BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Two Number Two May 2026 ISSN: 1350-7265

CONTENTS

(continued)

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

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Two Number Two May 2026 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 32 Number 2 May 2026 Pages 849–1685 ISI/BS

Volume 32 Number 2   May 2026   ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

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

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

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

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

---

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

In 2026 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **32**(2), 2026, 849–873
https://doi.org/10.3150/25-BEJ1882

Functional estimation in high-dimensional and infinite-dimensional models

VLADIMIR KOLTCHINSKII[a] and MINGHAO LI[b]

School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA, [a]*vlad@math.gatech.edu,* [b]*minghaoli@gatech.edu*

Let $\mathcal{P}$ be a family of probability measures on a measurable space $(S, \mathcal{A})$. Given a Banach space E, a functional $f : E \mapsto \mathbb{R}$ and a mapping $\theta : \mathcal{P} \mapsto E$, our goal is to estimate $f(\theta(P))$ based on i.i.d. observations $X_1, \dots, X_n \sim P, P \in \mathcal{P}$. Given a smooth functional f and estimators $\hat{\theta}_n(X_1, \dots, X_n), n \geq 1$ of $\theta(P)$, we use these estimators, the sample split and the Taylor expansion of $f(\theta(P))$ of a proper order to construct estimators $T_f(X_1, \dots, X_n)$ of $f(\theta(P))$. For these estimators and for a functional f of smoothness $s \geq 1$, we derive upper bounds on the L_p-errors of $T_f(X_1, \dots, X_n)$ with optimal dependence on sample size n, on the dimension or other complexity characteristics of parameter and on degree s of smoothness of the functional, and also study asymptotic normality and asymptotic efficiency of these estimators. The examples include functional estimation in high-dimensional models with many low dimensional components, functional estimation in high-dimensional exponential families and estimation of functionals of covariance operators in infinite-dimensional subgaussian models.

Keywords: Asymptotic efficiency; covariance operator; effective rank; exponential family; minimax optimality; smooth functionals

References

[1] Anastasiou, A. and Gaunt, R. (2021). Wasserstein distance error bounds for the multivariate normal approximation of the maximum likelihood estimator. *Electron. J. Stat.* **15** 5758–5810. MR4355697 https://doi.org/10.1214/21-ejs1920

[2] Bentkus, V., Bloznelis, M. and Götze, F. (1997). A Berry-Esseen bound for M-estimators. *Scand. J. Stat.* **24** 485–502. MR1615335 https://doi.org/10.1111/1467-9469.00076

[3] Bickel, P. and Ritov, Y. (1988). Estimating integrated square density derivatives: Sharp best order of convergence estimates. *Sankhyā* **50** 381–393. MR1065550

[4] Bickel, P., Klaasen, C.A.J., Ritov, Y. and Wellner, J. (1998). *Efficient and Adaptive Estimation for Semiparametric Models*. New York: Springer. MR1623559

[5] Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.* **23** 11–29. https://doi.org/10.1214/aos/1176324452

[6] Čencov, N.N. (1982). *Statistical Decision Rules and Optimal Inference*. American Mathematical Society. MR0645898

[7] Collier, O., Comminges, L. and Tsybakov, A. (2017). Minimax estimation of linear and quadratic functionals on sparsity classes. *Ann. Statist.* **45** 923–958. MR3662444 https://doi.org/10.1214/15-AOS1432

[8] de la Peña, V.H. and Giné, E. (1999). *Decoupling. From Dependence to Independence*. Springer. MR1666908 https://doi.org/10.1007/978-1-4612-0537-1

[9] Hall, P. and Martin, M.A. (1988). On bootstrap resampling and iteration. *Biometrika* **75** 661–671. MR0995110 https://doi.org/10.1093/biomet/75.4.661

[10] Ibragimov, I.A. and Khasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. New York: Springer-Verlag. MR0620321 https://doi.org/10.1007/978-1-4899-0027-2

[11] Ibragimov, I.A., Nemirovski, A.S. and Khasminskii, R.Z. (1987). Some problems of nonparametric estimation in Gaussian white noise. *Theory Probab. Appl.* **31** 391–406. MR0866866 https://doi.org/10.1137/1131054

