doi.org/10.1090/S0002-9947-07-04233-X>
- [16] Korolyuk, V.S. and Borovskikh, Yu.V. (2013). *Theory of U-Statistics*. NY: Springer.
- [17] Krivjakova, È.N., Martynov, G.V. and Tjurin, Ju.N. (1977). The distribution of the ω^2 statistic in the multi-dimensional case. *Teor. Veroyatn. Primen.* **22** 415–420. [MR0440785](#)
- [18] Lachal, A. (2002). Bridges of certain Wiener integrals. Prediction properties, relation with polynomial interpolation and differential equations. Application to goodness-of-fit testing. In *Limit Theorems in Probability and Statistics, Vol. II* (Balatonlelle, 1999) (I. Berkes, E. Csáki, M. Csörgő, eds.) 273–323. Budapest: János Bolyai Math. Soc. [MR1979998](#)
- [19] Li, W.V. and Shao, Q.-M. (2001). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes: Theory and Methods* (C.R. Rao and D. Shanbhag, eds.). *Handbook of Statist.* **19** 533–597. Amsterdam: North-Holland. [MR1861734](#) [https://doi.org/10.1016/S0169-7161\(01\)19019-X](https://doi.org/10.1016/S0169-7161(01)19019-X)
- [20] Lifshits, M.A. (1999). Asymptotic behavior of small ball probabilities. In *Probab. Theory and Math. Stat. Proc. VII International Vilnius Conference* (1998) (B. Grigelionis et al., eds.) 453–468. VSP/TEV.
- [21] Lifshits, M. (2012). *Lectures on Gaussian Processes. SpringerBriefs in Mathematics*. Heidelberg: Springer. [MR3024389](#) <https://doi.org/10.1007/978-3-642-24939-6>
- [22] Loève, M. (1977). *Probability Theory. I*, 4th ed. *Graduate Texts in Mathematics*, Vol. 45. New York: Springer. [MR0651017](#)
- [23] MacNeill, I.B. (1978). Properties of sequences of partial sums of polynomial regression residuals with applications to tests for change of regression at unknown times. *Ann. Statist.* **6** 422–433. [MR0474645](#)
- [24] Martynov, G.V. (1978). *Omega-square criteria*. Moscow: “Nauka”. (Russian) [MR0527912](#)
- [25] Nazarov, A.I. (2009). Exact L_2 -small ball asymptotics of Gaussian processes and the spectrum of boundary-value problems. *J. Theoret. Probab.* **22** 640–665. [MR2530107](#) <https://doi.org/10.1007/s10959-008-0173-7>
- [26] Nikitin, Y. (1995). *Asymptotic Efficiency of Nonparametric Tests*. Cambridge: Cambridge Univ. Press. [MR1335235](#) <https://doi.org/10.1017/CBO9780511530081>
- [27] Peccati, G. (2003). Explicit formulae for time-space Brownian chaos. *Bernoulli* **9** 25–48. [MR1963671](#) <https://doi.org/10.3150/bj/1068129009>
- [28] Peccati, G. and Yor, M. (2006). Identities in law between quadratic functionals of bivariate Gaussian processes, through Fubini theorems and symmetric projections. In *Approximation and Probability. Banach Center Publ.* **72** 235–250. Warsaw: Polish Acad. Sci. Inst. Math. [MR2325748](#) <https://doi.org/10.4064/bc72-0-15>
- [29] Petrova, Yu.P. (2017). Spectral asymptotics for problems with integral constraints. *Mat. Zametki* **102** 405–414. [MR3691705](#) <https://doi.org/10.4213/mzm11344>
- [30] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Berlin: Springer. [MR1725357](#) <https://doi.org/10.1007/978-3-662-06400-9>
- [31] Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.* **23** 470–472. [MR0049525](#) <https://doi.org/10.1214/aoms/1177729394>
- [32] Shorack, G.R. and Wellner, J.A. (2009). *Empirical Processes with Applications to Statistics. Classics in Applied Mathematics* **59**. Philadelphia, PA: SIAM. [MR3396731](#) <https://doi.org/10.1137/1.9780898719017.ch1>
- [33] Voinov, V., Nikulin, M. and Balakrishnan, N. (2013). *Chi-Squared Goodness of Fit Tests with Applications*. Amsterdam: Elsevier.

