

# BERNOULLI

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

Volume Thirty Number Four November 2024 ISSN: 1350-7265

## CONTENTS

PALMOWSKI, Z., SARIEV, H. and SAVOV, M.	2547
Moments of exponential functionals of Lévy processes on a deterministic horizon – identities and explicit expressions	
CHENG, X. and XIE, Y.	2572
Kernel two-sample tests for manifold data	
SCHMIDT, S.K.	2598
Detecting changes in the trend function of heteroscedastic time series	
SEI, T. and YANO, K.	2623
Minimum information dependence modeling	
NAULET, Z., ROUSSEAU, J. and CARON, F.	2644
Asymptotic analysis of statistical estimators related to MultiGraphex processes under misspecification	
HAZELTON, M., MCVEAGH, M., TUFFLEY, C. and VAN BRUNT, B.	2676
Some rapidly mixing hit-and-run samplers for latent counts in linear inverse problems	
KARGIN, V.	2700
On the joint distribution of the area and the number of peaks for Bernoulli excursions	
UTPALA, S. and SRIPERUMBUDUR, B.K.	2721
Shrinkage estimation of higher-order Bochner integrals	
AMEIJEIRAS-ALONSO, J. and GIJBELS, I.	2747
Smoothed circulars: Nonparametric estimation of circular cumulative distribution functions and circulars	
KURISU, D. and MATSUDA, Y.	2770
Local polynomial trend regression for spatial data on $\mathbb{R}^d$	
HE, S., HUANG, J.Z. and HE, K.	2795
Penalized spline estimation of principal components for sparse functional data: Rates of convergence	
KENDALL, W.S., MAJKA, M.B. and MIJATOVIĆ, A.	2821
Optimal Markovian coupling for finite activity Lévy processes	
HUNDRIESER, S., KLATT, M., MUNK, A. and STAUDT, T.	2846
A unifying approach to distributional limits for empirical optimal transport	
BUTLER, R.W.	2878
Approximate double-transform inversion when time is one of the variables	
DUMITRIU, I. and ZHU, Y.	2904
Extreme singular values of inhomogeneous sparse random rectangular matrices	
FRITZ, T.	2932
Characterizing the asymptotic and catalytic stochastic orders on topological abelian groups	

*(continued)*

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

# BERNOULLI

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

Volume Thirty Number Four November 2024 ISSN: 1350-7265

## CONTENTS

*(continued)*

STÉPHANOVITCH, A., TANIELIAN, U., CADRE, B., KLUTCHNIKOFF, N. and BIAU, G. Optimal 1-Wasserstein distance for WGANs	2955
LI, M.T.C., MARZOUK, Y. and ZAHM, O. Principal feature detection via $\phi$ -Sobolev inequalities	2979
BARBER, R.F., CANDÈS, E.J., RAMDAS, A. and TIBSHIRANI, R.J. De Finetti's theorem and related results for infinite weighted exchangeable sequences	3004
SHI, Z., BALASUBRAMANIAN, K. and POLONIK, W. A flexible approach for normal approximation of geometric and topological statistics	3029
YU, M., REN, Z. and ZHOU, W.-X. Gaussian differentially private robust mean estimation and inference	3059
KULCZYCKI, T., KULYK, O. and RYZNAR, M. Drift reduction method for SDEs driven by heterogeneous singular Lévy noise	3089
HU, Y., WANG, X., XIA, P. and ZHENG, J. Moment asymptotics for super-Brownian motions	3119
YANG, J. On the asymptotic behavior of a finite section of the optimal causal filter	3137
LI, W., WANG, Q. and YAO, J. Distance correlation test for high-dimensional independence	3165
CATTANEO, M.D., CHANDAK, R., JANSSON, M. and MA, X. Boundary adaptive local polynomial conditional density estimators	3193
LIU, Q. and PRIVAULT, N. Normal approximation of subgraph counts in the random-connection model	3224
SUN, D., GUO, Y., LI, Y., TU, W. and SUN, J. A robust approach for regression analysis of panel count data with time-varying covariates	3251
LEUNG, D. and SHAO, Q.-M. Nonuniform Berry-Esseen bounds for studentized U-statistics	3276
HUANG, X. and WANG, F.-Y. Regularities and exponential ergodicity in entropy for SDEs driven by distribution dependent noise	3303
CHANG, Y., DU, H. and LI, X. Percolation threshold for metric graph loop soup	3324
CHATLA, S.B. and LIU, R. Inverse regression for spatially distributed functional data	3334
KREBS, J. On the Bahadur representation of sample quantiles for score functionals	3356
<b>Erratum</b>	
IACOBELLI, G., RIBEIRO, R., VALLE, G. and ZUAZNÁBAR, L. Tree builder random walk: Recurrence, transience and ballisticity	3378

BERNOULLI

Volume 30 Number 4 November 2024 Pages 2547–3387

ISI/BS

Volume 30 Number 4 November 2024 ISSN 1350-7265

# BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

## Aims and Scope

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

## Bernoulli Society for Mathematical Statistics and Probability

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

## Meetings: <https://www.bernoullisociety.org/meetings>

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

## Executive Committee

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

---

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

©2024 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 2024 Bernoulli consists of 4 issues published in February, May, August and November.



**Bernoulli Society**  
for Mathematical Statistics  
and Probability

# Moments of exponential functionals of Lévy processes on a deterministic horizon – identities and explicit expressions

ZBIGNIEW PALMOWSKI<sup>1,a</sup>, HRISTO SARIEV<sup>2,3,b</sup> and MLADEN SAVOV<sup>3,2,c</sup>

<sup>1</sup>*Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, [zbigniew.palmowski@pwr.edu.pl](mailto:zbigniew.palmowski@pwr.edu.pl)*

<sup>2</sup>*Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Georgi Bonchev str., bl.8, Sofia 1113, Bulgaria, [h.sariev@math.bas.bg](mailto:h.sariev@math.bas.bg)*

<sup>3</sup>*Faculty of Mathematics and Informatics, Sofia University “St. Kliment Ohridski”, 5 James Bourchier blvd., Sofia 1164, Bulgaria, [msavov@fmi.uni-sofia.bg](mailto:msavov@fmi.uni-sofia.bg)*

In this work, we consider moments of exponential functionals of Lévy processes on a deterministic horizon. We derive two convolutional identities regarding these moments. The first one relates the complex moments of the exponential functional of a general Lévy process up to a deterministic time to those of the dual Lévy process. The second convolutional identity links the complex moments of the exponential functional of a Lévy process, which is not a compound Poisson process, to those of the exponential functionals of its ascending/descending ladder heights on a random horizon determined by the respective local times. As a consequence, we derive a universal expression for the half-negative moment of the exponential functional of any symmetric Lévy process, which resembles in its universality the passage time of symmetric random walks. The  $(n - 1/2)^{th}$ ,  $n \geq 0$  moments are also discussed. On the other hand, under extremely mild conditions, we obtain a series expansion for the complex moments (including those with negative real part) of the exponential functionals of subordinators. This significantly extends previous results and offers neat expressions for the negative real moments. In a special case, it turns out that the Riemann zeta function is the minus first moment of the exponential functional of the Gamma subordinator indexed in time.

*Keywords:* Bernstein function; exponential functional; Mellin transform

## References

- Alili, L., Matsumoto, H. and Shiraiishi, T. (2001). On a triplet of exponential Brownian functionals. In *Séminaire de Probabilités, XXXV. Lecture Notes in Math.* **1755** 396–415. Berlin: Springer. MR1837300 [https://doi.org/10.1007/978-3-540-44671-2\\_27](https://doi.org/10.1007/978-3-540-44671-2_27)
- Barker, A. and Savov, M. (2021). Bivariate Bernstein-gamma functions and moments of exponential functionals of subordinators. *Stochastic Process. Appl.* **131** 454–497. MR4165648 <https://doi.org/10.1016/j.spa.2020.09.017>
- 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>
- Behme, A., Lindner, A. and Reker, J. (2021). On the law of killed exponential functionals. *Electron. J. Probab.* **26** Paper No. 60, 35 pp. MR4254802 <https://doi.org/10.1214/21-ejp616>
- Bertoin, J. (1996). *Lévy Processes. Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press. MR1406564
- Bertoin, J. (2019). Ergodic aspects of some Ornstein-Uhlenbeck type processes related to Lévy processes. *Stochastic Process. Appl.* **129** 1443–1454. MR3926562 <https://doi.org/10.1016/j.spa.2018.05.007>
- Bertoin, J. and Yor, M. (2005). Exponential functionals of Lévy processes. *Probab. Surv.* **2** 191–212. MR2178044 <https://doi.org/10.1214/154957805100000122>

- Brignone, R. (2022). Moments of integrated exponential Lévy processes and applications to Asian options pricing. *Quant. Finance* **22** 1717–1729. [MR4464505](#) <https://doi.org/10.1080/14697688.2022.2070533>
- Fitzsimmons, P.J. and Pitman, J. (1999). Kac's moment formula and the Feynman-Kac formula for additive functionals of a Markov process. *Stochastic Process. Appl.* **79** 117–134. [MR1670526](#) [https://doi.org/10.1016/S0304-4149\(98\)00081-7](https://doi.org/10.1016/S0304-4149(98)00081-7)
- Grzywny, T., Leżaj, Ł. and Trojan, B. (2023). Transition densities of subordinators of positive order. *J. Inst. Math. Jussieu* **22** 1119–1179. [MR4574170](#) <https://doi.org/10.1017/S1474748021000360>
- Hackmann, D. and Kuznetsov, A. (2014). Asian options and meromorphic Lévy processes. *Finance Stoch.* **18** 825–844. [MR3255753](#) <https://doi.org/10.1007/s00780-014-0237-8>
- Hirsch, F. and Yor, M. (2013). On the Mellin transforms of the perpetuity and the remainder variables associated to a subordinator. *Bernoulli* **19** 1350–1377. [MR3102555](#) <https://doi.org/10.3150/12-BEJSP01>
- Kuznetsov, A., Pardo, J.C. and Savov, M. (2012). Distributional properties of exponential functionals of Lévy processes. *Electron. J. Probab.* **17** no. 8, 35 pp. [MR2878787](#) <https://doi.org/10.1214/EJP.v17-1755>
- Matsumoto, H. and Yor, M. (2005). Exponential functionals of Brownian motion. I. Probability laws at fixed time. *Probab. Surv.* **2** 312–347. [MR2203675](#) <https://doi.org/10.1214/154957805100000159>
- Maulik, K. and Zwart, B. (2006). Tail asymptotics for exponential functionals of Lévy processes. *Stochastic Process. Appl.* **116** 156–177. [MR2197972](#) <https://doi.org/10.1016/j.spa.2005.09.002>
- Minchev, M. and Savov, M. (2023). Asymptotics for densities of exponential functionals of subordinators. *Bernoulli* **29** 3307–3333. [MR4632140](#) <https://doi.org/10.3150/23-bej1584>
- Pardo, J.C., Patie, P. and Savov, M. (2012). A Wiener-Hopf type factorization for the exponential functional of Lévy processes. *J. Lond. Math. Soc. (2)* **86** 930–956. [MR3000836](#) <https://doi.org/10.1112/jlms/jds028>
- Patie, P. (2008).  $q$ -invariant functions for some generalizations of the Ornstein-Uhlenbeck semigroup. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 31–43. [MR2383732](#)
- Patie, P. (2013). Asian options under one-sided Lévy models. *J. Appl. Probab.* **50** 359–373. [MR3102485](#) <https://doi.org/10.1239/jap/1371648946>
- Patie, P. and Savov, M. (2012). Extended factorizations of exponential functionals of Lévy processes. *Electron. J. Probab.* **17** no. 38, 22 pp. [MR2928721](#) <https://doi.org/10.1214/EJP.v17-2057>
- Patie, P. and Savov, M. (2013). Exponential functional of Lévy processes: Generalized Weierstrass products and Wiener-Hopf factorization. *C. R. Math. Acad. Sci. Paris* **351** 393–396. [MR3072167](#) <https://doi.org/10.1016/j.crma.2013.04.023>
- Patie, P. and Savov, M. (2018). Bernstein-gamma functions and exponential functionals of Lévy processes. *Electron. J. Probab.* **23** Paper No. 75, 101 pp. [MR3835481](#) <https://doi.org/10.1214/18-EJP202>
- Patie, P. and Savov, M. (2021). Spectral expansions of non-self-adjoint generalized Laguerre semigroups. *Mem. Amer. Math. Soc.* **272** vii+182. [MR4320772](#) <https://doi.org/10.1090/memo/1336>
- Salminen, P. and Vostrikova, L. (2018). On exponential functionals of processes with independent increments. *Theory Probab. Appl.* **63** 267–291. [MR3796492](#) <https://doi.org/10.1137/s0040585x97f989040>
- Salminen, P. and Vostrikova, L. (2019). On moments of integral exponential functionals of additive processes. *Statist. Probab. Lett.* **146** 139–146. [MR3881609](#) <https://doi.org/10.1016/j.spl.2018.11.011>
- Urbanik, K. (1995). Infinite divisibility of some functionals on stochastic processes. *Probab. Math. Statist.* **15** 493–513. [MR1369817](#)
- Vostrikova, L. (2020). On distributions of exponential functionals of the processes with independent increments. *Mod. Stoch. Theory Appl.* **7** 291–313. [MR4159151](#) <https://doi.org/10.15559/20-vmsta159>
- Yor, M. (1992). On some exponential functionals of Brownian motion. *Adv. in Appl. Probab.* **24** 509–531. [MR1174378](#) <https://doi.org/10.2307/1427477>
- Yor, M. (2001). *Exponential Functionals of Brownian Motion and Related Processes*. Springer Finance. Berlin: Springer. [MR1854494](#) <https://doi.org/10.1007/978-3-642-56634-9>

# Kernel two-sample tests for manifold data

XIUYUAN CHENG<sup>1,a</sup> and YAO XIE<sup>2,b</sup>

<sup>1</sup>*Department of Mathematics, Duke University, Durham, NC, [xiuyuan.cheng@duke.edu](mailto:xiuyuan.cheng@duke.edu)*

<sup>2</sup>*H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA,*

<sup>b</sup>*[yao.xie@isye.gatech.edu](mailto:yao.xie@isye.gatech.edu)*

We present a study of a kernel-based two-sample test statistic related to the Maximum Mean Discrepancy (MMD) in the manifold data setting, assuming that high-dimensional observations are close to a low-dimensional manifold. We characterize the test level and power in relation to the kernel bandwidth, the number of samples, and the intrinsic dimensionality of the manifold. Specifically, when data densities  $p$  and  $q$  are supported on a  $d$ -dimensional sub-manifold  $\mathcal{M}$  embedded in an  $m$ -dimensional space and are Hölder with order  $\beta$  (up to 2) on  $\mathcal{M}$ , we prove a guarantee of the test power for finite sample size  $n$  that exceeds a threshold depending on  $d$ ,  $\beta$ , and  $\Delta_2$  the squared  $L^2$ -divergence between  $p$  and  $q$  on the manifold, and with a properly chosen kernel bandwidth  $\gamma$ . For small density departures, we show that with large  $n$  they can be detected by the kernel test when  $\Delta_2$  is greater than  $n^{-2\beta/(d+4\beta)}$  up to a certain constant and  $\gamma$  scales as  $n^{-1/(d+4\beta)}$ . The analysis extends to cases where the manifold has a boundary and the data samples contain high-dimensional additive noise. Our results indicate that the kernel two-sample test has no curse-of-dimensionality when the data lie on or near a low-dimensional manifold. We validate our theory and the properties of the kernel test for manifold data through a series of numerical experiments.

*Keywords:* Kernel methods; manifold data; Maximum Mean Discrepancy; two-sample test

## References

- Arcones, M.A. and Giné, E. (1992). On the bootstrap of  $U$  and  $V$  statistics. *Ann. Statist.* **20** 655–674. [MR1165586](https://doi.org/10.1214/aos/1176348650)
- Balasubramanian, K., Li, T. and Yuan, M. (2021). On the optimality of kernel-embedding based goodness-of-fit tests. *J. Mach. Learn. Res.* **22** 1–45. [MR4253694](https://doi.org/10.1214/19-AOS1913)
- Belkin, M. and Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15** 1373–1396.
- Belkin, M. and Niyogi, P. (2007). Convergence of Laplacian eigenmaps. In *Advances in Neural Information Processing Systems*. 129–136.
- Bhattacharya, B.B. (2020). Asymptotic distribution and detection thresholds for two-sample tests based on geometric graphs. *Ann. Statist.* **48** 2879–2903. [MR4152627](https://doi.org/10.1214/19-AOS1913)
- Bhuyan, M.H., Bhattacharyya, D.K. and Kalita, J.K. (2013). Network anomaly detection: Methods, systems and tools. *IEEE Commun. Surv. Tutor.* **16** 303–336.
- Bińkowski, M., Sutherland, D.J., Arbel, M. and Gretton, A. (2018). Demystifying MMD GANs. In *International Conference on Learning Representations*.
- Borgwardt, K.M., Gretton, A., Rasch, M.J., Kriegel, H.-P., Schölkopf, B. and Smola, A.J. (2006). Integrating structured biological data by kernel maximum mean discrepancy. *Bioinformatics* **22** e49–e57.
- Brito, M.R., Quiroz, A.J. and Yukich, J.E. (2013). Intrinsic dimension identification via graph-theoretic methods. *J. Multivariate Anal.* **116** 263–277. [MR3049904](https://doi.org/10.1016/j.jmva.2012.12.007)
- Buades, A., Coll, B. and Morel, J.-M. (2005). A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* **2** 60–65. IEEE.
- Calder, J. and García Trillos, N. (2022). Improved spectral convergence rates for graph Laplacians on  $\varepsilon$ -graphs and  $k$ -NN graphs. *Appl. Comput. Harmon. Anal.* **60** 123–175. [MR4393800](https://doi.org/10.1016/j.acha.2022.02.004)



- Cao, Y., Nemirovski, A., Xie, Y., Guigues, V. and Juditsky, A. (2018). Change detection via affine and quadratic detectors. *Electron. J. Stat.* **12** 1–57. [MR3743736 https://doi.org/10.1214/17-EJS1373](https://doi.org/10.1214/17-EJS1373)
- Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly detection: A survey. *ACM Comput. Surv.* **41**.
- Chandola, V., Banerjee, A. and Kumar, V. (2010). Anomaly detection for discrete sequences: A survey. *IEEE Trans. Knowl. Data Eng.* **24** 823–839.
- Chen, H. and Friedman, J.H. (2017). A new graph-based two-sample test for multivariate and object data. *J. Amer. Statist. Assoc.* **112** 397–409. [MR3646580 https://doi.org/10.1080/01621459.2016.1147356](https://doi.org/10.1080/01621459.2016.1147356)
- Cheng, X., Cloninger, A. and Coifman, R.R. (2020). Two-sample statistics based on anisotropic kernels. *Inf. Inference* **9** 677–719. [MR4146351 https://doi.org/10.1093/imaiai/iaz018](https://doi.org/10.1093/imaiai/iaz018)
- Cheng, X. and Wu, H.-T. (2022a). Convergence of graph Laplacian with kNN self-tuned kernels. *Inf. Inference* **11** 889–957. [MR4491976 https://doi.org/10.1093/imaiai/iaab019](https://doi.org/10.1093/imaiai/iaab019)
- Cheng, X. and Wu, N. (2022b). Eigen-convergence of Gaussian kernelized graph Laplacian by manifold heat interpolation. *Appl. Comput. Harmon. Anal.* **61** 132–190. [MR4452681 https://doi.org/10.1016/j.acha.2022.06.003](https://doi.org/10.1016/j.acha.2022.06.003)
- Cheng, X. and Xie, Y. (2024). Supplement to “Kernel two-sample tests for manifold data.” <https://doi.org/10.3150/23-BEJ1685SUPP>
- Chwialkowski, K., Strathmann, H. and Gretton, A. (2016). A kernel test of goodness of fit. In *JMLR: Workshop and Conference Proceedings*.
- Chwialkowski, K.P., Ramdas, A., Sejdinovic, D. and Gretton, A. (2015). Fast two-sample testing with analytic representations of probability measures. In *Advances in Neural Information Processing Systems* 1981–1989.
- Coifman, R.R. and Lafon, S. (2006). Diffusion maps. *Appl. Comput. Harmon. Anal.* **21** 5–30. [MR2238665 https://doi.org/10.1016/j.acha.2006.04.006](https://doi.org/10.1016/j.acha.2006.04.006)
- del Barrio, E., Cuesta-Albertos, J.A., Matrán, C. and Rodríguez-Rodríguez, J.M. (1999). Tests of goodness of fit based on the  $L_2$ -Wasserstein distance. *Ann. Statist.* **27** 1230–1239. [MR1740113 https://doi.org/10.1214/aos/1017938923](https://doi.org/10.1214/aos/1017938923)
- do Carmo, M.P. (1992). *Riemannian Geometry. Mathematics: Theory & Applications*. Boston, MA: Birkhäuser, Inc. Translated from the second Portuguese edition by Francis Flaherty. [MR1138207 https://doi.org/10.1007/978-1-4757-2201-7](https://doi.org/10.1007/978-1-4757-2201-7)
- Dunson, D.B., Wu, H.-T. and Wu, N. (2021). Spectral convergence of graph Laplacian and heat kernel reconstruction in  $L^\infty$  from random samples. *Appl. Comput. Harmon. Anal.* **55** 282–336. [MR4279237 https://doi.org/10.1016/j.acha.2021.06.002](https://doi.org/10.1016/j.acha.2021.06.002)
- Farahmand, A.M., Szepesvári, C. and Audibert, J.-Y. (2007). Manifold-adaptive dimension estimation. In *Proceedings of the 24th International Conference on Machine Learning* 265–272.
- Flynn, T. and Yoo, S. (2019). Change detection with the kernel cumulative sum algorithm. In *2019 IEEE 58th Conference on Decision and Control (CDC)*. 6092–6099.
- García Trillos, N., Gerlach, M., Hein, M. and Slepčev, D. (2020). Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace-Beltrami operator. *Found. Comput. Math.* **20** 827–887. [MR4130541 https://doi.org/10.1007/s10208-019-09436-w](https://doi.org/10.1007/s10208-019-09436-w)
- Gretton, A., Fukumizu, K., Harchaoui, Z. and Sriperumbudur, B.K. (2009). A fast, consistent kernel two-sample test. In *Advances in Neural Information Processing Systems* **22** 673–681. Curran Associates.
- Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773. [MR2913716 https://doi.org/10.1214/12-AOS1092](https://doi.org/10.1214/12-AOS1092)
- Györfi, L. and van der Meulen, E.C. (1991). A consistent goodness of fit test based on the total variation distance. In *Nonparametric Functional Estimation and Related Topics (Spetses, 1990)*. NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci. **335** 631–645. Dordrecht: Kluwer Academic. [MR1154355 https://doi.org/10.1007/978-1-4757-2201-7\\_32](https://doi.org/10.1007/978-1-4757-2201-7_32)
- Hein, M., Audibert, J.-Y. and von Luxburg, U. (2005). From graphs to manifolds—weak and strong pointwise consistency of graph Laplacians. In *Learning Theory. Lecture Notes in Computer Science* **3559** 470–485. Berlin: Springer. [MR2203281 https://doi.org/10.1007/11503415\\_32](https://doi.org/10.1007/11503415_32)
- Higgins, J.J. (2004). *An Introduction to Modern Nonparametric Statistics*. Pacific Grove, CA: Brooks/Cole.
- Hoëfding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. [MR0144363 https://doi.org/10.1080/01621459.1963.10501636](https://doi.org/10.1080/01621459.1963.10501636)
- Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications. Springer Series in Statistics*. New York: Springer. [MR2920735 https://doi.org/10.1007/978-1-4614-3655-3](https://doi.org/10.1007/978-1-4614-3655-3)

- Hotelling, H. (1931). The generalization of student's ratio. *Ann. Math. Stat.* **2** 360–378.
- Jitkrittum, W., Kanagawa, H. and Schölkopf, B. (2020). Testing goodness of fit of conditional density models with kernels. In *Conference on Uncertainty in Artificial Intelligence* 221–230. PMLR.
- Jitkrittum, W., Szabó, Z., Chwialkowski, K.P. and Gretton, A. (2016). Interpretable distribution features with maximum testing power. In *Advances in Neural Information Processing Systems* 181–189.
- Jitkrittum, W., Xu, W., Szabó, Z., Fukumizu, K. and Gretton, A. (2017). A linear-time kernel goodness-of-fit test. In *Advances in Neural Information Processing Systems* 262–271.
- Levina, E. and Bickel, P. (2004). Maximum likelihood estimation of intrinsic dimension. *Adv. Neural Inf. Process. Syst.* **17**.
- Li, T. and Yuan, M. (2019). On the optimality of Gaussian kernel based nonparametric tests against smooth alternatives. arXiv preprint. Available at [arXiv:1909.03302](https://arxiv.org/abs/1909.03302).
- Li, C.-L., Chang, W., Cheng, Y., Yang, Y. and Póczos, B. (2017). MMD GAN: Towards deeper understanding of moment matching network. In *Advances in Neural Information Processing Systems* 2203–2213.
- Lloyd, J.R. and Ghahramani, Z. (2015). Statistical model criticism using kernel two sample tests. In *Advances in Neural Information Processing Systems*. 829–837.
- Lopez-Paz, D. and Oquab, M. (2017). Revisiting classifier two-sample tests. In *International Conference on Learning Representations*.
- Massey, F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *J. Amer. Statist. Assoc.* **46** 68–78.
- Mordohai, P. and Medioni, G. (2010). Dimensionality estimation, manifold learning and function approximation using tensor voting. *J. Mach. Learn. Res.* **11** 411–450. [MR2591630](https://doi.org/10.1162/109974610455711)
- Ozakin, A. and Gray, A.G. (2009). Submanifold density estimation. In *Advances in Neural Information Processing Systems*. 1375–1382.
- Pettis, K.W., Bailey, T.A., Jain, A.K. and Dubes, R.C. (1979). An intrinsic dimensionality estimator from near-neighbor information. *IEEE Trans. Pattern Anal. Mach. Intell.* **1** 25–37.
- Peyré, G. (2009). Manifold models for signals and images. *Comput. Vis. Image Underst.* **113** 249–260.
- Pfanzagl, J. and Sheynin, O. (1996). Studies in the history of probability and statistics. XLIV. A forerunner of the  $t$ -distribution. *Biometrika* **83** 891–898. [MR1766040](https://doi.org/10.1093/biomet/83.4.891) <https://doi.org/10.1093/biomet/83.4.891>
- Pratt, J.W. and Gibbons, J.D. (1981). Kolmogorov-Smirnov two-sample tests. In *Concepts of Nonparametric Theory*. 318–344. New York, NY: Springer.
- Ramdas, A., García Trillos, N. and Cuturi, M. (2017). On Wasserstein two-sample testing and related families of nonparametric tests. *Entropy* **19** 47. [MR3608466](https://doi.org/10.3390/e19020047) <https://doi.org/10.3390/e19020047>
- Ramdas, A., Reddi, S.J., Póczos, B., Singh, A. and Wasserman, L. (2015). On the decreasing power of kernel and distance based nonparametric hypothesis tests in high dimensions. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.
- Ren, J., Liu, P.J., Fertig, E., Snoek, J., Poplin, R., Deprieto, M., Dillon, J. and Lakshminarayanan, B. (2019). Likelihood ratios for out-of-distribution detection. *Adv. Neural Inf. Process. Syst.* **32**.
- Saelens, W., Cannoodt, R., Todorov, H. and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37** 547–554. <https://doi.org/10.1038/s41587-019-0071-9>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 4510–4520.
- Serfling, R.J. (2009). *Approximation Theorems of Mathematical Statistics*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley. [MR0595165](https://doi.org/10.1112/jlms.12116)
- Shapiro, A., Xie, Y. and Zhang, R. (2021). Goodness-of-fit tests on manifolds. *IEEE Trans. Inf. Theory* **67** 2539–2553. [MR4282371](https://doi.org/10.1109/tit.2021.3050469) <https://doi.org/10.1109/tit.2021.3050469>
- Singer, A. (2006). From graph to manifold Laplacian: The convergence rate. *Appl. Comput. Harmon. Anal.* **21** 128–134. [MR2238670](https://doi.org/10.1016/j.acha.2006.03.004) <https://doi.org/10.1016/j.acha.2006.03.004>
- Sriperumbudur, B.K., Fukumizu, K., Gretton, A., Schölkopf, B. and Lanckriet, G.R.G. (2012). On the empirical estimation of integral probability metrics. *Electron. J. Stat.* **6** 1550–1599. [MR2988458](https://doi.org/10.1214/12-EJS722) <https://doi.org/10.1214/12-EJS722>
- Sutherland, D.J., Tung, H., Strathmann, H., De, S., Ramdas, A., Smola, A. and Gretton, A. (2017). Generative models and model criticism via optimized maximum mean discrepancy. In *International Conference on Learning Representations*.

- Tenenbaum, J.B., de Silva, V. and Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* **290** 2319–2323.
- Van den Berge, K., Roux de Bézieux, H.R., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., Dudoit, S. and Clement, L. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* **11** 1201. <https://doi.org/10.1038/s41467-020-14766-3>
- van der Maaten, L. (2014). Accelerating t-SNE using tree-based algorithms. *J. Mach. Learn. Res.* **15** 3221–3245. [MR3277169](https://doi.org/10.1162/jmlr.2014.15.1.3221)
- Wynne, G. and Duncan, A.B. (2022). A kernel two-sample test for functional data. *J. Mach. Learn. Res.* **23** 1–51. [MR4576658](https://doi.org/10.1162/jmlr.2022.23.1.4576658)
- Xie, Y. and Siegmund, D. (2013). Sequential multi-sensor change-point detection. *Ann. Statist.* **41** 670–692. [MR3099117](https://doi.org/10.1214/13-AOS1094) <https://doi.org/10.1214/13-AOS1094>
- Xie, L. and Xie, Y. (2021). Sequential change detection by optimal weighted  $\ell_2$  divergence. *IEEE J. Sel. Areas Inf. Theory* **2** 747–761.
- Zelnik-Manor, L. and Perona, P. (2005). Self-tuning spectral clustering. In *Advances in Neural Information Processing Systems* 1601–1608.
- Zhao, J., Jaffe, A., Li, H., Lindenbaum, O., Sefik, E., Jackson, R., Cheng, X., Flavell, R.A. and Kluger, Y. (2021). Detection of differentially abundant cell subpopulations in scRNA-seq data. *Proc. Natl. Acad. Sci. USA* **118**. <https://doi.org/10.1073/pnas.2100293118>
- Zhu, W., Qiu, Q., Huang, J., Calderbank, R., Sapiro, G. and Daubechies, I. (2018). LDMNet: Low dimensional manifold regularized neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2743–2751.

# Detecting changes in the trend function of heteroscedastic time series

SARA KRISTIN SCHMIDT<sup>a</sup>

*Department of Mathematics, Ruhr-Universität Bochum, Universitätsstraße 150, 44780 Bochum, Germany,*  
<sup>a</sup>[sara.schmidt@rub.de](mailto:sara.schmidt@rub.de)

We propose a new asymptotic test to assess the stationarity of a time series' mean that is applicable in the presence of both heteroscedasticity and short-range dependence. Our test statistic is composed of Gini's mean difference of local sample means. To analyse its asymptotic behaviour, we develop new limit theory for U-statistics of strongly mixing triangular arrays under non-stationarity. Most importantly, we show asymptotic normality of the test statistic under the hypothesis of a constant mean and prove the test's consistency against a very general class of alternatives, including both smooth and abrupt changes in the mean. We propose estimators for all parameters involved, including an adapted subsampling estimator for the long run variance, and show their consistency. Our procedure is practically evaluated in an extensive simulation study and in a data example considering the number of live births in Germany.

*Keywords:* Change point analysis; heteroscedasticity; short-range dependence; test for a constant mean; U-statistic of a non-stationary triangular array

## References

- An, H.Z. and Huang, F.C. (1996). The geometrical ergodicity of nonlinear autoregressive models. *Statist. Sinica* **6** 943–956. [MR1422412](#)
- Apostol, T.M. (1974). *Mathematical Analysis*, 2nd ed. Reading: Addison-Wesley. [MR0344384](#)
- Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. [MR0233396](#)
- Bradley, R.C. (2005). Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** 107–144. [MR2178042](#) <https://doi.org/10.1214/154957805100000104>
- Bücher, A., Dette, H. and Heinrichs, F. (2021). Are deviations in a gradually varying mean relevant? A testing approach based on sup-norm estimators. *Ann. Statist.* **49** 3583–3617. [MR4352542](#) <https://doi.org/10.1214/21-aos2098>
- Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14** 1171–1179. [MR0856813](#) <https://doi.org/10.1214/aos/1176350057>
- Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. [MR2743035](#)
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. [MR1429916](#) <https://doi.org/10.1214/aos/1034276620>
- Dehling, H., Fried, R., Sharipov, O.Sh., Vogel, D. and Wornowizki, M. (2013). Estimation of the variance of partial sums of dependent processes. *Statist. Probab. Lett.* **83** 141–147. [MR2998735](#) <https://doi.org/10.1016/j.spl.2012.08.012>
- Dette, H. and Wu, W. (2019). Detecting relevant changes in the mean of nonstationary processes—a mass excess approach. *Ann. Statist.* **47** 3578–3608. [MR4025752](#) <https://doi.org/10.1214/19-AOS1811>
- 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>
- Frück, K., Munk, A. and Sieling, H. (2014). Multiscale change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 495–580. [MR3210728](#) <https://doi.org/10.1111/rssb.12047>

- Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.* **42** 2243–2281. [MR3269979 https://doi.org/10.1214/14-AOS1245](https://doi.org/10.1214/14-AOS1245)
- Gerstenberger, C. and Vogel, D. (2015). On the efficiency of Gini’s mean difference. *Stat. Methods Appl.* **24** 569–596. [MR3421674 https://doi.org/10.1007/s10260-015-0315-x](https://doi.org/10.1007/s10260-015-0315-x)
- Górecki, T., Horváth, L. and Kokoszka, P. (2018). Change point detection in heteroscedastic time series. *Econom. Stat.* **7** 63–88. [MR3824127 https://doi.org/10.1016/j.ecosta.2017.07.005](https://doi.org/10.1016/j.ecosta.2017.07.005)
- Heinrichs, F. and Dette, H. (2021). A distribution free test for changes in the trend function of locally stationary processes. *Electron. J. Stat.* **15** 3762–3797. [MR4298981 https://doi.org/10.1214/21-ejs1871](https://doi.org/10.1214/21-ejs1871)
- Horváth, L. (1993). The maximum likelihood method for testing changes in the parameters of normal observations. *Ann. Statist.* **21** 671–680. [MR1232511 https://doi.org/10.1214/aos/1176349143](https://doi.org/10.1214/aos/1176349143)
- Juhl, T. and Xiao, Z. (2005). A nonparametric test for changing trends. *J. Econometrics* **127** 179–199. [MR2156332 https://doi.org/10.1016/j.jeconom.2004.05.014](https://doi.org/10.1016/j.jeconom.2004.05.014)
- Lindner, A.M. (2009). Stationarity, mixing, distributional properties and moments of GARCH(p,q)–processes. In *Handbook of Financial Time Series* 43–69. Berlin: Springer.
- Loader, C.R. (1996). Change point estimation using nonparametric regression. *Ann. Statist.* **24** 1667–1678. [MR1416655 https://doi.org/10.1214/aos/1032298290](https://doi.org/10.1214/aos/1032298290)
- McGonigle, E.T., Killick, R. and Nunes, M.A. (2021). Detecting changes in mean in the presence of time-varying autocovariance. *Stat* **10** Paper No. e351, 14. [MR4276025 https://doi.org/10.1002/sta4.351](https://doi.org/10.1002/sta4.351)
- Mies, F. (2023). Functional estimation and change detection for nonstationary time series. *J. Amer. Statist. Assoc.* **118** 1011–1022. [MR4595473 https://doi.org/10.1080/01621459.2021.1969239](https://doi.org/10.1080/01621459.2021.1969239)
- Mokkadem, A. (1988). Mixing properties of ARMA processes. *Stochastic Process. Appl.* **29** 309–315. [MR0958507 https://doi.org/10.1016/0304-4149\(88\)90045-2](https://doi.org/10.1016/0304-4149(88)90045-2)
- Pein, F., Sieling, H. and Munk, A. (2017). Heterogeneous change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1207–1227. [MR3689315 https://doi.org/10.1111/rssb.12202](https://doi.org/10.1111/rssb.12202)
- Peligrad, M. and Suresh, R. (1995). Estimation of variance of partial sums of an associated sequence of random variables. *Stochastic Process. Appl.* **56** 307–319. [MR1325225 https://doi.org/10.1016/0304-4149\(94\)00065-2](https://doi.org/10.1016/0304-4149(94)00065-2)
- Peligrad, M. (2002). Some remarks on coupling of dependent random variables. *Statist. Probab. Lett.* **60** 201–209. [MR1945442 https://doi.org/10.1016/S0167-7152\(02\)00318-8](https://doi.org/10.1016/S0167-7152(02)00318-8)
- Peligrad, M. and Shao, Q.M. (1995). Estimation of the variance of partial sums for  $\rho$ -mixing random variables. *J. Multivariate Anal.* **52** 140–157. [MR1325375 https://doi.org/10.1006/jmva.1995.1008](https://doi.org/10.1006/jmva.1995.1008)
- Pešta, M. and Wendler, M. (2020). Nuisance-parameter-free changepoint detection in non-stationary series. *TEST* **29** 379–408. [MR4095034 https://doi.org/10.1007/s11749-019-00659-1](https://doi.org/10.1007/s11749-019-00659-1)
- Petrov, V.V. (1975). *Sums of Independent Random Variables. Ergebnisse der Mathematik und Ihrer Grenzgebiete [Results in Mathematics and Related Areas], Band 82*. New York: Springer. [MR0388499](https://doi.org/10.1007/978-1-4939-9829-7)
- R CORE TEAM (2022). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.
- Schmidt, S.K. (2024). Supplement to “Detecting changes in the trend function of heteroscedastic time series.” <https://doi.org/10.3150/23-BEJ1686SUPP>
- Schmidt, S.K., Wornowizki, M., Fried, R. and Dehling, H. (2021). An asymptotic test for constancy of the variance under short-range dependence. *Ann. Statist.* **49** 3460–3481. [MR4352537 https://doi.org/10.1214/21-aos2092](https://doi.org/10.1214/21-aos2092)
- Tihomirov, A.N. (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Teor. Veroyatn. Primen.* **25** 800–818. [MR0595140](https://doi.org/10.1007/BF01069514)
- Vogt, M. (2015). Testing for structural change in time-varying nonparametric regression models. *Econometric Theory* **31** 811–859. [MR3377270 https://doi.org/10.1017/S0266466614000565](https://doi.org/10.1017/S0266466614000565)
- Wornowizki, M., Fried, R. and Meintanis, S.G. (2017). Fourier methods for analyzing piecewise constant volatilities. *AStA Adv. Stat. Anal.* **101** 289–308. [MR3679347 https://doi.org/10.1007/s10182-017-0288-1](https://doi.org/10.1007/s10182-017-0288-1)
- Wu, J.S. and Chu, C.K. (1993). Kernel-type estimators of jump points and values of a regression function. *Ann. Statist.* **21** 1545–1566. [MR1241278 https://doi.org/10.1214/aos/1176349271](https://doi.org/10.1214/aos/1176349271)
- Wu, W.B., Woodroffe, M. and Mentz, G. (2001). Isotonic regression: Another look at the changepoint problem. *Biometrika* **88** 793–804. [MR1859410 https://doi.org/10.1093/biomet/88.3.793](https://doi.org/10.1093/biomet/88.3.793)
- Wu, W.B. and Zhao, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. [MR2323759 https://doi.org/10.1111/j.1467-9868.2007.00594.x](https://doi.org/10.1111/j.1467-9868.2007.00594.x)
- Zhao, Z. and Li, X. (2013). Inference for modulated stationary processes. *Bernoulli* **19** 205–227. [MR3019492 https://doi.org/10.3150/11-BEJ399](https://doi.org/10.3150/11-BEJ399)

Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *J. Amer. Statist. Assoc.* **108** 726–740. MR3174655 <https://doi.org/10.1080/01621459.2013.787184>

# Minimum information dependence modeling

TOMONARI SEI<sup>1,a</sup> and KEISUKE YANO<sup>2,b</sup>

<sup>1</sup>The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan, <sup>a</sup>[sei@mist.i.u-tokyo.ac.jp](mailto:sei@mist.i.u-tokyo.ac.jp)

<sup>2</sup>The Institute of Statistical Mathematics, 10-3 Midori cho, Tachikawa City, Tokyo, 190-8562, Japan,  
<sup>b</sup>[yano@ism.ac.jp](mailto:yano@ism.ac.jp)

We propose a method to construct a joint statistical model for mixed-domain data to analyze their dependence. Multivariate Gaussian and log-linear models are particular examples of the proposed model. It is shown that the functional equation defining the model has a unique solution under fairly weak conditions. The model is characterized by two orthogonal parameters: the dependence parameter and the marginal parameter. To estimate the dependence parameter, a conditional inference together with a sampling procedure is proposed and is shown to provide a consistent estimator. Illustrative examples of data analyses involving penguins and earthquakes are presented.