[12] Jiao, J. and Han, Y. (2020). Bias correction with jackknife, bootstrap and Taylor series. *IEEE Trans. Inf. Theory* **66** 4392–4418. MR4130622 https://doi.org/10.1109/TIT.2020.2969439
[13] Koltchinskii, V. (2018). Asymptotic efficiency in high-dimensional covariance estimation. *Proc. ICM* **2018** 2891–2912. MR3966516 https://doi.org/10.1142/11060
[14] Koltchinskii, V. (2021). Asymptotically efficient estimation of smooth functionals of covariance operators. *J. Eur. Math. Soc.* **23** 765–843. MR4210724 https://doi.org/10.4171/jems/1023
[15] Koltchinskii, V. (2022). Estimation of smooth functionals in high-dimensional models: Bootstrap chains and Gaussian approximation. *Ann. Statist.* **50** 2386–2415. MR4474495 https://doi.org/10.1214/22-aos2197
[16] Koltchinskii, V. (2024). Estimation of smooth functionals of covariance operators: Jackknife bias reduction and bounds in terms of effective rank. *Ann. Inst. Henri Poincaré Probab. Stat.* **61** 665–712. MR4863057 https://doi.org/10.1214/23-aihp1433
[17] Koltchinskii, V. and Li, M. (2026). Supplement to "Functional estimation in high-dimensional and infinite-dimensional models." https://doi.org/10.3150/25-BEJ1882SUPP
[18] Koltchinskii, V., Löffler, M. and Nickl, R. (2020). Efficient estimation of linear functionals of principal components. *Ann. Statist.* **48** 464–490. MR4065170 https://doi.org/10.1214/19-AOS1816
[19] Koltchinskii, V. and Wahl, M. (2023). Functional estimation in log-concave location families. In *High Dimensional Probability IX* (R. Adamczak, N. Gozlan, K. Lounici and M. Madiman, eds.). *Progress in Probability* **80** 393–440. Birkhäuser. MR4633271 https://doi.org/10.1007/978-3-031-26979-0_15
[20] Koltchinskii, V. and Zhilova, M. (2021). Efficient estimation of smooth functionals in Gaussian shift models. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 351–386. MR4255178 https://doi.org/10.1214/20-aihp1081
[21] Koltchinskii, V. and Zhilova, M. (2021). Estimation of smooth functionals in normal models: Bias reduction and asymptotic efficiency. *Ann. Statist.* **49** 2577–2610. MR4338376 https://doi.org/10.1214/20-aos2047
[22] Koltchinskii, V. and Zhilova, M. (2021). Estimation of smooth functionals of location parameter in Gaussian and Poincaré random shift models. *Sankhya A* **83** 569–596. MR4284575 https://doi.org/10.1007/s13171-020-00232-1
[23] Laurent, B. (1996). Efficient estimation of integral functionals of a density. *Ann. Statist.* **24** 659–681. MR1394981 https://doi.org/10.1214/aos/1032894458
[24] Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the L_r norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. MR1670867 https://doi.org/10.1007/s004409970006
[25] Levit, B. (1975). On the efficiency of a class of non-parametric estimates. *Theory Probab. Appl.* **20** 723–740. MR0403052 https://doi.org/10.1137/1120081
[26] Levit, B. (1978). Asymptotically efficient estimation of nonlinear functionals. *Probl. Inf. Transm.* **14** 65–72. MR0533450
[27] Nemirovski, A. (1990). On necessary conditions for the efficient estimation of functionals of a nonparametric signal which is observed in white noise. *Theory Probab. Appl.* **35** 94–103. MR1050056 https://doi.org/10.1137/1135009
[28] Nemirovski, A. (2000). *Topics in Non-parametric Statistics. Ecole d'Ete de Probabilités de Saint-Flour. Lecture Notes in Mathematics* **1738**. New York: Springer. MR1775640 https://doi.org/10.1007/BFb0106703
[29] Pfanzagl, J. (1971). The Berry-Esseen bound for minimum contrast estimates. *Metrika* **17** 82–91. MR0295467 https://doi.org/10.1007/BF02613813
[30] Pinelis, I. (2017). Optimal order uniform and non-uniform bounds on the rates of convergence to normality for maximum likelihood estimators. *Electron. J. Stat.* **11** 1160–1179. MR3634332 https://doi.org/10.1214/17-EJS1264
[31] Rio, E. (2009). Upper bounds for minimal distances in the central limit theorem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 802–817. MR2548505 https://doi.org/10.1214/08-AIHP187
[32] Robins, J., Li, L., Tchetgen, E. and van der Vaart, A. (2008). Higher order influence functions and minimax estimation of nonlinear functionals. *IMS Collections Probability and Statistics: Essays in Honor of David A. Freedman* **2** 335–421. MR2459958 https://doi.org/10.1214/193940307000000527
[33] van der Vaart, A. (1998). *Asymptotic Statistics*. Cambridge University Press. MR1652247 https://doi.org/10.1017/CBO9780511802256
[34] Zhou, F., Li, P. and Zhang, C.-H. (2021). High-order statistical functional expansion and its application to some nonsmooth problems. arXiv:2112.15591. https://doi.org/10.48550/arXiv.2112.15591

Bernoulli **32**(2), 2026, 874–900
https://doi.org/10.3150/25-BEJ1883

Bernstein duality revisited: Frequency-dependent selection, coordinated mutation and opposing environments

FERNANDO CORDERO[1,2,a,b], SEBASTIAN HUMMEL[3,c] and GRÉGOIRE VÉCHAMBRE[4,d]

[1] *Institute of Mathematics, Department of Natural Sciences and Sustainable Ressources, BOKU University, Austria,* [a]*fernando.cordero@boku.ac.at*
[2] *Faculty of Technology, Bielefeld University, Germany,* [b]*fcordero@techfak.uni-bielefeld.de*
[3] *Department of Health Science and Technology, ETH Zürich, Zürich, Switzerland,* [c]*shummel@hest.ethz.ch*
[4] *State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China,* [d]*vechambre@amss.ac.cn*

This paper investigates the long-term behavior of a class of Λ-Wright–Fisher processes incorporating frequency-dependent selection, coordinated (bidirectional) selection, as well as individual and coordinated mutation. Our primary analytical tool is Bernstein duality, a generalization of moment duality. We introduce the corresponding dual process and establish the relevant duality relation. Without mutation, this work complements earlier studies that employed moment duality, Siegmund duality or other methods to classify the long-term behavior of similar processes. Notably, the current analysis encompasses parameter regimes that model bidirectional selection, a scenario that has proven challenging to analyze using moment duality. In the presence of mutation, we establish the ergodic properties of the process.