Posterior probabilities: Nonmonotonicity, asymptotic rates, log-concavity, and Turán’s inequality

SERGIU HART¹ and YOSEF RINOTT²

¹*The Hebrew University of Jerusalem, Federmann Center for the Study of Rationality, Einstein Institute of Mathematics, and Department of Economics.*

E-mail: hart@huji.ac.il

²*The Hebrew University of Jerusalem, Federmann Center for the Study of Rationality, and Department of Statistics.*

E-mail: yosef.rinott@mail.huji.ac.il

In the standard Bayesian framework data are assumed to be generated by a distribution parametrized by θ in a parameter space Θ , over which a prior distribution π is given. A Bayesian statistician quantifies the belief that the true parameter is θ_0 in Θ by its posterior probability given the observed data. We investigate the behavior of the posterior belief in θ_0 when the data are generated under some parameter θ_1 , which may or may not be the same as θ_0 . Starting from stochastic orders, specifically, likelihood ratio dominance, that obtain for resulting distributions of posteriors, we consider monotonicity properties of the posterior probabilities as a function of the sample size when data arrive sequentially. While the θ_0 -posterior is monotonically increasing (i.e., it is a submartingale) when the data are generated under that same θ_0 , it need not be monotonically decreasing in general, not even in terms of its overall expectation, when the data are generated under a different θ_1 . In fact, it may keep going up and down many times, even in simple cases such as iid coin tosses. We obtain precise asymptotic rates when the data come from the wide class of exponential families of distributions; these rates imply in particular that the expectation of the θ_0 -posterior under $\theta_1 \neq \theta_0$ is eventually strictly decreasing. Finally, we show that in a number of interesting cases this expectation is a log-concave function of the sample size, and thus unimodal. In the Bernoulli case we obtain this result by developing an inequality that is related to Turán’s inequality for Legendre polynomials.

Keywords: Bayesian analysis; stochastic and likelihood ratio orders; sequential observations; expected posteriors; unimodality; Legendre polynomials; exponential families

References

- [1] Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards Applied Mathematics Series, No. 55. 10th Printing. Washington, D. C.: U. S. Government Printing Office.
- [2] Arfken, G. (1985). *Mathematical Methods for Physicists*, 3rd ed. New York: Academic Press.
- [3] Barron, A.R. (1998). Information-theoretic characterization of Bayes performance and the choice of priors in parametric and nonparametric problems. In *Bayesian Statistics 6*. (J.M. Bernardo, J.O. Berger, and A. Smith (Eds.)) 27–52. New York: Oxford Univ. Press. [MR1723492](#)
- [4] Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR0804611](#) <https://doi.org/10.1007/978-1-4757-4286-2>
- [5] Berk, R.H. (1972). Consistency and asymptotic normality of MLE’s for exponential models. *Ann. Math. Stat.* **43** 193–204. [MR0298810](#)
- [6] Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* **35** 99–109. [MR0010358](#)
- [7] Billingsley, P. (2013). *Convergence of Probability Measures*. New York: Wiley.

- [8] Chernoff, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Stat.* **23** 493–507. [MR0057518](#) <https://doi.org/10.1214/aoms/1177729330>
- [9] Ferguson, T.S. (1967). *Mathematical Statistics: A Decision Theoretic Approach*. Probability and Mathematical Statistics, Vol. 1. New York: Academic Press. [MR0215390](#)
- [10] Ferguson, T.S. (1996). *A Course in Large Sample Theory*. Texts in Statistical Science Series. London: CRC Press. [MR1699953](#) <https://doi.org/10.1007/978-1-4899-4549-5>
- [11] Francetich, A. and Kreps, D. (2014). Bayesian inference does not lead you astray... on average. *Econom. Lett.* **125** 444–446. [MR3281735](#) <https://doi.org/10.1016/j.econlet.2014.10.022>
- [12] Grünwald, P.D. and Halpern, J.Y. (2014). When ignorance is bliss. Preprint. Available at [arXiv:1407.7188](#).
- [13] Hart, S. and Rinott, Y. (2020). Posterior probabilities: Dominance and optimism. *Econom. Lett.* **194** 109352, 3. [MR4121083](#) <https://doi.org/10.1016/j.econlet.2020.109352>
- [14] Jewitt, I. (1988). Justifying the first-order approach to principal-agent problems. *Econometrica* **56** 1177–1190. [MR0964151](#) <https://doi.org/10.2307/1911363>
- [15] Johnson, N.L., Kemp, A.W. and Kotz, S. (2005). *Univariate Discrete Distributions*, 3rd ed. Wiley Series in Probability and Statistics **444**. Hoboken, NJ: Wiley. [MR2163227](#) <https://doi.org/10.1002/0471715816>
- [16] Karlin, S. (1968). *Total Positivity*. Stanford: Stanford Univ. Press.
- [17] Mailath, G.J. and Samuelson, L. (2006). *Repeated Games and Reputations*. London: Oxford Univ. Press.
- [18] Miller, J.W. (2018). A detailed treatment of Doob’s theorem. Available at [arXiv:1801.03122v1](#) [math.ST].
- [19] Nielsen, F. (2011). Chernoff information of exponential families. Preprint. Available at [arXiv:1102.2684](#).
- [20] Schervish, M.J. (2012). *Theory of Statistics*. Springer Series in Statistics. New York: Springer.
- [21] Segura, J. (2011). Bounds for ratios of modified Bessel functions and associated Turán-type inequalities. *J. Math. Anal. Appl.* **374** 516–528. [MR2729238](#) <https://doi.org/10.1016/j.jmaa.2010.09.030>
- [22] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer Series in Statistics. New York: Springer. [MR2265633](#) <https://doi.org/10.1007/978-0-387-34675-5>
- [23] Szegő, G. (1948). On an inequality of P. Turán concerning Legendre polynomials. *Bull. Amer. Math. Soc.* **54** 401–405. [MR0023954](#) <https://doi.org/10.1090/S0002-9904-1948-09017-6>
- [24] Szwarc, R. (1998). Positivity of Turán determinants for orthogonal polynomials. In *Harmonic Analysis and Hypergroups (Delhi, 1995)*. Trends Math. 165–182. Boston, MA: Birkhäuser. [MR1616253](#)
- [25] 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>