**Keywords:** Conditional inference; copula; earthquake data; graphical model; mixed-domain; Monte Carlo method

## References

- Japan Meteorological Agency (2022). The seismological bulletin of Japan. Available at [https://www.data.jma.go.jp/svd/eqev/data/bulletin/index\\_e.html](https://www.data.jma.go.jp/svd/eqev/data/bulletin/index_e.html).
- Agrawal, R. (2020). Finite-sample concentration of the multinomial in relative entropy. *IEEE Trans. Inf. Theory* **66** 6297–6302. MR4173539 <https://doi.org/10.1109/TIT.2020.2996134>
- Albert, A. and Anderson, J.A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71** 1–10. MR0738319 <https://doi.org/10.1093/biomet/71.1.1>
- Amari, S. (2001). Information geometry on hierarchy of probability distributions. *IEEE Trans. Inf. Theory* **47** 1701–1711. MR1842511 <https://doi.org/10.1109/18.930911>
- Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry. Translations of Mathematical Monographs* **191**. Providence, RI: Amer. Math. Soc.. MR1800071 <https://doi.org/10.1090/mmono/191>
- Bedford, T., Daneshkhah, A. and Wilson, K.J. (2016). Approximate uncertainty modeling in risk analysis with vine copulas. *Risk Anal.* **36** 792–815.
- Bedford, T. and Wilson, K.J. (2014). On the construction of minimum information bivariate copula families. *Ann. Inst. Statist. Math.* **66** 703–723. MR3224606 <https://doi.org/10.1007/s10463-013-0422-0>
- Borwein, J.M., Lewis, A.S. and Nussbaum, R.D. (1994). Entropy minimization, DAD problems, and doubly stochastic kernels. *J. Funct. Anal.* **123** 264–307. MR1283029 <https://doi.org/10.1006/jfan.1994.1089>
- Bose, A. and Chatterjee, S. (2018). *U-Statistics,  $M_m$ -Estimators and Resampling. Texts and Readings in Mathematics* **75**. Singapore: Springer. MR3837543 <https://doi.org/10.1007/978-981-13-2248-8>
- Buja, A., Hastie, T. and Tibshirani, R. (1989). Linear smoothers and additive models. *Ann. Statist.* **17** 453–555. MR0994249 <https://doi.org/10.1214/aos/1176347115>
- Chen, Y. and Sei, T. (2022). A proper scoring rule for minimum information copulas. Available at [arXiv:2204.03118](https://arxiv.org/abs/2204.03118).
- Choi, L., Blume, J.D. and Dupont, W.D. (2015). Elucidating the foundations of statistical inference with 2 x 2 tables. *PLoS ONE* **10** e0121263. <https://doi.org/10.1371/journal.pone.0121263>
- Cox, D.R. and Hinkley, D.V. (1974). *Theoretical Statistics*. London: CRC Press. MR0370837
- Csiszár, I. (1975).  $I$ -divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **3** 146–158. MR0365798 <https://doi.org/10.1214/aop/1176996454>
- Dempster, A.P. (1972). Covariance selection. *Biometrics* **28** 157–175. MR3931974

- de la Peña, V.H. and Giné, E. (1999). *Decoupling: From Dependence to Independence. Probability and Its Applications (New York)*. New York: Springer. MR1666908 <https://doi.org/10.1007/978-1-4612-0537-1>
- Diaconis, P. and Sturmfels, B. (1998). Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26** 363–397. MR1608156 <https://doi.org/10.1214/aos/1030563990>
- Eckstein, S. and Nutz, M. (2022). Quantitative stability of regularized optimal transport and convergence of Sinkhorn’s algorithm. *SIAM J. Math. Anal.* **54** 5922–5948. MR4506579 <https://doi.org/10.1137/21M145505X>
- Falconer, K. (2014). *Fractal Geometry: Mathematical Foundations and Applications*, 3rd ed. Chichester: Wiley. MR3236784
- Geenens, G. (2020). Copula modeling for discrete random vectors. *Depend. Model.* **8** 417–440. MR4193472 <https://doi.org/10.1515/demo-2020-0022>
- Genest, C. and Nešlehová, J. (2007). A primer on copulas for count data. *Astin Bull.* **37** 475–515. MR2422797 <https://doi.org/10.2143/AST.37.2.2024077>
- Geyer, C.J. and Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependent data. *J. Roy. Statist. Soc. Ser. B* **54** 657–699. MR1185217
- Haasler, I., Singh, R., Zhang, Q., Karlsson, J. and Chen, Y. (2021). Multi-marginal optimal transport and probabilistic graphical models. *IEEE Trans. Inf. Theory* **67** 4647–4668. MR4306289 <https://doi.org/10.1109/tit.2021.3077465>
- Haberman, S.J. (1977). *The Analysis of Frequency Data: Statistical Research Monographs*. Chicago: Univ. Chicago Press.
- Hannan, J. and Harkness, W. (1963). Normal approximation to the distribution of two independent binomials, conditional on fixed sum. *Ann. Math. Stat.* **34** 1593–1595. MR0160281 <https://doi.org/10.1214/aoms/1177703893>
- Harkness, W.L. (1965). Properties of the extended hypergeometric distribution. *Ann. Math. Stat.* **36** 938–945. MR0182073 <https://doi.org/10.1214/aoms/1177700066>
- Hinton, G.E. (2002). Training products of experts by minimizing contrastive divergence. *Neural Comput.* **14** 1771–1800. <https://doi.org/10.1162/089976602760128018>
- Holland, P.W. and Wang, Y.J. (1987). Dependence function for continuous bivariate densities. *Comm. Statist. Theory Methods* **16** 863–876. MR0886560 <https://doi.org/10.1080/03610928708829408>
- Hyvärinen, A. (2006). Consistency of pseudolikelihood estimation of fully visible Boltzmann machines. *Neural Comput.* **18** 2283–2292. MR2256106 <https://doi.org/10.1162/neco.2006.18.10.2283>
- Jansen, M.J.W. (1997). Maximum entropy distributions with prescribed marginals and normal score correlations. In *Distributions with Given Marginals and Moment Problems* (V. Beneš and J. Štěpán, eds.), 87–92. Springer.
- Jaynes, E.T. (1957). Information theory and statistical mechanics. *Phys. Rev. (2)* **106** 620–630. MR0087305
- Jones, M.C., Pewsey, A. and Kato, S. (2015). On a class of circulars: Copulas for circular distributions. *Ann. Inst. Statist. Math.* **67** 843–862. MR3390169 <https://doi.org/10.1007/s10463-014-0493-6>
- Kou, S.G. and Ying, Z. (1996). Asymptotics for a  $2 \times 2$  table with fixed margins. *Statist. Sinica* **6** 809–829. MR1422405
- Kurowicka, D. and van Horsen, W.T. (2015). On an interaction function for copulas. *J. Multivariate Anal.* **138** 127–142. MR3348837 <https://doi.org/10.1016/j.jmva.2014.12.012>
- Lauritzen, S.L. (1996). *Graphical Models. Oxford Statistical Science Series* **17**. New York: Oxford University Press. MR1419991
- Léonard, C. (2012). From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.* **262** 1879–1920. MR2873864 <https://doi.org/10.1016/j.jfa.2011.11.026>
- Little, R.J.A. (1989). Testing the equality of two independent binomial proportions. *Amer. Statist.* **43** 283–288.
- Meeuwissen, A.M.H. and Bedford, T. (1997). Minimally informative distributions with given rank correlation for use in uncertainty analysis. *J. Stat. Comput. Simul.* **57** 143–174. MR1463104 <https://doi.org/10.1080/00949659708811806>
- Mukherjee, S. (2016). Estimation in exponential families on permutations. *Ann. Statist.* **44** 853–875. MR3476619 <https://doi.org/10.1214/15-AOS1389>
- Peyré, G. and Cuturi, M. (2019). Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** 355–607.
- Piantadosi, J., Howlett, P. and Borwein, J. (2012). Copulas with maximum entropy. *Optim. Lett.* **6** 99–125. MR2886586 <https://doi.org/10.1007/s11590-010-0254-2>



- Reid, N. (1995). The roles of conditioning in inference. *Statist. Sci.* **10** 138–157. [MR1368097](#)
- Rinaldo, A., Fienberg, S.E. and Zhou, Y. (2009). On the geometry of discrete exponential families with application to exponential random graph models. *Electron. J. Stat.* **3** 446–484. [MR2507456](#) <https://doi.org/10.1214/08-EJS350>
- Robbins, H. (1955). A remark on Stirling’s formula. *Amer. Math. Monthly* **62** 26–29. [MR0069328](#) <https://doi.org/10.2307/2308012>
- Rüschendorf, L. and Thomsen, W. (1993). Note on the Schrödinger equation and  $I$ -projections. *Statist. Probab. Lett.* **17** 369–375. [MR1237783](#) [https://doi.org/10.1016/0167-7152\(93\)90257-J](https://doi.org/10.1016/0167-7152(93)90257-J)
- Sei, T. and Yano, K. (2024). Supplement to “Minimum information dependence modeling.” <https://doi.org/10.3150/23-BEJ1687SUPP>
- Sinkhorn, R. and Knopp, P. (1967). Concerning nonnegative matrices and doubly stochastic matrices. *Pacific J. Math.* **21** 343–348. [MR0210731](#)
- van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. Springer Series in Statistics. New York: Springer. [MR1385671](#) <https://doi.org/10.1007/978-1-4757-2545-2>
- Watanabe, S. (1960). Information theoretical analysis of multivariate correlation. *IBM J. Res. Develop.* **4** 66–82. [MR0109755](#) <https://doi.org/10.1147/rd.41.0066>
- Weissman, T., Ordentlich, E., Seroussi, G., Verdu, S. and Weinberger, M. (2003). Inequalities for the  $L_1$  deviation of the empirical distribution. Technical Report HPL-2003-97R1, Hewlett-Packard Labs.
- Wessel, P. and Smith, W. (1998). New, improved version of generic mapping tools released. *Eos Trans. AGU* **79** 409–410.
- Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. Chichester: Wiley. [MR1112133](#)
- Yang, E., Ravikumar, P., Allen, G.I. and Liu, Z. (2015). Graphical models via univariate exponential family distributions. *J. Mach. Learn. Res.* **16** 3813–3847. [MR3450553](#)

# Asymptotic analysis of statistical estimators related to MultiGraphex processes under misspecification

ZACHARIE NAULET<sup>1,a</sup>, JUDITH ROUSSEAU<sup>2,b</sup> and FRANÇOIS CARON<sup>2,c</sup>

<sup>1</sup>Université Paris-Saclay, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France,

<sup>a</sup>[zacharie.naulet@universite-paris-saclay.fr](mailto:zacharie.naulet@universite-paris-saclay.fr)

<sup>2</sup>University of Oxford, Department of Statistics, Oxford, UK, <sup>b</sup>[judith.rousseau@stats.ox.ac.uk](mailto:judith.rousseau@stats.ox.ac.uk),

<sup>c</sup>[caron@stats.ox.ac.uk](mailto:caron@stats.ox.ac.uk)

This article studies the asymptotic properties of Bayesian or frequentist estimators of a vector of parameters related to structural properties of sequences of graphs. The estimators studied originate from a particular class of graphex model introduced by Caron and Fox (*J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** (2017) 1295–1366). The analysis is however performed here under very weak assumptions on the underlying data generating process, which may be different from the model of (*J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** (2017) 1295–1366) or from a graphex model. In particular, we consider generic sparse graph models, with unbounded degree, whose degree distribution satisfies some assumptions. We show that one can relate the limit of the estimator of one of the parameters to the sparsity constant of the true graph generating process. When taking a Bayesian approach, we also show that the posterior distribution is asymptotically normal. We discuss situations where classical random graphs models, such as configuration models, satisfy our assumptions.

**Keywords:** Bayesian nonparametrics; networks; random graphs; sparsity; caron and fox model; inference; maximum likelihood estimation; bayesian estimation; misspecification

## References

- [1] Bickel, P.J. and Chen, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci. USA* **106** 21068–21073.
- [2] Bickel, P.J., Chen, A. and Levina, E. (2011). The method of moments and degree distributions for network models. *Ann. Statist.* **39** 2280–2301. [MR2906868](#) <https://doi.org/10.1214/11-AOS904>
- [3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. [MR0898871](#) <https://doi.org/10.1017/CBO9780511721434>
- [4] Bollobás, B. and Riordan, O. (2009). Metrics for sparse graphs. In *Surveys in Combinatorics 2009. London Mathematical Society Lecture Note Series* **365** 211–287. Cambridge: Cambridge Univ. Press. [MR2588543](#)
- [5] Borgs, C., Chayes, J.T., Cohn, H. and Ganguly, S. (2021). Consistent nonparametric estimation for heavy-tailed sparse graphs. *Ann. Statist.* **49** 1904–1930. [MR4319235](#) <https://doi.org/10.1214/20-aos1985>
- [6] Borgs, C., Chayes, J.T., Cohn, H. and Holden, N. (2017). Sparse exchangeable graphs and their limits via graphon processes. *J. Mach. Learn. Res.* **18** Paper No. 210, 71. [MR3827098](#)
- [7] Borgs, C., Chayes, J.T., Cohn, H. and Zhao, Y. (2014). An  $L^p$  theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. Available at [arXiv:1401.2906](https://arxiv.org/abs/1401.2906).
- [8] Borgs, C., Chayes, J.T., Cohn, H. and Zhao, Y. (2018). An  $L^p$  theory of sparse graph convergence II: LD convergence, quotients and right convergence. *Ann. Probab.* **46** 337–396. [MR3758733](#) <https://doi.org/10.1214/17-AOP1187>

- [9] Borgs, C., Chayes, J.T., Dhara, S. and Sen, S. (2021). Limits of sparse configuration models and beyond: Graphexes and multigraphexes. *Ann. Probab.* **49** 2830–2873. MR4348680 <https://doi.org/10.1214/21-aop1508>
- [10] Boucheron, S. and Thomas, M. (2015). Tail index estimation, concentration and adaptivity. *Electron. J. Stat.* **9** 2751–2792. MR3435810 <https://doi.org/10.1214/15-EJS1088>
- [11] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.* **31** 929–953. MR1747450 <https://doi.org/10.1239/aap/1029955251>
- [12] Cai, D., Campbell, T. and Broderick, T. (2016). Edge-exchangeable graphs and sparsity. In *Advances in Neural Information Processing Systems* 4249–4257.
- [13] Caron, F. and Fox, E.B. (2014). Bayesian nonparametric models of sparse and exchangeable random graphs. Available at [arXiv:1401.1137](https://arxiv.org/abs/1401.1137).
- [14] Caron, F. and Fox, E.B. (2017). Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1295–1366. MR3731666 <https://doi.org/10.1111/rssb.12233>
- [15] Caron, F., Panero, F. and Rousseau, J. (2017). On sparsity, power-law and clustering properties of graphex processes. Available at [arXiv:1708.03120](https://arxiv.org/abs/1708.03120).
- [16] Carpentier, A. and Kim, A.K.H. (2015). Adaptive and minimax optimal estimation of the tail coefficient. *Statist. Sinica* **25** 1133–1144. MR3410301
- [17] Crane, H. and Dempsey, W. (2018). Edge exchangeable models for interaction networks. *J. Amer. Statist. Assoc.* **113** 1311–1326. MR3862359 <https://doi.org/10.1080/01621459.2017.1341413>
- [18] Di Benedetto, G., Caron, F. and Teh, Y.W. (2021). Nonexchangeable random partition models for microclustering. *Ann. Statist.* **49** 1931–1957. MR4319236 <https://doi.org/10.1214/20-aos2003>
- [19] Favaro, S. and Naulet, Z. (2023). Near-optimal estimation of the unseen under regularly varying tail populations. *Bernoulli* **29** 3423–3442. MR4632144 <https://doi.org/10.3150/23-bej1589>
- [20] Gao, C., Lu, Y., Ma, Z. and Zhou, H.H. (2016). Optimal estimation and completion of matrices with biclustering structures. *J. Mach. Learn. Res.* **17** Paper No. 161, 29. MR3569248
- [21] Gao, C. and Ma, Z. (2021). Minimax rates in network analysis: Graphon estimation, community detection and hypothesis testing. *Statist. Sci.* **36** 16–33. MR4194201 <https://doi.org/10.1214/19-ST5736>
- [22] Gnedin, A., Hansen, B. and Pitman, J. (2007). Notes on the occupancy problem with infinitely many boxes: General asymptotics and power laws. *Probab. Surv.* **4** 146–171. MR2318403 <https://doi.org/10.1214/07-PS092>
- [23] Gopalan, P., Hofman, J.M. and Blei, D.M. (2015). Scalable recommendation with hierarchical Poisson factorization. In *UAI* 326–335.
- [24] Janson, S. (2018). On edge exchangeable random graphs. *J. Stat. Phys.* **173** 448–484. MR3876897 <https://doi.org/10.1007/s10955-017-1832-9>
- [25] Karlin, S. (1967). Central limit theorems for certain infinite urn schemes. *J. Math. Mech.* **17** 373–401. MR0216548 <https://doi.org/10.1512/iumj.1968.17.17020>
- [26] 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>
- [27] Klopp, O. and Verzelen, N. (2019). Optimal graphon estimation in cut distance. *Probab. Theory Related Fields* **174** 1033–1090. MR3980311 <https://doi.org/10.1007/s00440-018-0878-1>
- [28] Molloy, M. and Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6** 161–179. MR1370952 <https://doi.org/10.1002/rsa.3240060204>
- [29] Naulet, Z., Rousseau, J. and Caron, F. (2024). Supplement to “Asymptotic analysis of statistical estimators related to MultiGraphex processes under misspecification.” <https://doi.org/10.3150/23-BEJ1689SUPP>
- [30] Naulet, Z., Roy, D.M., Sharma, E. and Veitch, V. (2021). Bootstrap estimators for the tail-index and for the count statistics of graphex processes. *Electron. J. Stat.* **15** 282–325. MR4195774 <https://doi.org/10.1214/20-EJS1789>
- [31] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249 <https://doi.org/10.1007/BF01213386>
- [32] Pitman, J. (2003). Poisson-Kingman partitions. In *Statistics and Science: A Festschrift for Terry Speed. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **40** 1–34. Beachwood, OH: IMS. MR2004330 <https://doi.org/10.1214/inms/1215091133>
- [33] Rouault, A. (1978). Lois de Zipf et sources markoviennes. *Ann. Inst. Henri Poincaré Sect. B (N.S.)* **14** 169–188. MR0507732

- [34] van den Esker, H., van der Hofstad, R., Hooghiemstra, G. and Znamenski, D. (2005). Distances in random graphs with infinite mean degrees. *Extremes* **8** 111–141. [MR2275914](#) <https://doi.org/10.1007/s10687-006-7963-z>
- [35] van der Hofstad, R. (2014). Random graphs and complex networks. Vol. I. Technical Report Eindhoven Univ. Technology.
- [36] Veitch, V. and Roy, D. M. (2015). The class of random graphs arising from exchangeable random measures. Available at [arXiv:1512.03099](#).
- [37] Veitch, V. and Roy, D.M. (2019). Sampling and estimation for (sparse) exchangeable graphs. *Ann. Statist.* **47** 3274–3299. [MR4025742](#) <https://doi.org/10.1214/18-AOS1778>
- [38] Williamson, S.A. (2016). Nonparametric network models for link prediction. *J. Mach. Learn. Res.* **17** Paper No. 202, 21. [MR3580355](#)
- [39] Wolfe, P.J. and Olhede, S.C. (2013). Nonparametric graphon estimation. Available at [arXiv:1309.5936](#).

# Some rapidly mixing hit-and-run samplers for latent counts in linear inverse problems

MARTIN HAZELTON<sup>1,a</sup>, MICHAEL MCVEAGH<sup>1,b</sup>, CHRISTOPHER TUFFLEY<sup>2,c</sup>  
and BRUCE VAN BRUNT<sup>2,d</sup>

<sup>1</sup>*University of Otago, Dunedin, New Zealand, [a](mailto:martin.hazelton@otago.ac.nz)[martin.hazelton@otago.ac.nz](mailto:martin.hazelton@otago.ac.nz), [b](mailto:mcveagh@gmail.com)[mcveagh@gmail.com](mailto:mcveagh@gmail.com)*

<sup>2</sup>*Massey University, Palmerston North, New Zealand, [c](mailto:C.Tuffley@massey.ac.nz)[C.Tuffley@massey.ac.nz](mailto:C.Tuffley@massey.ac.nz), [d](mailto:B.vanBrunt@massey.ac.nz)[B.vanBrunt@massey.ac.nz](mailto:B.vanBrunt@massey.ac.nz)*

Linear inverse problems for count data arise in a myriad of settings. The latent counts lie on a fibre that is too large to enumerate in most practical problems, but inference can proceed by sampling the fibre. We examine the mixing properties of hit-and-run samplers in this context. In general convergence can be arbitrarily slow. However, there is a class of linear inverse problems for which rapid mixing for uniform fibre sampling is possible, using Markov sub-bases that are of minimum size but yet provide a sufficiently rich range of sampling directions to avoid the need for zig-zagging walks to ensure connectivity. Focussing on such problems, we study a particular class of bases that enjoy these properties under certain easily checkable conditions on the configuration matrix. We also examine the mixing properties of these bases when employing commonly used Poisson models. Our theoretical results provide practical guidance on optimizing these Markov sub-bases.

*Keywords:* Augmenting path; Eulerian matrix; fibre sampler; Markov basis; mixing time; random walk; second largest eigenvalue modulus

## References

- 4ti2team (2015). 4ti2—A software package for algebraic, geometric and combinatorial problems on linear spaces (version 1.6.7). Available at [www.4ti2.de](http://www.4ti2.de).
- Airoldi, E.M. and Blocker, A.W. (2013). Estimating latent processes on a network from indirect measurements. *J. Amer. Statist. Assoc.* **108** 149–164. [MR3174609 https://doi.org/10.1080/01621459.2012.756328](https://doi.org/10.1080/01621459.2012.756328)
- Airoldi, E.M. and Haas, B. (2011). Polytope samplers for inference in ill-posed inverse problems. In *International Conference on Artificial Intelligence and Statistics* **15** 110–118.
- Aoki, S., Hara, H. and Takemura, A. (2012). *Markov Bases in Algebraic Statistics*. Springer Series in Statistics. New York: Springer. [MR2961912 https://doi.org/10.1007/978-1-4614-3719-2](https://doi.org/10.1007/978-1-4614-3719-2)
- Baumert, S., Ghate, A., Kiatsupaibul, S., Shen, Y., Smith, R.L. and Zabinsky, Z.B. (2009). Discrete hit-and-run for sampling points from arbitrary distributions over subsets of integer hyperrectangles. *Oper. Res.* **57** 727–739. [MR2554262 https://doi.org/10.1287/opre.1080.0600](https://doi.org/10.1287/opre.1080.0600)
- Brooks, S.P. and Roberts, G.O. (1998). Diagnosing convergence of Markov chain Monte Carlo algorithms. *Stat. Comput.* **8** 319–335.
- Castro, R., Coates, M., Liang, G., Nowak, R. and Yu, B. (2004). Network tomography: Recent developments. *Statist. Sci.* **19** 499–517. [MR2185628 https://doi.org/10.1214/088342304000000422](https://doi.org/10.1214/088342304000000422)
- Chen, Y., Dinwoodie, I.H. and Sullivant, S. (2006). Sequential importance sampling for multiway tables. *Ann. Statist.* **34** 523–545. [MR2275252 https://doi.org/10.1214/009053605000000822](https://doi.org/10.1214/009053605000000822)
- De Loera, J.A., Hemmecke, R. and Lee, J. (2015). On augmentation algorithms for linear and integer-linear programming: From Edmonds-Karp to Bland and beyond. *SIAM J. Optim.* **25** 2494–2511. [MR3429745 https://doi.org/10.1137/151002915](https://doi.org/10.1137/151002915)
- Diaconis, P. and Sturmfels, B. (1998). Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26** 363–397. [MR1608156 https://doi.org/10.1214/aos/1030563990](https://doi.org/10.1214/aos/1030563990)
- Dobra, A. (2003). Markov bases for decomposable graphical models. *Bernoulli* **9** 1093–1108. [MR2046819 https://doi.org/10.3150/bj/1072215202](https://doi.org/10.3150/bj/1072215202)

- Fang, J., Vardi, Y. and Zhang, C.-H. (2007). An iterative tomography algorithm for the estimation of network traffic. In *Complex Datasets and Inverse Problems. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **54** 12–23. Beachwood, OH: IMS. [MR2459176](#) <https://doi.org/10.1214/074921707000000030>
- Ghouila-Houri, A. (1962). Caractérisation des matrices totalement unimodulaires. *C. R. Acad. Sci. Paris* **254** 1192–1194. [MR0132752](#)
- Hazelton, M.L. (2015). Network tomography for integer-valued traffic. *Ann. Appl. Stat.* **9** 474–506. [MR3341124](#) <https://doi.org/10.1214/15-AOAS805>
- Hazelton, M.L. and Bilton, T.P. (2017). Polytope samplers for network tomography. *Aust. N. Z. J. Stat.* **59** 495–511. [MR3760156](#) <https://doi.org/10.1111/anzs.12216>
- Hazelton, M.L., McVeagh, M.R. and van Brunt, B. (2021). Geometrically aware dynamic Markov bases for statistical linear inverse problems. *Biometrika* **108** 609–626. [MR4298767](#) <https://doi.org/10.1093/biomet/asaa083>
- Hazelton, M., McVeagh, M., Tuffley, C. and van Brunt, B. (2024). Supplement to “Some rapidly mixing hit-and-run samplers for latent counts in linear inverse problems.” <https://doi.org/10.3150/23-BEJ1690SUPP>
- Hemmecke, R. and Malkin, P. (2005). Computing generating sets of lattice ideals. arXiv preprint [arXiv:math/0508359](https://arxiv.org/abs/math/0508359).
- Kaipio, J. and Somersalo, E. (2006). *Statistical and Computational Inverse Problems* **160**. Berlin: Springer.
- Levin, D.A., Peres, Y. and Wilmer, E.L. (2009). *Markov Chains and Mixing Times* **107**. Providence: American Mathematical Soc.
- Link, W.A., Yoshizaki, J., Bailey, L.L. and Pollock, K.H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* **66** 178–185. [MR2756704](#) <https://doi.org/10.1111/j.1541-0420.2009.01244.x>
- Rapallo, F. (2003). Algebraic Markov bases and MCMC for two-way contingency tables. *Scand. J. Stat.* **30** 385–397. [MR1983132](#) <https://doi.org/10.1111/1467-9469.00337>
- Schofield, M.R. and Bonner, S.J. (2015). Connecting the latent multinomial. *Biometrics* **71** 1070–1080. [MR3436732](#) <https://doi.org/10.1111/biom.12333>
- Sinclair, A. (1992). Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.* **1** 351–370. [MR1211324](#) <https://doi.org/10.1017/S0963548300000390>
- 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>
- Stanley, C. and Windisch, T. (2018). Heat-bath random walks with Markov bases. *Adv. in Appl. Math.* **92** 122–143. [MR3699116](#) <https://doi.org/10.1016/j.aam.2017.08.002>
- Takemura, A. and Aoki, S. (2004). Some characterizations of minimal Markov basis for sampling from discrete conditional distributions. *Ann. Inst. Statist. Math.* **56** 1–17. [MR2053726](#) <https://doi.org/10.1007/BF02530522>
- R Core Team (2022). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
- Tebaldi, C. and West, M. (1998). Bayesian inference on network traffic using link count data (with discussion). *J. Amer. Statist. Assoc.* **93** 557–576. [MR1631325](#) <https://doi.org/10.2307/2670105>
- Vardi, Y. (1996). Network tomography: Estimating source-destination traffic intensities from link data. *J. Amer. Statist. Assoc.* **91** 365–377. [MR1394093](#) <https://doi.org/10.2307/2291416>
- Windisch, T. (2016). Rapid mixing and Markov bases. *SIAM J. Discrete Math.* **30** 2130–2145. [MR3573306](#) <https://doi.org/10.1137/15M1022045>

# On the joint distribution of the area and the number of peaks for Bernoulli excursions

VLADISLAV KARGIN<sup>a</sup>

*Department of Mathematics and Statistics, Binghamton University, Binghamton, USA, <sup>a</sup>[vkargin@binghamton.edu](mailto:vkargin@binghamton.edu)*

Let  $P_n$  be a random Bernoulli excursion of length  $2n$ . We show that the area under  $P_n$  and the number of peaks of  $P_n$  are asymptotically independent. We also show that these statistics have the correlation coefficient asymptotic to  $c/\sqrt{n}$  for large  $n$ , where  $c < 0$ , and explicitly compute the coefficient  $c$ .

*Keywords:* Airy distribution; Bernoulli excursion; dominant balance method

## References

- Bailey, W.N. (1935). *Generalized Hypergeometric Series. Cambridge Tracts in Mathematics and Mathematical Physics* **32**. New York: Stechert-Hafner, Inc. [MR0185155](#)
- Berndt, B.C. and Yee, A.J. (2003). On the generalized Rogers–Ramanujan continued fraction. *Ramanujan J.* **7** 321–331. [MR2035809](#) <https://doi.org/10.1023/A:1026215700375>
- Blanco, S.A. and Petersen, T.K. (2014). Counting Dyck paths by area and rank. *Ann. Comb.* **18** 171–197. [MR3206149](#) <https://doi.org/10.1007/s00026-014-0218-9>
- Chassaing, P., Marckert, J.F. and Yor, M. (2000). The height and width of simple trees. In *Mathematics and Computer Science (Versailles, 2000)* (D. Gardy and A. Mokkadem, eds.). *Trends Math.* 17–30. Basel: Birkhäuser. [MR1798284](#)
- Csörgő, M., Shi, Z. and Yor, M. (1999). Some asymptotic properties of the local time of the uniform empirical process. *Bernoulli* **5** 1035–1058. [MR1735784](#) <https://doi.org/10.2307/3318559>
- Drmotá, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Vienna: SpringerWien-NewYork. [MR2484382](#) <https://doi.org/10.1007/978-3-211-75357-6>
- Fill, J.A. and Janson, S. (2009). Precise logarithmic asymptotics for the right tails of some limit random variables for random trees. *Ann. Comb.* **12** 403–416. [MR2496125](#) <https://doi.org/10.1007/s00026-009-0006-0>
- Flajolet, P. and Louchard, G. (2001). Analytic variations on the Airy distribution. *Algorithmica* **31** 361–377. [MR1855255](#) <https://doi.org/10.1007/s00453-001-0056-0>
- Flajolet, P. and Sedgewick, R. (2009). *Analytic Combinatorics*. Cambridge: Cambridge Univ. Press.
- Janson, S. (2001). Moment convergence in conditional limit theorems. *J. Appl. Probab.* **38** 421–437. [MR1834751](#) <https://doi.org/10.1017/s002190020001994x>
- Janson, S. (2007). Brownian excursion area, Wright’s constants in graph enumeration, and other Brownian areas. *Probab. Surv.* **4** 80–145. [MR2318402](#) <https://doi.org/10.1214/07-PS104>
- Janson, S. (2008). On the asymptotic joint distribution of height and width in random trees. *Studia Sci. Math. Hungar.* **45** 451–467. [MR2641443](#) <https://doi.org/10.1556/SScMath.2007.1064>
- Kolchin, V.F. (1986). *Random Mappings. Translation Series in Mathematics and Engineering*. New York: Optimization Software, Inc. [MR0865130](#)
- Labarbe, J.-M. and Marckert, J.-F. (2007). Asymptotics of Bernoulli random walks, bridges, excursions and meanders with a given number of peaks. *Electron. J. Probab.* **12** 229–261. [MR2299918](#) <https://doi.org/10.1214/EJP.v12-397>
- Louchard, G. (1984a). Kac’s formula, Levy’s local time and Brownian excursion. *J. Appl. Probab.* **21** 479–499. [MR0752014](#) <https://doi.org/10.2307/3213611>
- Louchard, G. (1984b). The Brownian excursion area: A numerical analysis. *Comput. Math. Appl.* **10** 413–417. [MR0783514](#) [https://doi.org/10.1016/0898-1221\(84\)90071-3](https://doi.org/10.1016/0898-1221(84)90071-3)

- Narayana, T.V. (1959). A partial order and its applications to probability theory. *Sankhyā* **21** 91–98. [MR0106498](#)
- Nguyen The, M. (2004). Area and inertial moment of Dyck paths. *Combin. Probab. Comput.* **13** 697–716. [MR2095979](#) <https://doi.org/10.1017/S0963548304006339>
- Petersen, T.K. (2015). *Eulerian Numbers. Birkhäuser Advanced Texts: Basler Lehrbücher [Birkhäuser Advanced Texts: Basel Textbooks]*. New York: Birkhäuser/Springer. [MR3408615](#) <https://doi.org/10.1007/978-1-4939-3091-3>
- Richard, C. (2002). Scaling behaviour of two-dimensional polygon models. *J. Stat. Phys.* **108** 459–493. [MR1913305](#) <https://doi.org/10.1023/A:1015773723188>
- Richard, C. (2009). On  $q$ -functional equations and excursion moments. *Discrete Math.* **309** 207–230. [MR2475013](#) <https://doi.org/10.1016/j.disc.2007.12.072>
- Richard, C. and Guttman, A.J. (2001).  $q$ -linear approximants: Scaling functions for polygon models. *J. Phys. A* **34** 4783–4796. [MR1850753](#) <https://doi.org/10.1088/0305-4470/34/23/301>
- Richard, C., Guttman, A.J. and Jensen, I. (2001). Scaling function and universal amplitude combinations for self-avoiding polygons. *J. Phys. A* **34** L495–L501. [MR1862759](#) <https://doi.org/10.1088/0305-4470/34/36/102>
- Richard, C., Jensen, I. and Guttman, A.J. (2008). Area distribution and scaling function for punctured polygons. *Electron. J. Combin.* **15** Research paper 53, 50 pp. [MR2398845](#) <https://doi.org/10.37236/777>
- Schwerdtfeger, U., Richard, C. and Thatte, B. (2010). Area limit laws for symmetry classes of staircase polygons. *Combin. Probab. Comput.* **19** 441–461. [MR2607376](#) <https://doi.org/10.1017/S0963548309990629>
- Stanley, R.P. (1999). *Enumerative Combinatorics. Vol. 2. Cambridge Studies in Advanced Mathematics* **62**. Cambridge: Cambridge Univ. Press. [MR1676282](#) <https://doi.org/10.1017/CBO9780511609589>
- Stanley, R.P. (2015). *Catalan Numbers*. New York: Cambridge Univ. Press. [MR3467982](#) <https://doi.org/10.1017/CBO9781139871495>
- Stein, E.M. and Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces. Princeton Mathematical Series* **32**. Princeton, NJ: Princeton Univ. Press. [MR0304972](#)
- Takács, L. (1991). A Bernoulli excursion and its various applications. *Adv. in Appl. Probab.* **23** 557–585. [MR1122875](#) <https://doi.org/10.2307/1427622>
- Takács, L. (1992). Random walk processes and their applications in order statistics. *Ann. Appl. Probab.* **2** 435–459. [MR1161061](#)
- Takács, L. (1994). On the total heights of random rooted binary trees. *J. Combin. Theory Ser. B* **61** 155–166. [MR1280604](#) <https://doi.org/10.1006/jctb.1994.1041>



# Shrinkage estimation of higher-order Bochner integrals

SAITEJA UTPALA<sup>1,a</sup> and BHARATH K. SRIPERUMBUDUR<sup>2,b</sup>

<sup>1</sup>Wadhvani AI, New Delhi, India, [saitejautpala@gmail.com](mailto:saitejautpala@gmail.com)

<sup>2</sup>Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA, [bks18@psu.edu](mailto:bks18@psu.edu)

We consider shrinkage estimation of higher-order Hilbert space-valued Bochner integrals in a non-parametric setting. We propose estimators that shrink the  $U$ -statistic estimator of the Bochner integral towards a pre-specified target element in the Hilbert space. Depending on the degeneracy of the kernel of the  $U$ -statistic, we construct consistent shrinkage estimators and develop oracle inequalities comparing the risks of the  $U$ -statistic estimator and its shrinkage version. Surprisingly, we show that the shrinkage estimator designed by assuming complete degeneracy of the kernel of the  $U$ -statistic is a consistent estimator even when the kernel is not completely degenerate. This work subsumes and improves upon Muandet et al. (*J. Mach. Learn. Res.* **17** (2016) 48) and Zhou, Chen and Huang (*J. Multivariate Anal.* **169** (2019) 166–178), which only handle mean element and covariance operator estimation in a reproducing kernel Hilbert space. We also specialize our results to normal mean estimation and show that for  $d \geq 3$ , the proposed estimator strictly improves upon the sample mean in terms of the mean squared error.

**Keywords:** Bernstein’s inequality; Bochner integral; completely degenerate; James-Stein estimator; shrinkage estimation; SURE;  $U$ -statistics

## References

- Arcones, M.A. and Giné, E. (1993). Limit theorems for  $U$ -processes. *Ann. Probab.* **21** 1494–1542. [MR1235426](#)
- Aronszajn, N. (1950). Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** 337–404. [MR0051437](#) <https://doi.org/10.2307/1990404>
- Balashubramanian, K., Li, T. and Yuan, M. (2021). On the optimality of kernel-embedding based goodness-of-fit tests. *J. Mach. Learn. Res.* **22** Paper No. 1. [MR4253694](#)
- Brandwein, A.C. and Strawderman, W.E. (1990). Stein estimation: The spherically symmetric case. *Statist. Sci.* **5** 356–369. [MR1080957](#)
- Brandwein, A.C. and Strawderman, W.E. (2012). Stein estimation for spherically symmetric distributions: Recent developments. *Statist. Sci.* **27** 11–23. [MR2953492](#) <https://doi.org/10.1214/10-STS323>
- Chen, Y., Wiesel, A., Eldar, Y.C. and Hero, A.O. (2010). Shrinkage algorithms for MMSE covariance estimation. *IEEE Trans. Signal Process.* **58** 5016–5029. [MR2722661](#) <https://doi.org/10.1109/TSP.2010.2053029>
- de la Peña, V.H. and Giné, E. (2012). *Decoupling: From Dependence to Independence*. New York: Springer Science & Business Media. [MR1666908](#) <https://doi.org/10.1007/978-1-4612-0537-1>
- Dinculeanu, N. (2000). *Vector Integration and Stochastic Integration in Banach Spaces. Pure and Applied Mathematics (New York)*. New York: Wiley Interscience. [MR1782432](#) <https://doi.org/10.1002/9781118033012>
- Fisher, T.J. and Sun, X. (2011). Improved Stein-type shrinkage estimators for the high-dimensional multivariate normal covariance matrix. *Comput. Statist. Data Anal.* **55** 1909–1918. [MR2765053](#) <https://doi.org/10.1016/j.csda.2010.12.006>
- Fukumizu, K., Bach, F.R. and Jordan, M.I. (2004). Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces. *J. Mach. Learn. Res.* **5** 73–99. [MR2247974](#) <https://doi.org/10.1162/153244303768966111>
- Gretton, A., Fukumizu, K., Teo, C., Song, L., Schölkopf, B. and Smola, A. (2007). A kernel statistical test of independence. In *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer and S. Roweis, eds.) **20**. Curran Associates.

- Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773. [MR2913716](#)
- James, W. and Stein, C. (1960). Estimation with quadratic loss. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. I* 361–379. Berkeley-Los Angeles, CA: Univ. California Press. [MR0133191](#)
- 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>
- Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivariate Anal.* **88** 365–411. [MR2026339](#) [https://doi.org/10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)
- Ledoit, O. and Wolf, M. (2018). Optimal estimation of a large-dimensional covariance matrix under Stein’s loss. *Bernoulli* **24** 3791–3832. [MR3788189](#) <https://doi.org/10.3150/17-BEJ979>
- Lee, A.J. (2019). *U-Statistics: Theory and Practice*. Routledge.
- Muandet, K., Sriperumbudur, B., Fukumizu, K., Gretton, A. and Schölkopf, B. (2016). Kernel mean shrinkage estimators. *J. Mach. Learn. Res.* **17** Paper No. 48. [MR3504608](#)
- Song, L., Smola, A., Gretton, A., Bedo, J. and Borgwardt, K. (2012). Feature selection via dependence maximization. *J. Mach. Learn. Res.* **13** 1393–1434. [MR2930643](#)
- Stein, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. I* 197–206. Berkeley-Los Angeles, CA: Univ. California Press. [MR0084922](#)
- Stein, C. (1975). Estimation of a covariance matrix. Rietz Lecture, 39th Annual Meeting, Atlanta, GA.
- Touloumis, A. (2015). Nonparametric Stein-type shrinkage covariance matrix estimators in high-dimensional settings. *Comput. Statist. Data Anal.* **83** 251–261. [MR3281809](#) <https://doi.org/10.1016/j.csda.2014.10.018>
- Utpala, S. and Sriperumbudur, B.K. (2024). Supplement to “Shrinkage estimation of higher-order Bochner integrals.” <https://doi.org/10.3150/23-BEJ1692SUPP>
- Zhou, Y., Chen, D.-R. and Huang, W. (2019). A class of optimal estimators for the covariance operator in reproducing kernel Hilbert spaces. *J. Multivariate Anal.* **169** 166–178. [MR3875593](#) <https://doi.org/10.1016/j.jmva.2018.09.003>

# Smoothed circulars: Nonparametric estimation of circular cumulative distribution functions and circulars

JOSE AMEIJERAS-ALONSO<sup>1,a</sup> and IRÈNE GIJBELS<sup>2,b</sup>

<sup>1</sup>*CITMaga, Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Spain, [jose.ameijeiras@usc.es](mailto:jose.ameijeiras@usc.es)*

<sup>2</sup>*Department of Mathematics, KU Leuven, Belgium, [irene.gijbels@kuleuven.be](mailto:irene.gijbels@kuleuven.be)*

Copulas are an important tool to study dependencies for data on the real line (or multivariate extensions of this), referred to as linear data. The analogue of copulas for circular data and data on the torus are circulars. This paper studies kernel estimation of circulars, and discusses important issues such as choice of circular kernels and ‘smoothing’ parameters. This leads to some new insights, and some contrasts with results for linear data. Since a circular is a multivariate cumulative distribution with circular uniform marginals, the paper also contributes to kernel estimation of cumulative distributions for toroidal data.