Keywords: Branching-coalescing particle system; coordination; duality; frequency-dependent selection; Λ-Wright–Fisher processes; random environment

References

[1] Baake, E., Esercito, L. and Hummel, S. (2023). Lines of descent in a Moran model with frequency-dependent selection and mutation. *Stochastic Process. Appl.* **160** 409–457. MR4569301 https://doi.org/10.1016/j.spa.2023.03.004
[2] Baake, E. and Wakolbinger, A. (2018). Lines of descent under selection. *J. Stat. Phys.* **172** 156–174. MR3810541 https://doi.org/10.1007/s10955-017-1921-9
[3] Bah, B. and Pardoux, E. (2015). The Λ-lookdown model with selection. *Stochastic Process. Appl.* **125** 1089–1126. MR3303969 https://doi.org/10.1016/j.spa.2014.10.014
[4] Bansaye, V., Caballero, M.-E. and Méléard, S. (2019). Scaling limits of population and evolution processes in random environment. *Electron. J. Probab.* **24** 1–38. MR3925459 https://doi.org/10.1214/19-EJP262
[5] Bleuven, C. and Landry, C.R. (2016). Molecular and cellular bases of adaptation to a changing environment in microorganisms. *Proc. R. Soc. B* **283** 20161458. https://doi.org/10.1098/rspb.2016.1458
[6] Cordero, F., Hummel, S. and Schertzer, E. (2022). General selection models: Bernstein duality and minimal ancestral structures. *Ann. Appl. Probab.* **32** 1499–1556. MR4429994 https://doi.org/10.1214/21-aap1683
[7] Cordero, F., Hummel, S. and Véchambre, G. (2025). Λ-Wright–Fisher processes with general selection and opposing environmental effects: Fixation and coexistence. *Ann. Appl. Probab.* **35** 393–457. MR4871712 https://doi.org/10.1214/24-aap2117
[8] Cordero, F. and Möhle, M. (2019). On the stationary distribution of the block counting process for population models with mutation and selection. *J. Math. Anal. Appl.* **474** 1049–1081. MR3926155 https://doi.org/10.1016/j.jmaa.2019.02.004

[9] Cordero, F. and Véchambre, G. (2023). Moran models and Wright–Fisher diffusions with selection and mutation in a one-sided random environment. *Adv. Appl. Probab.* 1–67. MR4624027 https://doi.org/10.1017/apr.2022.54

[10] Donihue, C.M., Herrel, A., Fabre, A.-C., Kamath, A., Geneva, A.J., Schoener, T.W., Kolbe, J.J. and Losos, J.B. (2018). Hurricane-induced selection on the morphology of an island lizard. *Nature* **560** 88–91. https://doi.org/10.1038/s41586-018-0352-3

[11] Foucart, C. (2013). The impact of selection in the Λ-Wright–Fisher model. *Electron. Commun. Probab.* **18**. 72. MR3101637 https://doi.org/10.1214/ECP.v18-2838

[12] González Casanova, A., Kurt, N. and Tobias, A. (2021). Particle systems with coordination. *ALEA Lat. Amer. J. Probab. Math. Stat.* **18** 1817–1844. MR4332221 https://doi.org/10.30757/alea.v18-68

[13] González Casanova, A. and Spanò, D. (2018). Duality and fixation in Ξ-Wright–Fisher processes with frequency-dependent selection. *Ann. Appl. Probab.* **28** 250–284. MR3770877 https://doi.org/10.1214/17-AAP1305

[14] González Casanova, A., Spanò, D. and Wilke-Berenguer, M. (2025). The effective strength of selection in random environment. *Ann. Appl. Probab.* **35** 701–748. MR4871719 https://doi.org/10.1214/24-aap2127

[15] Griffiths, R.C. (2014). The Λ-Fleming–Viot process and a connection with Wright–Fisher diffusion. *Adv. Appl. Probab.* **46** 1009–1035. MR3290427 https://doi.org/10.1239/aap/1418396241

[16] Kimura, M. (1962). On the probability of fixation of mutant genes in a population. *Genetics* **47** 713–719. https://doi.org/10.1093/genetics/47.6.713

[17] Koskela, J., Łatuszyński, K. and Spanò, D. (2024). Bernoulli factories and duality in Wright–Fisher and Allen–Cahn models of population genetics. *Theor. Popul. Biol.* **156** 40–45. https://doi.org/10.1016/j.tpb.2024.01.002

[18] Krone, S.M. and Neuhauser, C. (1997). Ancestral processes with selection. *Theor. Popul. Biol.* **51** 210–237.