Convergence of jump processes with stochastic intensity to Brownian motion with inert drift

CLAYTON BARNES

Technion-Israel's Institute of Technology.
E-mail: cbarnes@campus.technion.ac.il

Consider a random walker on the nonnegative lattice, moving in continuous time, whose positive transition intensity is proportional to the time the walker spends at the origin. In this way, the walker is a jump process with a stochastic and adapted jump intensity. We show that, upon Brownian scaling, the sequence of such processes converges to Brownian motion with inert drift (BMID). BMID was introduced by Frank Knight in 2001 and generalized by White in 2007. This confirms a conjecture of Burdzy and White in 2008 in the one-dimensional setting.

Keywords: Brownian motion; discrete approximation; random walk; local time

References

- [1] Barnes, C., Burdzy, K. and Gauthier, C.-E. (2019). Billiards with Markovian reflection laws. *Electron. J. Probab.* **24** Paper No. 147, 32. [MR4049083](#) <https://doi.org/10.1214/19-ejp398>
- [2] Barnes, C.L. (2020). Hydrodynamic limit and propagation of chaos for Brownian particles reflecting from a Newtonian barrier. *Ann. Appl. Probab.* **30** 1582–1613. [MR4132635](#) <https://doi.org/10.1214/19-AAP1536>
- [3] Bass, R.F., Burdzy, K., Chen, Z.-Q. and Hairer, M. (2010). Stationary distributions for diffusions with inert drift. *Probab. Theory Related Fields* **146** 1–47. [MR2550357](#) <https://doi.org/10.1007/s00440-008-0182-6>
- [4] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. [MR0233396](#)
- [5] Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. *Springer Series in Statistics*. New York-Berlin: Springer. [MR0636252](#)
- [6] Burdzy, K. and White, D. (2008). Markov processes with product-form stationary distribution. *Electron. Commun. Probab.* **13** 614–627. [MR2461535](#) <https://doi.org/10.1214/ECP.v13-1428>
- [7] 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>
- [8] Gauthier, C.-E. (2018). Central limit theorem for one and two dimensional self-repelling diffusions. *ALEA Lat. Am. J. Probab. Math. Stat.* **15** 691–703. [MR3808901](#) <https://doi.org/10.30757/alea.v15-26>
- [9] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. [MR1943877](#) <https://doi.org/10.1007/978-3-662-05265-5>
- [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] Knight, F.B. (2001). On the path of an inert object impinged on one side by a Brownian particle. *Probab. Theory Related Fields* **121** 577–598. [MR1872429](#) <https://doi.org/10.1007/s004400100160>
- [12] Mörters, P. and Peres, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge: Cambridge Univ. Press. [MR2604525](#) <https://doi.org/10.1017/CBO9780511750489>
- [13] White, D. (2007). Processes with inert drift. *Electron. J. Probab.* **12** 1509–1546. [MR2365876](#) <https://doi.org/10.1214/EJP.v12-465>