*Keywords:* Asymptotic bias and variance; circular data; copula; kernel estimation; mean squared error; optimal smoothing parameter; toroidal data

## References

- Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.
- Ameijeiras-Alonso, J. and Gijbels, I. (2024). Supplement to “Smoothed circulars: Nonparametric estimation of circular cumulative distribution functions and circulars.” <https://doi.org/10.3150/23-BEJ1693SUPPB>
- Bowman, A., Hall, P. and Prvan, T. (1998). Bandwidth selection for the smoothing of distribution functions. *Biometrika* **85** 799–808. [MR1666695 https://doi.org/10.1093/biomet/85.4.799](https://doi.org/10.1093/biomet/85.4.799)
- Card, G. and Dickinson, M.H. (2008). Visually mediated motor planning in the escape response of *Drosophila*. *Curr. Biol.* **18** 1300–1307. <https://doi.org/10.1016/j.cub.2008.07.094>
- Chen, S.X. and Huang, T.-M. (2007). Nonparametric estimation of copula functions for dependence modelling. *Canad. J. Statist.* **35** 265–282. [MR2393609 https://doi.org/10.1002/cjs.5550350205](https://doi.org/10.1002/cjs.5550350205)
- Di Marzio, M., Panzera, A. and Taylor, C.C. (2009). Local polynomial regression for circular predictors. *Statist. Probab. Lett.* **79** 2066–2075. [MR2571770 https://doi.org/10.1016/j.spl.2009.06.014](https://doi.org/10.1016/j.spl.2009.06.014)
- Di Marzio, M., Panzera, A. and Taylor, C.C. (2011). Kernel density estimation on the torus. *J. Statist. Plann. Inference* **141** 2156–2173. [MR2772221 https://doi.org/10.1016/j.jspi.2011.01.002](https://doi.org/10.1016/j.jspi.2011.01.002)
- Di Marzio, M., Panzera, A. and Taylor, C.C. (2012). Smooth estimation of circular cumulative distribution functions and quantiles. *J. Nonparametr. Stat.* **24** 935–949. [MR2995485 https://doi.org/10.1080/10485252.2012.721517](https://doi.org/10.1080/10485252.2012.721517)
- Duong, T. (2016). Non-parametric smoothed estimation of multivariate cumulative distribution and survival functions, and receiver operating characteristic curves. *J. Korean Statist. Soc.* **45** 33–50. [MR3456320 https://doi.org/10.1016/j.jkss.2015.06.002](https://doi.org/10.1016/j.jkss.2015.06.002)
- Fernández-Durán, J.J. (2007). Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums. *Biometrics* **63** 579–585. [MR2370817 https://doi.org/10.1111/j.1541-0420.2006.00716.x](https://doi.org/10.1111/j.1541-0420.2006.00716.x)
- Fisher, N.I. and Lee, A.J. (1983). A correlation coefficient for circular data. *Biometrika* **70** 327–332. [MR0712021 https://doi.org/10.1093/biomet/70.2.327](https://doi.org/10.1093/biomet/70.2.327)

- Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 1096–1127. [MR2469337](#) <https://doi.org/10.1214/07-AIHP148>
- Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.* **44** 199–213. [MR2517885](#) <https://doi.org/10.1016/j.insmatheco.2007.10.005>
- Heredia-Zavoni, E. and Montes-Iturrizaga, R. (2019). Modeling directional environmental contours using three dimensional vine copulas. *Ocean Eng.* **187** 106102.
- Jammalamadaka, S.R. and Ramakrishna Sarma, Y. (1988). A correlation coefficient for angular variables. In *Statistical Theory and Data Analysis, II (Tokyo, 1986)* 349–364. Amsterdam: North-Holland. [MR0999196](#)
- Johnson, R.A. and Wehrly, T. (1977). Measures and models for angular correlation and angular-linear correlation. *J. Roy. Statist. Soc. Ser. B* **39** 222–229. [MR0494693](#)
- Jones, M.C. (1990). The performance of kernel density functions in kernel distribution function estimation. *Statist. Probab. Lett.* **9** 129–132. [MR1045173](#) [https://doi.org/10.1016/0167-7152\(92\)90006-Q](https://doi.org/10.1016/0167-7152(92)90006-Q)
- Jones, M.C., Pewsey, A. and Kato, S. (2015). On a class of circulas: Copulas for circular distributions. *Ann. Inst. Statist. Math.* **67** 843–862. [MR3390169](#) <https://doi.org/10.1007/s10463-014-0493-6>
- Jupp, P.E. (2015). Copulae on products of compact Riemannian manifolds. *J. Multivariate Anal.* **140** 92–98. [MR3372555](#) <https://doi.org/10.1016/j.jmva.2015.04.008>
- Jupp, P.E. and Mardia, K.V. (1980). A general correlation coefficient for directional data and related regression problems. *Biometrika* **67** 163–173. [MR0570518](#) <https://doi.org/10.1093/biomet/67.1.163>
- Justel, A., Peña, D. and Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statist. Probab. Lett.* **35** 251–259. [MR1484961](#) [https://doi.org/10.1016/S0167-7152\(97\)00020-5](https://doi.org/10.1016/S0167-7152(97)00020-5)
- Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics. Wiley Series in Probability and Statistics*. Chichester: Wiley. [MR1828667](#)
- Oliveira, M., Crujeiras, R.M. and Rodríguez-Casal, A. (2012). A plug-in rule for bandwidth selection in circular density estimation. *Comput. Statist. Data Anal.* **56** 3898–3908. [MR2957840](#) <https://doi.org/10.1016/j.csda.2012.05.021>
- Omelka, M., Gijbels, I. and Veraverbeke, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. *Ann. Statist.* **37** 3023–3058. [MR2541454](#) <https://doi.org/10.1214/08-AOS666>
- Pewsey, A., Lewis, T. and Jones, M.C. (2007). The wrapped  $t$  family of circular distributions. *Aust. N. Z. J. Stat.* **49** 79–91. [MR2345412](#) <https://doi.org/10.1111/j.1467-842X.2006.00465.x>
- Rivest, L.-P. (1982). Some statistical methods for bivariate circular data. *J. Roy. Statist. Soc. Ser. B* **44** 81–90. [MR0655377](#)
- Shieh, G.S., Zheng, S., Johnson, R.A., Chang, Y.-F., Shimizu, K., Wang, C.-C. and Tang, S.-L. (2011). Modeling and comparing the organization of circular genomes. *Bioinformatics* **27** 912–918.
- Shimizu, K. and Iida, K. (2002). Pearson type VII distributions on spheres. *Comm. Statist. Theory Methods* **31** 513–526. [MR1902308](#) <https://doi.org/10.1081/STA-120003131>
- Taylor, C.C. (2008). Automatic bandwidth selection for circular density estimation. *Comput. Statist. Data Anal.* **52** 3493–3500. [MR2427365](#) <https://doi.org/10.1016/j.csda.2007.11.003>
- Tsuruta, Y. and Sagae, M. (2017). Asymptotic property of wrapped Cauchy kernel density estimation on the circle. *Bull. Inform. Cybernet.* **49** 1–10. [MR3728316](#)

# Local polynomial trend regression for spatial data on $\mathbb{R}^d$

DAISUKE KURISU<sup>1,a</sup> and YASUMASA MATSUDA<sup>2,b</sup>

<sup>1</sup>Center for Spatial Information Science, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan, <sup>a</sup>[daisukekurisu@csis.u-tokyo.ac.jp](mailto:daisukekurisu@csis.u-tokyo.ac.jp)

<sup>2</sup>Graduate School of Economics and Management, Tohoku University, Sendai 980-8576, Japan,

<sup>b</sup>[yasumasa.matsuda.a4@tohoku.ac.jp](mailto:yasumasa.matsuda.a4@tohoku.ac.jp)

This paper develops a general asymptotic theory of local polynomial (LP) regression for spatial data observed at irregularly spaced locations in a sampling region  $R_n \subset \mathbb{R}^d$ . We adopt a stochastic sampling design that can generate irregularly spaced sampling sites in a flexible manner including both pure increasing and mixed increasing domain frameworks. We first introduce a nonparametric regression model for spatial data defined on  $\mathbb{R}^d$  and then establish the asymptotic normality of LP estimators with general order  $p \geq 1$ . We also propose methods for constructing confidence intervals and establish uniform convergence rates of LP estimators. Our dependence structure conditions on the underlying processes cover a wide class of random fields such as Lévy-driven continuous autoregressive moving average random fields. As an application of our main results, we discuss a two-sample testing problem for mean functions and their partial derivatives.

**Keywords:** Irregularly spaced spatial data; Lévy-driven moving average random field; local polynomial regression; two-sample test

## References

- 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) <https://doi.org/10.1214/14-AOS1291>
- Bertoin, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press. [MR1406564](https://doi.org/10.1017/CBO9780511526378)
- 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) <https://doi.org/10.1111/rssb.12197>
- Calonico, S., Cattaneo, M.D. and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* **82** 2295–2326. [MR3301169](https://doi.org/10.3982/ECTA11757) <https://doi.org/10.3982/ECTA11757>
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. [MR1429916](https://doi.org/10.1214/aos/1034276620) <https://doi.org/10.1214/aos/1034276620>
- Dahlhaus, R. (2012). Locally stationary processes. *Handb. Statist.* **30** 351–413.
- Diggle, P.J., Tawn, J.A. and Moyeed, R.A. (1998). Model-based geostatistics. *J. R. Stat. Soc., Ser. C* **47** 299–350. With discussion and a reply by the authors. [MR1626544](https://doi.org/10.1111/1467-9876.00113) <https://doi.org/10.1111/1467-9876.00113>
- Ehrlich, M. and Seidel, T. (2018). The persistent effects of place-based policy: Evidence from the West-German Zonenrandgebiet. *Amer. Econ. J.: Econ. Policy* **10** 344–74.
- El Machkouri, M., Es-Sebאי, K. and Ouassou, I. (2017). On local linear regression for strongly mixing random fields. *J. Multivariate Anal.* **156** 103–115. [MR3624688](https://doi.org/10.1016/j.jmva.2017.02.002) <https://doi.org/10.1016/j.jmva.2017.02.002>
- El Machkouri, M. and Stoica, R. (2010). Asymptotic normality of kernel estimates in a regression model for random fields. *J. Nonparametr. Stat.* **22** 955–971. [MR2738877](https://doi.org/10.1080/10485250903505893) <https://doi.org/10.1080/10485250903505893>
- Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods*. *Springer Series in Statistics*. New York: Springer. [MR1964455](https://doi.org/10.1007/b97702) <https://doi.org/10.1007/b97702>
- Hahn, J., Todd, P. and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* **69** 201–209.

- 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) <https://doi.org/10.1214/009053604000000850>
- 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) <https://doi.org/10.3150/08-BEJ168>
- 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) <https://doi.org/10.1017/S0266466608080304>
- 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) <https://doi.org/10.1016/j.jeconom.2011.11.008>
- Keele, L.J. and Titiunik, R. (2015). Geographic boundaries as regression discontinuities. *Polit. Anal.* **23** 127–155.
- 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) <https://doi.org/10.1017/S0266466609090744>
- Kurusu, 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) <https://doi.org/10.1016/j.spl.2019.06.019>
- Kurusu, D. (2022). Nonparametric regression for locally stationary random fields under stochastic sampling design. *Bernoulli* **28** 1250–1275. [MR4388937](https://doi.org/10.3150/21-bej1385) <https://doi.org/10.3150/21-bej1385>
- Kurusu, D., Kato, K. and Shao, X. (2023). Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. *J. Amer. Statist. Assoc.* in press. <https://doi.org/10.1080/01621459.2023.2218578>
- Kurusu, D. and Matsuda, Y. (2024). Supplement to “Local polynomial trend regression for spatial data on  $\mathbb{R}^d$ .” <https://doi.org/10.3150/23-BEJ1694SUPP>
- Lahiri, S.N. (1996). On inconsistency of estimators based on spatial data under infill asymptotics. *Sankhyā Ser. A* **58** 403–417. [MR1659130](https://doi.org/10.2307/2346310)
- Lahiri, S.N. (1999). Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probab. Theory Related Fields* **114** 55–84. [MR1697139](https://doi.org/10.1007/s004400050221) <https://doi.org/10.1007/s004400050221>
- Lahiri, S.N. (2003a). Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **65** 356–388. [MR2028905](https://doi.org/10.2307/2346310)
- Lahiri, S.N. (2003b). *Resampling Methods for Dependent Data*. Springer Series in Statistics. New York: Springer. [MR2001447](https://doi.org/10.1007/978-1-4757-3803-2) <https://doi.org/10.1007/978-1-4757-3803-2>
- 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/10.1214/009053606000000551) <https://doi.org/10.1214/009053606000000551>
- Lu, Z. and Chen, X. (2002). Spatial nonparametric regression estimation: Non-isotropic case. *Acta Math. Appl. Sin. Engl. Ser.* **18** 641–656. [MR2012328](https://doi.org/10.1007/s102550200067) <https://doi.org/10.1007/s102550200067>
- Lu, Z. and Chen, X. (2004). Spatial kernel regression estimation: Weak consistency. *Statist. Probab. Lett.* **68** 125–136. [MR2066167](https://doi.org/10.1016/j.spl.2003.08.014) <https://doi.org/10.1016/j.spl.2003.08.014>
- 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) <https://doi.org/10.1080/01621459.2014.947376>
- Mardia, K.V. and Marshall, R.J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* **71** 135–146. [MR0738334](https://doi.org/10.1093/biomet/71.1.135) <https://doi.org/10.1093/biomet/71.1.135>
- Masry, E. (1996a). 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) <https://doi.org/10.1111/j.1467-9892.1996.tb00294.x>
- Masry, E. (1996b). Multivariate regression estimation local polynomial fitting for time series. *Stochastic Process. Appl.* **65** 81–101. [MR1422881](https://doi.org/10.1016/S0304-4149(96)00095-6) [https://doi.org/10.1016/S0304-4149\(96\)00095-6](https://doi.org/10.1016/S0304-4149(96)00095-6)
- Masry, E. and Fan, J. (1997). Local polynomial estimation of regression functions for mixing processes. *Scand. J. Stat.* **24** 165–179. [MR1455865](https://doi.org/10.1111/1467-9469.00056) <https://doi.org/10.1111/1467-9469.00056>
- 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) <https://doi.org/10.1111/j.1467-9868.2008.00685.x>
- 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) <https://doi.org/10.1007/s42081-018-0003-9>
- 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) <https://doi.org/10.1007/BF00339998>
- 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) <https://doi.org/10.1016/j.jeconom.2011.05.002>

- Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. Translated from the 1990 Japanese original, Revised by the author. [MR1739520](#)
- Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. [MR3097614](#) <https://doi.org/10.1214/12-AOS1043>
- Yu, B. (1994). Rates of convergence for empirical processes of stationary mixing sequences. *Ann. Probab.* **22** 94–116. [MR1258867](#)
- Zhang, H. (2002). On estimation and prediction for spatial generalized linear mixed models. *Biometrics* **58** 129–136. [MR1891051](#) <https://doi.org/10.1111/j.0006-341X.2002.00129.x>
- 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>
- 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>
- 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>

# Penalized spline estimation of principal components for sparse functional data: Rates of convergence

SHIYUAN HE<sup>1,2,a</sup>, JIANHUA Z. HUANG<sup>3,c</sup> and KEJUN HE<sup>2,b</sup>

<sup>1</sup>*School of Mathematics and Statistics, Beijing Technology and Business University, Beijing, China,*

<sup>a</sup>[20240101@btbu.edu.cn](mailto:20240101@btbu.edu.cn)

<sup>2</sup>*Center for Applied Statistics, Institute of Statistics and Big Data, Renmin University of China, Beijing, China,*

<sup>b</sup>[kejunhe@ruc.edu.cn](mailto:kejunhe@ruc.edu.cn)

<sup>3</sup>*School of Data Science, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Shenzhen, China,*

<sup>c</sup>[jhuang@cuhk.edu.cn](mailto:jhuang@cuhk.edu.cn)

This paper gives a comprehensive treatment of the convergence rates of penalized spline estimators for simultaneously estimating several leading principal component functions, when the functional data is sparsely observed. The penalized spline estimators are defined as the solution of a penalized empirical risk minimization problem, where the loss function belongs to a general class of loss functions motivated by the matrix Bregman divergence, and the penalty term is the integrated squared derivative. The theory reveals that the asymptotic behavior of penalized spline estimators depends on the interesting interplay between several factors, i.e., the smoothness of the unknown functions, the spline degree, the spline knot number, the penalty order, and the penalty parameter. The theory also classifies the asymptotic behavior into seven scenarios and characterizes whether and how the minimax optimal rates of convergence are achievable in each scenario.

**Keywords:** Functional principal component analysis; manifold geometry; matrix Bregman divergence; roughness penalty

## References

- [1] Bhatia, R. (2007). *Positive Definite Matrices. Princeton Series in Applied Mathematics.* Princeton, NJ: Princeton Univ. Press. [MR2284176](#)
- [2] Bhatia, R. (2013). *Matrix Analysis.* New York: Springer.
- [3] Cai, T.T. and Yuan, M. (2010). Nonparametric covariance function estimation for functional and longitudinal data. Technical Report.
- [4] Chen, L. and Huang, J.Z. (2012). Sparse reduced-rank regression for simultaneous dimension reduction and variable selection. *J. Amer. Statist. Assoc.* **107** 1533–1545. [MR3036414](#) <https://doi.org/10.1080/01621459.2012.734178>
- [5] Chiou, J.-M., Chen, Y.-T. and Yang, Y.-F. (2014). Multivariate functional principal component analysis: A normalization approach. *Statist. Sinica* **24** 1571–1596. [MR3308652](#)
- [6] Chiou, J.-M. and Li, P.-L. (2007). Functional clustering and identifying substructures of longitudinal data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 679–699. [MR2370075](#) <https://doi.org/10.1111/j.1467-9868.2007.00605.x>
- [7] Claeskens, G., Krivobokova, T. and Opsomer, J.D. (2009). Asymptotic properties of penalized spline estimators. *Biometrika* **96** 529–544. [MR2538755](#) <https://doi.org/10.1093/biomet/asp035>
- [8] Dhillon, I.S. and Tropp, J.A. (2007). Matrix nearness problems with Bregman divergences. *SIAM J. Matrix Anal. Appl.* **29** 1120–1146. [MR2369287](#) <https://doi.org/10.1137/060649021>
- [9] Ding, F., He, S., Jones, D.E. and Huang, J.Z. (2022). Functional PCA with covariate-dependent mean and covariance structure. *Technometrics* **64** 335–345. [MR4457327](#) <https://doi.org/10.1080/00401706.2021.2008502>



- [10] Edelman, A., Arias, T.A. and Smith, S.T. (1999). The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.* **20** 303–353. [MR1646856](#) <https://doi.org/10.1137/S0895479895290954>
- [11] Eilers, P.H.C. and Marx, B.D. (1996). Flexible smoothing with  $B$ -splines and penalties. *Statist. Sci.* **11** 89–121. [MR1435485](#) <https://doi.org/10.1214/ss/1038425655>
- [12] Gu, C. (2013). *Smoothing Spline ANOVA Models*, 2nd ed. *Springer Series in Statistics* **297**. New York: Springer. [MR3025869](#) <https://doi.org/10.1007/978-1-4614-5369-7>
- [13] Hall, P., Müller, H.-G. and Wang, J.-L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *Ann. Statist.* **34** 1493–1517. [MR2278365](#) <https://doi.org/10.1214/009053606000000272>
- [14] Happ, C. and Greven, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J. Amer. Statist. Assoc.* **113** 649–659. [MR3832216](#) <https://doi.org/10.1080/01621459.2016.1273115>
- [15] He, S., Huang, J.Z. and He, K. (2024). Supplement to “Penalized spline estimation of principal components for sparse functional data: rates of convergence.” <https://doi.org/10.3150/23-BEJ1695SUPP>
- [16] He, S., Wang, L. and Huang, J.Z. (2018). Characterization of type Ia supernova light curves using principal component analysis of sparse functional data. *Astrophys. J.* **857** 110.
- [17] He, S., Ye, H. and He, K. (2022). Spline estimation of functional principal components via manifold conjugate gradient algorithm. *Stat. Comput.* **32** 106. [MR4507157](#) <https://doi.org/10.1007/s11222-022-10175-2>
- [18] Hiai, F. and Petz, D. (2014). *Introduction to Matrix Analysis and Applications*. *Universitext*. New York: Springer. [MR3184500](#) <https://doi.org/10.1007/978-3-319-04150-6>
- [19] Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. [MR3379106](#) <https://doi.org/10.1002/9781118762547>
- [20] Huang, J.Z. (2003). Local asymptotics for polynomial spline regression. *Ann. Statist.* **31** 1600–1635. [MR2012827](#) <https://doi.org/10.1214/aos/1065705120>
- [21] Huang, J.Z. and Su, Y. (2021). Asymptotic properties of penalized spline estimators in concave extended linear models: Rates of convergence. *Ann. Statist.* **49** 3383–3407. [MR4352534](#) <https://doi.org/10.1214/21-aos2088>
- [22] James, G.M. and Hastie, T.J. (2001). Functional linear discriminant analysis for irregularly sampled curves. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 533–550. [MR1858401](#) <https://doi.org/10.1111/1467-9868.00297>
- [23] James, G.M., Hastie, T.J. and Sugar, C.A. (2000). Principal component models for sparse functional data. *Biometrika* **87** 587–602. [MR1789811](#) <https://doi.org/10.1093/biomet/87.3.587>
- [24] James, G.M. and Sugar, C.A. (2003). Clustering for sparsely sampled functional data. *J. Amer. Statist. Assoc.* **98** 397–408. [MR1995716](#) <https://doi.org/10.1198/0162145030000189>
- [25] Kauermann, G., Krivobokova, T. and Fahrmeir, L. (2009). Some asymptotic results on generalized penalized spline smoothing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 487–503. [MR2649606](#) <https://doi.org/10.1111/j.1467-9868.2008.00691.x>
- [26] Kulis, B., Sustik, M.A. and Dhillon, I.S. (2009). Low-rank kernel learning with Bregman matrix divergences. *J. Mach. Learn. Res.* **10** 341–376. [MR2485986](#)
- [27] Li, Y. and Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. [MR2766854](#) <https://doi.org/10.1214/10-AOS813>
- [28] Li, Y. and Ruppert, D. (2008). On the asymptotics of penalized splines. *Biometrika* **95** 415–436. [MR2521591](#) <https://doi.org/10.1093/biomet/asn010>
- [29] Paul, D. and Peng, J. (2009). Consistency of restricted maximum likelihood estimators of principal components. *Ann. Statist.* **37** 1229–1271. [MR2509073](#) <https://doi.org/10.1214/08-AOS608>
- [30] Peng, J. and Paul, D. (2009). A geometric approach to maximum likelihood estimation of the functional principal components from sparse longitudinal data. *J. Comput. Graph. Statist.* **18** 995–1015. [MR2598035](#) <https://doi.org/10.1198/jcgs.2009.08011>
- [31] Pitrik, J. and Virosztek, D. (2015). On the joint convexity of the Bregman divergence of matrices. *Lett. Math. Phys.* **105** 675–692. [MR3339204](#) <https://doi.org/10.1007/s11005-015-0757-y>
- [32] Ramsay, J.O. and Dalzell, C.J. (1991). Some tools for functional data analysis. *J. Roy. Statist. Soc. Ser. B* **53** 539–572. [MR1125714](#)

- [33] Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR2168993](#)
- [34] Rice, J.A. (2004). Functional and longitudinal data analysis: Perspectives on smoothing. *Statist. Sinica* **14** 631–647. [MR2087966](#)
- [35] Rice, J.A. and Silverman, B.W. (1991). Estimating the mean and covariance structure nonparametrically when the data are curves. *J. Roy. Statist. Soc. Ser. B* **53** 233–243. [MR1094283](#)
- [36] Ruppert, D., Wand, M.P. and Carroll, R.J. (2003). *Semiparametric Regression*. *Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge: Cambridge Univ. Press. [MR1998720](#) <https://doi.org/10.1017/CBO9780511755453>
- [37] Sang, P., Kong, D. and Yang, S. (2022). Functional principal component analysis for longitudinal observations with sampling at random. Available at <https://arxiv.org/abs/2203.14760>.
- [38] Shi, H., Yang, Y., Wang, L., Ma, D., Beg, M.F., Pei, J. and Cao, J. (2022). Two-dimensional functional principal component analysis for image feature extraction. *J. Comput. Graph. Statist.* **31** 1127–1140. [MR4513375](#) <https://doi.org/10.1080/10618600.2022.2035738>
- [39] Shin, Y.E., Zhou, L. and Ding, Y. (2022). Joint estimation of monotone curves via functional principal component analysis. *Comput. Statist. Data Anal.* **166** 107343. [MR4313495](#) <https://doi.org/10.1016/j.csda.2021.107343>
- [40] Silverman, B.W. (1996). Smoothed functional principal components analysis by choice of norm. *Ann. Statist.* **24** 1–24. [MR1389877](#) <https://doi.org/10.1214/aos/1033066196>
- [41] Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053. [MR0673642](#)
- [42] Talvila, E. (2001). Necessary and sufficient conditions for differentiating under the integral sign. *Amer. Math. Monthly* **108** 544–548. [MR1840661](#) <https://doi.org/10.2307/2695709>
- [43] Wang, X., Shen, J. and Ruppert, D. (2011). On the asymptotics of penalized spline smoothing. *Electron. J. Stat.* **5** 1–17. [MR2763795](#) <https://doi.org/10.1214/10-EJS593>
- [44] Xiao, L. (2019). Asymptotic theory of penalized splines. *Electron. J. Stat.* **13** 747–794. [MR3925516](#) <https://doi.org/10.1214/19-ejs1541>
- [45] Xiao, L. (2020). Asymptotic properties of penalized splines for functional data. *Bernoulli* **26** 2847–2875. [MR4140531](#) <https://doi.org/10.3150/20-BEJ1209>
- [46] Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. [MR2253106](#) <https://doi.org/10.1214/009053605000000660>
- [47] Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. [MR2160561](#) <https://doi.org/10.1198/016214504000001745>
- [48] Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321. [MR3546451](#) <https://doi.org/10.1214/16-AOS1446>
- [49] Zhou, L., Huang, J.Z. and Carroll, R.J. (2008). Joint modelling of paired sparse functional data using principal components. *Biometrika* **95** 601–619. [MR2443178](#) <https://doi.org/10.1093/biomet/asn035>
- [50] Zhou, L. and Pan, H. (2014). Principal component analysis of two-dimensional functional data. *J. Comput. Graph. Statist.* **23** 779–801. [MR3224656](#) <https://doi.org/10.1080/10618600.2013.827986>

# Optimal Markovian coupling for finite activity Lévy processes

WILFRID S. KENDALL<sup>1,a</sup>, MATEUSZ B. MAJKA<sup>2,c</sup> and  
ALEKSANDAR MIJATOVIĆ<sup>1,b</sup>

<sup>1</sup>*Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK, [w.s.kendall@warwick.ac.uk](mailto:w.s.kendall@warwick.ac.uk),*

<sup>b</sup>*[a.mijatovic@warwick.ac.uk](mailto:a.mijatovic@warwick.ac.uk)*

<sup>2</sup>*School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK,*

<sup>c</sup>*[m.majka@hw.ac.uk](mailto:m.majka@hw.ac.uk)*

We study optimal Markovian couplings of Markov processes, where the optimality is understood in terms of minimization of concave transport costs between evaluations of the coupled processes at corresponding times. We provide explicit constructions of such optimal couplings for one-dimensional finite-activity Lévy processes (continuous-time random walks) whose jump distributions are unimodal but not necessarily symmetric. Remarkably, the optimal Markovian coupling does not depend on the specific concave transport cost. To this end, we combine McCann’s results on optimal transport and Rogers’ results on random walks with a novel uniformization construction that allows us to characterize all Markovian couplings of finite-activity Lévy processes. In particular, we show that the optimal Markovian coupling for finite-activity Lévy processes with non-symmetric unimodal Lévy measures has to allow for non-simultaneous jumps of the two coupled processes.

*Keywords:* Concave transport cost; continuous-time random walk; finite activity Lévy process; immersion coupling; Lévy process; Markovian coupling; maximal coupling; optimal coupling; simultaneous optimality; unimodal distribution; Wasserstein distance

## References

- [1] Banerjee, S. and Kendall, W. (2018). Coupling polynomial Stratonovich integrals: The two-dimensional Brownian case. *Electron. J. Probab.* **23** 1–43. [MR3771761 https://doi.org/10.1214/18-EJP150](https://doi.org/10.1214/18-EJP150)
- [2] Banerjee, S. and Kendall, W.S. (2017). Rigidity for Markovian maximal couplings of elliptic diffusions. *Probab. Theory Related Fields* **168** 55–112. [MR3651049 https://doi.org/10.1007/s00440-016-0706-4](https://doi.org/10.1007/s00440-016-0706-4)
- [3] Böttcher, B. (2017). Markovian Maximal Coupling of Markov Processes. arXiv e-prints. Available at [arXiv: 1710.09654](https://arxiv.org/abs/1710.09654).
- [4] Bou-Rabee, N., Eberle, A. and Zimmer, R. (2020). Coupling and convergence for Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **30** 1209–1250. [MR4133372 https://doi.org/10.1214/19-AAP1528](https://doi.org/10.1214/19-AAP1528)
- [5] Brémaud, P. (1975). An extension of Watanabe’s theorem of characterization of Poisson processes over the positive real half line. *J. Appl. Probab.* **12** 396–399. [MR0380970 https://doi.org/10.2307/3212457](https://doi.org/10.2307/3212457)
- [6] Burdzy, K. and Kendall, W.S. (2000). Efficient Markovian couplings: Examples and counterexamples. *Ann. Appl. Probab.* **10** 362–409. [MR1768241 https://doi.org/10.1214/aoap/1019487348](https://doi.org/10.1214/aoap/1019487348)
- [7] Chau, N.H., Moulines, É., Rásonyi, M., Sabanis, S. and Zhang, Y. (2021). On stochastic gradient Langevin dynamics with dependent data streams: The fully nonconvex case. *SIAM J. Math. Data Sci.* **3** 959–986. [MR4313846 https://doi.org/10.1137/20M1355392](https://doi.org/10.1137/20M1355392)
- [8] Chen, M. (1994). Optimal Markovian couplings and applications. *Acta Math. Sin. New Ser.* **10** 260–275. A Chinese summary appears in *Acta Math. Sinica* **38** (1995), no. 4, 575. [MR1415697 https://doi.org/10.1007/BF02560717](https://doi.org/10.1007/BF02560717)
- [9] Chen, M.-F. (1994). Optimal couplings and application to Riemannian geometry. In *Probability Theory and Mathematical Statistics (Vilnius, 1993)* 121–142. Vilnius: TEV. [MR1649575](https://doi.org/10.1007/BF02560717)

- [10] Chen, M.-F. (2005). *Eigenvalues, Inequalities, and Ergodic Theory. Probability and Its Applications (New York)*. London: Springer London, Ltd. [MR2105651](#)
- [11] Cheng, X., Chatterji, N., Abbasi-Yadkori, Y., Bartlett, P.L. and Jordan, M.I. (2018). Sharp convergence rates for Langevin dynamics in the nonconvex setting. arXiv e-prints. Available at [arXiv:1805.01648](#).
- [12] Cheng, X., Chatterji, N.S., Bartlett, P.L. and Jordan, M.I. (2018). Underdamped Langevin MCMC: A non-asymptotic analysis. In *Proceedings of the 31st Conference on Learning Theory* (S. Bubeck, V. Perchet and P. Rigollet, eds.). *Proceedings of Machine Learning Research* **75** 300–323. PMLR.
- [13] Cheng, X., Yin, D., Bartlett, P. and Jordan, M. (2020). Stochastic gradient and Langevin processes. In *Proceedings of the 37th International Conference on Machine Learning* (H. Daumé III and A. Singh, eds.). *Proceedings of Machine Learning Research* **119** 1810–1819. PMLR, Virtual.
- [14] Connor, S. (2013). Optimal coadapted coupling for a random walk on the hyper-complete graph. *J. Appl. Probab.* **50** 1117–1130. [MR3161377](#) <https://doi.org/10.1239/jap/1389370103>
- [15] Connor, S. and Jacka, S. (2008). Optimal co-adapted coupling for the symmetric random walk on the hypercube. *J. Appl. Probab.* **45** 703–713. [MR2455179](#) <https://doi.org/10.1239/jap/1222441824>
- [16] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC *Financial Mathematics Series*. Boca Raton, FL: CRC Press/CRC. [MR2042661](#)
- [17] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I: Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications (New York)*. New York: Springer. [MR1950431](#)
- [18] Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. [MR3568041](#) <https://doi.org/10.1007/s00440-015-0673-1>
- [19] Eberle, A., Guillin, A. and Zimmer, R. (2019). Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Trans. Amer. Math. Soc.* **371** 7135–7173. [MR3939573](#) <https://doi.org/10.1090/tran/7576>
- [20] Eberle, A. and Majka, M.B. (2019). Quantitative contraction rates for Markov chains on general state spaces. *Electron. J. Probab.* **24** 1–36. [MR3933205](#) <https://doi.org/10.1214/19-EJP287>
- [21] Ernst, P.A., Kendall, W.S., Roberts, G.O. and Rosenthal, J.S. (2019). MEXIT: Maximal un-coupling times for stochastic processes. *Stochastic Process. Appl.* **129** 355–380. [MR3907003](#) <https://doi.org/10.1016/j.spa.2018.03.001>
- [22] Gangbo, W. and McCann, R.J. (1996). The geometry of optimal transportation. *Acta Math.* **177** 113–161. [MR1440931](#) <https://doi.org/10.1007/BF02392620>
- [23] Goldstein, S. (1979). Maximal coupling. *Z. Wahrsch. Verw. Gebiete* **46** 193–204. [MR0516740](#) <https://doi.org/10.1007/BF00533259>
- [24] Griffeath, D. (1975). A maximal coupling for Markov chains. *Z. Wahrsch. Verw. Gebiete* **31** 95–106. [MR0370771](#) <https://doi.org/10.1007/BF00539434>
- [25] Hammersley, J.M. and Morton, K.W. (1956). A new Monte Carlo technique: Antithetic variates. *Proc. Camb. Philos. Soc.* **52** 449–475. [MR0080984](#) <https://doi.org/10.1017/s0305004100031455>
- [26] Hsu, E.P. and Sturm, K.-T. (2013). Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.* **1** 93–104. [MR3197874](#) <https://doi.org/10.1007/s40304-013-0007-5>
- [27] Hu, K., Ren, Z., Šiška, D. and Szpruch, Ł. (2021). Mean-field Langevin dynamics and energy landscape of neural networks. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 2043–2065. [MR4328560](#) <https://doi.org/10.1214/20-aihp1140>
- [28] Huang, L.-J., Majka, M.B. and Wang, J. (2022). Strict Kantorovich contractions for Markov chains and Euler schemes with general noise. *Stochastic Process. Appl.* **151** 307–341. [MR4445709](#) <https://doi.org/10.1016/j.spa.2022.06.011>
- [29] Jacka, S.D. and Mijatović, A. (2015). Coupling and tracking of regime-switching martingales. *Electron. J. Probab.* **20** 1–39. [MR3335829](#) <https://doi.org/10.1214/EJP.v20-2307>
- [30] Jacka, S.D., Mijatović, A. and Širaj, D. (2014). Mirror and synchronous couplings of geometric Brownian motions. *Stochastic Process. Appl.* **124** 1055–1069. [MR3138606](#) <https://doi.org/10.1016/j.spa.2013.10.003>
- [31] Kellerer, H.G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** 399–432. [MR0761565](#) <https://doi.org/10.1007/BF00532047>
- [32] Kendall, W.S. (2015). Coupling, local times, immersions. *Bernoulli* **21** 1014–1046. [MR3338655](#) <https://doi.org/10.3150/14-BEJ596>

- [33] Liang, M., Majka, M.B. and Wang, J. (2021). Exponential ergodicity for SDEs and McKean–Vlasov processes with Lévy noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1665–1701. [MR4291453](#) <https://doi.org/10.1214/20-aihp1123>
- [34] Liang, M., Schilling, R.L. and Wang, J. (2020). A unified approach to coupling SDEs driven by Lévy noise and some applications. *Bernoulli* **26** 664–693. [MR4036048](#) <https://doi.org/10.3150/19-BEJ1148>
- [35] Liang, M. and Wang, J. (2020). Gradient estimates and ergodicity for SDEs driven by multiplicative Lévy noises via coupling. *Stochastic Process. Appl.* **130** 3053–3094. [MR4080736](#) <https://doi.org/10.1016/j.spa.2019.09.001>
- [36] Lindvall, T. (1992). *Lectures on the Coupling Method*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. A Wiley-Interscience Publication. [MR1180522](#)
- [37] Luo, D. and Wang, J. (2019). Refined basic couplings and Wasserstein-type distances for SDEs with Lévy noises. *Stochastic Process. Appl.* **129** 3129–3173. [MR3985558](#) <https://doi.org/10.1016/j.spa.2018.09.003>
- [38] Majka, M.B. (2017). Coupling and exponential ergodicity for stochastic differential equations driven by Lévy processes. *Stochastic Process. Appl.* **127** 4083–4125. [MR3718107](#) <https://doi.org/10.1016/j.spa.2017.03.020>
- [39] Majka, M.B. (2019). Transportation inequalities for non-globally dissipative SDEs with jumps via Malliavin calculus and coupling. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 2019–2057. [MR4029147](#) <https://doi.org/10.1214/18-AIHP941>
- [40] Majka, M.B., Mijatović, A. and Szpruch, L. (2020). Nonasymptotic bounds for sampling algorithms without log-concavity. *Ann. Appl. Probab.* **30** 1534–1581. [MR4132634](#) <https://doi.org/10.1214/19-AAP1535>
- [41] McCann, R.J. (1999). Exact solutions to the transportation problem on the line. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **455** 1341–1380. [MR1701760](#) <https://doi.org/10.1098/rspa.1999.0364>
- [42] Pitman, J.W. (1976). On coupling of Markov chains. *Z. Wahrsch. Verw. Gebiete* **35** 315–322. [MR0415775](#) <https://doi.org/10.1007/BF00532957>
- [43] 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>
- [44] Rogers, L.C.G. (1999). Fastest coupling of random walks. *J. Lond. Math. Soc. (2)* **60** 630–640. [MR1724813](#) <https://doi.org/10.1112/S0024610799008017>
- [45] Rosenthal, J.S. (1997). Faithful couplings of Markov chains: Now equals forever. *Adv. in Appl. Math.* **18** 372–381. [MR1436487](#) <https://doi.org/10.1006/aama.1996.0515>
- [46] Ross, S.M. (2014). *Introduction to Probability Models*, Eleventh ed. Amsterdam: Elsevier/Academic Press. [MR3307945](#)
- [47] Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration*. *Probability and Its Applications (New York)*. New York: Springer. [MR1741181](#) <https://doi.org/10.1007/978-1-4612-1236-2>
- [48] Villani, C. (2009). *Optimal Transport: Old and New*. Springer.
- [49] Watanabe, S. (1964). On discontinuous additive functionals and Lévy measures of a Markov process. *Jpn. J. Math.* **34** 53–70. [MR0185675](#) [https://doi.org/10.4099/jjm1924.34.0\\_53](https://doi.org/10.4099/jjm1924.34.0_53)
- [50] Williams, D. (1991). *Probability with Martingales*. *Cambridge Mathematical Textbooks*. Cambridge: Cambridge Univ. Press. [MR1155402](#) <https://doi.org/10.1017/CBO9780511813658>

# A unifying approach to distributional limits for empirical optimal transport

SHAYAN HUNDRIESER<sup>1,a</sup>, MARCEL KLATT<sup>1,b</sup>, AXEL MUNK<sup>1,2,3,c</sup> and THOMAS STAUDT<sup>1,3,d</sup>

<sup>1</sup>*Institute for Mathematical Stochastics, University of Göttingen, Göttingen, Germany,*

<sup>a</sup>*s.hundrieser@math.uni-goettingen.de,* <sup>b</sup>*mklatt@mathematik.uni-goettingen.de,* <sup>c</sup>*munk@math.uni-goettingen.de,*

<sup>d</sup>*thomas.staudt@uni-goettingen.de*

<sup>2</sup>*Max Planck Institute for Multidisciplinary Sciences, University of Göttingen, Göttingen, Germany*

<sup>3</sup>*Cluster of Excellence “Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells” (MBExC), University of Göttingen, University Medical Center, Göttingen, Germany*

We provide a unifying approach to *central limit type theorems* for empirical optimal transport (OT). The limit distribution is given by a supremum of a centered Gaussian process, and we explicitly characterize when it is centered normal or degenerates to a Dirac measure. Moreover, in contrast to recent contributions to distributional limit laws for empirical OT on Euclidean spaces which require centering around its expectation, the limits obtained here are centered around the *population* quantity, which is well-suited for statistical applications such as goodness-of-fit testing and randomized OT computation. Overall, our distributional limits are valid if one of the population probability measures is of intrinsic dimension at most three. At the heart of our theory lies the Kantorovich duality which represents the OT cost as a supremum over a function class  $\mathcal{F}_c$  for an underlying sufficiently regular and possibly unbounded cost function  $c$ . In this regard, OT is considered as a functional defined on  $\ell^\infty(\mathcal{F}_c)$ , the Banach space of bounded functionals from  $\mathcal{F}_c$  to  $\mathbb{R}$  and equipped with uniform norm. We prove the OT functional to be *Hadamard directionally differentiable* and conclude distributional convergence for increasing sample size via a *functional delta method* that necessitates weak convergence of an underlying empirical process in  $\ell^\infty(\mathcal{F}_c)$ . The latter can be dealt with *empirical process theory* and requires  $\mathcal{F}_c$  to be a *Donsker* class. We give sufficient conditions depending on the dimension of the ground space, the underlying cost function and the probability measures under consideration to guarantee the Donsker property. Altogether, our approach reveals a noteworthy trade-off inherent in central limit theorems for empirical OT: Kantorovich duality requires  $\mathcal{F}_c$  to be sufficiently rich, while weak convergence of the underlying empirical processes only occurs if  $\mathcal{F}_c$  is not too complex.