[19] Meyn, S.P. and Tweedie, R.L. (1993). A survey of Foster-Lyapunov techniques for general state space Markov processes. In *Proceedings of the Workshop on Stochastic Stability and Stochastic Stabilization*. Metz, France. MR1287609 https://doi.org/10.1007/978-1-4471-3267-7

[20] Mitchell, A., Romano, G.H., Groisman, B., Yona, A., Dekel, E., Kupiec, M., Dahan, O. and Pilpel, Y. (2009). Adaptive prediction of environmental changes by microorganisms. *Nature* **460** 220–224. https://doi.org/10.1038/nature08112

[21] Neuhauser, C. (1999). The ancestral graph and gene genealogy under frequency-dependent selection. *Theor. Popul. Biol.* **56** 203–214.

[22] Neuhauser, C. and Krone, S.M. (1997). The genealogy of samples in models with selection. *Genetics* **145** 519–534.

[23] Norris, J.R. (1998). *Markov Chains*, 2nd ed. Cambridge: Cambridge Univ. Press. MR1600720

[24] Siegmund, D. (1976). The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* **4** 914–924. MR0431386 https://doi.org/10.1214/aop/1176995936

[25] Taylor, J.E. (2007). The common ancestor process for a Wright-Fisher diffusion. *Electron. J. Probab.* **12** 808–847. MR2318411 https://doi.org/10.1214/EJP.v12-418

[26] Véchambre, G. (2023). Combinatorics of ancestral lines for a Wright–Fisher diffusion with selection in a Lévy environment. *Ann. Appl. Probab.* **33** 4875–4935. MR4674067 https://doi.org/10.1214/23-aap1936

[27] Voss-Bohme, A., Schenk, W. and Koellner, A.K. (2011). On the equivalence between Liggett duality of Markov processes and the duality relation between their generators. *Markov Process. Related Fields* **17** 315–346. MR2895555

Bernoulli **32**(2), 2026, 901–925
https://doi.org/10.3150/25-BEJ1885

Theory and inference for multivariate autoregressive binary models and dynamical modeling of absence-presence data in ecology

GUILLAUME FRANCHI[a] and LIONEL TRUQUET[b]

Univ Rennes, Ensai, CNRS, CREST – UMR 9194, F-35000 Rennes, France, [a]*guillaume.franchi@ensai.fr,* [b]*lionel.truquet@ensai.fr*

We introduce a general class of autoregressive models for studying the dynamic of multivariate binary time series with stationary exogenous covariates. We first show that existence of a stationary path for such models is almost automatic and does not require parameter restrictions when the noise term is not compactly supported. We then study in details statistical inference in a dynamic version of a multivariate probit type model, as a particular case of our general construction. To avoid a complex likelihood optimization, we combine pseudo-likelihood and pairwise likelihood methods for which asymptotic results are obtained for a single path analysis and also for panel data, using ergodic theorems for multi-indexed partial sums. The latter scenario is particularly important for analyzing absence-presence of species in ecology, a field where data are often collected from surveys at various locations. Our results also give a theoretical background for such models which are often used by the practitioners but without a probabilistic framework.

Keywords: Binary time series; ergodic properties of multiple parameters processes; iterated random maps

References

Ashford, J. and Sowden, R. (1970). Multi-variate probit analysis. *Biometrics* **26** 535–546.

Banys, P., Davydov, Y. and Paulauskas, V. (2010). Remarks on the SLLN for linear random fields. *Statist. Probab. Lett.* **80** 489–496. MR2593590 https://doi.org/10.1016/j.spl.2009.11.026

Candelon, B., Dumitrescu, E.-I., Hurlin, C. and Palm, F.C. (2013). Multivariate dynamic probit models: An application to financial crises mutation. In *VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims* 395–427. Limited: Emerald Group Publishing. MR3496842 https://doi.org/10.1108/S0731-905320130000031011

Chaubert, F., Mortier, F. and Saint André, L. (2008). Multivariate dynamic model for ordinal outcomes. *J. Multivariate Anal.* **99** 1717–1732. MR2526172 https://doi.org/10.1016/j.jmva.2008.01.011

Chib, S. and Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika* **85** 347–361.

De Jong, R.M. and Woutersen, T. (2011). Dynamic time series binary choice. *Econometric Theory* **27** 673–702. MR2822362 https://doi.org/10.1017/S0266466610000472

Debaly, Z.M., Neumann, M.H. and Truquet, L. (2025). Mixing properties of nonstationary multivariate count processes. *J. Time Series Anal.* **46** 552–581. MR4888809 https://doi.org/10.1111/jtsa.12775

Debaly, Z.M. and Truquet, L. (2021). Iterations of dependent random maps and exogeneity in nonlinear dynamics. *Econometric Theory* **37** 1135–1172. MR4348399 https://doi.org/10.1017/S0266466620000559

Fernández-Val, I. and Weidner, M. (2016). Individual and time effects in nonlinear panel models with large N, T. *J. Econometrics* **192** 291–312. MR3463676 https://doi.org/10.1016/j.jeconom.2015.12.014

Fokianos, K. and Truquet, L. (2019). On categorical time series models with covariates. *Stochastic Process. Appl.* **129** 3446–3462. MR3985569 https://doi.org/10.1016/j.spa.2018.09.012

Franchi, G. and Truquet, L. (2026). Supplement to "Theory and inference for multivariate autoregressive binary models and dynamical modeling of absence-presence data in ecology." https://doi.org/10.3150/25-BEJ1885SUPP

Giap, D.X. and Van Quang, N. (2016). Multidimensional and multivalued ergodic theorems for measure-preserving transformations. *Set-Valued Var. Anal.* **24** 637–658. MR3570348 https://doi.org/10.1007/s11228-016-0361-z

Greene, W. (2009). Discrete choice modeling. In *In Palgrave Handbook of Econometrics*. *Applied Econometrics* **2** 473–556. Springer.

Guanche, Y., Minguez, R. and Méndez, F.J. (2014). Autoregressive logistic regression applied to atmospheric circulation patterns. *Clim. Dyn.* **42** 537–552.

Hahn, J. and Kuersteiner, G. (2011). Bias reduction for dynamic nonlinear panel models with fixed effects. *Econometric Theory* **27** 1152–1191. MR2868837 https://doi.org/10.1017/S0266466611000028

Hajivassiliou, V.A. and Ruud, P.A. (1994). Classical estimation methods for LDV models using simulation. *Handb. Econom.* **4** 2383–2441. MR1315975

Hsiao, C. (2022). *Analysis of Panel Data 64*. Cambridge University Press. MR4615778 https://doi.org/10.1017/9781009057745

Jentsch, C. and Reichmann, L. (2022). Generalized binary vector autoregressive processes. *J. Time Series Anal.* **43** 285–311. MR4400295 https://doi.org/10.1111/jtsa.12614

Kenne Pagui, E.C. and Canale, A. (2016). Pairwise likelihood inference for multivariate ordinal responses with applications to customer satisfaction. *Appl. Stoch. Models Bus. Ind.* **32** 273–282. MR3488020 https://doi.org/10.1002/asmb.2147

Klesov, O. and Molchanov, I. (2017). Moment conditions in strong laws of large numbers for multiple sums and random measures. *Statist. Probab. Lett.* **131** 56–63. MR3706696 https://doi.org/10.1016/j.spl.2017.08.007

Krengel, U. (2011). Ergodic Theorems 6. Walter de Gruyter. MR0797411 https://doi.org/10.1515/9783110844641

Liesenfeld, R. and Richard, J.-F. (2010). Efficient estimation of probit models with correlated errors. *J. Econometrics* **156** 367–376. MR2609939 https://doi.org/10.1016/j.jeconom.2009.11.006

Manner, H., Türk, D. and Eichler, M. (2016). Modeling and forecasting multivariate electricity price spikes. *Energy Econ.* **60** 255–265.

Manski, C.F. (1975). Maximum score estimation of the stochastic utility model of choice. *J. Econometrics* **3** 205–228. MR0436905 https://doi.org/10.1016/0304-4076(75)90032-9

Masarotto, G. and Varin, C. (2012). Gaussian copula marginal regression. *Electron. J. Stat.* **6** 1517–1549. MR2988457 https://doi.org/10.1214/12-EJS721

Müller, G. and Czado, C. (2005). An autoregressive ordered probit model with application to high-frequency financial data. *J. Comput. Graph. Statist.* **14** 320–338. MR2160816 https://doi.org/10.1198/106186005X48687

Nguyen, X.X. and Zessin, H. (1979). Ergodic theorems for spatial processes. *Z. Wahrsch. Verw. Gebiete* **48** 133–158. MR0534841 https://doi.org/10.1007/BF01886869

Park, J.Y. and Phillips, P.C. (2000). Nonstationary binary choice. *Econometrica* **68** 1249–1280. MR1779149 https://doi.org/10.1111/1468-0262.00157

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J.S. and Thuiller, W. (2021). On the interpretations of joint modeling in community ecology. *Trends Ecol. Evol.* **36** 391–401.

Propp, J.G. and Wilson, D.B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9** 223–252. MR1611693 https://doi.org/10.1002/(sici)1098-2418(199608/09)9:1/2<223::aid-rsa14>3.0.co;2-o

Samorodnitsky, G. et al. (2016). *Stochastic Processes and Long Range Dependence 26*. Springer. MR3561100 https://doi.org/10.1007/978-3-319-45575-4

Sebastian-Gonzalez, E., Sanchez-Zapata, J.A., Botella, F. and Ovaskainen, O. (2010). Testing the heterospecific attraction hypothesis with time-series data on species co-occurrence. *Proc. R. Soc. Lond., B Biol. Sci.* **277** 2983–2990.