*Keywords:* Bootstrap; central limit theorem; empirical processes; Kantorovich potential; optimal transport; regularity theory; Wasserstein distance

## References

- Ajtai, M., Komlós, J. and Tusnády, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264. [MR0779885 https://doi.org/10.1007/BF02579135](https://doi.org/10.1007/BF02579135)
- Altschuler, J., Niles-Weed, J. and Rigollet, P. (2017). Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. In *Adv. Neural Inf. Process. Syst. (NeurIPS)* (I. Guyon et al., eds.) **30**. Curran Associates.
- Aurenhammer, F., Hoffmann, F. and Aronov, B. (1998). Minkowski-type theorems and least-squares clustering. *Algorithmica* **20** 61–76. [MR1483422 https://doi.org/10.1007/PL00009187](https://doi.org/10.1007/PL00009187)
- Barthe, F. and Bordenave, C. (2013). Combinatorial optimization over two random point sets. In *Séminaire de Probabilités XLV. Lecture Notes in Math.* **2078** 483–535. Cham: Springer. [MR3185927 https://doi.org/10.1007/978-3-319-00321-4\\_19](https://doi.org/10.1007/978-3-319-00321-4_19)
- Bernton, E., Ghosal, P. and Nutz, M. (2022). Entropic optimal transport: Geometry and large deviations. *Duke Math. J.* **171** 3363–3400. [MR4505361 https://doi.org/10.1215/00127094-2022-0035](https://doi.org/10.1215/00127094-2022-0035)

- Berthet, P. and Fort, J.-C. (2019). Weak convergence of empirical Wasserstein type distances. Preprint. Available at [arXiv:1911.02389](https://arxiv.org/abs/1911.02389).
- Berthet, P., Fort, J.-C. and Klein, T. (2020). A central limit theorem for Wasserstein type distances between two distinct univariate distributions. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 954–982. [MR4076772 https://doi.org/10.1214/19-AIHP990](https://doi.org/10.1214/19-AIHP990)
- Bobkov, S. and Ledoux, M. (2019). One-dimensional empirical measures, order statistics, and Kantorovich transport distances. *Mem. Amer. Math. Soc.* **261** v+126. [MR4028181 https://doi.org/10.1090/memo/1259](https://doi.org/10.1090/memo/1259)
- Bobkov, S.G. and Ledoux, M. (2021). A simple Fourier analytic proof of the AKT optimal matching theorem. *Ann. Appl. Probab.* **31** 2567–2584. [MR4350968 https://doi.org/10.1214/20-aap1656](https://doi.org/10.1214/20-aap1656)
- Boissard, E. and Le Gouic, T. (2014). On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 539–563. [MR3189084 https://doi.org/10.1214/12-AIHP517](https://doi.org/10.1214/12-AIHP517)
- Bonneel, N., Van De Panne, M., Paris, S. and Heidrich, W. (2011). Displacement interpolation using Lagrangian mass transport. In *Proceedings of SIGGRAPH ASIA 2011* 1–12.
- Bühlmann, P. (1995). The blockwise bootstrap for general empirical processes of stationary sequences. *Stochastic Process. Appl.* **58** 247–265. [MR1348377 https://doi.org/10.1016/0304-4149\(95\)00019-4](https://doi.org/10.1016/0304-4149(95)00019-4)
- Cárcamo, J., Cuevas, A. and Rodríguez, L.-A. (2020). Directional differentiability for supremum-type functionals: Statistical applications. *Bernoulli* **26** 2143–2175. [MR4091104 https://doi.org/10.3150/19-BEJ1188](https://doi.org/10.3150/19-BEJ1188)
- Chizat, L., Roussillon, P., Léger, F., Vialard, F.-X. and Peyré, G. (2020). Faster Wasserstein distance estimation with the Sinkhorn divergence. In *Adv. Neural Inf. Process. Syst. (NeurIPS)* (H. Larochelle et al., eds.) **33** 2257–2269.
- Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Chichester: Wiley. [MR1215046](https://doi.org/10.1002/9781118111111)
- Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Adv. Neural Inf. Process. Syst. (NeurIPS)* (C.J. Burges et al., eds.) **26**. Curran Associates.
- Dedecker, J. and Louhichi, S. (2002). Maximal inequalities and empirical central limit theorems. In *Empirical Process Techniques for Dependent Data* 137–159. Boston, MA: Birkhäuser. [MR1958779](https://doi.org/10.1007/978-1-4614-0000-0_10)
- del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. [MR1698999 https://doi.org/10.1214/aop/1022677394](https://doi.org/10.1214/aop/1022677394)
- del Barrio, E., Giné, E. and Utzet, F. (2005). Asymptotics for  $L_2$  functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* **11** 131–189. [MR2121458 https://doi.org/10.3150/bj/1110228245](https://doi.org/10.3150/bj/1110228245)
- del Barrio, E., González-Sanz, A. and Loubes, J.-M. (2024). Central limit theorems for semi-discrete Wasserstein distances. *Bernoulli* **30** 554–580. [MR4665589 https://doi.org/10.3150/23-bej1608](https://doi.org/10.3150/23-bej1608)
- del Barrio, E., González-Sanz, A. and Loubes, J.-M. (2023). Central limit theorems for general transportation costs. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at [arXiv:2102.06379](https://arxiv.org/abs/2102.06379).
- del Barrio, E., Gordaliza, P. and Loubes, J.-M. (2019). A central limit theorem for  $L_p$  transportation cost on the real line with application to fairness assessment in machine learning. *Inf. Inference* **8** 817–849. [MR4045479 https://doi.org/10.1093/imaiai/iaz016](https://doi.org/10.1093/imaiai/iaz016)
- del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** 926–951. [MR3916938 https://doi.org/10.1214/18-AOP1275](https://doi.org/10.1214/18-AOP1275)
- del Barrio, E., González-Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2023). An improved central limit theorem and fast convergence rates for entropic transportation costs. *SIAM J. Math. Data Sci.* **5** 639–669. [MR4616887 https://doi.org/10.1137/22M149260X](https://doi.org/10.1137/22M149260X)
- Dobrić, V. and Yukich, J.E. (1995). Asymptotics for transportation cost in high dimensions. *J. Theoret. Probab.* **8** 97–118. [MR1308672 https://doi.org/10.1007/BF02213456](https://doi.org/10.1007/BF02213456)
- Doukhan, P., Massart, P. and Rio, E. (1995). Invariance principles for absolutely regular empirical processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 393–427. [MR1324814](https://doi.org/10.1007/978-1-4614-0000-0_10)
- Dudley, R.M. (1968). The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Stat.* **40** 40–50. [MR0236977 https://doi.org/10.1214/aoms/1177697802](https://doi.org/10.1214/aoms/1177697802)
- Dudley, R.M. (2014). *Uniform Central Limit Theorems*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **142**. New York: Cambridge Univ. Press. [MR3445285](https://doi.org/10.1017/C9780521876223)

- Dümbgen, L. (1993). On nondifferentiable functions and the bootstrap. *Probab. Theory Related Fields* **95** 125–140. [MR1207311](#) <https://doi.org/10.1007/BF01197342>
- Fang, Z. and Santos, A. (2019). Inference on directionally differentiable functions. *Rev. Econ. Stud.* **86** 377–412. [MR3936869](#) <https://doi.org/10.1093/restud/rdy049>
- Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. [MR3383341](#) <https://doi.org/10.1007/s00440-014-0583-7>
- Galichon, A. (2016). *Optimal Transport Methods in Economics*. Princeton, NJ: Princeton Univ. Press. [MR3586373](#) <https://doi.org/10.1515/9781400883592>
- Gangbo, W. and McCann, R.J. (1996). The geometry of optimal transportation. *Acta Math.* **177** 113–161. [MR1440931](#) <https://doi.org/10.1007/BF02392620>
- Giné, E. and Zinn, J. (1986). Empirical processes indexed by Lipschitz functions. *Ann. Probab.* **14** 1329–1338. [MR0866353](#)
- Goldfeld, Z. and Greenewald, K. (2020). Gaussian-smoothed optimal transport: Metric structure and statistical efficiency. In *International Conference on Artificial Intelligence and Statistics (AISTATS)* 3327–3337. PMLR.
- Goldfeld, Z., Kato, K., Nietert, S. and Rioux, G. (2022a). Limit distribution theory for smooth  $p$ -Wasserstein distances. Preprint. Available at [arXiv:2203.00159](#).
- Goldfeld, Z., Kato, K., Rioux, G. and Sadhu, R. (2022b). Statistical inference with regularized optimal transport. Preprint. Available at [arXiv:2205.04283](#).
- Goldman, M. and Trevisan, D. (2021). Convergence of asymptotic costs for random Euclidean matching problems. *Probab. Math. Phys.* **2** 341–362. [MR4408015](#) <https://doi.org/10.2140/pmp.2021.2.341>
- González-Delgado, J., González-Sanz, A., Cortés, J. and Neuvial, P. (2023). Two-sample goodness-of-fit tests on the flat torus based on Wasserstein distance and their relevance to structural biology. *Electron. J. Stat.* **17** 1547–1586. [MR4598874](#) <https://doi.org/10.1214/23-ejs2135>
- González-Sanz, A. and Hundrieser, S. (2023). Weak limits for empirical entropic optimal transport: Beyond smooth costs. Preprint. Available at [arXiv:2305.09745](#).
- Hallin, M., Mordant, G. and Segers, J. (2021). Multivariate goodness-of-fit tests based on Wasserstein distance. *Electron. J. Stat.* **15** 1328–1371. [MR4255302](#) <https://doi.org/10.1214/21-ejs1816>
- Hallin, M., del Barrio, E., Cuesta-Albertos, J. and Matrán, C. (2021). Distribution and quantile functions, ranks and signs in dimension  $d$ : A measure transportation approach. *Ann. Statist.* **49** 1139–1165. [MR4255122](#) <https://doi.org/10.1214/20-aos1996>
- Hartmann, V. and Schuhmacher, D. (2020). Semi-discrete optimal transport: A solution procedure for the unsquared Euclidean distance case. *Math. Methods Oper. Res.* **92** 133–163. [MR4152920](#) <https://doi.org/10.1007/s00186-020-00703-z>
- Heinemann, F., Munk, A. and Zemel, Y. (2022). Randomized Wasserstein barycenter computation: Resampling with statistical guarantees. *SIAM J. Math. Data Sci.* **4** 229–259. [MR4386483](#) <https://doi.org/10.1137/20M1385263>
- Hundrieser, S., Klatt, M. and Munk, A. (2021). Limit distributions and sensitivity analysis for entropic optimal transport on countable spaces. *Ann. Appl. Probab.* To appear. Available at [arXiv:2105.00049](#).
- Hundrieser, S., Klatt, M. and Munk, A. (2022). The statistics of circular optimal transport. In *Directional Statistics for Innovative Applications – A Bicentennial Tribute to Florence Nightingale. Forum Interdiscip. Math.* 57–82. Singapore: Springer. [MR4461115](#) [https://doi.org/10.1007/978-981-19-1044-9\\_4](https://doi.org/10.1007/978-981-19-1044-9_4)
- Hundrieser, S., Staudt, T. and Munk, A. (2023). Empirical optimal transport between different measures adapts to lower complexity. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at [arXiv:2202.10434](#).
- Hundrieser, S., Klatt, M., Munk, A. and Staudt, T. (2024). Supplement to “A unifying approach to distributional limits for empirical optimal transport.” <https://doi.org/10.3150/23-BEJ1697SUPP>
- Klatt, M., Munk, A. and Zemel, Y. (2022). Limit laws for empirical optimal solutions in random linear programs. *Ann. Oper. Res.* **315** 251–278. [MR4458612](#) <https://doi.org/10.1007/s10479-022-04698-0>
- Klatt, M., Tameling, C. and Munk, A. (2020). Empirical regularized optimal transport: Statistical theory and applications. *SIAM J. Math. Data Sci.* **2** 419–443. [MR4105566](#) <https://doi.org/10.1137/19M1278788>
- Kolouri, S., Park, S.R., Thorpe, M., Slepčev, D. and Rohde, G.K. (2017). Optimal mass transport: Signal processing and machine-learning applications. *IEEE Signal Process. Mag.* **34** 43–59.
- Le Cam, L. (1986). The central limit theorem around 1935. *Statist. Sci.* **1** 78–96. With comments, and a rejoinder by the author. [MR0833276](#)



- Lee, J.M. (2013). *Introduction to Smooth Manifolds*, 2nd ed. *Graduate Texts in Mathematics* **218**. New York: Springer. [MR2954043](#)
- Li, Q.-R., Santambrogio, F. and Wang, X.-J. (2014). Regularity in Monge's mass transfer problem. *J. Math. Pures Appl.* (9) **102** 1015–1040. [MR3277433](#) <https://doi.org/10.1016/j.matpur.2014.03.001>
- Liu, S., Bunea, F. and Niles-Weed, J. (2023). Asymptotic confidence sets for random linear programs. Preprint. Available at [arXiv:2302.12364](#).
- Ma, X.-N., Trudinger, N.S. and Wang, X.-J. (2005). Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.* **177** 151–183. [MR2188047](#) <https://doi.org/10.1007/s00205-005-0362-9>
- Manole, T. and Niles-Weed, J. (2023). Sharp convergence rates for empirical optimal transport with smooth costs. *Ann. Appl. Probab.* To appear. Preprint. Available at [arXiv:2106.13181](#).
- Manole, T., Balakrishnan, S., Niles-Weed, J. and Wasserman, L. (2021). Plugin estimation of smooth optimal transport maps. Preprint. Available at [arXiv:2107.12364](#).
- Mason, D.M. (2016). A weighted approximation approach to the study of the empirical Wasserstein distance. In *High Dimensional Probability VII. Progress in Probability* **71** 137–154. Cham: Springer. [MR3565262](#) [https://doi.org/10.1007/978-3-319-40519-3\\_6](https://doi.org/10.1007/978-3-319-40519-3_6)
- Méridot, Q. (2011). A multiscale approach to optimal transport. In *Comput. Graph. Forum* **30** 1583–1592. Wiley Online Library.
- Munk, A. and Czado, C. (1998). Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 223–241. [MR1625620](#) <https://doi.org/10.1111/1467-9868.00121>
- Munk, A. and Dette, H. (1998). Nonparametric comparison of several regression functions: Exact and asymptotic theory. *Ann. Statist.* **26** 2339–2368. [MR1700235](#) <https://doi.org/10.1214/aos/1024691474>
- Nietert, S., Goldfeld, Z. and Kato, K. (2021). Smooth  $p$ -Wasserstein distance: Structure, empirical approximation, and statistical applications. In *International Conference on Machine Learning (ICML)* (M. Meila and T. Zhang, eds.). *Proceedings of Machine Learning Research* **139** 8172–8183. PMLR.
- Niles-Weed, J. and Rigollet, P. (2022). Estimation of Wasserstein distances in the spiked transport model. *Bernoulli* **28** 2663–2688. [MR4474558](#) <https://doi.org/10.3150/21-bej1433>
- Panaretos, V.M. and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** 405–431. [MR3939527](#) <https://doi.org/10.1146/annurev-statistics-030718-104938>
- Peyré, G. and Cuturi, M. (2019). Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** 355–607.
- Phandoidaen, N. and Richter, S. (2022). Empirical process theory for locally stationary processes. *Bernoulli* **28** 453–480. [MR4337712](#) <https://doi.org/10.3150/21-bej1351>
- Rachev, S.T. and Rüschendorf, L. (1998a). *Mass Transportation Problems. Vol. I: Theory. Probability and Its Applications (New York)*. New York: Springer. [MR1619170](#)
- Rachev, S.T. and Rüschendorf, L. (1998b). *Mass Transportation Problems. Vol. II: Applications. Probability and Its Applications (New York)*. New York: Springer. [MR1619171](#)
- Radulović, D. (1996). The bootstrap for empirical processes based on stationary observations. *Stochastic Process. Appl.* **65** 259–279. [MR1425360](#) [https://doi.org/10.1016/S0304-4149\(96\)00102-0](https://doi.org/10.1016/S0304-4149(96)00102-0)
- Radulović, D. (2002). On the bootstrap and empirical processes for dependent sequences. In *Empirical Process Techniques for Dependent Data* 345–364. Boston, MA: Birkhäuser. [MR1958789](#) [https://doi.org/10.1007/978-1-4612-0099-4\\_13](https://doi.org/10.1007/978-1-4612-0099-4_13)
- Rio, E. (2017). *Asymptotic Theory of Weakly Dependent Random Processes. Probability Theory and Stochastic Modelling* **80**. Berlin: Springer. [MR3642873](#) <https://doi.org/10.1007/978-3-662-54323-8>
- Römsch, W. (2006). Delta method, infinite dimensional. In *Encyclopedia of Statistical Sciences* **16** 1575–1583. New York: Wiley.
- Sadhu, R., Goldfeld, Z. and Kato, K. (2021). Limit distribution theory for the smooth 1-Wasserstein distance with applications. Preprint. Available at [arXiv:2107.13494](#).
- Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Progress in Nonlinear Differential Equations and Their Applications* **87**. Cham: Birkhäuser/Springer. [MR3409718](#) <https://doi.org/10.1007/978-3-319-20828-2>
- Schiebinger, G., Shu, J., Tabaka, M., Cleary, B., Subramanian, V., Solomon, A., Gould, J., Liu, S., Lin, S., Berube, P. et al. (2019). Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* **176** 928–943.

- Singh, S. and Póczos, B. (2018). Minimax distribution estimation in Wasserstein distance. Preprint. Available at [arXiv:1802.08855](https://arxiv.org/abs/1802.08855).
- Sommerfeld, M. and Munk, A. (2018). Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 219–238. [MR3744719 https://doi.org/10.1111/rssb.12236](https://doi.org/10.1111/rssb.12236)
- Sommerfeld, M., Schrieber, J., Zemel, Y. and Munk, A. (2019). Optimal transport: Fast probabilistic approximation with exact solvers. *J. Mach. Learn. Res.* **20** Paper No. 105, 23 pp. [MR3990459](https://arxiv.org/abs/1906.07026)
- Staudt, T., Hundrieser, S. and Munk, A. (2022). On the uniqueness of Kantorovich potentials. Preprint. Available at [arXiv:2201.08316](https://arxiv.org/abs/2201.08316).
- Talagrand, M. (1994a). Matching theorems and empirical discrepancy computations using majorizing measures. *J. Amer. Math. Soc.* **7** 455–537. [MR1227476 https://doi.org/10.2307/2152764](https://doi.org/10.2307/2152764)
- Talagrand, M. (1994b). The transportation cost from the uniform measure to the empirical measure in dimension  $\geq 3$ . *Ann. Probab.* **22** 919–959. [MR1288137](https://doi.org/10.2307/2288137)
- Tameling, C., Sommerfeld, M. and Munk, A. (2019). Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications. *Ann. Appl. Probab.* **29** 2744–2781. [MR4019874 https://doi.org/10.1214/19-AAP1463](https://doi.org/10.1214/19-AAP1463)
- Tameling, C., Stoldt, S., Stephan, T., Naas, J., Jakobs, S. and Munk, A. (2021). Colocalization for super-resolution microscopy via optimal transport. *Nat. Comput. Sci.* **1** 199–211. <https://doi.org/10.1038/s43588-021-00050-x>
- van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes – With Applications to Statistics*. Springer Series in Statistics. Cham: Springer. [MR4628026 https://doi.org/10.1007/978-3-031-29040-4](https://doi.org/10.1007/978-3-031-29040-4)
- 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)
- Villani, C. (2008). *Optimal Transport: Old and New*. A Series of Comprehensive Studies in Mathematics. Berlin: Springer.
- Weed, J. and Bach, F. (2019). Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** 2620–2648. [MR4003560 https://doi.org/10.3150/18-BEJ1065](https://doi.org/10.3150/18-BEJ1065)
- Weitkamp, C.A., Proksch, K., Tameling, C. and Munk, A. (2022). Distribution of distances based object matching: Asymptotic inference. *J. Amer. Statist. Assoc.* To appear. Available at [arXiv:2006.12287](https://arxiv.org/abs/2006.12287).
- Whitney, H. (1934). Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36** 63–89. [MR1501735 https://doi.org/10.2307/1989708](https://doi.org/10.2307/1989708)
- 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](https://doi.org/10.1073/pnas.0506715102)
- Yang, Y., Nurbekyan, L., Negrini, E., Martin, R. and Pasha, M. (2023). Optimal transport for parameter identification of chaotic dynamics via invariant measures. *SIAM J. Appl. Dyn. Syst.* **22** 269–310. [MR4547709 https://doi.org/10.1137/21M1421337](https://doi.org/10.1137/21M1421337)

# Approximate double-transform inversion when time is one of the variables

RONALD W. BUTLER<sup>a</sup>

Fort Collins, Colorado 80525 USA, <sup>a</sup>[brennerweg@yahoo.com](mailto:brennerweg@yahoo.com)

For a continuous-time (integer-time) stochastic process, its distribution at arbitrary time  $t$  ( $n$ ) is often a difficult computation. To use a saddlepoint approximation, its time-indexed moment generating function (MGF) is needed and seldomly is that available. What is often readily available, however, is the Laplace transform (generating function) in time  $t$  ( $n$ ) of this time-indexed MGF which we call a double transform. Such double transforms often take a simple analytic form and we show how they may be inverted to determine the survival function for the process at time  $t$  or  $n$ . Two general approaches are considered. First, we show that the double-saddlepoint methods initiated by Skovgaard (*J. Appl. Probab.* **24** (1987) 875–887) may be used by treating the time variable  $t$  or  $n$  as a random variable with an improper distribution. The second method inverts the double transform in two stages. First, it uses a residue expansion (Butler (*J. Appl. Probab.* **56** (2019) 307–338; *Stoch. Models* **39** (2023) 469–501)) to invert it in  $t$  or  $n$  which is then followed by a single-saddlepoint approximation of the Lugannani-Rice (*Adv. in Appl. Probab.* **12** (1980) 475–490) type. Applications from renewal theory and renewal reward (cumulative) processes illustrate the remarkable accuracy that results from both of these saddlepoint approximation methods.

**Keywords:** Double transform; renewal theory; renewal reward process; residue approximation; saddlepoint approximation

## References

- Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Wiley Series in Probability and Mathematical Statistics. Chichester: Wiley. [MR0489333](#)
- Bleistein, N. (1966). Uniform asymptotic expansions of integrals with stationary point near algebraic singularity. *Comm. Pure Appl. Math.* **19** 353–370. [MR0204943](#) <https://doi.org/10.1002/cpa.3160190403>
- Booth, J.G. and Butler, R.W. (1990). Randomization distributions and saddlepoint approximations in generalized linear models. *Biometrika* **77** 787–796. [MR1086689](#) <https://doi.org/10.1093/biomet/77.4.787>
- Butler, R.W. (2007). *Saddlepoint Approximations with Applications*. Cambridge Series in Statistical and Probabilistic Mathematics **22**. Cambridge: Cambridge Univ. Press. [MR2357347](#) <https://doi.org/10.1017/CBO9780511619083>
- Butler, R.W. (2017). Asymptotic expansions and hazard rates for compound and first-passage distributions. *Bernoulli* **23** 3508–3536. [MR3654814](#) <https://doi.org/10.3150/16-BEJ854>
- Butler, R.W. (2019). Asymptotic expansions and saddlepoint approximations using the analytic continuation of moment generating functions. *J. Appl. Probab.* **56** 307–338. [MR3981159](#) <https://doi.org/10.1017/jpr.2019.19>
- Butler, R.W. (2023). Residue expansions and saddlepoint approximations in stochastic models using the analytic continuation of generating functions. *Stoch. Models* **39** 469–501. [MR4619887](#) <https://doi.org/10.1080/15326349.2022.2114496>
- Cox, D.R. (1962). *Renewal Theory*. London: Methuen & Co., Ltd.. [MR0153061](#)
- Daniels, H.E. (1954). Saddlepoint approximations in statistics. *Ann. Math. Stat.* **25** 631–650. [MR0066602](#) <https://doi.org/10.1214/aoms/1177728652>
- Daniels, H.E. (1987). Tail probability approximations. *Int. Stat. Rev.* **55** 37–48. [MR0962940](#) <https://doi.org/10.2307/1403269>
- Doetsch, G. (1974). *Introduction to the Theory and Application of the Laplace Transformation*. New York: Springer. [MR0344810](#)

- Feller, W. (1967). *An Introduction to Probability Theory and Its Applications, Vol. I.*, New York: Wiley.
- Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. [MR0270403](#)
- Henrici, P. (1977). *Applied and Computational Complex Analysis. Vol. 2.* New York: Wiley Interscience. [MR0453984](#)
- Howard, R.A. (1971). *Dynamic Probabilistic Systems, Volume I: Markov Models.* New York: Wiley.
- Lugannani, R. and Rice, S. (1980). Saddle point approximation for the distribution of the sum of independent random variables. *Adv. in Appl. Probab.* **12** 475–490. [MR0569438](#) <https://doi.org/10.2307/1426607>
- Skovgaard, I.M. (1987). Saddlepoint expansions for conditional distributions. *J. Appl. Probab.* **24** 875–887. [MR0913828](#) <https://doi.org/10.2307/3214212>
- Tijms, H.C. (2003). *A First Course in Stochastic Models.* Chichester: Wiley. [MR2190630](#) <https://doi.org/10.1002/047001363X>
- Wilf, H.S. (2006). *Generatingfunctionology*, 3rd ed. Wellesley, MA: A K Peters, Ltd. [MR2172781](#)

# Extreme singular values of inhomogeneous sparse random rectangular matrices

IOANA DUMITRIU<sup>1,a</sup> and YIZHE ZHU<sup>2,b</sup>

<sup>1</sup>*Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA,*

<sup>a</sup>*idumitriu@ucsd.edu*

<sup>2</sup>*Department of Mathematics, University of California Irvine, Irvine, CA 92697, USA, <sup>b</sup>yizhe.zhu@uci.edu*

We develop a unified approach to bounding the largest and smallest singular values of an inhomogeneous random rectangular matrix, based on the non-backtracking operator and the Ihara-Bass formula for general random Hermitian matrices with a bipartite block structure. We obtain probabilistic upper (respectively, lower) bounds for the largest (respectively, smallest) singular values of a large rectangular random matrix  $X$ . These bounds are given in terms of the maximal and minimal  $\ell_2$ -norms of the rows and columns of the variance profile of  $X$ . The proofs involve finding probabilistic upper bounds on the spectral radius of an associated non-backtracking matrix  $B$ . The two-sided bounds can be applied to the centered adjacency matrix of sparse inhomogeneous Erdős-Rényi bipartite graphs for a wide range of sparsity, down to criticality. In particular, for Erdős-Rényi bipartite graphs  $\mathcal{G}(n, m, p)$  with  $p = \omega(\log n)/n$ , and  $m/n \rightarrow y \in (0, 1)$ , our sharp bounds imply that there are no outliers outside the support of the Marčenko-Pastur law almost surely. This result extends the Bai-Yin theorem to sparse rectangular random matrices.

*Keywords:* Extreme singular value; inhomogeneous random matrix; non-backtracking operator; random bipartite graph

## References

- Alt, J., Ducatez, R. and Knowles, A. (2021). Extremal eigenvalues of critical Erdős-Rényi graphs. *Ann. Probab.* **49** 1347–1401. [MR4255147](#) <https://doi.org/10.1214/20-aop1483>
- Auffinger, A. and Tang, S. (2016). Extreme eigenvalues of sparse, heavy tailed random matrices. *Stochastic Process. Appl.* **126** 3310–3330. [MR3549709](#) <https://doi.org/10.1016/j.spa.2016.04.029>
- Avron, H., Druinsky, A. and Toledo, S. (2019). Spectral condition-number estimation of large sparse matrices. *Numer. Linear Algebra Appl.* **26** e2235, 19. [MR3946070](#) <https://doi.org/10.1002/nla.2235>
- 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>
- Bai, Z.D. and Yin, Y.Q. (1993). Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.* **21** 1275–1294. [MR1235416](#)
- Bandeira, A.S., Boedihardjo, M.T. and van Handel, R. (2023). Matrix concentration inequalities and free probability. *Invent. Math.* **234** 419–487. [MR4635836](#) <https://doi.org/10.1007/s00222-023-01204-6>
- Bandeira, A.S. and van Handel, R. (2016). Sharp nonasymptotic bounds on the norm of random matrices with independent entries. *Ann. Probab.* **44** 2479–2506. [MR3531673](#) <https://doi.org/10.1214/15-AOP1025>
- Basak, A. and Rudelson, M. (2017). Invertibility of sparse non-Hermitian matrices. *Adv. Math.* **310** 426–483. [MR3620692](#) <https://doi.org/10.1016/j.aim.2017.02.009>
- Basak, A. and Rudelson, M. (2021). Sharp transition of the invertibility of the adjacency matrices of sparse random graphs. *Probab. Theory Related Fields* **180** 233–308. [MR4265022](#) <https://doi.org/10.1007/s00440-021-01038-4>
- Bass, H. (1992). The Ihara-Selberg zeta function of a tree lattice. *Internat. J. Math.* **3** 717–797. [MR1194071](#) <https://doi.org/10.1142/S0129167X92000357>
- Benaych-Georges, F., Bordenave, C. and Knowles, A. (2019). Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs. *Ann. Probab.* **47** 1653–1676. [MR3945756](#) <https://doi.org/10.1214/18-AOP1293>

- Benaych-Georges, F., Bordenave, C. and Knowles, A. (2020). Spectral radii of sparse random matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 2141–2161. [MR4116720](#) <https://doi.org/10.1214/19-AIHP1033>
- Bhojanapalli, S. and Jain, P. (2014). Universal matrix completion. In *International Conference on Machine Learning* 1881–1889. PMLR.
- Bordenave, C. (2020). A new proof of Friedman’s second eigenvalue theorem and its extension to random lifts. *Ann. Sci. Éc. Norm. Supér. (4)* **53** 1393–1439. [MR4203039](#) <https://doi.org/10.24033/asens.2450>
- Bordenave, C., Coste, S. and Nadakuditi, R.R. (2023). Detection thresholds in very sparse matrix completion. *Found. Comput. Math.* **23** 1619–1743. [MR4649432](#) <https://doi.org/10.1007/s10208-022-09568-6>
- Bordenave, C., Lelarge, M. and Massoulié, L. (2018). Nonbacktracking spectrum of random graphs: Community detection and nonregular Ramanujan graphs. *Ann. Probab.* **46** 1–71. [MR3758726](#) <https://doi.org/10.1214/16-AOP1142>
- Borodin, A., Corwin, I. and Guionnet, A., eds. (2019). *Random Matrices. IAS/Park City Mathematics Series* **26**. Providence, RI: Amer. Math. Soc.
- 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>
- Brailovskaya, T. and van Handel, R. (2022). Universality and sharp matrix concentration inequalities. Available at [arXiv:2201.05142](https://arxiv.org/abs/2201.05142).
- Brito, G., Dumitriu, I. and Harris, K.D. (2022). Spectral gap in random bipartite biregular graphs and applications. *Combin. Probab. Comput.* **31** 229–267. [MR4390294](#) <https://doi.org/10.1017/s0963548321000249>
- Cai, T.T., Han, R. and Zhang, A.R. (2022). On the non-asymptotic concentration of heteroskedastic Wishart-type matrix. *Electron. J. Probab.* **27** Paper No. 29, 40. [MR4385832](#) <https://doi.org/10.1214/22-ejp758>
- Chafaï, D. and Tikhomirov, K. (2018). On the convergence of the extremal eigenvalues of empirical covariance matrices with dependence. *Probab. Theory Related Fields* **170** 847–889. [MR3773802](#) <https://doi.org/10.1007/s00440-017-0778-9>
- Che, Z. and Lopatto, P. (2019). Universality of the least singular value for sparse random matrices. *Electron. J. Probab.* **24** Paper No. 9, 53. [MR3916329](#) <https://doi.org/10.1214/19-EJP269>
- Cook, N. (2018). Lower bounds for the smallest singular value of structured random matrices. *Ann. Probab.* **46** 3442–3500. [MR3857860](#) <https://doi.org/10.1214/17-AOP1251>
- Coste, S. and Zhu, Y. (2021). Eigenvalues of the non-backtracking operator detached from the bulk. *Random Matrices Theory Appl.* **10** Paper No. 2150028, 21. [MR4302283](#) <https://doi.org/10.1142/S2010326321500283>
- Deshpande, Y., Montanari, A., O’Donnell, R., Schramm, T. and Sen, S. (2019). The threshold for SDP-refutation of random regular NAE-3SAT. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* 2305–2321. Philadelphia, PA: SIAM. [MR3909609](#) <https://doi.org/10.1137/1.9781611975482.140>
- Dumitriu, I. and Zhu, Y. (2021). Spectra of random regular hypergraphs. *Electron. J. Combin.* **28** Paper No. 3.36, 25. [MR4295659](#) <https://doi.org/10.37236/8741>
- Feldheim, O.N. and Sodin, S. (2010). A universality result for the smallest eigenvalues of certain sample covariance matrices. *Geom. Funct. Anal.* **20** 88–123. [MR2647136](#) <https://doi.org/10.1007/s00039-010-0055-x>
- Florescu, L. and Perkins, W. (2016). Spectral thresholds in the bipartite stochastic block model. In *Conference on Learning Theory* 943–959. PMLR.
- Gordon, Y. (1985). Some inequalities for Gaussian processes and applications. *Israel J. Math.* **50** 265–289. [MR0800188](#) <https://doi.org/10.1007/BF02759761>
- Götze, F. and Tikhomirov, A. (2023). On the largest and the smallest singular value of sparse rectangular random matrices. *Electron. J. Probab.* **28** Paper No. 27, 18. [MR4548419](#) <https://doi.org/10.1214/23-ejp919>
- Guédon, O., Litvak, A.E. and Tatarko, K. (2020). Random polytopes obtained by matrices with heavy-tailed entries. *Commun. Contemp. Math.* **22** 1950027, 28. [MR4106820](#) <https://doi.org/10.1142/S0219199719500275>
- Guédon, O., Litvak, A.E., Pajor, A. and Tomczak-Jaegermann, N. (2017). On the interval of fluctuation of the singular values of random matrices. *J. Eur. Math. Soc. (JEMS)* **19** 1469–1505. [MR3635358](#) <https://doi.org/10.4171/JEMS/697>
- Guruswami, V., Manohar, P. and Mosheiff, J. (2022).  $\ell_p$ -spread and restricted isometry properties of sparse random matrices. In *37th Computational Complexity Conference. LIPIcs. Leibniz Int. Proc. Inform.* **234** Art. No. 7, 17. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. [MR4481615](#) <https://doi.org/10.4230/lipics.ccc.2022.7>

- Han, Q. (2022). Exact spectral norm error of sample covariance. Available at [arXiv:2207.13594](https://arxiv.org/abs/2207.13594).
- Heiny, J. and Mikosch, T. (2018). Almost sure convergence of the largest and smallest eigenvalues of high-dimensional sample correlation matrices. *Stochastic Process. Appl.* **128** 2779–2815. [MR3811704 https://doi.org/10.1016/j.spa.2017.10.002](https://doi.org/10.1016/j.spa.2017.10.002)
- Janwa, H. and Lal, A.K. (2003). On Tanner codes: Minimum distance and decoding. *Appl. Algebra Engrg. Comm. Comput.* **13** 335–347. [MR1959169 https://doi.org/10.1007/s00200-003-0098-4](https://doi.org/10.1007/s00200-003-0098-4)
- Koltchinskii, V. and Mendelson, S. (2015). Bounding the smallest singular value of a random matrix without concentration. *Int. Math. Res. Not. IMRN* **23** 12991–13008. [MR3431642 https://doi.org/10.1093/imrn/rnv096](https://doi.org/10.1093/imrn/rnv096)
- Kotani, M. and Sunada, T. (2000). Zeta functions of finite graphs. *J. Math. Sci. Univ. Tokyo* **7** 7–25. [MR1749978 https://doi.org/10.1007/s00033-000-0007-0](https://doi.org/10.1007/s00033-000-0007-0)
- Krivelevich, M. and Sudakov, B. (2003). The largest eigenvalue of sparse random graphs. *Combin. Probab. Comput.* **12** 61–72. [MR1967486 https://doi.org/10.1017/S0963548302005424](https://doi.org/10.1017/S0963548302005424)
- Latała, R., van Handel, R. and Youssef, P. (2018). The dimension-free structure of nonhomogeneous random matrices. *Invent. Math.* **214** 1031–1080. [MR3878726 https://doi.org/10.1007/s00222-018-0817-x](https://doi.org/10.1007/s00222-018-0817-x)
- Le, C.M., Levina, E. and Vershynin, R. (2017). Concentration and regularization of random graphs. *Random Structures Algorithms* **51** 538–561. [MR3689343 https://doi.org/10.1002/rsa.20713](https://doi.org/10.1002/rsa.20713)
- Litvak, A.E. and Rivasplata, O. (2012). Smallest singular value of sparse random matrices. *Studia Math.* **212** 195–218. [MR3009072 https://doi.org/10.4064/sm212-3-1](https://doi.org/10.4064/sm212-3-1)
- Livshyts, G.V. (2021). The smallest singular value of heavy-tailed not necessarily i.i.d. random matrices via random rounding. *J. Anal. Math.* **145** 257–306. [MR4361906 https://doi.org/10.1007/s11854-021-0183-2](https://doi.org/10.1007/s11854-021-0183-2)
- Livshyts, G.V., Tikhomirov, K. and Vershynin, R. (2021). The smallest singular value of inhomogeneous square random matrices. *Ann. Probab.* **49** 1286–1309. [MR4255145 https://doi.org/10.1214/20-aop1481](https://doi.org/10.1214/20-aop1481)
- Rudelson, M. and Vershynin, R. (2009). Smallest singular value of a random rectangular matrix. *Comm. Pure Appl. Math.* **62** 1707–1739. [MR2569075 https://doi.org/10.1002/cpa.20294](https://doi.org/10.1002/cpa.20294)
- Rudelson, M. and Vershynin, R. (2010). Non-asymptotic theory of random matrices: Extreme singular values. In *Proceedings of the International Congress of Mathematicians. Volume III* 1576–1602. New Delhi: Hindustan Book Agency. [MR2827856 https://doi.org/10.1007/978-81-322-0391-3\\_10](https://doi.org/10.1007/978-81-322-0391-3_10)
- Stephan, L. and Massoulié, L. (2022). Non-backtracking spectra of weighted inhomogeneous random graphs. *Math. Stat. Learn.* **5** 201–271. [MR4526300 https://doi.org/10.4171/msl/34](https://doi.org/10.4171/msl/34)
- Stephan, L. and Zhu, Y. (2022). Sparse random hypergraphs: Non-backtracking spectra and community detection. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science—FOCS 2022* 567–575. Los Alamitos, CA: IEEE Computer Soc. [MR4537236 https://doi.org/10.1109/FOCS52361.2022](https://doi.org/10.1109/FOCS52361.2022)
- Tikhomirov, K. (2015). The limit of the smallest singular value of random matrices with i.i.d. entries. *Adv. Math.* **284** 1–20. [MR3391069 https://doi.org/10.1016/j.aim.2015.07.020](https://doi.org/10.1016/j.aim.2015.07.020)
- Tikhomirov, K.E. (2016). The smallest singular value of random rectangular matrices with no moment assumptions on entries. *Israel J. Math.* **212** 289–314. [MR3504328 https://doi.org/10.1007/s11856-016-1287-8](https://doi.org/10.1007/s11856-016-1287-8)
- Tikhomirov, K. (2018). Sample covariance matrices of heavy-tailed distributions. *Int. Math. Res. Not. IMRN* **20** 6254–6289. [MR3872323 https://doi.org/10.1093/imrn/rnx067](https://doi.org/10.1093/imrn/rnx067)
- Tikhomirov, K. and Youssef, P. (2021). Outliers in spectrum of sparse Wigner matrices. *Random Structures Algorithms* **58** 517–605. [MR4234995 https://doi.org/10.1002/rsa.20982](https://doi.org/10.1002/rsa.20982)
- Tropp, J. A. (2015). An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.* **8** 1–230.
- van Handel, R. (2017). On the spectral norm of Gaussian random matrices. *Trans. Amer. Math. Soc.* **369** 8161–8178. [MR3695857 https://doi.org/10.1090/tran/6922](https://doi.org/10.1090/tran/6922)
- Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. [MR2963170 https://doi.org/10.1017/9781108231596](https://doi.org/10.1017/9781108231596)
- Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge Series in Statistical and Probabilistic Mathematics **47**. Cambridge: Cambridge Univ. Press. [MR3837109 https://doi.org/10.1017/9781108231596](https://doi.org/10.1017/9781108231596)
- Wan, Y. and Meila, M. (2015). A class of network models recoverable by spectral clustering. *Adv. Neural Inf. Process* **28**.
- Wang, K. and Wood, P.M. (2023). Limiting empirical spectral distribution for the non-backtracking matrix of an Erdős-Rényi random graph. *Combin. Probab. Comput.* **32** 956–973. [MR4653733 https://doi.org/10.1017/S096354832300024x](https://doi.org/10.1017/S096354832300024x)

- Watanabe, Y. and Fukumizu, K. (2009). Graph Zeta function in the Bethe free energy and loopy belief propagation. *Adv. Neural Inf. Process* **22**.
- Zhivotovskiy, N. (2021). Dimension-free bounds for sums of independent matrices and simple tensors via the variational principle. Available at [arXiv:2108.08198](https://arxiv.org/abs/2108.08198).
- Zhou, Z. and Amini, A.A. (2019). Analysis of spectral clustering algorithms for community detection: The general bipartite setting. *J. Mach. Learn. Res.* **20** Paper No. 47, 47. [MR3948087](https://arxiv.org/abs/1905.07832)
- Zhu, Y. (2020). A graphon approach to limiting spectral distributions of Wigner-type matrices. *Random Structures Algorithms* **56** 251–279. [MR4052853](https://arxiv.org/abs/1905.07832) <https://doi.org/10.1002/rsa.20894>
- Zhu, Y. (2023). On the second eigenvalue of random bipartite biregular graphs. *J. Theoret. Probab.* **36** 1269–1303. [MR4591874](https://arxiv.org/abs/2301.09590) <https://doi.org/10.1007/s10959-022-01190-0>



# Characterizing the asymptotic and catalytic stochastic orders on topological abelian groups

TOBIAS FRITZ<sup>a</sup>

*Department of Mathematics, University of Innsbruck, Innsbruck, Austria, <sup>a</sup>[tobias.fritz@uibk.ac.at](mailto:tobias.fritz@uibk.ac.at)*

We study the usual stochastic order between probability measures on preordered topological abelian groups, focusing on asymptotic and catalytic versions thereof. In the asymptotic version,  $\mu$  dominates  $\nu$  if the i.i.d. random walk generated by  $\mu$  first-order dominates the one generated by  $\nu$  at late times. In the catalytic version,  $\mu$  dominates  $\nu$  if there is a third measure  $\tau$  such that the convolution  $\mu * \tau$  first-order dominates  $\nu * \tau$ . Provided that the preorder on  $G$  is induced by a suitably large positive cone and that both measures are compactly supported Radon, our main result gives a sufficient condition for asymptotic and catalytic dominance to hold in terms of a family of inequalities closely related to the cumulant-generating functions. While this sufficient condition requires these inequalities to be strict, the non-strict versions of these inequalities are necessary. This result has been known for  $G = \mathbb{R}$ , but is new already for  $\mathbb{R}^n$  with  $n > 1$ . It is a direct application of a recently proven theorem of real algebra, namely a *Vergleichsstellensatz* for preordered semirings. We finally use our result to derive a formula for the rate at which the probabilities of a random walk decay *relative* to those of another, now for walks on any preordered topological vector space with compactly supported Radon steps. Taking one of these walks to be deterministic reproduces a version of Cramér’s large deviation theorem for infinite dimensions.