Smythe, R.T. (1973). Strong laws of large numbers for r-dimensional arrays of random variables. *Ann. Probab.* **1** 164–170. MR0346881 https://doi.org/10.1214/aop/1176997031

Talhouk, A., Doucet, A. and Murphy, K. (2012). Efficient Bayesian inference for multivariate probit models with sparse inverse correlation matrices. *J. Comput. Graph. Statist.* **21** 739–757. MR2970917 https://doi.org/10.1080/10618600.2012.679239

Ting, B., Wright, F. and Zhou, Y.-H. (2022). Fast multivariate probit estimation via a two-stage composite likelihood. *Stat. Biosci.* **14** 533–549.

Truquet, L. (2020). Coupling and perturbation techniques for categorical time series. *Bernoulli* **26** 3249–3279. MR4140544 https://doi.org/10.3150/20-BEJ1225

Truquet, L. (2023). Strong mixing properties of discrete-valued time series with exogenous covariates. *Stochastic Process. Appl.* **160** 294–317. MR4567527 https://doi.org/10.1016/j.spa.2023.03.006

Tuzcuoglu, K. (2023). Composite likelihood estimation of an autoregressive panel ordered probit model with random effects. *J. Bus. Econom. Statist.* **41** 593–607. MR4568045 https://doi.org/10.1080/07350015.2022.2044829

Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R. and McCarthy, M.A. (2019). A comparison of joint species distribution models for presence–absence data. *Methods Ecol. Evol.* **10** 198–211.

Young, G., Valdez, E.A. and Kohn, R. (2009). Multivariate probit models for conditional claim-types. *Insurance Math. Econom.* **44** 214–228. MR2517886 https://doi.org/10.1016/j.insmatheco.2008.11.004

Zhao, Y. and Joe, H. (2005). Composite likelihood estimation in multivariate data analysis. *Canad. J. Statist.* **33** 335–356. MR2193979 https://doi.org/10.1002/cjs.5540330303

Bernoulli **32**(2), 2026, 926–951
https://doi.org/10.3150/25-BEJ1886

Volatility and jump activity estimation in a stable Cox-Ingersoll-Ross model

ELISE BAYRAKTAR[a] and EMMANUELLE CLÉMENT[b]

LAMA UMR8050, Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, F-77447 Marne-la-Vallée, France, [a]*elise.bayraktar@univ-eiffel.fr,* [b]*emmanuelle.clement@univ-eiffel.fr*

We consider the parametric estimation of the volatility and jump activity in a stable Cox-Ingersoll-Ross (α-stable CIR) model driven by a standard Brownian Motion and a non-symmetric stable Lévy process with jump activity $\alpha \in (1, 2)$. The main difficulties to obtain rate efficiency in estimating these quantities arise from the superposition of the diffusion component with jumps of infinite variation. Extending the approach proposed in Mies (*Electron. J. Stat.* **14** (2020) 4165–4206), we address the joint estimation of the volatility, scaling and jump activity parameters from high-frequency observations of the process and prove that the proposed estimators are rate optimal up to a logarithmic factor.

Keywords: Cox-Ingersoll-Ross model; Lévy process; parametric inference; stable process; stochastic differential equation

References

Aït-Sahalia, Y. and Jacod, J. (2009). Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.* **37** 2202–2244. MR2543690 https://doi.org/10.1214/08-AOS640

Aït-Sahalia, Y. and Jacod, J. (2012). Identifying the successive Blumenthal-Getoor indices of a discretely observed process. *Ann. Statist.* **40** 1430–1464. MR3015031 https://doi.org/10.1214/12-AOS976

Amorino, C. and Gloter, A. (2020a). Unbiased truncated quadratic variation for volatility estimation in jump diffusion processes. *Stochastic Process. Appl.* **130** 5888–5939. MR4140022 https://doi.org/10.1016/j.spa.2020.04.010

Amorino, C. and Gloter, A. (2020b). Contrast function estimation for the drift parameter of ergodic jump diffusion process. *Scand. J. Stat.* **47** 279–346. MR4157142 https://doi.org/10.1111/sjos.12406

Barczy, M., Ben Alaya, M., Kebaier, A. and Pap, G. (2019). Asymptotic properties of maximum likelihood estimator for the growth rate of a stable CIR process based on continuous time observations. *Statistics* **53** 533–568. MR3945312 https://doi.org/10.1080/02331888.2019.1579216

Barndorff-Nielsen, O.E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *J. Financ. Econom.* **2** 1–37.

Barndorff-Nielsen, O.E., Shephard, N. and Winkel, M. (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic Process. Appl.* **116** 796–806. MR2218336 https://doi.org/10.1016/j.spa.2006.01.007

Bayraktar, E. and Clément, E. (2025a). Estimation of a pure-jump stable Cox-Ingersoll-Ross process. *Bernoulli* **31** 484–508. MR4815993 https://doi.org/10.3150/24-bej1736

Bayraktar, E. and Clément, E. (2026b). Supplement to “Volatility and jump activity estimation in a stable Cox-Ingersoll-Ross model.” https://doi.org/10.3150/25-BEJ1886SUPP

Boniece, B.C., Figueroa-López, J.E. and Han, Y. (2024). Efficient integrated volatility estimation in the presence of infinite variation jumps via debiased truncated realized variations. *Stochastic Process. Appl.* **176** 104429. MR4774175 https://doi.org/10.1016/j.spa.2024.104429

Bull, A.D. (2016). Near-optimal estimation of jump activity in semimartingales. *Ann. Statist.* **44** 58–86. MR3449762 https://doi.org/10.1214/15-AOS1349

1350-7265

Dawson, D.A. and Li, Z. (2006). Skew convolution semigroups and affine Markov processes. *Ann. Probab.* **34** 1103–1142. MR2243880 https://doi.org/10.1214/009117905000000747

Fu, Z. and Li, Z. (2010). Stochastic equations of non-negative processes with jumps. *Stochastic Process. Appl.* **120** 306–330. MR2584896 https://doi.org/10.1016/j.spa.2009.11.005

Gloter, A., Loukianova, D. and Mai, H. (2018). Jump filtering and efficient drift estimation for Lévy-driven SDEs. *Ann. Statist.* **46** 1445–1480. MR3819106 https://doi.org/10.1214/17-AOS1591

Huang, X. and Tauchen, G. (2005). The relative contribution of jumps to total price variance. *J. Financ. Econom.* **3** 456–499.

Jacod, J. and Protter, P. (2012). *Discretization of Processes. Stochastic Modelling and Applied Probability 67*. Heidelberg: Springer. MR2859096 https://doi.org/10.1007/978-3-642-24127-7

Jacod, J. and Reiss, M. (2014). A remark on the rates of convergence for integrated volatility estimation in the presence of jumps. *Ann. Statist.* **42** 1131–1144. MR3224283 https://doi.org/10.1214/13-AOS1179

Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes 288*, Second ed. Berlin: Springer-Verlag. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

Jacod, J. and Sørensen, M. (2018). A review of asymptotic theory of estimating functions. *Stat. Inference Stoch. Process.* **21** 415–434. MR3824976 https://doi.org/10.1007/s11203-018-9178-8

Jacod, J. and Todorov, V. (2014). Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Statist.* **42** 1029–1069. MR3210995 https://doi.org/10.1214/14-AOS1213

Jacod, J. and Todorov, V. (2016). Efficient estimation of integrated volatility in presence of infinite variation jumps with multiple activity indices. In *The Fascination of Probability, Statistics and Their Applications* 317–341. Cham: Springer. MR3495691

Jiao, Y., Ma, C. and Scotti, S. (2017). Alpha-CIR model with branching processes in sovereign interest rate modeling. *Finance Stoch.* **21** 789–813. MR3663644 https://doi.org/10.1007/s00780-017-0333-7

Jiao, Y., Ma, C., Scotti, S. and Zhou, C. (2021). The alpha-Heston stochastic volatility model. *Math. Finance* **31** 943–978. MR4273541 https://doi.org/10.1111/mafi.12306

Jing, B.-Y., Kong, X.-B., Liu, Z. and Mykland, P. (2012). On the jump activity index for semimartingales. *J. Econometrics* **166** 213–223. MR2862961 https://doi.org/10.1016/j.jeconom.2011.09.036

Li, Z. (2020). Continuous-state branching processes with immigration. In *From Probability to Finance—Lecture Notes of. BICMR Summer School on Financial Mathematics. Math. Lect. Peking Univ.* 1–69. Singapore: Springer. MR4339413 https://doi.org/10.1007/978-981-15-1576-7_1

Li, Z. and Ma, C. (2015). Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stochastic Process. Appl.* **125** 3196–3233. MR3343292 https://doi.org/10.1016/j.spa.2015.03.002

Li, L. and Taguchi, D. (2019). On a positivity preserving numerical scheme for jump-extended CIR process: The alpha-stable case. *BIT* **59** 747–774. MR4001784 https://doi.org/10.1007/s10543-019-00753-8

Mancini, C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. Giornale dell'Istituto Italiano degli Attuari LXIV.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scand. J. Stat.* **36** 270–296. MR2528985 https://doi.org/10.1111/j.1467-9469.2008.00622.x

Mancini, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Process. Appl.* **121** 845–855. MR2770909 https://doi.org/10.1016/j.spa.2010.12.001

Mies, F. (2020). Rate-optimal estimation of the Blumenthal-Getoor index of a Lévy process. *Electron. J. Stat.* **14** 4165–4206. MR4175392 https://doi.org/10.1214/20-EJS1769

Mies, F. and Podolskij, M. (2023). Estimation of mixed fractional stable processes using high-frequency data. *Ann. Statist.* **51** 1946–1964. MR4678791 https://doi.org/10.1214/23-aos2312

Sato, K.-I. (2013). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge University Press. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR3185174

Yang, X. (2017). Maximum likelihood type estimation for discretely observed CIR model with small α-stable noises. *Statist. Probab. Lett.* **120** 18–27. MR3567917 https://doi.org/10.1016/j.spl.2016.09.014

Bernoulli **32**(2), 2026, 952–977
https://doi.org/10.3150/25-BEJ1887

Nonparametric logistic regression with deep learning

ATSUTOMO YARA[1,a] and YOSHIKAZU TERADA[1,2,b]