*Keywords:* Large deviation theory; preordered semiring; random walks; real algebra; usual stochastic order; Vergleichsstellensatz

## References

- [1] Aliprantis, C.D. and Tourky, R. (2007). *Cones and Duality. Graduate Studies in Mathematics* **84**. Providence, RI: Amer. Math. Soc. [MR2317344](#) <https://doi.org/10.1090/gsm/084>
- [2] Aubrun, G. and Nechita, I. (2009). Stochastic domination for iterated convolutions and catalytic majorization. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 611–625. [MR2548496](#) <https://doi.org/10.1214/08-AIHP175>
- [3] Bahadur, R.R. and Zabell, S.L. (1979). Large deviations of the sample mean in general vector spaces. *Ann. Probab.* **7** 587–621. [MR0537209](#)
- [4] Campbell, L.L. (1965). A coding theorem and Rényi’s entropy. *Inf. Control* **8** 423–429. [MR0180403](#)
- [5] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics (New York)* **38**. New York: Springer. [MR1619036](#) <https://doi.org/10.1007/978-1-4612-5320-4>
- [6] Edwards, D.A. (1978). On the existence of probability measures with given marginals. *Ann. Inst. Fourier (Grenoble)* **28** 53–78. [MR0513882](#)
- [7] Farooq, M.U., Fritz, T., Haapasalo, E. and Tomamichel, M. (2023). Asymptotic and catalytic matrix majorization. Available at [arXiv:2301.07353](https://arxiv.org/abs/2301.07353).
- [8] Folland, G.B. (2016). *A Course in Abstract Harmonic Analysis*, 2nd ed. *Textbooks in Mathematics*. Boca Raton, FL: CRC Press. [MR3444405](#)
- [9] Fremlin, D.H. (2006). *Measure Theory. Vol. 4*. Colchester: Torres Fremlin. Topological measure spaces. Part I, II, Corrected second printing of the 2003 original. [MR2462372](#)
- [10] Fritz, T. (2019). Antisymmetry of the stochastic order on all ordered topological spaces. *Anal. Geom. Metric Spaces* **7** 250–252. [MR4048385](#) <https://doi.org/10.1515/agms-2019-0012>
- [11] Fritz, T. (2023). Abstract Vergleichsstellensätze for preordered semifields and semirings I. *SIAM J. Appl. Algebra Geom.* **7** 505–547. [MR4609385](#) <https://doi.org/10.1137/22M1498413>

- [12] Fritz, T. (2023). Abstract Vergleichsstellensätze for preordered semifields and semirings II. Available at [arXiv:2112.05949](https://arxiv.org/abs/2112.05949).
- [13] Fritz, T. and Perrone, P. (2018). Bimonoidal structure of probability monads. In *Proceedings of the Thirty-Fourth Conference on the Mathematical Foundations of Programming Semantics (MFPS XXXIV)*. *Electron. Notes Theor. Comput. Sci.* **341** 121–149. Amsterdam: Elsevier Sci. B. V. [MR3899689 https://doi.org/10.1016/j.entcs.2018.11.007](https://doi.org/10.1016/j.entcs.2018.11.007)
- [14] Fritz, T. (2023). Asymptotic and catalytic containment of representations of  $SU(n)$ . *Algebraic Combin.* To appear.
- [15] Golan, J.S. (1999). *Semirings and Their Applications*. Dordrecht: Kluwer Academic. Updated and expanded version of it The theory of semirings, with applications to mathematics and theoretical computer science [Longman Sci. Tech., Harlow, 1992; MR1163371 (93b:16085)]. [MR1746739 https://doi.org/10.1007/978-94-015-9333-5](https://doi.org/10.1007/978-94-015-9333-5)
- [16] Goodearl, K.R. (1986). *Partially Ordered Abelian Groups with Interpolation*. *Mathematical Surveys and Monographs* **20**. Providence, RI: Amer. Math. Soc. [MR0845783 https://doi.org/10.1090/surv/020](https://doi.org/10.1090/surv/020)
- [17] Hadar, J. and Russell, W.R. (1969). Rules for ordering uncertain prospects. *Amer. Econ. Rev.* **59** 25–34.
- [18] Kellerer, H.G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** 399–432. [MR0761565 https://doi.org/10.1007/BF00532047](https://doi.org/10.1007/BF00532047)
- [19] Klenke, A. (2014). *Probability Theory*, 2nd ed. *Universitext*. London: Springer. German, A comprehensive course. [MR3112259 https://doi.org/10.1007/978-1-4471-5361-0](https://doi.org/10.1007/978-1-4471-5361-0)
- [20] Mas-Colell, A. and Whinston, M.D. (1995). *Microeconomic Theory*. Oxford: Oxford Univ. Press.
- [21] Pomatto, L., Strack, P. and Tamuz, O. (2020). Stochastic dominance under independent noise. *J. Polit. Econ.* **128**.
- [22] Schwartz, L. (1973). *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. *Tata Institute of Fundamental Research Studies in Mathematics, No. 6*. London: Bombay by Oxford University Press. Published for the Tata Institute of Fundamental Research. [MR0426084](https://doi.org/10.1017/9780521216768)
- [23] 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)

# Optimal 1-Wasserstein distance for WGANs

ARTHUR STÉPHANOVITCH<sup>1,a</sup>, UGO TANELIAN<sup>2,b</sup>, BENOÎT CADRE<sup>3,c</sup>,  
NICOLAS KLUTCHNIKOFF<sup>3,d</sup> and GÉRARD BIAU<sup>4,e</sup>

<sup>1</sup>Université Paris Cité, CNRS, LPSM, F-75013 Paris, France, <sup>a</sup>[stephanovitch@lpsm.paris](mailto:stephanovitch@lpsm.paris)

<sup>2</sup>Criteo AI Lab, Paris, France, <sup>b</sup>[ugo.tanielian@gmail.com](mailto:ugo.tanielian@gmail.com)

<sup>3</sup>Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France, <sup>c</sup>[benoit.cadre@univ-rennes2.fr](mailto:benoit.cadre@univ-rennes2.fr),

<sup>d</sup>[nicolas.klutchnikoff@univ-rennes2.fr](mailto:nicolas.klutchnikoff@univ-rennes2.fr)

<sup>4</sup>Sorbonne Université, CNRS, LPSM, F-75005 Paris, France, <sup>e</sup>[gerard.biau@sorbonne-universite.fr](mailto:gerard.biau@sorbonne-universite.fr)

The mathematical forces at work behind Generative Adversarial Networks raise challenging theoretical issues. Motivated by the important question of characterizing the geometrical properties of the generated distributions, we provide a thorough analysis of Wasserstein GANs (WGANs) in both the finite sample and asymptotic regimes. We study the specific case where the latent space is univariate and derive results valid regardless of the dimension of the output space. We show in particular that for a fixed sample size, the optimal WGANs are closely linked with connected paths minimizing the sum of the squared Euclidean distances between the sample points. We also highlight the fact that WGANs are able to approach (for the 1-Wasserstein distance) the target distribution as the sample size tends to infinity, at a given convergence rate and provided the family of generative Lipschitz functions grows appropriately. We derive in passing new results on optimal transport theory in the semi-discrete setting.

**Keywords:** Optimal distribution; optimal transport theory; rate of convergence; shortest path; Wasserstein distance; Wasserstein Generative Adversarial Networks

## References

- Arjovsky, M., Chintala, S. and Bottou, L. (2017). Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning* (D. Precup and Y.W. Teh, eds.) **70** 214–223. PMLR.
- Aurenhammer, F., Hoffmann, F. and Aronov, B. (1998). Minkowski-type theorems and least-squares clustering. *Algorithmica* **20** 61–76. [MR1483422 https://doi.org/10.1007/PL00009187](https://doi.org/10.1007/PL00009187)
- Biau, G., Sangnier, M. and Tanielian, U. (2021). Some theoretical insights into Wasserstein GANs. *J. Mach. Learn. Res.* **22** Paper No. 119, 45. [MR4279770](https://doi.org/10.48550/jmlr.2021.22.119)
- Borji, A. (2019). Pros and cons of GAN evaluation measures. *Comput. Vis. Image Underst.* **179** 41–65.
- Deheuvels, P. (1986). On the influence of the extremes of an i.i.d. sequence on the maximal spacings. *Ann. Probab.* **14** 194–208. [MR0815965](https://doi.org/10.2307/2346933)
- Facco, E., d’Errico, M., Rodriguez, A. and Laio, A. (2017). Estimating the intrinsic dimension of datasets by a minimal neighborhood information. *Sci. Rep.* **7** 12140. <https://doi.org/10.1038/s41598-017-11873-y>
- Fefferman, C., Mitter, S. and Narayanan, H. (2016). Testing the manifold hypothesis. *J. Amer. Math. Soc.* **29** 983–1049. [MR3522608 https://doi.org/10.1090/jams/852](https://doi.org/10.1090/jams/852)
- Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. [MR3383341 https://doi.org/10.1007/s00440-014-0583-7](https://doi.org/10.1007/s00440-014-0583-7)
- Geiß, D., Klein, R., Penninger, R. and Rote, G. (2013). Optimally solving a transportation problem using Voronoi diagrams. *Comput. Geom.* **46** 1009–1016. [MR3061462 https://doi.org/10.1016/j.comgeo.2013.05.005](https://doi.org/10.1016/j.comgeo.2013.05.005)
- Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence and K.Q. Weinberger, eds.) **27** 2672–2680. Curran Associates.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A.C. (2017). Improved training of Wasserstein GANs. In *Advances in Neural Information Processing Systems* (I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett, eds.) **30** 5767–5777. Curran Associates.

- Gulrajani, I., Raffel, C. and Metz, L. (2019). Towards GAN benchmarks which require generalization. In *International Conference on Learning Representations*.
- Hartmann, V. and Schuhmacher, D. (2020). Semi-discrete optimal transport: A solution procedure for the un-squared Euclidean distance case. *Math. Methods Oper. Res.* **92** 133–163. [MR4152920](#) <https://doi.org/10.1007/s00186-020-00703-z>
- Hur, Y., Guo, W. and Liang, T. (2021). Reversible Gromov–Monge sampler for simulation-based inference. [arXiv:2109.14090](#).
- Kantorovič, L.V. and Rubiňštejn, G.Š. (1958). On a space of completely additive functions. *Vestn. Leningr. Univ.* **13** 52–59. [MR0102006](#)
- Karras, T., Aila, T., Laine, S. and Lehtinen, J. (2018). Progressive growing of GANs for improved quality, stability, and variation. In *International Conference on Learning Representations*.
- Karras, T., Laine, S. and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition* 4396–4405. [MR4351661](#)
- Kodali, N., Abernethy, J., Hays, J. and Kira, Z. (2017). On convergence and stability of GANs. Available at [arXiv:1705.07215](#).
- Liang, T. (2021). How well generative adversarial networks learn distributions. *J. Mach. Learn. Res.* **22** Paper No. 228, 41. [MR4329807](#)
- Lucic, M., Kurach, K., Michalski, M., Gelly, S. and Bousquet, O. (2018). Are GANs created equal? A large-scale study. In *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi and R. Garnett, eds.) **31** 697–706. Curran Associates.
- Luise, G., Pontil, M. and Ciliberto, C. (2020). Generalization properties of optimal transport GANs with latent distribution learning. Available at [arXiv:2007.14641](#).
- Mescheder, L., Geiger, A. and Nowozin, S. (2018). Which training methods for GANs do actually converge? In *Proceedings of the 35th International Conference on Machine Learning* (J. Dy and A. Krause, eds.) **80** 3481–3490. PMLR.
- Müller, A. (1997). Integral probability metrics and their generating classes of functions. *Adv. in Appl. Probab.* **29** 429–443. [MR1450938](#) <https://doi.org/10.2307/1428011>
- Pratelli, A. (2007). On the equality between Monge’s infimum and Kantorovich’s minimum in optimal mass transportation. *Ann. Inst. Henri Poincaré Probab. Stat.* **43** 1–13. [MR2288266](#) <https://doi.org/10.1016/j.anihpb.2005.12.001>
- Radford, A., Metz, L. and Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. In *4th International Conference on Learning Representations* (Y. Bengio and Y. LeCun, eds.).
- Schreuder, N., Brunel, V.-E. and Dalalyan, A.S. (2021). Statistical guarantees for generative models without domination. In *Algorithmic Learning Theory. Proc. Mach. Learn. Res. (PMLR)* **132** 21. [place of publication not identified]. [MR4227353](#)
- Singh, S., Uppal, A., Li, B., Li, C.L., Zaheer, M. and Poczos, B. (2018). Nonparametric density estimation under adversarial losses. In *Advances in Neural Information Processing Systems* (S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi and R. Garnett, eds.) **31** 10225–10236. Curran Associates.
- Stéphanovitch, A., Tanielian, U., Cadre, B., Klutchnikoff, N. and Biau, G. (2024). Supplement to “Optimal 1-Wasserstein distance for WGANs.” <https://doi.org/10.3150/23-BEJ1701SUPP>
- Tanielian, U., Issenhuth, T., Dohmatob, E. and Mary, J. (2020). Learning disconnected manifolds: A no GAN’s land. In *Proceedings of the 37th International Conference on Machine Learning* (H. Daumé III and A. Singh, eds.) **119** 9418–9427. PMLR.
- Uppal, A., Singh, S. and Poczos, B. (2019). Nonparametric density estimation and convergence rates for GANs under Besov IPM losses. In *Advances in Neural Information Processing Systems* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d’Alché-Buc, E. Fox and R. Garnett, eds.) **32** 9089–9100. Curran Associates.
- Vaishnavh, N., Raffel, C. and Goodfellow, I.J. (2018). Theoretical insights into memorization in GANs. In *Neural Information Processing Systems 2018 – Integration of Deep Learning Theories Workshop*.
- 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>

- Vondrick, C., Pirsaviash, H. and Torralba, A. (2016). Generating videos with scene dynamics. In *Advances in Neural Information Processing Systems* (D. Lee, M. Sugiyama, U. von Luxburg, I. Guyon and R. Garnett, eds.) **29** 613–621. Curran Associates.
- Yu, L., Zhang, W., Wang, J. and Yu, Y. (2017). SeqGAN: Sequence generative adversarial nets with policy gradient. In *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* 2852–2858. AAAI Press.
- Zhou, Z., Liang, J., Song, Y., Yu, L., Wang, H., Zhang, W., Yu, Y. and Zhang, Z. (2019). Lipschitz generative adversarial nets. In *Proceedings of the 36th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.) **97** 7584–7593. PMLR.

# Principal feature detection via $\phi$ -Sobolev inequalities

MATTHEW T.C. LI<sup>1,a</sup>, YOUSSEF MARZOUK<sup>1,b</sup> and OLIVIER ZAHM<sup>2,c</sup>

<sup>1</sup>Massachusetts Institute of Technology, Cambridge, MA 02139, USA, <sup>a</sup>[mtcli@mit.edu](mailto:mtcli@mit.edu), <sup>b</sup>[ymarz@mit.edu](mailto:ymarz@mit.edu)

<sup>2</sup>Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France, <sup>c</sup>[olivier.zahm@inria.fr](mailto:olivier.zahm@inria.fr)

We investigate the approximation of high-dimensional target measures as low-dimensional updates of a dominating reference measure. This approximation class replaces the associated density with the composition of: (i) a feature map that identifies the leading principal components or features of the target measure, relative to the reference, and (ii) a low-dimensional profile function. When the reference measure satisfies a subspace  $\phi$ -Sobolev inequality, we construct a computationally tractable approximation that yields certifiable error guarantees with respect to the Amari  $\alpha$ -divergences. Our construction proceeds in two stages. First, for any feature map and any  $\alpha$ -divergence, we obtain an analytical expression for the optimal profile function. Second, for linear feature maps, the principal features are obtained from eigenvectors of a matrix involving gradients of the log-density. Neither step requires explicit access to normalizing constants. Notably, by leveraging the  $\phi$ -Sobolev inequalities, we demonstrate that these features universally certify approximation errors across the range of  $\alpha$ -divergences  $\alpha \in (0, 1]$ . We then propose an application to Bayesian inverse problems and provide an analogous construction with approximation guarantees that hold in expectation over the data. We conclude with an extension of the proposed dimension reduction strategy to nonlinear feature maps.

**Keywords:** Amari  $\alpha$ -divergences; Bayesian inference; feature detection; gradient-based dimension reduction; principal components;  $\phi$ -Sobolev inequalities

## References

- [1] Amari, S.-I. (2009).  $\alpha$ -divergence is unique, belonging to both  $f$ -divergence and Bregman divergence classes. *IEEE Trans. Inf. Theory* **55** 4925–4931. [MR2596950](#) <https://doi.org/10.1109/TIT.2009.2030485>
- [2] Andrieu, C., Lee, A.W.L., Power, S. and Wang, A. (2022). Explicit convergence bounds for Metropolis Markov chains: Isoperimetry, spectral gaps and profiles. Available at [arXiv:2211.08959](#).
- [3] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Cham: Springer. [MR3155209](#) <https://doi.org/10.1007/978-3-319-00227-9>
- [4] Bakry, D. and Ledoux, M. (2006). A logarithmic Sobolev form of the Li–Yau parabolic inequality. *Rev. Mat. Iberoam.* **22** 683–702. [MR2294794](#) <https://doi.org/10.4171/RMI/470>
- [5] Banerjee, A., Guo, X. and Wang, H. (2005). On the optimality of conditional expectation as a Bregman predictor. *IEEE Trans. Inf. Theory* **51** 2664–2669. [MR2246384](#) <https://doi.org/10.1109/TIT.2005.850145>
- [6] Baptista, R., Marzouk, Y. and Zahm, O. (2022). Gradient-based data and parameter dimension reduction for Bayesian models: An information theoretic perspective. Available at [arXiv:2207.08670](#).
- [7] Beckner, W. (1989). A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.* **105** 397–400. [MR0954373](#) <https://doi.org/10.2307/2046956>
- [8] Beskos, A., Girolami, M., Lan, S., Farrell, P.E. and Stuart, A.M. (2017). Geometric MCMC for infinite-dimensional inverse problems. *J. Comput. Phys.* **335** 327–351. [MR3612501](#) <https://doi.org/10.1016/j.jcp.2016.12.041>
- [9] Bigoni, D., Marzouk, Y., Prieur, C. and Zahm, O. (2022). Nonlinear dimension reduction for surrogate modeling using gradient information. *Inf. Inference* **11** 1597–1639. [MR4526330](#) <https://doi.org/10.1093/imaiai/iaac006>

- [10] Bolley, F. and Gentil, I. (2010). Phi-entropy inequalities for diffusion semigroups. *J. Math. Pures Appl.* **93** 449–473. [MR2609029 https://doi.org/10.1016/j.matpur.2010.02.004](https://doi.org/10.1016/j.matpur.2010.02.004)
- [11] Brennan, M., Bigoni, D., Zahm, O., Spantini, A. and Marzouk, Y. (2020). Greedy inference with structure-exploiting lazy maps. In *Adv. Neural Inf. Process.* **33** 8330–8342.
- [12] Canonne, C.L. (2022). A short note on an inequality between KL and TV. Available at [arXiv:2202.07198](https://arxiv.org/abs/2202.07198).
- [13] Chafaï, D. (2004). Entropies, convexity, and functional inequalities: On  $\Phi$ -entropies and  $\Phi$ -Sobolev inequalities. *J. Math. Kyoto Univ.* **44** 325–363. [MR2081075 https://doi.org/10.1215/kjm/1250283556](https://doi.org/10.1215/kjm/1250283556)
- [14] Chen, P. and Ghattas, O. (2020). Projected Stein variational gradient descent. In *Adv. Neural Inf. Process. Syst.* **33** 1947–1958.
- [15] Chewi, S., Erdogdu, M.A., Li, M., Shen, R. and Zhang, S. (2022). Analysis of Langevin Monte Carlo from Poincaré to Log-Sobolev. In *Proceedings of Thirty Fifth Conference on Learning Theory* (P.-L. Loh and M. Raginsky, eds.). *PMLR* **178** 1–2. Proc. Mach. Learn. Res.
- [16] Constantine, P.G. (2015). *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*. *SIAM Spotlights* **2**. Philadelphia, PA: SIAM. [MR3486165 https://doi.org/10.1137/1.9781611973860](https://doi.org/10.1137/1.9781611973860)
- [17] Constantine, P.G. and Diaz, P. (2017). Global sensitivity metrics from active subspaces. *Reliab. Eng. Syst. Saf.* **162** 1–13.
- [18] Constantine, P.G., Kent, C. and Bui-Thanh, T. (2016). Accelerating Markov chain Monte Carlo with active subspaces. *SIAM J. Sci. Comput.* **38** A2779–A2805. [MR3543164 https://doi.org/10.1137/15M1042127](https://doi.org/10.1137/15M1042127)
- [19] Cotter, S.L., Roberts, G.O., Stuart, A.M. and White, D. (2013). MCMC methods for functions: Modifying old algorithms to make them faster. *Statist. Sci.* **28** 424–446. [MR3135540 https://doi.org/10.1214/13-STS421](https://doi.org/10.1214/13-STS421)
- [20] Cui, T., Dolgov, S. and Zahm, O. (2023). Scalable conditional deep inverse Rosenblatt transports using tensor trains and gradient-based dimension reduction. *J. Comput. Phys.* **485** Paper No. 112103, 31. [MR4574525 https://doi.org/10.1016/j.jcp.2023.112103](https://doi.org/10.1016/j.jcp.2023.112103)
- [21] Cui, T., Law, K.J.H. and Marzouk, Y.M. (2016). Dimension-independent likelihood-informed MCMC. *J. Comput. Phys.* **304** 109–137. [MR3422405 https://doi.org/10.1016/j.jcp.2015.10.008](https://doi.org/10.1016/j.jcp.2015.10.008)
- [22] Cui, T., Martin, J., Marzouk, Y.M., Solonen, A. and Spantini, A. (2014). Likelihood-informed dimension reduction for nonlinear inverse problems. *Inverse Probl.* **30** 114015, 28. [MR3274599 https://doi.org/10.1088/0266-5611/30/11/114015](https://doi.org/10.1088/0266-5611/30/11/114015)
- [23] Cui, T. and Tong, X.T. (2022). A unified performance analysis of likelihood-informed subspace methods. *Bernoulli* **28** 2788–2815. [MR4474562 https://doi.org/10.3150/21-bej1437](https://doi.org/10.3150/21-bej1437)
- [24] Cui, T., Tong, X.T. and Zahm, O. (2022). Prior normalization for certified likelihood-informed subspace detection of Bayesian inverse problems. *Inverse Probl.* **38** Paper No. 124002, 36. [MR4500900 https://doi.org/10.1088/1361-6420/ac9582](https://doi.org/10.1088/1361-6420/ac9582)
- [25] Cui, T. and Zahm, O. (2021). Data-free likelihood-informed dimension reduction of Bayesian inverse problems. *Inverse Probl.* **37** Paper No. 045009, 41. [MR4234453 https://doi.org/10.1088/1361-6420/abeafb](https://doi.org/10.1088/1361-6420/abeafb)
- [26] Del Moral, P., Doucet, A. and Jasra, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 411–436. [MR2278333 https://doi.org/10.1111/j.1467-9868.2006.00553.x](https://doi.org/10.1111/j.1467-9868.2006.00553.x)
- [27] Ehre, M., Flock, R., Fußeder, M., Papaioannou, I. and Straub, D. (2023). Certified dimension reduction for Bayesian updating with the cross-entropy method. *SIAM/ASA J. Uncertain. Quantificat.* **11** 358–388. [MR4558760 https://doi.org/10.1137/22M1484031](https://doi.org/10.1137/22M1484031)
- [28] Fan, K. (1949). On a theorem of Weyl concerning eigenvalues of linear transformations. I. *Proc. Natl. Acad. Sci. USA* **35** 652–655. [MR0034519 https://doi.org/10.1073/pnas.35.11.652](https://doi.org/10.1073/pnas.35.11.652)
- [29] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. [MR0420249 https://doi.org/10.2307/2373688](https://doi.org/10.2307/2373688)
- [30] Guillin, A., Léonard, C., Wu, L. and Yao, N. (2009). Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields* **144** 669–695. [MR2496446 https://doi.org/10.1007/s00440-008-0159-5](https://doi.org/10.1007/s00440-008-0159-5)
- [31] Harremoës, P. and Vajda, I. (2011). On pairs of  $f$ -divergences and their joint range. *IEEE Trans. Inf. Theory* **57** 3230–3235. [MR2817015 https://doi.org/10.1109/TIT.2011.2137353](https://doi.org/10.1109/TIT.2011.2137353)
- [32] Kaipio, J. and Somersalo, E. (2005). *Statistical and Computational Inverse Problems*. *Applied Mathematical Sciences* **160**. New York: Springer. [MR2102218](https://doi.org/10.1007/978-1-4939-9832-9)
- [33] Kallenberg, O. (1997). *Foundations of Modern Probability. Probability and Its Applications (New York)*. New York: Springer. [MR1464694](https://doi.org/10.1007/978-1-4939-9832-9)

- [34] Kim, K.-T., Villa, U., Parno, M., Marzouk, Y., Ghattas, O. and Petra, N. (2023). HIPPLYlib-MUQ: A Bayesian inference software framework for integration of data with complex predictive models under uncertainty. *ACM Trans. Math. Software* **49** Art. 17, 31. [MR4616414 https://doi.org/10.1145/3580278](https://doi.org/10.1145/3580278)
- [35] Kingma, D.P. and Welling, M. (2014). Auto-encoding variational Bayes. In *2nd International Conference on Learning Representations, ICLR*.
- [36] Laparra, V., Camps-Valls, G. and Malo, J. (2011). Iterative Gaussianization: From ICA to random rotations. *IEEE Trans. Neural Netw.* **22** 537–549. <https://doi.org/10.1109/TNN.2011.2106511>
- [37] Latała, R. and Oleszkiewicz, K. (2000). Between Sobolev and Poincaré. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1745** 147–168. Berlin: Springer. [MR1796718 https://doi.org/10.1007/BFb0107213](https://doi.org/10.1007/BFb0107213)
- [38] Lee, M.R. (2019). Modified active subspaces using the average of gradients. *SIAM/ASA J. Uncertain. Quantificat.* **7** 53–66. [MR3895346 https://doi.org/10.1137/17M1140662](https://doi.org/10.1137/17M1140662)
- [39] Li, M.T.C., Marzouk, Y. and Zahm, O. (2024). Supplement to “Principal feature detection via  $\phi$ -Sobolev inequalities.” <https://doi.org/10.3150/23-BEJ1702SUPP>
- [40] Liese, F. and Vajda, I. (2006). On divergences and informations in statistics and information theory. *IEEE Trans. Inf. Theory* **52** 4394–4412. [MR2300826 https://doi.org/10.1109/TIT.2006.881731](https://doi.org/10.1109/TIT.2006.881731)
- [41] Liu, S. and Owen, A.B. (2023). Preintegration via active subspace. *SIAM J. Numer. Anal.* **61** 495–514. [MR4558754 https://doi.org/10.1137/22M1479129](https://doi.org/10.1137/22M1479129)
- [42] Liu, X., Zhu, H., Ton, J.-F., Wynne, G. and Duncan, A. (2022). Grassmann Stein variational gradient descent. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics. PMLR 2002–2021*. Proc. Mach. Learn. Res.
- [43] Mangoubi, O. and Vishnoi, N.K. (2018). Dimensionally tight bounds for second-order Hamiltonian Monte Carlo. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems. NIPS’18* 6030–6040. Red Hook, NY, USA: Curran Associates Inc.
- [44] Marzouk, Y., Moselhy, T., Parno, M. and Spantini, A. (2017). Sampling via measure transport: An introduction. In *Handbook of Uncertainty Quantification. Vol. 1, 2, 3* 785–825. Cham: Springer. [MR3821485](https://doi.org/10.1007/978-1-4939-9832-7_11)
- [45] Papamakarios, G., Nalisnick, E., Rezende, D.J., Mohamed, S. and Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.* **22** Paper No. 57, 64. [MR4253750](https://doi.org/10.48550/jmlr.2021.22.1)
- [46] Pillai, N.S., Stuart, A.M. and Thiéry, A.H. (2012). Optimal scaling and diffusion limits for the Langevin algorithm in high dimensions. *Ann. Appl. Probab.* **22** 2320–2356. [MR3024970 https://doi.org/10.1214/11-AAP828](https://doi.org/10.1214/11-AAP828)
- [47] Pillaud-Vivien, L., Bach, F., Lelièvre, T., Rudi, A. and Stoltz, G. (2020). Statistical estimation of the Poincaré constant and application to sampling multimodal distributions. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics. PMLR 2753–2763*. Proc. Mach. Learn. Res.
- [48] Polyanskiy, Y. and Wu, Y. (2023). *Information Theory: From Coding to Learning*. Cambridge University Press. Forthcoming.
- [49] Rebeschini, P. and van Handel, R. (2015). Can local particle filters beat the curse of dimensionality? *Ann. Appl. Probab.* **25** 2809–2866. [MR3375889 https://doi.org/10.1214/14-AAP1061](https://doi.org/10.1214/14-AAP1061)
- [50] Rezende, D. and Mohamed, S. (2015). Variational inference with normalizing flows. In *Proceedings of the 32nd International Conference on Machine Learning* (F. Bach and D. Blei, eds.). *PMLR* **37** 1530–1538.
- [51] Rezende, D.J., Mohamed, S. and Wierstra, D. (2014). Stochastic backpropagation and approximate inference in deep generative models. In *Proceedings of the 31st International Conference on Machine Learning* (E.P. Xing and T. Jebara, eds.). *PMLR* **32** 1278–1286.
- [52] Roberts, G.O. and Rosenthal, J.S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 255–268. [MR1625691 https://doi.org/10.1111/1467-9868.00123](https://doi.org/10.1111/1467-9868.00123)
- [53] 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](https://doi.org/10.1214/154957804100000024)
- [54] Roy, O. and Vetterli, M. (2007). The effective rank: A measure of effective dimensionality. In *2007 15th European Signal Processing Conference* 606–610. IEEE.
- [55] Rudolf, D. and Sprungk, B. (2018). On a generalization of the preconditioned Crank–Nicolson Metropolis algorithm. *Found. Comput. Math.* **18** 309–343. [MR3777781 https://doi.org/10.1007/s10208-016-9340-x](https://doi.org/10.1007/s10208-016-9340-x)
- [56] Samarov, A.M. (1993). Exploring regression structure using nonparametric functional estimation. *J. Amer. Statist. Assoc.* **88** 836–847. [MR1242934](https://doi.org/10.2307/22934)



- [57] Snyder, M., Bengtsson, T., Bickel, P. and Anderson, L. (2008). Obstacles to high-dimensional particle filtering. *Mon. Weather Rev.* 4629–4640.
- [58] Spantini, A., Solonen, A., Cui, T., Martin, J., Tenorio, L. and Marzouk, Y. (2015). Optimal low-rank approximations of Bayesian linear inverse problems. *SIAM J. Sci. Comput.* **37** A2451–A2487. [MR3418226](#) <https://doi.org/10.1137/140977308>
- [59] Stuart, A.M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. [MR2652785](#) <https://doi.org/10.1017/S0962492910000061>
- [60] Tabak, E.G. and Turner, C.V. (2013). A family of nonparametric density estimation algorithms. *Comm. Pure Appl. Math.* **66** 145–164. [MR2999294](#) <https://doi.org/10.1002/cpa.21423>
- [61] Tong, S. and Stadler, G. (2023). Large deviation theory-based adaptive importance sampling for rare events in high dimensions. *SIAM/ASA J. Uncertain. Quantificat.* **11** 788–813. [MR4612613](#) <https://doi.org/10.1137/22M1524758>
- [62] Uribe, F., Papaioannou, I., Marzouk, Y.M. and Straub, D. (2021). Cross-entropy-based importance sampling with failure-informed dimension reduction for rare event simulation. *SIAM/ASA J. Uncertain. Quantificat.* **9** 818–847. [MR4274842](#) <https://doi.org/10.1137/20M1344585>
- [63] Vempala, S.S. and Wibisono, A. (2023). Rapid convergence of the unadjusted Langevin algorithm: Isoperimetry suffices. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **2327** 381–438. Cham: Springer. [MR4651229](#) [https://doi.org/10.1007/978-3-031-26300-2\\_15](https://doi.org/10.1007/978-3-031-26300-2_15)
- [64] Zahm, O., Constantine, P.G., Prieur, C. and Marzouk, Y.M. (2020). Gradient-based dimension reduction of multivariate vector-valued functions. *SIAM J. Sci. Comput.* **42** A534–A558. [MR4069329](#) <https://doi.org/10.1137/18M1221837>
- [65] Zahm, O., Cui, T., Law, K., Spantini, A. and Marzouk, Y. (2022). Certified dimension reduction in nonlinear Bayesian inverse problems. *Math. Comp.* **91** 1789–1835. [MR4435948](#) <https://doi.org/10.1090/mcom/3737>

# De Finetti’s theorem and related results for infinite weighted exchangeable sequences

RINA FOYGEL BARBER<sup>1,a</sup>, EMMANUEL J. CANDÈS<sup>2,b</sup>, AADITYA RAMDAS<sup>3,c</sup> and RYAN J. TIBSHIRANI<sup>4,d</sup>

<sup>1</sup>*Department of Statistics, University of Chicago, Chicago, IL, USA, [rina@uchicago.edu](mailto:rina@uchicago.edu)*

<sup>2</sup>*Departments of Statistics and Mathematics, Stanford University, Stanford, CA, USA, [candes@stanford.edu](mailto:candes@stanford.edu)*

<sup>3</sup>*Departments of Statistics and Machine Learning, Carnegie Mellon University, Pittsburgh, PA, USA,*

*[aramdas@cmu.edu](mailto:aramdas@cmu.edu)*

<sup>4</sup>*Department of Statistics, University of California, Berkeley, CA, USA, [ryantibs@berkeley.edu](mailto:ryantibs@berkeley.edu)*

De Finetti’s theorem, also called the de Finetti–Hewitt–Savage theorem, is a foundational result in probability and statistics. Roughly, it says that an infinite sequence of exchangeable random variables can always be written as a mixture of independent and identically distributed (i.i.d.) sequences of random variables. In this paper, we consider a weighted generalization of exchangeability that allows for weight functions to modify the individual distributions of the random variables along the sequence, provided that – modulo these weight functions – there is still some common exchangeable base measure. We study conditions under which a de Finetti-type representation exists for weighted exchangeable sequences, as a mixture of distributions which satisfy a weighted form of the i.i.d. property. Our approach establishes a nested family of conditions that lead to weighted extensions of other well-known related results as well, in particular, extensions of the zero-one law and the law of large numbers.

*Keywords:* De Finetti; exchangeability

## References

- Alam, I. (2020). Generalizing the de Finetti–Hewitt–Savage theorem. arXiv preprint. Available at [arXiv:2008.08754](https://arxiv.org/abs/2008.08754).
- Aldous, D.J. (1985). Exchangeability and related topics. In *École D’été de Probabilités de Saint-Flour, XIII – 1983. Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. [MR0883646](https://doi.org/10.1007/BFb0099421) <https://doi.org/10.1007/BFb0099421>
- Barber, R.F., Candès, E.J., Ramdas, A. and Tibshirani, R.J. (2024). Supplement to “De Finetti’s theorem and related results for infinite weighted exchangeable sequences.” <https://doi.org/10.3150/23-BEJ1704SUPP>
- Bogachev, V.I. (2007). *Measure Theory. Vol. I, II.* Berlin: Springer. [MR2267655](https://doi.org/10.1007/978-3-540-34514-5) <https://doi.org/10.1007/978-3-540-34514-5>
- Candès, E., Lei, L. and Ren, Z. (2023). Conformalized survival analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 24–45.
- de Finetti, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. [MR1508036](https://doi.org/10.1016/S0021-8995(37)90040-1)
- de Finetti, B. (1929). Funzione caratteristica di un fenomeno aleatorio. In *Atti del Congresso Internazionale dei Matematici* 179–190.
- Diaconis, P. (1977). Finite forms of de Finetti’s theorem on exchangeability. *Synthese* **36** 271–281. Foundations of probability and statistics, II. [MR0517222](https://doi.org/10.1007/BF00486116) <https://doi.org/10.1007/BF00486116>
- Diaconis, P. and Freedman, D. (1980). Finite exchangeable sequences. *Ann. Probab.* **8** 745–764. [MR0577313](https://doi.org/10.2307/2346183)
- Diaconis, P. and Freedman, D. (1987). A dozen de Finetti-style results in search of a theory. *Ann. Inst. Henri Poincaré Probab. Stat.* **23** 397–423. [MR0898502](https://doi.org/10.1016/S0021-8995(87)90040-1)
- Dubins, L.E. and Freedman, D.A. (1979). Exchangeable processes need not be mixtures of independent, identically distributed random variables. *Z. Wahrsch. Verw. Gebiete* **48** 115–132. [MR0534840](https://doi.org/10.1007/BF01886868) <https://doi.org/10.1007/BF01886868>

- Dynkin, E.B. (1953). Classes of equivalent random quantities. *Uspekhi Mat. Nauk* **8** 125–130. [MR0055601](#)
- Faden, A.M. (1985). The existence of regular conditional probabilities: Necessary and sufficient conditions. *Ann. Probab.* **13** 288–298. [MR0770643](#)
- Fannjiang, C., Bates, S., Angelopoulos, A.N., Listgarten, J. and Jordan, M.I. (2022). Conformal prediction under feedback covariate shift for biomolecular design. *Proc. Natl. Acad. Sci. USA* **119** Paper No. e2204569119, 12. [MR4542869](#) <https://doi.org/10.1073/pnas.2204569119>
- Farrell, R.H. (1962). Representation of invariant measures. *Illinois J. Math.* **6** 447–467. [MR0150264](#)
- Fritz, T., Gonda, T. and Perrone, P. (2021). De Finetti's theorem in categorical probability. *J. Stoch. Anal.* **2** 4. [MR4344737](#)
- Hewitt, E. and Savage, L.J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501. [MR0076206](#) <https://doi.org/10.2307/1992999>
- Kingman, J.F.C. (1978). Uses of exchangeability. *Ann. Probab.* **6** 183–197. [MR0494344](#) <https://doi.org/10.1214/aop/1176995566>
- Knopp, K. (1990). *Theory and Application of Infinite Series*. Dover Publications. [MR0077665](#)
- Kolmogorov, A.N. (1930). Sur la loi forte des grands nombres. *C. R. Acad. Sci., Sér. I Math.* **191** 910–912.
- Lauritzen, S.L. (1988). *Extremal Families and Systems of Sufficient Statistics. Lecture Notes in Statistics* **49**. New York: Springer. [MR0971253](#) <https://doi.org/10.1007/978-1-4612-1023-8>
- Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J. and Wasserman, L. (2018). Distribution-free predictive inference for regression. *J. Amer. Statist. Assoc.* **113** 1094–1111. [MR3862342](#) <https://doi.org/10.1080/01621459.2017.1307116>
- Lei, L. and Candès, E.J. (2021). Conformal inference of counterfactuals and individual treatment effects. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 911–938. [MR4349122](#)
- Maitra, A. (1977). Integral representations of invariant measures. *Trans. Amer. Math. Soc.* **229** 209–225. [MR0442197](#) <https://doi.org/10.2307/1998506>
- Podkopaev, A. and Ramdas, A. (2021). Distribution-free uncertainty quantification for classification under label shift. In *Uncertainty in Artificial Intelligence* 844–853. PMLR.
- Schervish, M.J. (1995). *Theory of Statistics. Springer Series in Statistics*. New York: Springer. [MR1354146](#) <https://doi.org/10.1007/978-1-4612-4250-5>
- Tang, W. (2023). Finite and infinite weighted exchangeable sequences. arXiv preprint. Available at [arXiv:2306.11584](https://arxiv.org/abs/2306.11584).
- Tibshirani, R.J., Barber, R.F., Candès, E.J. and Ramdas, A. (2019). Conformal prediction under covariate shift. In *Advances in Neural Information Processing Systems*.
- Varadarajan, V.S. (1963). Groups of automorphisms of Borel spaces. *Trans. Amer. Math. Soc.* **109** 191–220. [MR0159923](#) <https://doi.org/10.2307/1993903>
- Vovk, V., Gammerman, A. and Shafer, G. (2005). *Algorithmic Learning in a Random World*. New York: Springer. [MR2161220](#)
- Williams, D. (1991). *Probability with Martingales. Cambridge Mathematical Textbooks*. Cambridge: Cambridge Univ. Press. [MR1155402](#) <https://doi.org/10.1017/CBO9780511813658>

# A flexible approach for normal approximation of geometric and topological statistics

ZHAOYANG SHI<sup>a</sup>, KRISHNAKUMAR BALASUBRAMANIAN<sup>b</sup> and WOLFGANG POLONIK<sup>c</sup>

*Department of Statistics, University of California, Davis, USA, <sup>a</sup>zysshi@ucdavis.edu, <sup>b</sup>kbala@ucdavis.edu, <sup>c</sup>wpolonik@ucdavis.edu*

We derive normal approximation results for a class of stabilizing functionals of binomial or Poisson point process, that are not necessarily expressible as sums of certain score functions. Our approach is based on a flexible notion of the add-one cost operator, which helps one to deal with the second-order cost operator via suitably appropriate first-order operators. We combine this flexible notion with the theory of strong stabilization to establish our results. We illustrate the applicability of our results by establishing normal approximation results for certain geometric and topological statistics arising frequently in practice. Several existing results also emerge as special cases of our approach.

*Keywords:* Central limit theorem; normal approximation; Poincaré inequality; Poisson and binomial point processes; Stein’s method; stochastic geometry; topological data analysis

## References

- Baryshnikov, Y. and Yukich, J.E. (2005). Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.* **15** 213–253. [MR2115042 https://doi.org/10.1214/105051604000000594](https://doi.org/10.1214/105051604000000594)
- Berrett, T.B., Samworth, R.J. and Yuan, M. (2019). Efficient multivariate entropy estimation via  $k$ -nearest neighbour distances. *Ann. Statist.* **47** 288–318. [MR3909934 https://doi.org/10.1214/18-AOS1688](https://doi.org/10.1214/18-AOS1688)
- Boissonnat, J.-D., Chazal, F. and Yvinec, M. (2018). *Geometric and Topological Inference. Cambridge Texts in Applied Mathematics*. Cambridge: Cambridge Univ. Press. [MR3837127 https://doi.org/10.1017/9781108297806](https://doi.org/10.1017/9781108297806)
- Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. [MR2435859 https://doi.org/10.1214/07-AOP370](https://doi.org/10.1214/07-AOP370)
- Chatterjee, S. (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Related Fields* **143** 1–40. [MR2449121 https://doi.org/10.1007/s00440-007-0118-6](https://doi.org/10.1007/s00440-007-0118-6)
- Chatterjee, S. and Sen, S. (2017). Minimal spanning trees and Stein’s method. *Ann. Appl. Probab.* **27** 1588–1645. [MR3678480 https://doi.org/10.1214/16-AAP1239](https://doi.org/10.1214/16-AAP1239)
- Edelsbrunner, H. and Harer, J.L. (2010). *Computational Topology: An Introduction*. Providence, RI: Amer. Math. Soc. [MR2572029 https://doi.org/10.1090/mbk/069](https://doi.org/10.1090/mbk/069)
- Englund, G. (1981). A remainder term estimate for the normal approximation in classical occupancy. *Ann. Probab.* **9** 684–692. [MR0624696](https://doi.org/10.1214/009051604000000594)
- Goel, A., Trinh, K.D. and Tsunoda, K. (2019). Strong law of large numbers for Betti numbers in the thermodynamic regime. *J. Stat. Phys.* **174** 865–892. [MR3913900 https://doi.org/10.1007/s10955-018-2201-z](https://doi.org/10.1007/s10955-018-2201-z)
- Kallenberg, O. (1997). *Foundations of Modern Probability. Probability and Its Applications (New York)*. New York: Springer. [MR1464694](https://doi.org/10.1007/978-1-4612-0086-9)
- Kesten, H. and Lee, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **6** 495–527. [MR1398055 https://doi.org/10.1214/aoap/1034968141](https://doi.org/10.1214/aoap/1034968141)
- Kozachenko, L.F. and Leonenko, N.N. (1987). Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii* **23** 9–16. [MR0908626](https://doi.org/10.1007/978-1-4612-0086-9)
- Krebs, J., Roycraft, B. and Polonik, W. (2021). On approximation theorems for the Euler characteristic with applications to the bootstrap. *Electron. J. Stat.* **15** 4462–4509. [MR4312855 https://doi.org/10.1214/21-ejs1898](https://doi.org/10.1214/21-ejs1898)

- Lachièze-Rey, R. and Peccati, G. (2017). New Berry-Esseen bounds for functionals of binomial point processes. *Ann. Appl. Probab.* **27** 1992–2031. MR3693518 <https://doi.org/10.1214/16-AAP1218>
- Lachièze-Rey, R., Peccati, G. and Yang, X. (2022). Quantitative two-scale stabilization on the Poisson space. *Ann. Appl. Probab.* **32** 3085–3145. MR4474528 <https://doi.org/10.1214/21-aap1768>
- Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 <https://doi.org/10.1214/18-AAP1405>
- Last, G., Peccati, G. and Schulte, M. (2016). Normal approximation on Poisson spaces: Mehler’s formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields* **165** 667–723. MR3520016 <https://doi.org/10.1007/s00440-015-0643-7>
- Last, G. and Penrose, M.D. (2011). Poisson process Fock space representation, chaos expansion and covariance inequalities. *Probab. Theory Related Fields* **150** 663–690. MR2824870 <https://doi.org/10.1007/s00440-010-0288-5>
- Nourdin, I., Peccati, G. and Reinert, G. (2009). Second order Poincaré inequalities and CLTs on Wiener space. *J. Funct. Anal.* **257** 593–609. MR2527030 <https://doi.org/10.1016/j.jfa.2008.12.017>
- Owada, T. and Thomas, A.M. (2020). Limit theorems for process-level Betti numbers for sparse and critical regimes. *Adv. in Appl. Probab.* **52** 1–31. MR4092806 <https://doi.org/10.1017/apr.2019.50>
- Peccati, G. and Reitzner, M., eds. (2016). *Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener–Itô Chaos Expansions and Stochastic Geometry*. Bocconi & Springer Series **7**. Cham: Springer. MR3444831 <https://doi.org/10.1007/978-3-319-05233-5>
- Peccati, G., Solé, J.L., Taqqu, M.S. and Utzet, F. (2010). Stein’s method and normal approximation of Poisson functionals. *Ann. Probab.* **38** 443–478. MR2642882 <https://doi.org/10.1214/09-AOP477>
- Peköz, E.A., Röllin, A. and Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann. Appl. Probab.* **23** 1188–1218. MR3076682 <https://doi.org/10.1214/12-AAP868>
- Penrose, M.D. (2005). Multivariate spatial central limit theorems with applications to percolation and spatial graphs. *Ann. Probab.* **33** 1945–1991. MR2165584 <https://doi.org/10.1214/009117905000000206>
- Penrose, M.D. (2007). Gaussian limits for random geometric measures. *Electron. J. Probab.* **12** 989–1035. MR2336596 <https://doi.org/10.1214/EJP.v12-429>
- Penrose, M.D. and Yukich, J.E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11** 1005–1041. MR1878288 <https://doi.org/10.1214/aoap/1015345393>
- Penrose, M.D. and Yukich, J.E. (2005). Normal approximation in geometric probability. In *Stein’s Method and Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 37–58. Singapore: Singapore Univ. Press. MR2201885 [https://doi.org/10.1142/9789812567673\\_0003](https://doi.org/10.1142/9789812567673_0003)
- Penrose, M.D. and Yukich, J.E. (2013). Limit theory for point processes in manifolds. *Ann. Appl. Probab.* **23** 2161–2211. MR3127932 <https://doi.org/10.1214/12-AAP897>
- Schreiber, T. (2010). Limit theorems in stochastic geometry. In *New Perspectives in Stochastic Geometry* 111–144. Oxford: Oxford Univ. Press. MR2654677
- Schulte, M. and Yukich, J.E. (2023). Rates of multivariate normal approximation for statistics in geometric probability. *Ann. Appl. Probab.* **33** 507–548. MR4551557 <https://doi.org/10.1214/22-aap1822>
- Shi, Z., Balasubramanian, K. and Polonik, W. (2024). Supplement to “A flexible approach for normal approximation of geometric and topological statistics.” <https://doi.org/10.3150/23-BEJ1705SUPP>
- Thomas, A.M. and Owada, T. (2021). Functional limit theorems for the Euler characteristic process in the critical regime. *Adv. in Appl. Probab.* **53** 57–80. MR4232749 <https://doi.org/10.1017/apr.2020.46>
- Trinh, K.D. (2017). A remark on the convergence of Betti numbers in the thermodynamic regime. *Pac. J. Math. Ind.* **9** 1–7. MR3620264 <https://doi.org/10.1186/s40736-017-0029-0>
- 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>
- Yukich, J.E. (2015). Surface order scaling in stochastic geometry. *Ann. Appl. Probab.* **25** 177–210. MR3297770 <https://doi.org/10.1214/13-AAP992>

# Gaussian differentially private robust mean estimation and inference

MYEONGHUN YU<sup>1,a</sup>, ZHAO REN<sup>2,b</sup> and WEN-XIN ZHOU<sup>3,c</sup>

<sup>1</sup>*Department of Mathematics, University of California, San Diego, La Jolla, CA, 92093, USA, [myyu@ucsd.edu](mailto:myyu@ucsd.edu)*

<sup>2</sup>*Department of Statistics, University of Pittsburgh, Pittsburgh, PA, 15260, USA, [zren@pitt.edu](mailto:zren@pitt.edu)*

<sup>3</sup>*Department of Information and Decision Sciences, University of Illinois at Chicago, Chicago, IL, 60607, USA, [wenxinz@uic.edu](mailto:wenxinz@uic.edu)*

In this paper, we propose differentially private algorithms for robust (multivariate) mean estimation and inference under heavy-tailed distributions, with a focus on Gaussian differential privacy. First, we provide a comprehensive analysis of the Huber mean estimator with increasing dimensions, including non-asymptotic deviation bound, Bahadur representation, and (uniform) Gaussian approximations. Secondly, we privatize the Huber mean estimator via noisy gradient descent, which is proven to achieve near-optimal statistical guarantees. The key is to characterize quantitatively the trade-off between statistical accuracy, degree of robustness and privacy level, governed by a carefully chosen robustification parameter. Finally, we construct private confidence intervals for the proposed estimator by incorporating a private and robust covariance estimator. Our findings are demonstrated by simulation studies.

*Keywords:* Confidence interval; differential privacy; heavy-tailed distribution; Huber loss; mean estimation

## References

- [1] Avella-Medina, M. (2021). Privacy-preserving parametric inference: A case for robust statistics. *J. Amer. Statist. Assoc.* **116** 969–983. [MR4270037 https://doi.org/10.1080/01621459.2019.1700130](https://doi.org/10.1080/01621459.2019.1700130)
- [2] Avella-Medina, M., Bradshaw, C. and Loh, P.-L. (2023). Differentially private inference via noisy optimization. *Ann. Statist.* **51** 2067–2092. [MR4678796 https://doi.org/10.1214/23-aos2321](https://doi.org/10.1214/23-aos2321)
- [3] Barber, R.F. and Duchi, J. (2014). Privacy: A few definitional aspects and consequences for minimax mean-squared error. In *53rd IEEE Conference on Decision and Control* 1365–1369.
- [4] Barzilai, J. and Borwein, J.M. (1988). Two-point step size gradient methods. *IMA J. Numer. Anal.* **8** 141–148. [MR0967848 https://doi.org/10.1093/imanum/8.1.141](https://doi.org/10.1093/imanum/8.1.141)
- [5] Bassily, R., Smith, A. and Thakurta, A. (2014). Private empirical risk minimization: Efficient algorithms and tight error bounds. In *55th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2014* 464–473. Los Alamitos, CA: IEEE Computer Soc. [MR3344896 https://doi.org/10.1109/FOCS.2014.56](https://doi.org/10.1109/FOCS.2014.56)
- [6] Bubeck, S., Cesa-Bianchi, N. and Lugosi, G. (2013). Bandits with heavy tail. *IEEE Trans. Inf. Theory* **59** 7711–7717. [MR3124669 https://doi.org/10.1109/TIT.2013.2277869](https://doi.org/10.1109/TIT.2013.2277869)
- [7] Bun, M. and Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography. Part I. Lecture Notes in Computer Science* **9985** 635–658. Berlin: Springer. [MR3591832 https://doi.org/10.1007/978-3-662-53641-4\\_24](https://doi.org/10.1007/978-3-662-53641-4_24)
- [8] Bun, M. and Steinke, T. (2019). Average-case averages: Private algorithms for smooth sensitivity and mean estimation. In *Advances in Neural Information Processing Systems* 181–191.
- [9] Cai, T.T., Wang, Y. and Zhang, L. (2020). The cost of privacy in generalized linear models: Algorithms and minimax lower bounds. ArXiv preprint. Available at [arXiv:2011.03900](https://arxiv.org/abs/2011.03900).
- [10] Cai, T.T., Wang, Y. and Zhang, L. (2021). The cost of privacy: Optimal rates of convergence for parameter estimation with differential privacy. *Ann. Statist.* **49** 2825–2850. [MR4338894 https://doi.org/10.1214/21-aos2058](https://doi.org/10.1214/21-aos2058)
- [11] Catoni, O. (2012). Challenging the empirical mean and empirical variance: A deviation study. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 1148–1185. [MR3052407 https://doi.org/10.1214/11-AIHP454](https://doi.org/10.1214/11-AIHP454)

- [12] Chen, M., Gao, C. and Ren, Z. (2018). Robust covariance and scatter matrix estimation under Huber's contamination model. *Ann. Statist.* **46** 1932–1960. MR3845006 <https://doi.org/10.1214/17-AOS1607>
- [13] Chen, X. and Zhou, W.-X. (2020). Robust inference via multiplier bootstrap. *Ann. Statist.* **48** 1665–1691. MR4124339 <https://doi.org/10.1214/19-AOS1863>
- [14] 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. MR3517099
- [15] Depersin, J. and Lecué, G. (2022). Robust sub-Gaussian estimation of a mean vector in nearly linear time. *Ann. Statist.* **50** 511–536. MR4382026 <https://doi.org/10.1214/21-aos2118>
- [16] Depersin, J. and Lecué, G. (2022). Optimal robust mean and location estimation via convex programs with respect to any pseudo-norms. *Probab. Theory Related Fields* **183** 997–1025. MR4453320 <https://doi.org/10.1007/s00440-022-01127-y>
- [17] Devroye, L., Lerasle, M., Lugosi, G. and Oliveira, R.I. (2016). Sub-Gaussian mean estimators. *Ann. Statist.* **44** 2695–2725. MR3576558 <https://doi.org/10.1214/16-AOS1440>
- [18] Diakonikolas, I. and Kane, D.M. (2019). Recent advances in algorithmic high-dimensional robust statistics. ArXiv preprint. Available at [arXiv:1911.05911](https://arxiv.org/abs/1911.05911).
- [19] Dong, J., Roth, A. and Su, W.J. (2022). Gaussian differential privacy. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 3–54. MR4400389
- [20] Duchi, J.C., Jordan, M.I. and Wainwright, M.J. (2018). Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* **113** 182–201. MR3803452 <https://doi.org/10.1080/01621459.2017.1389735>
- [21] Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I. and Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology—EUROCRYPT 2006. Lecture Notes in Computer Science* **4004** 486–503. Berlin: Springer. MR2423560 [https://doi.org/10.1007/11761679\\_29](https://doi.org/10.1007/11761679_29)
- [22] Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography. Lecture Notes in Computer Science* **3876** 265–284. Berlin: Springer. MR2241676 [https://doi.org/10.1007/11681878\\_14](https://doi.org/10.1007/11681878_14)
- [23] Dwork, C. and Rothblum, G.N. (2016). Concentrated differential privacy. ArXiv preprint. Available at [arXiv:1603.01887](https://arxiv.org/abs/1603.01887).
- [24] 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. MR4047307 <https://doi.org/10.1080/01621459.2018.1527700>
- [25] Fan, J., Li, Q. and Wang, Y. (2017). Estimation of high dimensional mean regression in the absence of symmetry and light tail assumptions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 247–265. MR3597972 <https://doi.org/10.1111/rssb.12166>
- [26] Hampel, F., Hennig, C. and Ronchetti, E. (2011). A smoothing principle for the Huber and other location  $M$ -estimators. *Comput. Statist. Data Anal.* **55** 324–337. MR2736558 <https://doi.org/10.1016/j.csda.2010.05.001>
- [27] Heyde, C.C. (1967). On the influence of moments on the rate of convergence to the normal distribution. *Z. Wahrsch. Verw. Gebiete* **8** 12–18. MR0215344 <https://doi.org/10.1007/BF00533941>
- [28] Hopkins, S.B. (2020). Mean estimation with sub-Gaussian rates in polynomial time. *Ann. Statist.* **48** 1193–1213. MR4102693 <https://doi.org/10.1214/19-AOS1843>
- [29] Hopkins, S.B., Kamath, G. and Majid, M. (2022). Efficient mean estimation with pure differential privacy via a sum-of-squares exponential mechanism. In *STOC '22—Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing* 1406–1417. New York: ACM. MR4490088
- [30] Hopkins, S.B., Li, J. and Zhang, F. (2020). Robust and heavy-tailed mean estimation made simple, via regret minimization. In *Advances in Neural Information Processing Systems* **33** 11902–11912.
- [31] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Stat.* **35** 73–101. MR0161415 <https://doi.org/10.1214/aoms/1177703732>
- [32] Kamath, G., Li, J., Singhal, V. and Ullman, J. (2019). Privately learning high-dimensional distributions. In *Conference on Learning Theory* **99** 1853–1902.
- [33] Kamath, G., Mouzakis, A. and Singhal, V. (2022). New lower bounds for private estimation and a generalized fingerprinting lemma. In *Advances in Neural Information Processing Systems* **35** 24405–24418.
- [34] Kamath, G., Singhal, V. and Ullman, J. (2020). Private mean estimation of heavy-tailed distributions. In *Conference on Learning Theory* **125** 2204–2235.

- [35] Karwa, V. and Vadhan, S. (2018). Finite sample differentially private confidence intervals. In *9th Innovations in Theoretical Computer Science. LIPIcs. Leibniz Int. Proc. Inform.* **94** Art. No. 44. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. [MR3761780](#)
- [36] Ke, Y., Minsker, S., Ren, Z., Sun, Q. and Zhou, W.-X. (2019). User-friendly covariance estimation for heavy-tailed distributions. *Statist. Sci.* **34** 454–471. [MR4017523](#) <https://doi.org/10.1214/19-STS711>
- [37] Liu, X., Kong, W., Kakade, S. and Oh, S. (2021). Robust and differentially private mean estimation. In *Advances in Neural Information Processing Systems* **34** 3887–3901.
- [38] Lugosi, G. and Mendelson, S. (2019). Sub-Gaussian estimators of the mean of a random vector. *Ann. Statist.* **47** 783–794. [MR3909950](#) <https://doi.org/10.1214/17-AOS1639>
- [39] 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>
- [40] Lugosi, G. and Mendelson, S. (2021). Robust multivariate mean estimation: The optimality of trimmed mean. *Ann. Statist.* **49** 393–410. [MR4206683](#) <https://doi.org/10.1214/20-AOS1961>
- [41] Mathieu, T. (2022). Concentration study of M-estimators using the influence function. *Electron. J. Stat.* **16** 3695–3750. [MR4444667](#) <https://doi.org/10.1214/22-ejs2030>
- [42] McSherry, F. and Talwar, K. (2007). Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science* 94–103.
- [43] Mendelson, S. and Zhivotovskiy, N. (2020). Robust covariance estimation under  $L_4$ - $L_2$  norm equivalence. *Ann. Statist.* **48** 1648–1664. [MR4124338](#) <https://doi.org/10.1214/19-AOS1862>
- [44] Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** 2308–2335. [MR3378468](#) <https://doi.org/10.3150/14-BEJ645>
- [45] Minsker, S. (2018). Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries. *Ann. Statist.* **46** 2871–2903. [MR3851758](#) <https://doi.org/10.1214/17-AOS1642>
- [46] Minsker, S. and Wei, X. (2020). Robust modifications of U-statistics and applications to covariance estimation problems. *Bernoulli* **26** 694–727. [MR4036049](#) <https://doi.org/10.3150/19-BEJ1149>
- [47] Mironov, I. (2017). Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)* 263–275.
- [48] Murtagh, J. and Vadhan, S. (2016). The complexity of computing the optimal composition of differential privacy. In *Theory of Cryptography. Part I. Lecture Notes in Computer Science* **9562** 157–175. Berlin: Springer. [MR3487659](#) [https://doi.org/10.1007/978-3-662-49096-9\\_7](https://doi.org/10.1007/978-3-662-49096-9_7)
- [49] Rohde, A. and Steinberger, L. (2020). Geometrizing rates of convergence under local differential privacy constraints. *Ann. Statist.* **48** 2646–2670. [MR4152116](#) <https://doi.org/10.1214/19-AOS1901>
- [50] Song, S., Chaudhuri, K. and Sarwate, A.D. (2013). Stochastic gradient descent with differentially private updates. In *2013 IEEE Global Conference on Signal and Information Processing* 245–248.
- [51] Spokoiny, V. and Zhilova, M. (2015). Bootstrap confidence sets under model misspecification. *Ann. Statist.* **43** 2653–2675. [MR3405607](#) <https://doi.org/10.1214/15-AOS1355>
- [52] Sun, Q., Zhou, W.-X. and Fan, J. (2020). Adaptive Huber regression. *J. Amer. Statist. Assoc.* **115** 254–265. [MR4078461](#) <https://doi.org/10.1080/01621459.2018.1543124>
- [53] Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science. Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. [MR3837109](#) <https://doi.org/10.1017/9781108231596>
- [54] Wang, Y., Kifer, D. and Lee, J. (2019). Differentially private confidence intervals for empirical risk minimization. *J. Priv. Confid.* **9** 1–36.
- [55] Wasserman, L. and Zhou, S. (2010). A statistical framework for differential privacy. *J. Amer. Statist. Assoc.* **105** 375–389. [MR2656057](#) <https://doi.org/10.1198/jasa.2009.tm08651>
- [56] Wei, X. and Minsker, S. (2017). Estimation of the covariance structure of heavy-tailed distributions. In *Advances in Neural Information Processing Systems* **30** 2855–2864.
- [57] Yu, M., Ren, Z. and Zhou, W.-X. (2024). Supplement to “Gaussian differentially private robust mean estimation and inference.” <https://doi.org/10.3150/23-BEJ1706SUPP>
- [58] Zhou, W.-X., Bose, K., Fan, J. and Liu, H. (2018). A new perspective on robust M-estimation: Finite sample theory and applications to dependence-adjusted multiple testing. *Ann. Statist.* **46** 1904–1931. [MR3845005](#) <https://doi.org/10.1214/17-AOS1606>



# Drift reduction method for SDEs driven by heterogeneous singular Lévy noise

TADEUSZ KULCZYCKI<sup>a</sup> , OLEKSII KULYK<sup>b</sup>  and MICHAŁ RYZNAR<sup>c</sup> 

Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland,  
<sup>a</sup>[tadeusz.kulczycki@pwr.edu.pl](mailto:tadeusz.kulczycki@pwr.edu.pl), <sup>b</sup>[oleksii.kulyk@pwr.edu.pl](mailto:oleksii.kulyk@pwr.edu.pl), <sup>c</sup>[michal.ryznar@pwr.edu.pl](mailto:michal.ryznar@pwr.edu.pl)

We study SDE

$$dX_t = b(X_t) dt + A(X_{t-}) dZ_t, \quad X_0 = x \in \mathbb{R}^d, \quad t \geq 0,$$

where  $Z = (Z^1, \dots, Z^d)^T$ , with  $Z^i, i = 1, \dots, d$  being independent one-dimensional symmetric jump Lévy processes, not necessarily identically distributed. In particular, we cover the case when each  $Z^i$  is one-dimensional symmetric  $\alpha_i$ -stable process, where  $\alpha_i \in (0, 2)$  are not necessarily equal but satisfy certain balance condition which prevents hypoelliptic effects. Under certain assumptions on  $b$ ,  $A$  and  $Z$  we show that the weak solution to the SDE is uniquely defined and is a Markov process. We also provide a representation of the transition probability density and establish Hölder regularity of the corresponding transition semigroup. The method we propose is based on a reduction of an SDE with a drift term to another SDE without such a term but with coefficients depending on time variable. Such a method has the same spirit as the classic characteristic method and seems to be of independent interest.

**Keywords:** Drift; Hölder regularity; Lévy process; stochastic differential equation; transition density

## References

- Bass, R.F. and Chen, Z.-Q. (2006). Systems of equations driven by stable processes. *Probab. Theory Related Fields* **134** 175–214. [MR2222382 https://doi.org/10.1007/s00440-004-0426-z](https://doi.org/10.1007/s00440-004-0426-z)
- Bogdan, K., Grzywny, T. and Ryznar, M. (2014). Density and tails of unimodal convolution semigroups. *J. Funct. Anal.* **266** 3543–3571. [MR3165234 https://doi.org/10.1016/j.jfa.2014.01.007](https://doi.org/10.1016/j.jfa.2014.01.007)
- Chaker, J. (2019). The martingale problem for a class of nonlocal operators of diagonal type. *Math. Nachr.* **292** 2316–2337. [MR4033009 https://doi.org/10.1002/mana.201800452](https://doi.org/10.1002/mana.201800452)
- Chaker, J. (2020). Regularity of solutions to anisotropic nonlocal equations. *Math. Z.* **296** 1135–1155. [MR4159826 https://doi.org/10.1007/s00209-020-02459-y](https://doi.org/10.1007/s00209-020-02459-y)
- Chaker, J. and Kassmann, M. (2020). Nonlocal operators with singular anisotropic kernels. *Comm. Partial Differential Equations* **45** 1–31. [MR4037095 https://doi.org/10.1080/03605302.2019.1651335](https://doi.org/10.1080/03605302.2019.1651335)
- Chen, Z.-Q., Hao, Z. and Zhang, X. (2020). Hölder regularity and gradient estimates for SDEs driven by cylindrical  $\alpha$ -stable processes. *Electron. J. Probab.* **25** Paper No. 137. [MR4179301 https://doi.org/10.1214/20-ejp542](https://doi.org/10.1214/20-ejp542)
- Chen, Z.-Q., Zhang, X. and Zhao, G. (2021). Supercritical SDEs driven by multiplicative stable-like Lévy processes. *Trans. Amer. Math. Soc.* **374** 7621–7655. [MR4328678 https://doi.org/10.1090/tran/8343](https://doi.org/10.1090/tran/8343)
- Coddington, E.A. and Levinson, N. (1955). *Theory of Ordinary Differential Equations*. New York-Toronto-London: McGraw-Hill, Inc. [MR0069338](https://doi.org/10.1007/978-1-4612-0630-0)
- Debussche, A. and Fournier, N. (2013). Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. *J. Funct. Anal.* **264** 1757–1778. [MR3022725 https://doi.org/10.1016/j.jfa.2013.01.009](https://doi.org/10.1016/j.jfa.2013.01.009)
- Friesen, M., Jin, P. and Rüdiger, B. (2021). Existence of densities for stochastic differential equations driven by Lévy processes with anisotropic jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 250–271. [MR4255174 https://doi.org/10.1214/20-aihp1077](https://doi.org/10.1214/20-aihp1077)
- Huang, L. and Menozzi, S. (2016). A parametrix approach for some degenerate stable driven SDEs. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1925–1975. [MR3573301 https://doi.org/10.1214/15-AIHP704](https://doi.org/10.1214/15-AIHP704)

- Knopova, V. and Kulik, A. (2018). Parametrix construction of the transition probability density of the solution to an SDE driven by  $\alpha$ -stable noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 100–140. [MR3765882](#) <https://doi.org/10.1214/16-AIHP796>
- Knopova, V., Kulik, A. and Schilling, R.L. (2021). Construction and heat kernel estimates of general stable-like Markov processes. *Dissertationes Math.* **569** 86. [MR4361582](#) <https://doi.org/10.4064/dm824-8-2021>
- Konakov, V.D. and Markova, A.R. (2017). Nonlinear trend exclusion procedure for models described by stochastic differential and difference equations. *Autom. Remote Control* **78** 1438–1448. [MR3710271](#) <https://doi.org/10.1134/s0005117917080057>
- Konakov, V., Menozzi, S. and Molchanov, S. (2010). Explicit parametrix and local limit theorems for some degenerate diffusion processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 908–923. [MR2744877](#) <https://doi.org/10.1214/09-AIHP207>
- Kulczycki, T., Kulik, A. and Ryznar, M. (2022). On weak solution of SDE driven by inhomogeneous singular Lévy noise. *Trans. Amer. Math. Soc.* **375** 4567–4618. [MR4439486](#) <https://doi.org/10.1090/tran/8612>
- Kulczycki, T. and Ryznar, M. (2020). Semigroup properties of solutions of SDEs driven by Lévy processes with independent coordinates. *Stochastic Process. Appl.* **130** 7185–7217. [MR4167204](#) <https://doi.org/10.1016/j.spa.2020.07.011>
- Kulczycki, T., Ryznar, M. and Sztonyk, P. (2021). Strong Feller property for SDEs driven by multiplicative cylindrical stable noise. *Potential Anal.* **55** 75–126. [MR4261305](#) <https://doi.org/10.1007/s11118-020-09850-8>
- Kulik, A.M. (2019). On weak uniqueness and distributional properties of a solution to an SDE with  $\alpha$ -stable noise. *Stochastic Process. Appl.* **129** 473–506. [MR3907007](#) <https://doi.org/10.1016/j.spa.2018.03.010>
- Menozzi, S. (2011). Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16** 234–250. [MR2802040](#) <https://doi.org/10.1214/ECP.v16-1619>
- Priola, E. and Zabczyk, J. (2011). Structural properties of semilinear SPDEs driven by cylindrical stable processes. *Probab. Theory Related Fields* **149** 97–137. [MR2773026](#) <https://doi.org/10.1007/s00440-009-0243-5>

# Moment asymptotics for super-Brownian motions

YAOZHONG HU<sup>1,a</sup>, XIONG WANG<sup>2,b</sup>, PANQIU XIA<sup>3,c</sup> and JIAYU ZHENG<sup>4,d</sup>

<sup>1</sup>*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB, T6G 2G1, Canada, [yaozhong@ualberta.ca](mailto:yaozhong@ualberta.ca)*

<sup>2</sup>*Department of Mathematics, Johns Hopkins University, Baltimore, MD 21218, USA, [xiong\\_wang@jhu.edu](mailto:xiong_wang@jhu.edu)*

<sup>3</sup>*Department of Mathematics and Statistics, Auburn University, Auburn, AL 36849, USA, [pqxia@auburn.edu](mailto:pqxia@auburn.edu)*

<sup>4</sup>*Faculty of Computational Mathematics and Cybernetics, Shenzhen MSU-BIT University, Shenzhen, Guangdong, 518172, China, [jyzheng@smbu.edu.cn](mailto:jyzheng@smbu.edu.cn)*

In this paper, long-time and high-order moment asymptotics for super-Brownian motions (sBm's) are studied. By using a moment formula for sBm's (e.g. (*Ann. Appl. Probab.* **33** (2023) 3872–3915, Theorem 3.1)), precise upper and lower bounds for all positive integer moments at any time  $t > 0$  of sBm's for certain initial conditions are achieved. Then, the moment asymptotics as time goes to infinity or as the moment order goes to infinity follow immediately. Additionally, as an application of the two-sided moment bounds, the tail probability estimates of sBm's are obtained.

**Keywords:** Intermittency; moment asymptotics; moment formula; super-Brownian motion; tail probability; two-sided moment bounds

## References

- Bertini, L. and Cancrini, N. (1995). The stochastic heat equation: Feynman-Kac formula and intermittency. *J. Stat. Phys.* **78** 1377–1401. [MR1316109 https://doi.org/10.1007/BF02180136](https://doi.org/10.1007/BF02180136)
- Bertini, L. and Giacomin, G. (1997). Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** 571–607. [MR1462228 https://doi.org/10.1007/s002200050044](https://doi.org/10.1007/s002200050044)
- Carmona, R.A. and Molchanov, S.A. (1994). Parabolic Anderson problem and intermittency. *Mem. Amer. Math. Soc.* **108** viii+125. [MR1185878 https://doi.org/10.1090/memo/0518](https://doi.org/10.1090/memo/0518)
- Chen, X. (2015). Precise intermittency for the parabolic Anderson equation with an  $(1 + 1)$ -dimensional time-space white noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1486–1499. [MR3414455 https://doi.org/10.1214/15-AIHP673](https://doi.org/10.1214/15-AIHP673)
- Chen, L. and Dalang, R.C. (2015). Moments and growth indices for the nonlinear stochastic heat equation with rough initial conditions. *Ann. Probab.* **43** 3006–3051. [MR3433576 https://doi.org/10.1214/14-AOP954](https://doi.org/10.1214/14-AOP954)
- Chen, L., Guo, Y. and Song, J. (2024+). Moments and asymptotics for a class of SPDEs with space-time white noise. *Trans. Amer. Math. Soc.* To appear in.
- Chen, L. and Xia, P. (2023). Asymptotic properties of stochastic partial differential equations in the sublinear regime. Preprint. Available at [arXiv:2306.06761](https://arxiv.org/abs/2306.06761).
- Conus, D., Joseph, M. and Khoshnevisan, D. (2013). On the chaotic character of the stochastic heat equation, before the onset of intermittency. *Ann. Probab.* **41** 2225–2260. [MR3098071 https://doi.org/10.1214/11-AOP717](https://doi.org/10.1214/11-AOP717)
- Conus, D. and Khoshnevisan, D. (2012). On the existence and position of the farthest peaks of a family of stochastic heat and wave equations. *Probab. Theory Related Fields* **152** 681–701. [MR2892959 https://doi.org/10.1007/s00440-010-0333-4](https://doi.org/10.1007/s00440-010-0333-4)
- Das, S. and Tsai, L.-C. (2021). Fractional moments of the stochastic heat equation. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 778–799. [MR4260483 https://doi.org/10.1214/20-aihp1095](https://doi.org/10.1214/20-aihp1095)
- Dawson, D.A. (1993). Measure-valued Markov processes. In *École D'Été de Probabilités de Saint-Flour XXI—1991. Lecture Notes in Math.* **1541** 1–260. Berlin: Springer. [MR1242575 https://doi.org/10.1007/BFb0084190](https://doi.org/10.1007/BFb0084190)

- Dawson, D.A. and Kurtz, T.G. (1982). Applications of duality to measure-valued diffusion processes. In *Advances in Filtering and Optimal Stochastic Control (Cocoyoc, 1982)*. *Lect. Notes Control Inf. Sci.* **42** 91–105. Berlin: Springer. [MR0794506](#) <https://doi.org/10.1007/BFb0004528>
- Dynkin, E.B. (2002). *Diffusions, Superdiffusions and Partial Differential Equations*. *American Mathematical Society Colloquium Publications* **50**. Providence, RI: Amer. Math. Soc. [MR1883198](#) <https://doi.org/10.1090/coll/050>
- Dynkin, E.B. (2004). *Superdiffusions and Positive Solutions of Nonlinear Partial Differential Equations*. *University Lecture Series* **34**. Providence, RI: Amer. Math. Soc. Appendix A by J.-F. Le Gall and Appendix B by I. E. Verbitsky. [MR2089791](#) <https://doi.org/10.1090/ulect/034>
- Etheridge, A.M. (2000). *An Introduction to Superprocesses*. *University Lecture Series* **20**. Providence, RI: Amer. Math. Soc. [MR1779100](#) <https://doi.org/10.1090/ulect/020>
- Foondun, M. and Khoshnevisan, D. (2009). Intermittence and nonlinear parabolic stochastic partial differential equations. *Electron. J. Probab.* **14** 548–568. [MR2480553](#) <https://doi.org/10.1214/EJP.v14-614>
- Hairer, M. (2013). Solving the KPZ equation. *Ann. of Math. (2)* **178** 559–664. [MR3071506](#) <https://doi.org/10.4007/annals.2013.178.2.4>
- Hu, Y., Nualart, D. and Xia, P. (2019). Hölder continuity of the solutions to a class of SPDE's arising from branching particle systems in a random environment. *Electron. J. Probab.* **24** Paper No. 105, 52. [MR4017123](#) <https://doi.org/10.1214/19-ejp357>
- Hu, Y. and Wang, X. (2024). Matching upper and lower moment bounds for a large class of stochastic PDEs driven by general space-time Gaussian noises. *Stoch. Partial Differ. Equ. Anal. Comput.* **12** 1–52. [MR4709538](#) <https://doi.org/10.1007/s40072-022-00278-2>
- Hu, Y., Kouritzin, M.A., Xia, P. and Zheng, J. (2023). On mean-field super-Brownian motions. *Ann. Appl. Probab.* **33** 3872–3915. [MR4663499](#) <https://doi.org/10.1214/22-aap1909>
- Iscoc, I. and Lee, T.-Y. (1993). Large deviations for occupation times of measure-valued branching Brownian motions. *Stoch. Stoch. Rep.* **45** 177–209. [MR1306931](#) <https://doi.org/10.1080/17442509308833861>
- Kardar, M., Parisi, G. and Zhang, Y.-C. (1986). Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** 889–892.
- Khoshnevisan, D. (2014). *Analysis of Stochastic Partial Differential Equations*. *CBMS Regional Conference Series in Mathematics* **119**. Washington, DC: Amer. Math. Soc.. [MR3222416](#) <https://doi.org/10.1090/cbms/119>
- Khoshnevisan, D., Kim, K. and Xiao, Y. (2017). Intermittency and multifractality: A case study via parabolic stochastic PDEs. *Ann. Probab.* **45** 3697–3751. [MR3729613](#) <https://doi.org/10.1214/16-AOP1147>
- Khoshnevisan, D., Kim, K. and Xiao, Y. (2018). A macroscopic multifractal analysis of parabolic stochastic PDEs. *Comm. Math. Phys.* **360** 307–346. [MR3795193](#) <https://doi.org/10.1007/s00220-018-3136-6>
- Konno, N. and Shiga, T. (1988). Stochastic partial differential equations for some measure-valued diffusions. *Probab. Theory Related Fields* **79** 201–225. [MR0958288](#) <https://doi.org/10.1007/BF00320919>
- Lee, T.-Y. and Remillard, B. (1995). Large deviations for the three-dimensional super-Brownian motion. *Ann. Probab.* **23** 1755–1771. [MR1379167](#)
- Li, Z. and Pu, F. (2023). Gaussian fluctuation for spatial average of super-Brownian motion. *Stoch. Anal. Appl.* **41** 752–769. [MR4601713](#) <https://doi.org/10.1080/07362994.2022.2079530>
- Mueller, C. (2009). Some tools and results for parabolic stochastic partial differential equations. In *A Minicourse on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1962** 111–144. Berlin: Springer. [MR2508775](#) [https://doi.org/10.1007/978-3-540-85994-9\\_4](https://doi.org/10.1007/978-3-540-85994-9_4)
- Perkins, E. (2002). Dawson–Watanabe superprocesses and measure-valued diffusions. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1999)*. *Lecture Notes in Math.* **1781** 125–324. Berlin: Springer. [MR1915445](#)
- Wang, H. (1998). A class of measure-valued branching diffusions in a random medium. *Stoch. Anal. Appl.* **16** 753–786. [MR1632574](#) <https://doi.org/10.1080/07362999808809560>

# On the asymptotic behavior of a finite section of the optimal causal filter

JUNHO YANG<sup>a</sup>

*Institute of Statistical Science, No. 128, Academia Rd. Sect. 2, Nangang, Taipei, Taiwan,*

<sup>a</sup>[junhoyang@stat.sinica.edu.tw](mailto:junhoyang@stat.sinica.edu.tw)

We establish an  $L_1$ -bound between the coefficients of the optimal causal filter applied to the data-generating process and its finite sample approximation. Here, we assume that the data-generating process is a second-order stationary time series with either short or long memory autocovariances. To derive the  $L_1$ -bound, we first provide an exact expression for the coefficients of the causal filter and their approximations in terms of the absolute convergent series of the multistep ahead infinite and finite predictor coefficients, respectively. Then, we prove a so-called uniform Baxter's inequality to obtain a bound for the difference between the infinite and finite multistep ahead predictor coefficients in both short and long memory time series. The  $L_1$ -approximation error bound for the causal filter coefficients can be used to evaluate the performance of the linear predictions of time series through the mean squared error criterion.

*Keywords:* Mean squared prediction error; predictor coefficients; short and long memory time series; uniform Baxter's inequality

## References

- Anderson, B.D.O. and Moore, J.B. (2012). *Optimal Filtering*. Chelmsford, MA: Courier Corporation.
- Baxter, G. (1962). An asymptotic result for the finite predictor. *Math. Scand.* **10** 137–144. [MR0149584 https://doi.org/10.7146/math.scand.a-10520](https://doi.org/10.7146/math.scand.a-10520)
- Baxter, G. (1963). A norm inequality for a “finite-section” Wiener-Hopf equation. *Illinois J. Math.* **7** 97–103. [MR0145285 https://doi.org/10.1214/11-PS178](https://doi.org/10.1214/11-PS178)
- Bhansali, R.J. (1996). Asymptotically efficient autoregressive model selection for multistep prediction. *Ann. Inst. Statist. Math.* **48** 577–602. [MR1424784 https://doi.org/10.1007/BF00050857](https://doi.org/10.1007/BF00050857)
- Bingham, N.H. (2012). Szegő's theorem and its probabilistic descendants. *Probab. Surv.* **9** 287–324. [MR2956573 https://doi.org/10.1214/11-PS178](https://doi.org/10.1214/11-PS178)
- Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. [MR0898871 https://doi.org/10.1017/CBO9780511721434](https://doi.org/10.1017/CBO9780511721434)
- Brockwell, P.J. and Davis, R.A. (2006). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR2839251 https://doi.org/10.1007/978-1-4939-9829-9](https://doi.org/10.1007/978-1-4939-9829-9)
- Cheng, R. and Pourahmadi, M. (1993). Baxter's inequality and convergence of finite predictors of multivariate stochastic processes. *Probab. Theory Related Fields* **95** 115–124. [MR1207310 https://doi.org/10.1007/BF01197341](https://doi.org/10.1007/BF01197341)
- Devinatz, A. (1964). Asymptotic estimates for the finite predictor. *Math. Scand.* **15** 111–120. [MR0188708 https://doi.org/10.7146/math.scand.a-10734](https://doi.org/10.7146/math.scand.a-10734)
- Doob, J.L. (1953). *Stochastic Processes*. New York: Wiley. [MR0058896 https://doi.org/10.1002/9781112512463](https://doi.org/10.1002/9781112512463)
- Durbin, J. (1960). The fitting of time-series models. *Rev. Inst. Int. Stat.* **28** 233–244.
- Findley, D.F. (1991). Convergence of finite multistep predictors from incorrect models and its role in model selection. *Note Mat.* **11** 145–155.
- Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Series Anal.* **1** 15–29. [MR0605572 https://doi.org/10.1111/j.1467-9892.1980.tb00297.x](https://doi.org/10.1111/j.1467-9892.1980.tb00297.x)

- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika* **68** 165–176. [MR0614953](#) <https://doi.org/10.1093/biomet/68.1.165>
- Ibragimov, I.A. (1964). On the asymptotic behavior of the prediction error. *Theory Probab. Appl.* **9** 627–634.
- Inoue, A. (2000). Asymptotics for the partial autocorrelation function of a stationary process. *J. Anal. Math.* **81** 65–109. [MR1785278](#) <https://doi.org/10.1007/BF02788986>
- Inoue, A. (2002). Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes. *Ann. Appl. Probab.* **12** 1471–1491. [MR1936600](#) <https://doi.org/10.1214/aoap/1037125870>
- Inoue, A. and Kasahara, Y. (2006). Explicit representation of finite predictor coefficients and its applications. *Ann. Statist.* **34** 973–993. [MR2283400](#) <https://doi.org/10.1214/009053606000000209>
- Inoue, A., Kasahara, Y. and Pourahmadi, M. (2018). Baxter’s inequality for finite predictor coefficients of multivariate long-memory stationary processes. *Bernoulli* **24** 1202–1232. [MR3706792](#) <https://doi.org/10.3150/16-BEJ897>
- Kokoszka, P.S. and Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stochastic Process. Appl.* **60** 19–47. [MR1362317](#) [https://doi.org/10.1016/0304-4149\(95\)00034-8](https://doi.org/10.1016/0304-4149(95)00034-8)
- 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>
- Levinson, N. (1947). The Wiener RMS (root mean square) error criterion in filter design and prediction. *J. Math. Phys. Mass. Inst. Tech.* **25** 261–278. [MR0019257](#) <https://doi.org/10.1002/sapm1946251261>
- McElroy, T.S. and Wildi, M. (2020). The multivariate linear prediction problem: Model-based and direct filtering solutions. *Econom. Stat.* **14** 112–130. [MR4080577](#) <https://doi.org/10.1016/j.ecosta.2019.12.004>
- Oppenheim, A.V., Schaffer, R.W. and Buck, J.R. (2001). *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, NJ: Prentice Hall.
- Poskitt, D.S. (2007). Autoregressive approximation in nonstandard situations: The fractionally integrated and non-invertible cases. *Ann. Inst. Statist. Math.* **59** 697–725. [MR2415731](#) <https://doi.org/10.1007/s10463-006-0074-4>
- Pourahmadi, M. (1989). On the convergence of finite linear predictors of stationary processes. *J. Multivariate Anal.* **30** 167–180. [MR1015366](#) [https://doi.org/10.1016/0047-259X\(89\)90033-X](https://doi.org/10.1016/0047-259X(89)90033-X)
- Pourahmadi, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*. *Wiley Series in Probability and Statistics: Applied Probability and Statistics*. New York: Wiley Interscience. [MR1849562](#)
- Rozañov, Y.A. (1967). *Stationary Random Processes*. San Francisco, CA: Holden-Day. [MR0214134](#)
- Subba Rao, S. and Yang, J. (2021). Reconciling the Gaussian and Whittle likelihood with an application to estimation in the frequency domain. *Ann. Statist.* **49** 2774–2802. [MR4338383](#) <https://doi.org/10.1214/21-aos2055>
- Subba Rao, S. and Yang, J. (2023). A prediction perspective on the Wiener-Hopf equations for time series. *J. Time Series Anal.* **44** 23–42. [MR4562410](#)
- Szegö, G. (1921). Über die Randwerte einer analytischen Funktion. *Math. Ann.* **84** 232–244. [MR1512033](#) <https://doi.org/10.1007/BF01459407>
- Wiener, N. (1949). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*. Cambridge, MA: Technology Press of The Massachusetts Institute of Technology. [MR0031213](#)
- Wiener, N. and Hopf, E. (1931). Über eine klasse singulärer integralgleichungen. *Sitzungsber. Preuss. Akad. Wiss. Berlin* **31** 696–706.
- Wiener, N. and Masani, P. (1958). The prediction theory of multivariate stochastic processes. II. The linear predictor. *Acta Math.* **99** 93–137. [MR0097859](#) <https://doi.org/10.1007/BF02392423>
- Wildi, M. and McElroy, T. (2016). Optimal real-time filters for linear prediction problems. *J. Time Ser. Econom.* **8** 155–192. [MR3518405](#) <https://doi.org/10.1515/jtse-2014-0019>

# Distance correlation test for high-dimensional independence

WEIMING LI<sup>1,a</sup>, QINWEN WANG<sup>2,b</sup> and JIANFENG YAO<sup>3,c</sup>

<sup>1</sup>*School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China,*

<sup>a</sup>*li.weiming@shufe.edu.cn*

<sup>2</sup>*School of Data Science, Fudan University, Shanghai, China,* <sup>b</sup>*wqw@fudan.edu.cn*

<sup>3</sup>*School of Data Science, The Chinese University of Hong Kong (Shenzhen), Shenzhen, China,*

<sup>c</sup>*jeff Yao@cuhk.edu.cn*

In this paper, a new self-normalized and scale invariant statistic  $T_n$ , which is based on distance correlations, is developed for testing mutual independence of a high-dimensional random vector. The asymptotic normality of the statistic is established under mild moment conditions by assuming both the dimension  $p$  of the vector and the sample size  $n$  grow to infinity. In particular, the test procedure has the consistency against sparse alternatives where the dependence can be nonlinear and non-monotonic. Technically, the asymptotic normality of the test statistic is built upon martingale decomposition and novel moment method with appropriate combinatorics.

*Keywords:* High-dimensional independent test; mutual independence; distance correlation; distance covariance; sparse alternatives

## References

- Bai, Z. and Silverman, 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>
- Bai, Z., Jiang, D., Yao, J.-F. and Zheng, S. (2009). Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* **37** 3822–3840. MR2572444 <https://doi.org/10.1214/09-AOS694>
- Bergsma, W. and Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall's tau. *Bernoulli* **20** 1006–1028. MR3178526 <https://doi.org/10.3150/13-BEJ514>
- Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR1324786
- Cai, T.T. and Jiang, T. (2011). Limiting laws of coherence of random matrices with applications to testing covariance structure and construction of compressed sensing matrices. *Ann. Statist.* **39** 1496–1525. MR2850210 <https://doi.org/10.1214/11-AOS879>
- Cai, T.T. and Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli* **19** 2359–2388. MR3160557 <https://doi.org/10.3150/12-BEJ455>
- Chen, S.X., Zhang, L.-X. and Zhong, P.-S. (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105** 810–819. MR2724863 <https://doi.org/10.1198/jasa.2010.tm09560>
- Drton, M., Han, F. and Shi, H. (2020). High-dimensional consistent independence testing with maxima of rank correlations. *Ann. Statist.* **48** 3206–3227. MR4185806 <https://doi.org/10.1214/19-AOS1926>
- Gao, J., Han, X., Pan, G. and Yang, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 677–693. MR3641402 <https://doi.org/10.1111/rssb.12189>
- Han, F., Chen, S. and Liu, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306 <https://doi.org/10.1093/biomet/asx050>
- He, Y., Xu, G., Wu, C. and Pan, W. (2021). Asymptotically independent U-statistics in high-dimensional testing. *Ann. Statist.* **49** 154–181. MR4206673 <https://doi.org/10.1214/20-AOS1951>

- Jiang, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880. [MR2052906](#) <https://doi.org/10.1214/105051604000000143>
- Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. [MR1863961](#) <https://doi.org/10.1214/aos/1009210544>
- Ledoit, O., Wolf, M. et al. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** 1081–1102. [MR1926169](#) <https://doi.org/10.1214/aos/1031689018>
- Leung, D. and Drton, M. (2018). Testing independence in high dimensions with sums of rank correlations. *Ann. Statist.* **46** 280–307. [MR3766953](#) <https://doi.org/10.1214/17-AOS1550>
- Leung, D. and Shao, Q. (2019). Asymptotic power of Rao’s score test for independence in high dimensions. *Bernoulli* **25** 241–263. [MR3892319](#) <https://doi.org/10.3150/17-bej985>
- Li, Z., Wang, Q. and Li, R. (2021). Central limit theorem for linear spectral statistics of large dimensional Kendall’s rank correlation matrices and its applications. *Ann. Statist.* **49** 1569–1593. [MR4298873](#) <https://doi.org/10.1214/20-aos2013>
- Li, W., Wang, Q. and Yao, J. (2024). Supplement to “Distance correlation test for high-dimensional independence.” <https://doi.org/10.3150/23-BEJ1710SUPP>
- Li, W. and Yao, J. (2018). On structure testing for component covariance matrices of a high dimensional mixture. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 293–318. [MR3763693](#) <https://doi.org/10.1111/rssb.12248>
- Onatski, A., Moreira, M.J., Hallin, M. et al. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.* **41** 1204–1231. [MR3113808](#) <https://doi.org/10.1214/13-AOS1100>
- Onatski, A., Moreira, M.J., Hallin, M. et al. (2014). Signal detection in high dimension: The multispiked case. *Ann. Statist.* **42** 225–254. [MR3189485](#) <https://doi.org/10.1214/13-AOS1181>
- Pan, W., Wang, X., Zhang, H., Zhu, H. and Zhu, J. (2020). Ball covariance: A generic measure of dependence in Banach space. *J. Amer. Statist. Assoc.* **115** 307–317. [MR4078465](#) <https://doi.org/10.1080/01621459.2018.1543600>
- Schott, J.R. (2005). Testing for complete independence in high dimensions. *Biometrika* **92** 951–956. [MR2234197](#) <https://doi.org/10.1093/biomet/92.4.951>
- Shi, H., Drton, M. and Han, F. (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *J. Amer. Statist. Assoc.* **117** 395–410. [MR4399094](#) <https://doi.org/10.1080/01621459.2020.1782223>
- Shi, H., Hallin, M., Drton, M. and Han, F. (2022). On universally consistent and fully distribution-free rank tests of vector independence. *Ann. Statist.* **50** 1933–1959. [MR4474478](#) <https://doi.org/10.1214/21-aos2151>
- Székely, G.J. and Rizzo, M.L. (2013). The distance correlation  $t$ -test of independence in high dimension. *J. Multivariate Anal.* **117** 193–213. [MR3053543](#) <https://doi.org/10.1016/j.jmva.2013.02.012>
- Székely, G.J., Rizzo, M.L. et al. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.* **42** 2382–2412. [MR3269983](#) <https://doi.org/10.1214/14-AOS1255>
- Székely, G.J., Rizzo, M.L., Bakirov, N.K. et al. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. [MR2382665](#) <https://doi.org/10.1214/009053607000000505>
- Wang, Q. and Yao, J. (2013). On the sphericity test with large-dimensional observations. *Electron. J. Stat.* **7** 2164–2192. [MR3104916](#) <https://doi.org/10.1214/13-EJS842>
- Weihhs, L., Drton, M. and Leung, D. (2016). Efficient computation of the Bergsma-Dassios sign covariance. *Comput. Statist.* **31** 315–328. [MR3481807](#) <https://doi.org/10.1007/s00180-015-0639-x>
- Yao, S., Zhang, X. and Shao, X. (2018). Testing mutual independence in high dimension via distance covariance. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 455–480. [MR3798874](#) <https://doi.org/10.1111/rssb.12259>



# Boundary adaptive local polynomial conditional density estimators

MATIAS D. CATTANEO<sup>1,a</sup>, RAJITA CHANDAK<sup>1,b</sup>, MICHAEL JANSSON<sup>2,c</sup> and XINWEI MA<sup>3,d</sup> 

<sup>1</sup>*Department of Operations Research and Financial Engineering, Princeton University, Princeton NJ, United States,* <sup>a</sup>[cattaneo@princeton.edu](mailto:cattaneo@princeton.edu), <sup>b</sup>[rchandak@princeton.edu](mailto:rchandak@princeton.edu)

<sup>2</sup>*Department of Economics, UC Berkeley, Berkeley CA, United States,* <sup>c</sup>[mjansson@berkeley.edu](mailto:mjansson@berkeley.edu)

<sup>3</sup>*Department of Economics, UC San Diego, La Jolla CA, United States,* <sup>d</sup>[x1ma@ucsd.edu](mailto:x1ma@ucsd.edu)

We begin by introducing a class of conditional density estimators based on local polynomial techniques. The estimators are boundary adaptive and easy to implement. We then study the (pointwise and) uniform statistical properties of the estimators, offering characterizations of both probability concentration and distributional approximation. In particular, we establish uniform convergence rates in probability and valid Gaussian distributional approximations for the Studentized  $t$ -statistic process. We also discuss implementation issues such as consistent estimation of the covariance function for the Gaussian approximation, optimal integrated mean squared error bandwidth selection, and valid robust bias-corrected inference. We illustrate the applicability of our results by constructing valid confidence bands and hypothesis tests for both parametric specification and shape constraints, explicitly characterizing their approximation errors. A companion R software package implementing our main results is provided.

**Keywords:** Conditional density estimation; confidence bands; local polynomial methods; specification testing; strong approximation; uniform inference

## References

- [1] Calonico, S., Cattaneo, M.D. and Farrell, M.H. (2018). On the effect of bias estimation on coverage accuracy in nonparametric inference. *J. Amer. Statist. Assoc.* **113** 767–779. MR3832225 <https://doi.org/10.1080/01621459.2017.1285776>
- [2] Calonico, S., Cattaneo, M.D. and Farrell, M.H. (2022). Coverage error optimal confidence intervals for local polynomial regression. *Bernoulli* **28** 2998–3022. MR4474570 <https://doi.org/10.3150/21-bej1445>
- [3] Cattaneo, M.D., Chandak, R., Jansson, M. and Ma, X. (2022). `lpcde`: Local polynomial conditional density estimation and inference. Working paper, [arXiv:2204.10375](https://arxiv.org/abs/2204.10375).
- [4] Cattaneo, M.D., Chandak, R., Jansson, M. and Ma, X. (2024). Supplement to “Boundary adaptive local polynomial conditional density estimators.” <https://doi.org/10.3150/23-BEJ1711SUPP>
- [5] Cattaneo, M.D., Jansson, M. and Ma, X. (2020). Simple local polynomial density estimators. *J. Amer. Statist. Assoc.* **115** 1449–1455. MR4143477 <https://doi.org/10.1080/01621459.2019.1635480>
- [6] Cheng, M.-Y. (1994). On boundary effects of smooth curve estimators Ph.D. thesis University of North Carolina Press, Chapel Hill.
- [7] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Anti-concentration and honest, adaptive confidence bands. *Ann. Statist.* **42** 1787–1818. MR3262468 <https://doi.org/10.1214/14-AOS1235>
- [8] Chernozhukov, V., Chetverikov, D. and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Ann. Statist.* **42** 1564–1597. MR3262461 <https://doi.org/10.1214/14-AOS1230>
- [9] Chernozhukov, V., Chetverikov, D., Kato, K. and Koike, Y. (2022). Improved central limit theorem and bootstrap approximations in high dimensions. *Ann. Statist.* **50** 2562–2586. MR4500619 <https://doi.org/10.1214/22-aos2193>

- [10] De Gooijer, J.G. and Zerom, D. (2003). On conditional density estimation. *Stat. Neerl.* **57** 159–176. [MR2028911](#) <https://doi.org/10.1111/1467-9574.00226>
- [11] de la Peña, V.H. and Montgomery-Smith, S.J. (1995). Decoupling inequalities for the tail probabilities of multivariate  $U$ -statistics. *Ann. Probab.* **23** 806–816. [MR1334173](#)
- [12] Einmahl, U. and Mason, D.M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.* **13** 1–37. [MR1744994](#) <https://doi.org/10.1023/A:1007769924157>
- [13] Einmahl, U. and Mason, D.M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.* **33** 1380–1403. [MR2195639](#) <https://doi.org/10.1214/009053605000000129>
- [14] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. London: CRC Press. [MR1383587](#)
- [15] Fan, J., Yao, Q. and Tong, H. (1996). Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. *Biometrika* **83** 189–206. [MR1399164](#) <https://doi.org/10.1093/biomet/83.1.189>
- [16] Ferrigno, S., Maumy-Bertrand, M. and Muller, A. (2010). Uniform law of the logarithm for the local linear estimator of the conditional distribution function. *C. R. Math. Acad. Sci. Paris* **348** 1015–1019. [MR2721792](#) <https://doi.org/10.1016/j.crma.2010.08.003>
- [17] Giné, E., Latała, R. and Zinn, J. (2000). Exponential and moment inequalities for  $U$ -statistics. In *High Dimensional Probability, II (Seattle, WA, 1999)*. *Progress in Probability* **47** 13–38. Boston, MA: Birkhäuser. [MR1857312](#)
- [18] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics*. New York: Cambridge Univ. Press. [MR3588285](#) <https://doi.org/10.1017/CBO9781107337862>
- [19] Hall, P. (1979). On the rate of convergence of normal extremes. *J. Appl. Probab.* **16** 433–439. [MR0531778](#) <https://doi.org/10.1017/s0021900200046647>
- [20] Hall, P. (1993). On Edgeworth expansion and bootstrap confidence bands in nonparametric curve estimation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **55** 291–304.
- [21] Hall, P., Racine, J. and Li, Q. (2004). Cross-validation and the estimation of conditional probability densities. *J. Amer. Statist. Assoc.* **99** 1015–1026. [MR2109491](#) <https://doi.org/10.1198/016214504000000548>
- [22] Hall, P., Wolff, R.C.L. and Yao, Q. (1999). Methods for estimating a conditional distribution function. *J. Amer. Statist. Assoc.* **94** 154–163. [MR1689221](#) <https://doi.org/10.2307/2669691>
- [23] Khas'minskii, R.Z. (1979). A lower bound on the risks of non-parametric estimates of densities in the uniform metric. *Theory Probab. Appl.* **23** 794–798.
- [24] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z. Wahrsch. Verw. Gebiete* **32** 111–131. [MR0375412](#) <https://doi.org/10.1007/BF00533093>
- [25] Rio, E. (1994). Local invariance principles and their application to density estimation. *Probab. Theory Related Fields* **98** 21–45. [MR1254823](#) <https://doi.org/10.1007/BF01311347>
- [26] Scott, D.W. (2015). *Multivariate Density Estimation: Theory, Practice, and Visualization*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. [MR3329609](#)
- [27] Simonoff, J.S. (1996). *Smoothing Methods in Statistics*. *Springer Series in Statistics*. New York: Springer. [MR1391963](#) <https://doi.org/10.1007/978-1-4612-4026-6>
- [28] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. [MR1385671](#) <https://doi.org/10.1007/978-1-4757-2545-2>
- [29] Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing*. *Monographs on Statistics and Applied Probability* **60**. London: CRC Press. [MR1319818](#) <https://doi.org/10.1007/978-1-4899-4493-1>
- [30] Wasserman, L. (2006). *All of Nonparametric Statistics*. *Springer Texts in Statistics*. New York: Springer. [MR2172729](#)

# Normal approximation of subgraph counts in the random-connection model

QINGWEI LIU<sup>a</sup> and NICOLAS PRIVAULT<sup>b</sup> 

*Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371, Singapore, <sup>a</sup>[qingwei.liu@ntu.edu.sg](mailto:qingwei.liu@ntu.edu.sg), <sup>b</sup>[nprivault@ntu.edu.sg](mailto:nprivault@ntu.edu.sg)*

This paper derives normal approximation results for subgraph counts written as multiparameter stochastic integrals in a random-connection model based on a Poisson point process. By combinatorial arguments we express the cumulants of general subgraph counts using sums over connected partition diagrams, after cancellation of terms obtained by Möbius inversion. Using the Statulevičius condition, we deduce convergence rates in the Kolmogorov distance by studying the growth of subgraph count cumulants as the intensity of the underlying Poisson point process tends to infinity. Our analysis covers general subgraphs in the dilute and full random graph regimes, and tree-like subgraphs in the sparse random graph regime.

*Keywords:* Cumulant method; Kolmogorov distance; normal approximation; Poisson point process; random-connection model; random graphs; subgraph count

## References

- Bachmann, S. and Reitzner, M. (2018). Concentration for Poisson  $U$ -statistics: Subgraph counts in random geometric graphs. *Stochastic Process. Appl.* **128** 3327–3352. [MR3849811](#) <https://doi.org/10.1016/j.spa.2017.11.001>
- Balakrishnan, R. and Ranganathan, K. (2012). *A Textbook of Graph Theory*, 2nd ed. *Universitext*. New York: Springer. [MR2977757](#) <https://doi.org/10.1007/978-1-4614-4529-6>
- Barbour, A.D., Karoński, M. and Ruciński, A. (1989). A central limit theorem for decomposable random variables with applications to random graphs. *J. Combin. Theory Ser. B* **47** 125–145. [MR1047781](#) [https://doi.org/10.1016/0095-8956\(89\)90014-2](https://doi.org/10.1016/0095-8956(89)90014-2)
- Bender, E.A., Odlyzko, A.M. and Richmond, L.B. (1985). The asymptotic number of irreducible partitions. *European J. Combin.* **6** 1–6. [MR0793481](#) [https://doi.org/10.1016/S0195-6698\(85\)80015-9](https://doi.org/10.1016/S0195-6698(85)80015-9)
- Bogdan, K., Rosiński, J., Serafin, G. and Wojciechowski, Ł. (2017). Lévy systems and moment formulas for mixed Poisson integrals. In *Stochastic Analysis and Related Topics. Progress in Probability* **72** 139–164. Cham: Birkhäuser/Springer. [MR3737628](#) [https://doi.org/10.1007/978-3-319-59671-6\\_7](https://doi.org/10.1007/978-3-319-59671-6_7)
- Can, V.H. and Trinh, K.D. (2022). Random connection models in the thermodynamic regime: Central limit theorems for add-one cost stabilizing functionals. *Electron. J. Probab.* **27** 1–40. [MR4388562](#) <https://doi.org/10.1214/22-EJP759>
- Döring, H. and Eichelsbacher, P. (2013). Moderate deviations via cumulants. *J. Theoret. Probab.* **26** 360–385. [MR3055808](#) <https://doi.org/10.1007/s10959-012-0437-0>
- Döring, H., Jansen, S. and Schubert, K. (2022). The method of cumulants for the normal approximation. *Probab. Surv.* **19** 185–270. [MR4408127](#) <https://doi.org/10.1214/22-ps7>
- Eichelsbacher, P. and Rednoss, B. (2023). Kolmogorov bounds for decomposable random variables and subgraph counting by the Stein-Tikhomirov method. *Bernoulli* **29** 1821–1848. [MR4580898](#) <https://doi.org/10.3150/22-bej1522>
- Eichelsbacher, P. and Thäle, C. (2014). New Berry-Esseen bounds for non-linear functionals of Poisson random measures. *Electron. J. Probab.* **19** 1–25. [MR3275854](#) <https://doi.org/10.1214/EJP.v19-3061>
- Erdős, P. and Rényi, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. [MR0120167](#) <https://doi.org/10.5486/pmd.1959.6.3-4.12>

- Féray, V., Méliot, P.-L. and Nikeghbali, A. (2016). *Mod- $\phi$  Convergence: Normality Zones and Precise Deviations*. SpringerBriefs in Probability and Mathematical Statistics. Cham: Springer. MR3585777 <https://doi.org/10.1007/978-3-319-46822-8>
- Gilbert, E.N. (1959). Random graphs. *Ann. Math. Stat.* **30** 1141–1144. MR0108839 <https://doi.org/10.1214/aoms/1177706098>
- Götze, F., Heinrich, L. and Hipp, C. (1995).  $m$ -dependent random fields with analytic cumulant generating function. *Scand. J. Stat.* **22** 183–195. MR1339750
- Grote, J. and Thäle, C. (2018a). Concentration and moderate deviations for Poisson polytopes and polyhedra. *Bernoulli* **24** 2811–2841. MR3779703 <https://doi.org/10.3150/17-BEJ946>
- Grote, J. and Thäle, C. (2018b). Gaussian polytopes: A cumulant-based approach. *J. Complexity* **47** 1–41. MR3804598 <https://doi.org/10.1016/j.jco.2018.03.001>
- Gusakova, A. and Thäle, C. (2021). The volume of simplices in high-dimensional Poisson-Delaunay tessellations. *Ann. Henri Lebesgue* **4** 121–153. MR4213157 <https://doi.org/10.5802/ahl.68>
- Heinrich, L. (2007). An almost-Markov-type mixing condition and large deviations for Boolean models in the line. *Acta Appl. Math.* **96** 247–262. MR2327539 <https://doi.org/10.1007/s10440-007-9105-2>
- Heinrich, L. and Spiess, M. (2009). Berry-Esseen bounds and Cramér-type large deviations for the volume distribution of Poisson cylinder processes. *Lith. Math. J.* **49** 381–398. MR2591874 <https://doi.org/10.1007/s10986-009-9061-9>
- Jansen, S. (2019). Cluster expansions for Gibbs point processes. *Adv. in Appl. Probab.* **51** 1129–1178. MR4032174 <https://doi.org/10.1017/apr.2019.46>
- Janson, S. (1988). Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *Ann. Probab.* **16** 305–312. MR0920273
- Khorunzhiy, O. (2008). On connected diagrams and cumulants of Erdős-Rényi matrix models. *Comm. Math. Phys.* **282** 209–238. MR2415478 <https://doi.org/10.1007/s00220-008-0533-2>
- Krokowski, K., Reichenbachs, A. and Thäle, C. (2017). Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.* **45** 1071–1109. MR3630293 <https://doi.org/10.1214/15-AOP1081>
- Lachièze-Rey, R. and Peccati, G. (2013). Fine Gaussian fluctuations on the Poisson space II: Rescaled kernels, marked processes and geometric  $U$ -statistics. *Stochastic Process. Appl.* **123** 4186–4218. MR3096352 <https://doi.org/10.1016/j.spa.2013.06.004>
- Lachièze-Rey, R. and Reitzner, M. (2016).  $U$ -statistics in stochastic geometry. In *Stochastic Analysis for Poisson Point Processes*. Bocconi Springer Ser. **7** 229–253. Bocconi Univ. Press. MR3585402
- Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 <https://doi.org/10.1214/18-AAP1405>
- Last, G., Nestmann, F. and Schulte, M. (2021). The random connection model and functions of edge-marked Poisson processes: Second order properties and normal approximation. *Ann. Appl. Probab.* **31** 128–168. MR4254476 <https://doi.org/10.1214/20-aap1585>
- Last, G., Peccati, G. and Schulte, M. (2016). Normal approximation on Poisson spaces: Mehler’s formula, second order Poincaré inequalities and stabilization. *Probab. Theory Related Fields* **165** 667–723. MR3520016 <https://doi.org/10.1007/s00440-015-0643-7>
- Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks **7**. Cambridge: Cambridge Univ. Press. MR3791470
- Malyshev, V.A. and Minlos, R.A. (1991). *Gibbs Random Fields: Cluster Expansions. Mathematics and Its Applications (Soviet Series)* **44**. Dordrecht: Kluwer Academic. Translated from the Russian by R. Kotecký and P. Holický. MR1191166 <https://doi.org/10.1007/978-94-011-3708-9>
- Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos: Moments, Cumulants and Diagrams: A Survey with Computer Implementation*. Bocconi & Springer Series **1**. Milan: Springer; Bocconi Univ. Press. Supplementary material available online. MR2791919 <https://doi.org/10.1007/978-88-470-1679-8>
- Penrose, M. (2003). *Random Geometric Graphs*. Oxford Studies in Probability **5**. Oxford: Oxford Univ. Press. MR1986198 <https://doi.org/10.1093/acprof:oso/9780198506263.001.0001>
- Penrose, M.D. and Yukich, J.E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11** 1005–1041. MR1878288 <https://doi.org/10.1214/aoap/1015345393>

- Penrose, M.D. and Yukich, J.E. (2005). Normal approximation in geometric probability. In *Stein's Method and Applications. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **5** 37–58. Singapore: Singapore Univ. Press. [MR2201885 https://doi.org/10.1142/9789812567673\\_0003](https://doi.org/10.1142/9789812567673_0003)
- Privault, N. (2012). Moments of Poisson stochastic integrals with random integrands. *Probab. Math. Statist.* **32** 227–239. [MR3021456 https://doi.org/10.1007/s11464-012-0227-2](https://doi.org/10.1007/s11464-012-0227-2)
- Privault, N. (2019). Moments of  $k$ -hop counts in the random-connection model. *J. Appl. Probab.* **56** 1106–1121. [MR4041451 https://doi.org/10.1017/jpr.2019.63](https://doi.org/10.1017/jpr.2019.63)
- Privault, N. (2022). Asymptotic analysis of  $k$ -hop connectivity in the 1D unit disk random graph model. Preprint, 40 pages. Available at [arXiv:2203.14535](https://arxiv.org/abs/2203.14535).
- Privault, N. and Serafin, G. (2020). Normal approximation for sums of weighted  $U$ -statistics—application to Kolmogorov bounds in random subgraph counting. *Bernoulli* **26** 587–615. [MR4036045 https://doi.org/10.3150/19-BEJ1141](https://doi.org/10.3150/19-BEJ1141)
- Privault, N. and Serafin, G. (2022). Berry-Esseen bounds for functionals of independent random variables. *Electron. J. Probab.* **27** 1–37. [MR4440064 https://doi.org/10.1214/22-ejp795](https://doi.org/10.1214/22-ejp795)
- Reitzner, M. and Schulte, M. (2013). Central limit theorems for  $U$ -statistics of Poisson point processes. *Ann. Probab.* **41** 3879–3909. [MR3161465 https://doi.org/10.1214/12-AOP817](https://doi.org/10.1214/12-AOP817)
- Röllin, A. (2022). Kolmogorov bounds for the normal approximation of the number of triangles in the Erdős-Rényi random graph. *Probab. Engrg. Inform. Sci.* **36** 747–773. [MR4459700 https://doi.org/10.1017/S0269964821000061](https://doi.org/10.1017/S0269964821000061)
- Rota, G.-C. (1964). On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrsch. Verw. Gebiete* **2** 340–368. [MR0174487 https://doi.org/10.1007/BF00531932](https://doi.org/10.1007/BF00531932)
- Ruciński, A. (1988). When are small subgraphs of a random graph normally distributed? *Probab. Theory Related Fields* **78** 1–10. [MR0940863 https://doi.org/10.1007/BF00718031](https://doi.org/10.1007/BF00718031)
- Rudzkis, R., Saulis, L. and Statulevičius, V. A. (1978). A general lemma on probabilities of large deviations. *Liet. Mat. Rink.* **18** 99–116, 217. [MR0501287 https://doi.org/10.1007/BF00718031](https://doi.org/10.1007/BF00718031)
- Saulis, L. and Statulevičius, V.A. (1991). *Limit Theorems for Large Deviations. Mathematics and Its Applications (Soviet Series)* **73**. Dordrecht: Kluwer Academic. Translated and revised from the 1989 Russian original. [MR1171883 https://doi.org/10.1007/978-94-011-3530-6](https://doi.org/10.1007/978-94-011-3530-6)
- Schulte, M. (2016). Normal approximation of Poisson functionals in Kolmogorov distance. *J. Theoret. Probab.* **29** 96–117. [MR3463079 https://doi.org/10.1007/s10959-014-0576-6](https://doi.org/10.1007/s10959-014-0576-6)
- Schulte, M. and Thäle, C. (2023). Moderate deviations on Poisson chaos. Preprint. Available at [arXiv:2304.00876v1](https://arxiv.org/abs/2304.00876v1).
- Zhang, Z.-S. (2022). Berry-Esseen bounds for generalized  $U$ -statistics. *Electron. J. Probab.* **27** 1–36. [MR4492985 https://doi.org/10.1214/22-ejp860](https://doi.org/10.1214/22-ejp860)

# A robust approach for regression analysis of panel count data with time-varying covariates

DAYU SUN<sup>1,a</sup>, YUANYUAN GUO<sup>2,b</sup>, YANG LI<sup>3,c</sup>, WANZHU TU<sup>3,d</sup> and  
JIANGUO SUN<sup>4,e</sup>

<sup>1</sup>Department of Biostatistics and Bioinformatics, Emory University, GA 30322, U.S.A., <sup>a</sup>[dayu.sun@emory.edu](mailto:dayu.sun@emory.edu)

<sup>2</sup>Department of Biostatistics and Bioinformatics, Duke University, NC 27710, U.S.A., <sup>b</sup>[yuanyuan.guo@duke.edu](mailto:yuanyuan.guo@duke.edu)

<sup>3</sup>Department of Biostatistics and Health Data Science, Indiana University, Indianapolis, IN 46202, U.S.A.,

<sup>c</sup>[yli18@iu.edu](mailto:yli18@iu.edu), <sup>d</sup>[wtu1@iu.edu](mailto:wту1@iu.edu)

<sup>4</sup>Department of Statistics, University of Missouri, MO 65211, U.S.A., <sup>e</sup>[sunj@missouri.edu](mailto:sunj@missouri.edu)

The validity of statistical inference for panel count data with time-varying covariates depends on the correct specification of within-subject correlation structures; misspecification often leads to questionable inference. To alleviate, robust inference has been proposed for mean models, which implicitly assume monotone mean functions. When covariate values fluctuate with time, however, the assumed monotonicity becomes unrealistic. In this research, we propose a robust inference based on rate models that are free of such constraints. Since the asymptotic variance has no closed form under the rate model, we further develop computationally efficient robust variance estimators using the Expectation-Maximization (EM) algorithm, thus sidestepping the need for computationally intensive numerical methods, which could undermine the robustness. Rigorous theoretical development is provided in support of parameter estimation and inference. Extensive simulation studies demonstrate the superiority of the proposed method. We present a real clinical application to illustrate the use of the proposed method.

**Keywords:** EM algorithm; panel count data; robust variance estimation; semiparametric model; time-to-event data

## References

- [1] Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Statist.* **10** 1100–1120. [MR0673646 https://doi.org/10.1214/aos/1176345976](https://doi.org/10.1214/aos/1176345976)
- [2] Balakrishnan, N. and Zhao, X. (2009). New multi-sample nonparametric tests for panel count data. *Ann. Statist.* **37** 1112–1149. [MR2509069 https://doi.org/10.1214/08-AOS599](https://doi.org/10.1214/08-AOS599)
- [3] Balakrishnan, N. and Zhao, X. (2011). A class of multi-sample nonparametric tests for panel count data. *Ann. Inst. Statist. Math.* **63** 135–156. [MR2748938 https://doi.org/10.1007/s10463-008-0209-x](https://doi.org/10.1007/s10463-008-0209-x)
- [4] Batteiger, B.E., Tu, W., Ofner, S., Van Der Pol, B., Stothard, D.R., Orr, D.P., Katz, B.P. and Fortenberry, J.D. (2010). Repeated Chlamydia trachomatis genital infections in adolescent women. *J. Infect. Dis.* **201** 42–51. <https://doi.org/10.1086/648734>
- [5] Ciarlet, P.G. (2013). *Linear and Nonlinear Functional Analysis with Applications*. Philadelphia, PA: SIAM. [MR3136903](https://doi.org/10.1137/12M0903)
- [6] Cook, R.J. and Lawless, J.F. (2007). *The Statistical Analysis of Recurrent Events. Statistics for Biology and Health*. New York: Springer. [MR3822124](https://doi.org/10.1002/9781118133111)
- [7] Deng, S. (2013). Semiparametric regression analysis of panel count data with time-dependent covariates and informative observation and censoring times. *Comm. Statist. Theory Methods* **42** 4170–4183. [MR3170990 https://doi.org/10.1080/03610926.2011.642922](https://doi.org/10.1080/03610926.2011.642922)
- [8] Du, M. and Sun, J. (2021). Statistical analysis of interval-censored failure time data. *Chinese J. Appl. Probab. Statist.* **37** 627–654. [MR4410679 https://doi.org/10.3969/j.issn.1001-4268.2021.06.006](https://doi.org/10.3969/j.issn.1001-4268.2021.06.006)
- [9] Du, M. and Sun, J. (2022). Variable selection for interval-censored failure time data. *Int. Stat. Rev.* **90** 193–215. [MR4481432 https://doi.org/10.1111/insr.12480](https://doi.org/10.1111/insr.12480)

- [10] Du, M. and Zhao, H. (2021). A unified approach to variable selection for Cox's proportional hazards model with interval-censored failure time data. *Stat. Methods Med. Res.* **30** 1833–1849. MR4307785 <https://doi.org/10.1177/09622802211009259>
- [11] Gao, F. and Chan, K.C.G. (2019). Semiparametric regression analysis of length-biased interval-censored data. *Biometrics* **75** 121–132. MR3953713 <https://doi.org/10.1111/biom.12970>
- [12] Ge, L., Zhu, L. and Sun, J. (2021). Regression analysis of mixed panel count data with informative indicator processes. *Stat. Med.* **40** 1262–1271. MR4384374 <https://doi.org/10.1002/sim.8839>
- [13] Ghosh, P. and Tu, W. (2009). Assessing sexual attitudes and behaviors of young women: A joint model with nonlinear time effects, time varying covariates, and dropouts. *J. Amer. Statist. Assoc.* **104** 474–485. MR2751432 <https://doi.org/10.1198/jasa.2009.0013>
- [14] Hu, X.J., Lagakos, S.W. and Lockhart, R.A. (2009). Marginal analysis of panel counts through estimating functions. *Biometrika* **96** 445–456. MR2507154 <https://doi.org/10.1093/biomet/asp010>
- [15] Hu, X.J., Sun, J. and Wei, L.-J. (2003). Regression parameter estimation from panel counts. *Scand. J. Stat.* **30** 25–43. MR1963891 <https://doi.org/10.1111/1467-9469.00316>
- [16] Hua, L., Zhang, Y. and Tu, W. (2014). A spline-based semiparametric sieve likelihood method for over-dispersed panel count data. *Canad. J. Statist.* **42** 217–245. MR3208337 <https://doi.org/10.1002/cjs.11208>
- [17] Jiang, H., Su, W. and Zhao, X. (2020). Robust estimation for panel count data with informative observation times and censoring times. *Lifetime Data Anal.* **26** 65–84. MR4051838 <https://doi.org/10.1007/s10985-018-09457-7>
- [18] Kosorok, M.R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. Springer Series in Statistics. New York: Springer. MR2724368 <https://doi.org/10.1007/978-0-387-74978-5>
- [19] Li, N., Sun, L. and Sun, J. (2010). Semiparametric transformation models for panel count data with dependent observation process. *Stat. Biosci.* **2** 191–210. <https://doi.org/10.1007/s12561-010-9029-7>
- [20] Li, N., Zhao, H. and Sun, J. (2013). Semiparametric transformation models for panel count data with correlated observation and follow-up times. *Stat. Med.* **32** 3039–3054. MR3073834 <https://doi.org/10.1002/sim.5724>
- [21] Li, S., Hu, T., Zhao, S. and Sun, J. (2020). Regression analysis of multivariate current status data with semiparametric transformation frailty models. *Statist. Sinica* **30** 1117–1134. MR4214176
- [22] Li, S., Hu, T., Zhao, X. and Sun, J. (2019). A class of semiparametric transformation cure models for interval-censored failure time data. *Comput. Statist. Data Anal.* **133** 153–165. MR3926472 <https://doi.org/10.1016/j.csda.2018.09.008>
- [23] Li, Z., Liu, H. and Tu, W. (2015). A sexually transmitted infection screening algorithm based on semiparametric regression models. *Stat. Med.* **34** 2844–2857. MR3375984 <https://doi.org/10.1002/sim.6515>
- [24] Lin, D.Y., Wei, L.J., Yang, I. and Ying, Z. (2000). Semiparametric regression for the mean and rate functions of recurrent events. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 711–730. MR1796287 <https://doi.org/10.1111/1467-9868.00259>
- [25] Liu, L., Su, W., Yin, G., Zhao, X. and Zhang, Y. (2022). Nonparametric inference for reversed mean models with panel count data. *Bernoulli* **28** 2968–2997. MR4474569 <https://doi.org/10.3150/21-bej1444>
- [26] Louis, T.A. (1982). Finding the observed information matrix when using the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **44** 226–233. MR0676213
- [27] Mao, L. (2020). A unified approach to the calculation of information operators in semiparametric models. *Biometrika* **107** 983–995. MR4186500 <https://doi.org/10.1093/biomet/asaa037>
- [28] Murphy, S.A., Rossini, A.J. and van der Vaart, A.W. (1997). Maximum likelihood estimation in the proportional odds model. *J. Amer. Statist. Assoc.* **92** 968–976. MR1482127 <https://doi.org/10.2307/2965560>
- [29] Murphy, S.A. and van der Vaart, A.W. (2000). On profile likelihood. *J. Amer. Statist. Assoc.* **95** 449–465. With comments and a rejoinder by the authors. MR1803168 <https://doi.org/10.2307/2669386>
- [30] Oakes, D. (1999). Direct calculation of the information matrix via the EM algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 479–482. MR1680298 <https://doi.org/10.1111/1467-9868.00188>
- [31] Su, Y.-R. and Wang, J.-L. (2016). Semiparametric efficient estimation for shared-frailty models with doubly-censored clustered data. *Ann. Statist.* **44** 1298–1331. MR3485961 <https://doi.org/10.1214/15-AOS1406>
- [32] Sun, D., Guo, Y., Li, Y., Tu, W. and Sun, J. (2024). Supplement to “A robust approach for regression analysis of panel count data with time-varying covariates.” <https://doi.org/10.3150/23-BEJ1713SUPP>

- [33] Sun, J. and Wei, L.J. (2000). Regression analysis of panel count data with covariate-dependent observation and censoring times. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 293–302. MR1749540 <https://doi.org/10.1111/1467-9868.00232>
- [34] Sun, J. and Zhao, X. (2013). *Statistical Analysis of Panel Count Data. Statistics for Biology and Health*. New York: Springer. MR3136574 <https://doi.org/10.1007/978-1-4614-8715-9>
- [35] Torrone, E., Papp, J. and Weinstock, H. (2014). Prevalence of Chlamydia trachomatis genital infection among persons aged 14–39 years—United States, 2007–2012. *Morb. Mort. Wkly. Rep.* **63** 834–838.
- [36] Tu, W., Batteiger, B.E., Wiehe, S., Ofner, M.S.S., Van Der Pol, B., Katz, B.P., Orr, D.P. and Fortenberry, J.D. (2009). Time from first intercourse to first sexually transmitted infection diagnosis among adolescent women. *Arch. Pediatr. Adolesc. Med.* **163** 1106–1111. <https://doi.org/10.1001/archpediatrics.2009.203>
- [37] van der Vaart, A. (2002). Semiparametric statistics. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1999). Lecture Notes in Math.* **1781** 331–457. Berlin: Springer. MR1915446
- [38] 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>
- [39] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes. Springer Series in Statistics*. New York: Springer. With applications to statistics. MR1385671 <https://doi.org/10.1007/978-1-4757-2545-2>
- [40] van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems* (E.A. Cator, G. Jongbloed, C. Kraaikamp, H.P. Lopuhaä and J.A. Wellner, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942 <https://doi.org/10.1214/074921707000000382>
- [41] Wang, Y. and Yu, Z. (2021). A kernel regression model for panel count data with time-varying coefficients. *Statist. Sinica* **31** 1707–1725. MR4328837 <https://doi.org/10.5705/ss.202019.0220>
- [42] Wellner, J.A. and Zhang, Y. (2000). Two estimators of the mean of a counting process with panel count data. *Ann. Statist.* **28** 779–814. MR1792787 <https://doi.org/10.1214/aos/1015951998>
- [43] Wellner, J.A. and Zhang, Y. (2007). Two likelihood-based semiparametric estimation methods for panel count data with covariates. *Ann. Statist.* **35** 2106–2142. MR2363965 <https://doi.org/10.1214/009053607000000181>
- [44] Yu, G., Zhu, L., Li, Y., Sun, J. and Robison, L.L. (2017). Regression analysis of mixed panel count data with dependent terminal events. *Stat. Med.* **36** 1669–1680. MR3631986 <https://doi.org/10.1002/sim.7217>
- [45] Zeng, D. and Lin, D.Y. (2021). Maximum likelihood estimation for semiparametric regression models with panel count data. *Biometrika* **108** 947–963. MR4341361 <https://doi.org/10.1093/biomet/asaa091>
- [46] Zeng, D., Mao, L. and Lin, D.Y. (2016). Maximum likelihood estimation for semiparametric transformation models with interval-censored data. *Biometrika* **103** 253–271. MR3509885 <https://doi.org/10.1093/biomet/asw013>
- [47] Zhao, H., Tu, W. and Yu, Z. (2018). A nonparametric time-varying coefficient model for panel count data. *J. Nonparametr. Stat.* **30** 640–661. MR3843044 <https://doi.org/10.1080/10485252.2018.1458982>
- [48] Zhao, X. and Zhang, Y. (2017). Asymptotic normality of nonparametric  $M$ -estimators with applications to hypothesis testing for panel count data. *Statist. Sinica* **27** 931–950. MR3675037
- [49] Zhu, L., Tong, X., Sun, J., Chen, M., Srivastava, D.K., Leisenring, W. and Robison, L.L. (2014). Regression analysis of mixed recurrent-event and panel-count data. *Biostatistics* **15** 555–568. <https://doi.org/10.1093/biostatistics/kxu009>



# Nonuniform Berry-Esseen bounds for studentized U-statistics

DENNIS LEUNG<sup>1,a</sup> and QI-MAN SHAO<sup>2,b</sup>

<sup>1</sup>*School of Mathematics and Statistics, University of Melbourne, Victoria, Australia,*

<sup>a</sup>[dennis.leung@unimelb.edu.au](mailto:dennis.leung@unimelb.edu.au)

<sup>2</sup>*Department of Statistics and Data Science, SICM, National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology, Shenzhen, China,* <sup>b</sup>[shaoqm@sustech.edu.cn](mailto:shaoqm@sustech.edu.cn)

We establish *nonuniform* Berry-Esseen (B-E) bounds for Studentized U-statistics of the rate  $1/\sqrt{n}$  under a third-moment assumption, which covers the t-statistic that corresponds to a kernel of degree 1 as a special case. While an interesting data example raised by Novak (*Theory Probab. Appl.* **49** (2005) 336–344) can show that the form of the nonuniform bound for standardized U-statistics is actually *invalid* for their Studentized counterparts, our main results suggest that, the validity of such a bound can be restored by minimally augmenting it with an additive correction term that decays exponentially in  $n$ . To our best knowledge, this is the first time that valid nonuniform B-E bounds for Studentized U-statistics have appeared in the literature.

*Keywords:* Exponential lower tail bound of non-negative kernel U-statistics; nonlinear statistics; nonuniform Berry-Esseen bound; Stein’s method; Studentization; U-statistics; variable censoring

## References

- Arvesen, J.N. (1969). Jackknifing  $U$ -statistics. *Ann. Math. Stat.* **40** 2076–2100. [MR0264805](#) <https://doi.org/10.1214/aoms/1177697287>
- Bentkus, V., Götze, F. and Zitikis, R. (1994). Lower estimates of the convergence rate for  $U$ -statistics. *Ann. Probab.* **22** 1707–1714. [MR1331199](#)
- Berry, A.C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** 122–136. [MR0003498](#) <https://doi.org/10.2307/1990053>
- Bickel, P.J. (1974). Edgeworth expansions in non parametric statistics. *Ann. Statist.* **2** 1–20. [MR0350952](#)
- Callaert, H. and Janssen, P. (1978). The Berry-Esseen theorem for  $U$ -statistics. *Ann. Statist.* **6** 417–421. [MR0464359](#)
- Callaert, H. and Veraverbeke, N. (1981). The order of the normal approximation for a studentized  $U$ -statistic. *Ann. Statist.* **9** 194–200. [MR0600547](#)
- Chan, Y.-K. and Wierman, J. (1977). On the Berry-Esseen theorem for  $U$ -statistics. *Ann. Probab.* **5** 136–139. [MR0433551](#) <https://doi.org/10.1214/aop/1176995897>
- Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein’s Method. Probability and Its Applications (New York)*. Heidelberg: Springer. [MR2732624](#) <https://doi.org/10.1007/978-3-642-15007-4>
- Chen, L.H.Y. and Shao, Q.-M. (2007). Normal approximation for nonlinear statistics using a concentration inequality approach. *Bernoulli* **13** 581–599. [MR2331265](#) <https://doi.org/10.3150/07-BEJ5164>
- de la Peña, V.H., Lai, T.L. and Shao, Q.-M. (2009). *Self-Normalized Processes: Limit Theory and Statistical Applications. Probability and Its Applications (New York)*. Berlin: Springer. [MR2488094](#) <https://doi.org/10.1007/978-3-540-85636-8>
- Efron, B. (1969). Student’s  $t$ -test under symmetry conditions. *J. Amer. Statist. Assoc.* **64** 1278–1302. [MR0251826](#)
- Esseen, C.-G. (1942). On the Liapounoff limit of error in the theory of probability. *Ark. Mat. Astron. Fys.* **28A** 19. [MR0011909](#)
- Esseen, C.G. (1956). A moment inequality with an application to the central limit theorem. *Skand. Aktuarietidskr.* **39** 160–170. [MR0090166](#) <https://doi.org/10.1080/03461238.1956.10414946>

- Filippova, A. (1962). Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. *Theory Probab. Appl.* **7** 24–57.
- Friedrich, K.O. (1989). A Berry-Esseen bound for functions of independent random variables. *Ann. Statist.* **17** 170–183. MR0981443 <https://doi.org/10.1214/aos/1176347009>
- Grams, W.F. and Serfling, R.J. (1973). Convergence rates for  $U$ -statistics and related statistics. *Ann. Statist.* **1** 153–160. MR0336788
- Helmers, R. (1985). The Berry-Esseen bound for Studentized  $U$ -statistics. *Canad. J. Statist.* **13** 79–82. MR0792555 <https://doi.org/10.2307/3315169>
- 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>
- Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR0144363
- 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>
- Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93.
- Lai, T.L., Shao, Q.-M. and Wang, Q. (2011). Cramér type moderate deviations for Studentized  $U$ -statistics. *ESAIM Probab. Stat.* **15** 168–179. MR2870510 <https://doi.org/10.1051/ps/2009014>
- Leung, D. and Shao, Q.-M. (2023). Another look at Stein's method for Studentized nonlinear statistics with an application to  $U$ -statistics. ArXiv preprint. Available at [arXiv:2301.02098](https://arxiv.org/abs/2301.02098).
- Leung, D. and Shao, Q.-M. (2024). Supplement to “Nonuniform Berry-Esseen bounds for studentized  $U$ -statistics.” <https://doi.org/10.3150/23-BEJ1714SUPP>
- Novak, S.Y. (2005). On self-normalized sums and Student's statistic. *Theory Probab. Appl.* **49** 336–344.
- Rosenthal, H.P. (1970). On the subspaces of  $L^p$  ( $p > 2$ ) spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303. MR0271721 <https://doi.org/10.1007/BF02771562>
- 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>
- Shevtsova, I. (2011). On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands. ArXiv preprint. Available at [arXiv:1111.6554](https://arxiv.org/abs/1111.6554).
- Student (1908). The probable error of a mean. *Biometrika* **6** 1–25.
- van Zwet, W.R. (1984). A Berry-Esseen bound for symmetric statistics. *Z. Wahrsch. Verw. Gebiete* **66** 425–440. MR0751580 <https://doi.org/10.1007/BF00533707>
- Wang, Q. and Jing, B.-Y. (1999). An exponential nonuniform Berry-Esseen bound for self-normalized sums. *Ann. Probab.* **27** 2068–2088. MR1742902 <https://doi.org/10.1214/aop/1022677562>
- Wang, Q., Jing, B.-Y. and Zhao, L. (2000). The Berry-Esseen bound for Studentized statistics. *Ann. Probab.* **28** 511–535. MR1756015 <https://doi.org/10.1214/aop/1019160129>
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics* **1** 80–83.
- Zhao, L.C. (1983). The rate of the normal approximation for a Studentized  $U$ -statistic. *Sci. Exploration* **3** 45–52. MR0746581

# Regularities and exponential ergodicity in entropy for SDEs driven by distribution dependent noise

XING HUANG<sup>a</sup> and FENG-YU WANG<sup>b</sup>

*Center for Applied Mathematics, Tianjin University, Tianjin 300072, China, <sup>a</sup>xinghuang@tju.edu.cn, <sup>b</sup>wangfy@tju.edu.cn*

As two crucial tools characterizing regularity properties of stochastic systems, the log-Harnack inequality and Bismut formula have been intensively studied for distribution dependent (McKean-Vlasov) SDEs. However, due to technical difficulties, existing results mainly focus on the case with distribution free noise. In this paper, we introduce a noise decomposition argument to establish the log-Harnack inequality and Bismut formula for SDEs with distribution dependent noise, in both non-degenerate and degenerate situations. As an application, the exponential ergodicity in entropy is investigated.

*Keywords:* Bismut formula; distribution dependent SDE; exponential ergodicity in entropy; log-Harnack inequality

## References

- Arnaudon, M., Thalmaier, A. and Wang, F.-Y. (2009). Gradient estimates and Harnack inequalities on non-compact Riemannian manifolds. *Stochastic Process. Appl.* **119** 3653–3670. MR2568290 <https://doi.org/10.1016/j.spa.2009.07.001>
- Bai, Y. and Huang, X. (2023). Log-Harnack inequality and exponential ergodicity for distribution dependent Chan-Karolyi-Longstaff-Sanders and Vasicek models. *J. Theoret. Probab.* **36** 1902–1921. MR4621087 <https://doi.org/10.1007/s10959-022-01210-z>
- Baños, D. (2018). The Bismut-Elworthy-Li formula for mean-field stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 220–233. MR3765887 <https://doi.org/10.1214/16-AIHP801>
- Bao, J., Ren, P. and Wang, F.-Y. (2021). Bismut formula for Lions derivative of distribution-path dependent SDEs. *J. Differ. Equ.* **282** 285–329. MR4219322 <https://doi.org/10.1016/j.jde.2021.02.019>
- Bauer, M., Meyer-Brandis, T. and Proske, F. (2018). Strong solutions of mean-field stochastic differential equations with irregular drift. *Electron. J. Probab.* **23** Paper No. 132, 35. MR3896869 <https://doi.org/10.1214/18-EJP259>
- Bismut, J.-M. (1984). *Large Deviations and the Malliavin Calculus. Progress in Mathematics* **45**. Boston, MA: Birkhäuser, Inc. MR0755001
- Bogachev, V.I., Röckner, M. and Shaposhnikov, S.V. (2016). Distances between transition probabilities of diffusions and applications to nonlinear Fokker-Planck-Kolmogorov equations. *J. Funct. Anal.* **271** 1262–1300. MR3522009 <https://doi.org/10.1016/j.jfa.2016.05.016>
- Crisan, D. and McMurray, E. (2018). Smoothing properties of McKean-Vlasov SDEs. *Probab. Theory Related Fields* **171** 97–148. MR3800831 <https://doi.org/10.1007/s00440-017-0774-0>
- Elworthy, K.D. and Li, X.-M. (1994). Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** 252–286. MR1297021 <https://doi.org/10.1006/jfan.1994.1124>
- Guillin, A. and Wang, F.-Y. (2012). Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality. *J. Differ. Equ.* **253** 20–40. MR2917400 <https://doi.org/10.1016/j.jde.2012.03.014>
- Huang, X., Ren, P. and Wang, F.-Y. (2021). Distribution dependent stochastic differential equations. *Front. Math. China* **16** 257–301. MR4254653 <https://doi.org/10.1007/s11464-021-0920-y>

- Huang, X., Röckner, M. and Wang, F.-Y. (2019). Nonlinear Fokker-Planck equations for probability measures on path space and path-distribution dependent SDEs. *Discrete Contin. Dyn. Syst.* **39** 3017–3035. [MR3959419 https://doi.org/10.3934/dcds.2019125](https://doi.org/10.3934/dcds.2019125)
- Huang, X. and Song, Y. (2021). Well-posedness and regularity for distribution dependent SPDEs with singular drifts. *Nonlinear Anal.* **203** Paper No. 112167, 18. [MR4164554 https://doi.org/10.1016/j.na.2020.112167](https://doi.org/10.1016/j.na.2020.112167)
- Huang, X. and Wang, F.-Y. (2019). Distribution dependent SDEs with singular coefficients. *Stochastic Process. Appl.* **129** 4747–4770. [MR4013879 https://doi.org/10.1016/j.spa.2018.12.012](https://doi.org/10.1016/j.spa.2018.12.012)
- Huang, X. and Wang, F.-Y. (2021). Derivative estimates on distributions of McKean-Vlasov SDEs. *Electron. J. Probab.* **26** Paper No. 15, 12. [MR4225923 https://doi.org/10.1214/21-EJP582](https://doi.org/10.1214/21-EJP582)
- Huang, X. and Wang, F.-Y. (2022). Log-Harnack inequality and Bismut formula for singular McKean-Vlasov SDEs. Available at [arXiv:2207.11536](https://arxiv.org/abs/2207.11536).
- Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236** 244–264. [MR2227134 https://doi.org/10.1016/j.jfa.2005.12.010](https://doi.org/10.1016/j.jfa.2005.12.010)
- Ren, P. and Wang, F.-Y. (2019). Bismut formula for Lions derivative of distribution dependent SDEs and applications. *J. Differ. Equ.* **267** 4745–4777. [MR3983053 https://doi.org/10.1016/j.jde.2019.05.016](https://doi.org/10.1016/j.jde.2019.05.016)
- Ren, P. and Wang, F.-Y. (2021a). Exponential convergence in entropy and Wasserstein for McKean-Vlasov SDEs. *Nonlinear Anal.* **206** Paper No. 112259, 20. [MR4206077 https://doi.org/10.1016/j.na.2021.112259](https://doi.org/10.1016/j.na.2021.112259)
- Ren, P. and Wang, F.-Y. (2021b). Derivative formulas in measure on Riemannian manifolds. *Bull. Lond. Math. Soc.* **53** 1786–1800. [MR4379563 https://doi.org/10.1112/blms.12542](https://doi.org/10.1112/blms.12542)
- Seidman, T.I. (1988). How violent are fast controls? *Math. Control Signals Systems* **1** 89–95. [MR0923278 https://doi.org/10.1007/BF02551238](https://doi.org/10.1007/BF02551238)
- Song, Y. (2020). Gradient estimates and exponential ergodicity for mean-field SDEs with jumps. *J. Theoret. Probab.* **33** 201–238. [MR4064299 https://doi.org/10.1007/s10959-018-0845-x](https://doi.org/10.1007/s10959-018-0845-x)
- Tahmasebi, M. (2022). The Bismut-Elworthy-Li formula for semi-linear distribution-dependent SDEs driven by fractional Brownian motion. Available at [arXiv:2209.05586](https://arxiv.org/abs/2209.05586).
- Wang, F.-Y. (1997). Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields* **109** 417–424. [MR1481127 https://doi.org/10.1007/s004400050137](https://doi.org/10.1007/s004400050137)
- Wang, F.-Y. (2010). Harnack inequalities on manifolds with boundary and applications. *J. Math. Pures Appl. (9)* **94** 304–321. [MR2679029 https://doi.org/10.1016/j.matpur.2010.03.001](https://doi.org/10.1016/j.matpur.2010.03.001)
- Wang, F.-Y. (2013). *Harnack Inequalities for Stochastic Partial Differential Equations. SpringerBriefs in Mathematics*. New York: Springer. [MR3099948 https://doi.org/10.1007/978-1-4614-7934-5](https://doi.org/10.1007/978-1-4614-7934-5)
- Wang, F.-Y. (2017). Hypercontractivity and applications for stochastic Hamiltonian systems. *J. Funct. Anal.* **272** 5360–5383. [MR3639531 https://doi.org/10.1016/j.jfa.2017.03.015](https://doi.org/10.1016/j.jfa.2017.03.015)
- Wang, F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.* **128** 595–621. [MR3739509 https://doi.org/10.1016/j.spa.2017.05.006](https://doi.org/10.1016/j.spa.2017.05.006)
- Wang, F.-Y. (2023). Derivative formula for singular McKean-Vlasov SDEs. *Commun. Pure Appl. Anal.* **22** 1866–1898. [MR4599393 https://doi.org/10.3934/cpaa.2023050](https://doi.org/10.3934/cpaa.2023050)
- Wang, F.-Y. and Zhang, X.-C. (2013). Derivative formula and applications for degenerate diffusion semigroups. *J. Math. Pures Appl. (9)* **99** 726–740. [MR3055216 https://doi.org/10.1016/j.matpur.2012.10.007](https://doi.org/10.1016/j.matpur.2012.10.007)

# Percolation threshold for metric graph loop soup

YINSHAN CHANG<sup>1,a</sup>, HANG DU<sup>2,b</sup> and XINYI LI<sup>3,c</sup>

<sup>1</sup>College of Mathematics, Sichuan University, Chengdu, China, [ychang@scu.edu.cn](mailto:ychang@scu.edu.cn)

<sup>2</sup>School of Mathematical Sciences, Peking University, Beijing, China, [duhang@pku.edu.cn](mailto:duhang@pku.edu.cn)

<sup>3</sup>Beijing International Center for Mathematical Research, Peking University, Beijing, China,

[xinyili@bicmr.pku.edu.cn](mailto:xinyili@bicmr.pku.edu.cn)

In this short note, we show that the critical threshold for the percolation of metric graph loop soup on a large class of transient metric graphs (including quasi-transitive graphs such as  $\mathbb{Z}^d$ ,  $d \geq 3$ ) is  $1/2$ .

*Keywords:* Percolation threshold; loop soup; metric graph

## References

- Camia, F. (2017). Scaling limits, Brownian loops, and conformal fields. In *Advances in Disordered Systems, Random Processes and Some Applications* 205–269. Cambridge: Cambridge Univ. Press. [MR3644279](#)
- Chang, Y. (2017). Supercritical loop percolation on  $\mathbb{Z}^d$  for  $d \geq 3$ . *Stochastic Process. Appl.* **127** 3159–3186. [MR3692311](#) <https://doi.org/10.1016/j.spa.2017.02.003>
- Chang, Y. and Sapozhnikov, A. (2016). Phase transition in loop percolation. *Probab. Theory Related Fields* **164** 979–1025. [MR3477785](#) <https://doi.org/10.1007/s00440-015-0624-x>
- Drewitz, A., Prévost, A. and Rodriguez, P.-F. (2022). Cluster capacity functionals and isomorphism theorems for Gaussian free fields. *Probab. Theory Related Fields* **183** 255–313. [MR4421175](#) <https://doi.org/10.1007/s00440-021-01090-0>
- Duminil-Copin, H. and Tassion, V. (2016). A new proof of the sharpness of the phase transition for Bernoulli percolation and the Ising model. *Comm. Math. Phys.* **343** 725–745. [MR3477351](#) <https://doi.org/10.1007/s00220-015-2480-z>
- Grimmett, G.R. and Marstrand, J.M. (1990). The supercritical phase of percolation is well behaved. *Proc. R. Soc. Lond. Ser. A* **430** 439–457. [MR1068308](#) <https://doi.org/10.1098/rspa.1990.0100>
- Janson, S. (1984). Bounds on the distributions of extremal values of a scanning process. *Stochastic Process. Appl.* **18** 313–328. [MR0770197](#) [https://doi.org/10.1016/0304-4149\(84\)90303-X](https://doi.org/10.1016/0304-4149(84)90303-X)
- Lawler, G.F. and Trujillo Ferreras, J.A. (2007). Random walk loop soup. *Trans. Amer. Math. Soc.* **359** 767–787. [MR2255196](#) <https://doi.org/10.1090/S0002-9947-06-03916-X>
- Lawler, G.F. and Werner, W. (2004). The Brownian loop soup. *Probab. Theory Related Fields* **128** 565–588. [MR2045953](#) <https://doi.org/10.1007/s00440-003-0319-6>
- Le Jan, Y. (2011). *Markov Paths, Loops and Fields. Lecture Notes in Math.* **2026**. Heidelberg: Springer. [MR2815763](#) <https://doi.org/10.1007/978-3-642-21216-1>
- Le Jan, Y. and Lemaire, S. (2013). Markovian loop clusters on graphs. *Illinois J. Math.* **57** 525–558. [MR3263044](#)
- Lupu, T. (2016a). Loop percolation on discrete half-plane. *Electron. Commun. Probab.* **21** Paper No. 30, 9. [MR3485399](#) <https://doi.org/10.1214/16-ECP4571>
- Lupu, T. (2016b). From loop clusters and random interacements to the free field. *Ann. Probab.* **44** 2117–2146. [MR3502602](#) <https://doi.org/10.1214/15-AOP1019>
- Prévost, A. (2023). Percolation for the Gaussian free field on the cable system: Counterexamples. *Electron. J. Probab.* **28** Paper No. 62, 43. [MR4583069](#) <https://doi.org/10.1214/23-ejp949>
- Rozanov, Yu.A. (1982). *Markov Random Fields. Applications of Mathematics.* New York-Berlin: Springer. [MR0676644](#)
- Sheffield, S. and Werner, W. (2012). Conformal loop ensembles: The Markovian characterization and the loop-soup construction. *Ann. of Math. (2)* **176** 1827–1917. [MR2979861](#) <https://doi.org/10.4007/annals.2012.176.3.8>

Symanzik, K. (1967). Euclidean quantum field theory. In *Local Quantum Theory*, (R. Jost, ed.) 152–226. New York: Acad. Press.

# Inverse regression for spatially distributed functional data

SUNEEL BABU CHATLA<sup>1,a</sup> and RUIQI LIU<sup>2,b</sup>

<sup>1</sup>*Department of Mathematical Sciences, University of Texas at El Paso, TX 79968, USA,* [asbchatla@utep.edu](mailto:asbchatla@utep.edu)

<sup>2</sup>*Department of Mathematics and Statistics, Texas Tech University, TX 79409, USA,* [bruiqliu@ttu.edu](mailto:bruiqliu@ttu.edu)

Spatially distributed functional data are prevalent in many statistical applications such as meteorology, energy forecasting, census data, disease mapping, and neurological studies. Given their complex and high-dimensional nature, functional data often require dimension reduction methods to extract meaningful information. Inverse regression is one such approach that has become very popular in the past two decades. We study the inverse regression in the framework of functional data observed at irregularly positioned spatial sites. The functional predictor is the sum of a spatially dependent functional effect and a spatially independent functional nugget effect, while the relation between the scalar response and the functional predictor is modeled using the inverse regression framework. For estimation, we consider local linear smoothing with a general weighting scheme, which includes as special cases the schemes under which equal weights are assigned to each observation or to each subject. This framework enables us to present the asymptotic results for different types of sampling plans over time such as non-dense, dense, and ultra-dense. We discuss the domain-expanding infill (DEI) framework for spatial asymptotics, which is a mix of the traditional expanding domain and infill frameworks. The DEI framework overcomes the limitations of traditional spatial asymptotics in the existing literature. Under this unified framework, we develop asymptotic theory and identify conditions that are necessary for the estimated eigen-directions to achieve optimal rates of convergence. Our asymptotic results include pointwise and  $L_2$  convergence rates. Simulation studies using synthetic data and an application to a real-world dataset confirm the effectiveness of our methods.

**Keywords:** Covariance operator; domain-expanding infill asymptotics; irregularly positioned; local linear smoothing; nugget effect; unified framework

## References

- Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*, 2nd ed. *Lecture Notes in Statistics* **110**. New York: Springer. [MR1640691](#) <https://doi.org/10.1007/978-1-4612-1718-3>
- Chatla, S.B. and Liu, R. (2024). Supplement to “Inverse regression for spatially distributed functional data.” <https://doi.org/10.3150/23-BEJ1717SUPP>
- Chen, D., Hall, P. and Müller, H.-G. (2011). Single and multiple index functional regression models with nonparametric link. *Ann. Statist.* **39** 1720–1747. [MR2850218](#) <https://doi.org/10.1214/11-AOS882>
- Chouaf, A. and Laksaci, A. (2012). On the functional local linear estimate for spatial regression. *Stat. Risk Model.* **29** 189–214. [MR2972527](#) <https://doi.org/10.1524/strm.2012.1114>
- Cook, R.D. (2007). Fisher lecture: Dimension reduction in regression. *Statist. Sci.* **22** 1–26. [MR2408655](#) <https://doi.org/10.1214/088342306000000682>
- Cook, R.D. (2018). *An Introduction to Envelopes: Dimension Reduction for Efficient Estimation in Multivariate Statistics*: John Wiley & Sons. [MR3774758](#) <https://doi.org/10.1146/annurev-statistics-031017-100257>
- Cook, R.D., Forzani, L. and Yao, A.F. (2010). Necessary and sufficient conditions for consistency of a method for smoothed functional inverse regression. *Statist. Sinica* **20** 235–238. [MR2640692](#)
- Cook, R.D. and Li, B. (2002). Dimension reduction for conditional mean in regression. *Ann. Statist.* **30** 455–474. [MR1902895](#) <https://doi.org/10.1214/aos/1021379861>
- Cook, R.D. and Weisberg, S. (1991). Sliced inverse regression for dimension reduction: Comment. *J. Amer. Statist. Assoc.* **86** 328–332.

- Cressie, N.A.C. (2015). *Statistics for Spatial Data*, Revised ed. *Wiley Classics Library*. New York: Wiley. Paperback edition of the 1993 edition [MR1239641]. [MR3559472](#)
- Dalenius, T., Hájek, J. and Zubrzycki, S. (1960). On plane sampling and related geometrical problems. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob.* **1** 125–150. Berkeley-Los Angeles, Calif.: Univ. California Press. [MR0133943](#)
- Delicado, P., Giraldo, R., Comas, C. and Mateu, J. (2010). Statistics for spatial functional data: Some recent contributions. *Environmetrics* **21** 224–239. [MR2842240](#) <https://doi.org/10.1002/env.1003>
- Duan, N. and Li, K.-C. (1991). Slicing regression: A link-free regression method. *Ann. Statist.* **19** 505–530. [MR1105834](#) <https://doi.org/10.1214/aos/1176348109>
- Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications. Monographs on Statistics and Applied Probability* **66**. London: CRC Press. [MR1383587](#)
- Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods. Springer Series in Statistics*. New York: Springer. [MR1964455](#) <https://doi.org/10.1007/b97702>
- Ferré, L. and Yao, A.F. (2003). Functional sliced inverse regression analysis. *Statistics* **37** 475–488. [MR2022235](#) <https://doi.org/10.1080/0233188031000112845>
- Ferré, L. and Yao, A.-F. (2005). Smoothed functional inverse regression. *Statist. Sinica* **15** 665–683. [MR2233905](#)
- Ferré, L. and Yao, A.F. (2007). Reply to the paper by Liliana Forzani and R. Dennis Cook: “A note on smoothed functional inverse regression”. *Statist. Sinica* **17** 1683–1687. [MR2413540](#)
- Forzani, L. and Cook, R.D. (2007). A note on smoothed functional inverse regression. *Statist. Sinica* **17** 1677–1681. [MR2413539](#)
- Gao, J., Lu, Z. and Tjøstheim, D. (2006). Estimation in semiparametric spatial regression. *Ann. Statist.* **34** 1395–1435. [MR2278362](#) <https://doi.org/10.1214/009053606000000317>
- Giraldo, R., Delicado, P. and Mateu, J. (2010). Continuous time-varying Kriging for spatial prediction of functional data: An environmental application. *J. Agric. Biol. Environ. Stat.* **15** 66–82. [MR2755385](#) <https://doi.org/10.1007/s13253-009-0012-z>
- Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. [MR2332269](#) <https://doi.org/10.1214/009053606000000957>
- 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>
- 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>
- Hörmann, S. and Kokoszka, P. (2013). Consistency of the mean and the principal components of spatially distributed functional data. *Bernoulli* **19** 1535–1558. [MR3129024](#) <https://doi.org/10.3150/12-BEJ418>
- Hsing, T. and Ren, H. (2009). An RKHS formulation of the inverse regression dimension-reduction problem. *Ann. Statist.* **37** 726–755. [MR2502649](#) <https://doi.org/10.1214/07-AOS589>
- Jiang, C.-R. and Wang, J.-L. (2015). Correction to “Inverse regression for longitudinal data” [MR3210979]. *Ann. Statist.* **43** 2326–2329. [MR3396987](#) <https://doi.org/10.1214/15-AOS1326>
- Jiang, C.-R., Yu, W. and Wang, J.-L. (2014). Inverse regression for longitudinal data. *Ann. Statist.* **42** 563–591. [MR3210979](#) <https://doi.org/10.1214/13-AOS1193>
- Kim, K., Li, B., Yu, Z. and Li, L. (2020). On post dimension reduction statistical inference. *Ann. Statist.* **48** 1567–1592. [MR4124335](#) <https://doi.org/10.1214/19-AOS1859>
- Kurusu, D. (2022a). Nonparametric regression for locally stationary random fields under stochastic sampling design. *Bernoulli* **28** 1250–1275. [MR4388937](#) <https://doi.org/10.3150/21-bej1385>
- Kurusu, D. (2022b). Nonparametric regression for locally stationary functional time series. *Electron. J. Stat.* **16** 3973–3995. [MR4456781](#) <https://doi.org/10.1214/22-ejs2041>
- Lahiri, S.N. (1996). On inconsistency of estimators based on spatial data under infill asymptotics. *Sankhyā Ser. A* **58** 403–417. [MR1659130](#)
- Laksaci, A., Rachdi, M. and Rahmani, S. (2013). Spatial modelization: Local linear estimation of the conditional distribution for functional data. *Spat. Stat.* **6** 1–23.
- Lee, K.-Y. and Li, L. (2022). Functional sufficient dimension reduction through average Fréchet derivatives. *Ann. Statist.* **50** 904–929. [MR4404923](#) <https://doi.org/10.1214/21-aos2131>
- Li, K.-C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86** 316–327. [MR1137117](#)



- Li, B. (2018). *Sufficient Dimension Reduction: Methods and Applications with R. Monographs on Statistics and Applied Probability* **161**. Boca Raton, FL: CRC Press. [MR3838449](#) <https://doi.org/10.1201/9781315119427>
- Li, Y. and Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *Ann. Statist.* **38** 3321–3351. [MR2766854](#) <https://doi.org/10.1214/10-AOS813>
- Li, B. and Song, J. (2017). Nonlinear sufficient dimension reduction for functional data. *Ann. Statist.* **45** 1059–1095. [MR3662448](#) <https://doi.org/10.1214/16-AOS1475>
- Li, B. and Song, J. (2022). Dimension reduction for functional data based on weak conditional moments. *Ann. Statist.* **50** 107–128. [MR4382010](#) <https://doi.org/10.1214/21-aos2091>
- Liu, C., Ray, S. and Hooker, G. (2017). Functional principal component analysis of spatially correlated data. *Stat. Comput.* **27** 1639–1654. [MR3687330](#) <https://doi.org/10.1007/s11222-016-9708-4>
- 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>
- Masry, E. (1986). Recursive probability density estimation for weakly dependent stationary processes. *IEEE Trans. Inf. Theory* **32** 254–267. [MR0838413](#) <https://doi.org/10.1109/TIT.1986.1057163>
- Matérn, B. (1986). *Spatial Variation*, 2nd ed. *Lecture Notes in Statistics* **36**. Berlin: Springer. With a Swedish summary. [MR0867886](#) <https://doi.org/10.1007/978-1-4615-7892-5>
- Müller, H.-G. and Stadtmüller, U. (2005). Generalized functional linear models. *Ann. Statist.* **33** 774–805. [MR2163159](#) <https://doi.org/10.1214/009053604000001156>
- Novak, E. (1988). *Deterministic and Stochastic Error Bounds in Numerical Analysis. Lecture Notes in Math.* **1349**. Berlin: Springer. [MR0971255](#) <https://doi.org/10.1007/BFb0079792>
- Quenouille, M.H. (1949). Approximate tests of correlation in time-series. *J. Roy. Statist. Soc. Ser. B* **11** 68–84. [MR0032176](#)
- Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. [MR2168993](#)
- Song, J. (2019). On sufficient dimension reduction for functional data: Inverse moment-based methods. *Wiley Interdiscip. Rev.: Comput. Stat.* **11** e1459, 13. [MR3999524](#) <https://doi.org/10.1002/wics.1459>
- Tran, L.T. (1990). Kernel density estimation on random fields. *J. Multivariate Anal.* **34** 37–53. [MR1062546](#) [https://doi.org/10.1016/0047-259X\(90\)90059-Q](https://doi.org/10.1016/0047-259X(90)90059-Q)
- Traub, J.F. and Werschulz, A.G. (1998). *Complexity and Information. Lezioni Lincee. [Lincei Lectures]*. Cambridge: Cambridge Univ. Press. [MR1692462](#) <https://doi.org/10.1080/16073606.1997.9631861>
- Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. [MR2160561](#) <https://doi.org/10.1198/016214504000001745>
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. [MR2054303](#) <https://doi.org/10.1198/016214504000000241>
- Zhang, H. and Li, Y. (2022). Unified principal component analysis for sparse and dense functional data under spatial dependency. *J. Bus. Econom. Statist.* **40** 1523–1537. [MR4492051](#) <https://doi.org/10.1080/07350015.2021.1938085>
- Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321. [MR3546451](#) <https://doi.org/10.1214/16-AOS1446>
- Zhang, X. and Wang, J.-L. (2018). Optimal weighting schemes for longitudinal and functional data. *Statist. Probab. Lett.* **138** 165–170. [MR3788733](#) <https://doi.org/10.1016/j.spl.2018.03.007>

# On the Bahadur representation of sample quantiles for score functionals

JOHANNES KREBS<sup>a</sup>

*KU Eichstätt-Ingolstadt, Ostenstraße 28, 85072 Eichstätt, Germany, <sup>a</sup>[johannes.krebs@ku.de](mailto:johannes.krebs@ku.de)*

We establish the Bahadur representation of sample quantiles for stabilizing score functionals in stochastic geometry and study local fluctuations of the corresponding empirical distribution function. The scores are obtained from a Poisson process. We apply the results to trimmed and Winsorized means of the score functionals and establish a law of the iterated logarithm for the sample quantiles of the scores.

*Keywords:* Bahadur representation; law of the iterated logarithm; Poisson process; stochastic geometry; strong stabilization

## 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>
- [2] Barbour, A.D. and Xia, A. (2006). Normal approximation for random sums. *Adv. in Appl. Probab.* **38** 693–728. [MR2256874](#) <https://doi.org/10.1239/aap/1158684998>
- [3] Baryshnikov, Y. and Yukich, J.E. (2005). Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.* **15** 213–253. [MR2115042](#) <https://doi.org/10.1214/10505160400000594>
- [4] Bickel, P.J. and Breiman, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. [MR0682809](#)
- [5] Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Stat.* **42** 1656–1670. [MR0383482](#) <https://doi.org/10.1214/aoms/1177693164>
- [6] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. [MR0233396](#)
- [7] Błaszczyszyn, B., Yogeshwaran, D. and Yukich, J.E. (2019). Limit theory for geometric statistics of point processes having fast decay of correlations. *Ann. Probab.* **47** 835–895. [MR3916936](#) <https://doi.org/10.1214/18-AOP1273>
- [8] Brakke, K.A. (1987). Statistics of random plane Voronoi tessellations. Department of Mathematical Sciences, Susquehanna University (Manuscript 1987a).
- [9] Chatterjee, S. (2008). A new method of normal approximation. *Ann. Probab.* **36** 1584–1610. [MR2435859](#) <https://doi.org/10.1214/07-AOP370>
- [10] Chatterjee, S. and Sen, S. (2017). Minimal spanning trees and Stein’s method. *Ann. Appl. Probab.* **27** 1588–1645. [MR3678480](#) <https://doi.org/10.1214/16-AAP1239>
- [11] Deheuvels, P. (1997). Strong laws for local quantile processes. *Ann. Probab.* **25** 2007–2054. [MR1487444](#) <https://doi.org/10.1214/aop/1023481119>
- [12] Devroye, L., Györfi, L., Lugosi, G. and Walk, H. (2017). On the measure of Voronoi cells. *J. Appl. Probab.* **54** 394–408. [MR3668473](#) <https://doi.org/10.1017/jpr.2017.7>
- [13] Dutta, K. and Sen, P.K. (1971). On the Bahadur representation of sample quantiles in some stationary multivariate autoregressive processes. *J. Multivariate Anal.* **1** 186–198. [MR0301881](#) [https://doi.org/10.1016/0047-259X\(71\)90010-8](https://doi.org/10.1016/0047-259X(71)90010-8)
- [14] Einmahl, J.H.J. (1996). A short and elementary proof of the main Bahadur-Kiefer theorem. *Ann. Probab.* **24** 526–531. [MR1387649](#) <https://doi.org/10.1214/aop/1042644730>
- [15] Flimm, D., Pawlas, Z. and Yukich, J.E. (2020). Limit theory for unbiased and consistent estimators of statistics of random tessellations. *J. Appl. Probab.* **57** 679–702. [MR4125471](#) <https://doi.org/10.1017/jpr.2020.4>

- [16] Hayen, A. and Quine, M.P. (2002). Areas of components of a Voronoi polygon in a homogeneous Poisson process in the plane. *Adv. in Appl. Probab.* **34** 281–291. MR1909915 <https://doi.org/10.1239/aap/1025131218>
- [17] Hesse, C.H. (1990). A Bahadur-type representation for empirical quantiles of a large class of stationary, possibly infinite-variance, linear processes. *Ann. Statist.* **18** 1188–1202. MR1062705 <https://doi.org/10.1214/aos/1176347746>
- [18] Ho, H.-C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24** 992–1024. MR1401834 <https://doi.org/10.1214/aos/1032526953>
- [19] Hug, D., Reitzner, M. and Schneider, R. (2004). Large Poisson-Voronoi cells and Crofton cells. *Adv. in Appl. Probab.* **36** 667–690. MR2079908 <https://doi.org/10.1239/aap/1093962228>
- [20] Kiefer, J. (1967). On Bahadur's representation of sample quantiles. *Ann. Math. Stat.* **38** 1323–1342. MR0217844 <https://doi.org/10.1214/aoms/1177698690>
- [21] Kiefer, J. (1970). Deviations between the sample quantile process and the sample  $df$ . In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)* (M. Puri, ed.). 299–319. London: Cambridge Univ. Press. MR0277071
- [22] Kiefer, J. (1970). Old and new methods for studying order statistics and sample quantiles. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969)* (M. Puri, ed.). 349–357. London: Cambridge Univ. Press. MR0286228
- [23] Krebs, J. (2021). On the law of the iterated logarithm and strong invariance principles in stochastic geometry. *Bernoulli* **27** 1695–1723. MR4260506 <https://doi.org/10.3150/20-bej1288>
- [24] Krebs, J. (2024). Supplement to “On the Bahadur representation of sample quantiles for score functionals.” <https://doi.org/10.3150/24-BEJ1718SUPP>
- [25] Kulik, R. (2007). Bahadur-Kiefer theory for sample quantiles of weakly dependent linear processes. *Bernoulli* **13** 1071–1090. MR2364227 <https://doi.org/10.3150/07-BEJ6086>
- [26] Lachièze-Rey, R. and Peccati, G. (2017). New Berry-Esseen bounds for functionals of binomial point processes. *Ann. Appl. Probab.* **27** 1992–2031. MR3693518 <https://doi.org/10.1214/16-AAP1218>
- [27] Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 <https://doi.org/10.1214/18-AAP1405>
- [28] McGivney, K. and Yukich, J.E. (1999). Asymptotics for Voronoi tessellations on random samples. *Stochastic Process. Appl.* **83** 273–288. MR1708209 [https://doi.org/10.1016/S0304-4149\(99\)00035-6](https://doi.org/10.1016/S0304-4149(99)00035-6)
- [29] Miles, R.E. and Maillardet, R.J. (1982). The basic structures of Voronoi and generalized Voronoi polygons. *J. Appl. Probab.* **19** 97–111. Essays in statistical science. MR0633183
- [30] Penrose, M.D. (2007). Laws of large numbers in stochastic geometry with statistical applications. *Bernoulli* **13** 1124–1150. MR2364229 <https://doi.org/10.3150/07-BEJ5167>
- [31] Penrose, M.D. and Yukich, J.E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Probab.* **11** 1005–1041. MR1878288 <https://doi.org/10.1214/aoap/1015345393>
- [32] Penrose, M.D. and Yukich, J.E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Probab.* **13** 277–303. MR1952000 <https://doi.org/10.1214/aoap/1042765669>
- [33] Polonik, W. (1997). Minimum volume sets and generalized quantile processes. *Stochastic Process. Appl.* **69** 1–24. MR1464172 [https://doi.org/10.1016/S0304-4149\(97\)00028-8](https://doi.org/10.1016/S0304-4149(97)00028-8)
- [34] Reitzner, M. and Schulte, M. (2013). Central limit theorems for  $U$ -statistics of Poisson point processes. *Ann. Probab.* **41** 3879–3909. MR3161465 <https://doi.org/10.1214/12-AOP817>
- [35] Sen, P.K. (1968). Asymptotic normality of sample quantiles for  $m$ -dependent processes. *Ann. Math. Stat.* **39** 1724–1730. MR0232522 <https://doi.org/10.1214/aoms/1177698155>
- [36] Sen, P.K. (1972). On the Bahadur representation of sample quantiles for sequences of  $\phi$ -mixing random variables. *J. Multivariate Anal.* **2** 77–95. MR0303581 [https://doi.org/10.1016/0047-259X\(72\)90011-5](https://doi.org/10.1016/0047-259X(72)90011-5)
- [37] Shang, Z. (2010). Convergence rate and Bahadur type representation of general smoothing spline M-estimates. *Electron. J. Stat.* **4** 1411–1442. MR2741207 <https://doi.org/10.1214/10-EJS588>
- [38] Steele, J.M. (1988). Growth rates of Euclidean minimal spanning trees with power weighted edges. *Ann. Probab.* **16** 1767–1787. MR0958215
- [39] Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.* **33** 1934–1963. MR2166566 <https://doi.org/10.1214/009053605000000291>
- [40] 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>

- [41] Yukich, J.E. (2000). Asymptotics for weighted minimal spanning trees on random points. *Stochastic Process. Appl.* **85** 123–138. [MR1730615](#) [https://doi.org/10.1016/S0304-4149\(99\)00068-X](https://doi.org/10.1016/S0304-4149(99)00068-X)
- [42] Zuyev, S.A. (1992). Estimates for distributions of the Voronoï polygon's geometric characteristics. *Random Structures Algorithms* **3** 149–162. [MR1151358](#) <https://doi.org/10.1002/rsa.3240030205>

# Erratum: Tree builder random walk: Recurrence, transience and ballisticity

GIULIO IACOBELLI<sup>1,a</sup>, RODRIGO RIBEIRO<sup>2,c</sup>, GLAUCO VALLE<sup>1,b</sup> and  
LEONEL ZUAZNÁBAR<sup>3,d</sup>

<sup>1</sup>*UFRJ - Instituto de Matemática. Caixa Postal 68530, 21945-970, Rio de Janeiro, Brasil, <sup>a</sup>[giulio@im.ufrj.br](mailto:giulio@im.ufrj.br),  
<sup>b</sup>[glauco.valle@im.ufrj.br](mailto:glauco.valle@im.ufrj.br)*

<sup>2</sup>*PUC Chile - Facultad de Matemáticas. Av. Vicuña Mackenna 4860, Macul, La Florida, Región Metropolitana,  
Chile, <sup>c</sup>[rribeiro@impa.br](mailto:rribeiro@impa.br)*

<sup>3</sup>*USP - Instituto de Matemática e Estatística. R. do Matão, 1010 - Butantã, São Paulo - SP, CEP: 05508-090,  
Brasil, <sup>d</sup>[lzuaznabar@ime.usp.br](mailto:lzuaznabar@ime.usp.br)*

## References

- Figueiredo, D., Iacobelli, G., Oliveira, R., Reed, B. and Ribeiro, R. (2021). On a random walk that grows its own tree. *Electron. J. Probab.* **26** Paper No. 6. MR4216519 <https://doi.org/10.1214/20-ejp574>
- Iacobelli, G., Ribeiro, R., Valle, G. and Zuaznabar, L. (2022). Tree builder random walk: Recurrence, transience and ballisticity. *Bernoulli* **28** 150–180. MR4337701 <https://doi.org/10.3150/21-bej1337>
- Palacios, J.L. (2009). On the moments of hitting times for random walks on trees. *J. Probab. Stat.* Art. ID 241539. MR2602886 <https://doi.org/10.1155/2009/241539>
- Palacios, J.L., Gómez, E. and Río, M.D. (2014). Hitting times of walks on graphs through voltages. *J. Probab. Stat.* **2014**.