[1]*Graduate School of Engineering Science, The University of Osaka, Osaka, Japan,*
[a]*a-yara@sigmath.es.osaka-u.ac.jp*
[2]*Center for Advanced Intelligence Project (AIP), RIKEN, Tokyo, Japan,* [b]*terada.yoshikazu.es@osaka-u.ac.jp*

Consider the nonparametric logistic regression problem. In the logistic regression, we usually consider the maximum likelihood estimator, and the excess risk is the expectation of the Kullback-Leibler (KL) divergence between the true and estimated conditional class probabilities. However, in the nonparametric logistic regression, the KL divergence could diverge easily, and thus, the convergence of the excess risk is difficult to prove or does not hold. Several existing studies show the convergence of the KL divergence under strong assumptions. In most cases, our goal is to estimate the true conditional class probabilities. Thus, instead of analyzing the excess risk itself, it suffices to show the consistency of the maximum likelihood estimator in some suitable metric. In this paper, using a simple unified approach for analyzing the nonparametric maximum likelihood estimator (NPMLE), we directly derive convergence rates of the NPMLE in the Hellinger distance under mild assumptions. Although our results are similar to the results in some existing studies, we provide simple and more direct proofs for these results. As an important application, we derive convergence rates of the NPMLE with fully connected deep neural networks and show that the derived rate nearly achieves the minimax optimal rate.

Keywords: Classification; conditional probability estimation; deep neural networks; nonparametric estimation

References

Bauer, B. and Kohler, M. (2019). On deep learning as a remedy for the curse of dimensionality in nonparametric regression. *Ann. Statist.* **47** 2261–2285. MR3953451 https://doi.org/10.1214/18-AOS1747

Bilodeau, B., Foster, D.J. and Roy, D.M. (2023). Minimax rates for conditional density estimation via empirical entropy. *Ann. Statist.* **51** 762–790. MR4601001 https://doi.org/10.1214/23-aos2270

Bos, T. and Schmidt-Hieber, J. (2022). Convergence rates of deep ReLU networks for multiclass classification. *Electron. J. Stat.* **16** 2724–2773. MR4406243 https://doi.org/10.1214/22-ejs2011

Chen, M., Jiang, H., Liao, W. and Zhao, T. (2022). Nonparametric regression on low-dimensional manifolds using deep ReLU networks: Function approximation and statistical recovery. *Inf. Inference* **11** 1203–1253. MR4526322 https://doi.org/10.1093/imaiai/iaac001

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **21** C1–C68. MR3769544 https://doi.org/10.1111/ectj.12097

Eckle, K. and Schmidt-Hieber, J. (2019). A comparison of deep networks with ReLU activation function and linear spline-type methods. *Neural Netw.* **110** 232–242. https://doi.org/10.1016/j.neunet.2018.11.005

Feng, H., Huang, S. and Zhou, D.-X. (2021). Generalization analysis of CNNs for classification on spheres. *IEEE Trans. Neural Netw. Learn. Syst.* **34** 6200–6213. MR4637020 https://doi.org/10.1109/tnnls.2021.3134675

Hayakawa, S. and Suzuki, T. (2020). On the minimax optimality and superiority of deep neural network learning over sparse parameter spaces. *Neural Netw.* **123** 343–361. https://doi.org/10.1016/j.neunet.2019.12.014

Hu, T., Shang, Z. and Cheng, G. (2020). Sharp rate of convergence for deep neural network classifiers under the teacher-student setting. https://doi.org/10.48550/arXiv.2001.06892.

Imaizumi, M. and Fukumizu, K. (2019). Deep neural networks learn non-smooth functions effectively. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*. **89** 869–878. PMLR.

Kim, Y., Ohn, I. and Kim, D. (2021). Fast convergence rates of deep neural networks for classification. *Neural Netw.* **138** 179–197. MR4285303 https://doi.org/10.1016/j.neunet.2021.02.012

Ko, H., Suh, N. and Huo, X. (2023). On excess risk convergence rates of neural network classifiers. https://doi.org/10.48550/arXiv.2309.15075.

Kohler, M., Krzyżak, A. and Walter, B. (2022). On the rate of convergence of image classifiers based on convolutional neural networks. *Ann. Inst. Statist. Math.* **74** 1085–1108. MR4498394 https://doi.org/10.1007/s10463-022-00828-4

Kohler, M. and Langer, S. (2020). Statistical theory for image classification using deep convolutional neural networks with cross-entropy loss. https://doi.org/10.48550/arXiv.2011.13602.

Kohler, M. and Langer, S. (2021). On the rate of convergence of fully connected deep neural network regression estimates. *Ann. Statist.* **49** 2231–2249. MR4319248 https://doi.org/10.1214/20-aos2034

Lee, Y., Lin, Y. and Wahba, G. (2004). Multicategory support vector machines: Theory and application to the classification of microarray data and satellite radiance data. *J. Amer. Statist. Assoc.* **99** 67–81. MR2054287 https://doi.org/10.1198/016214504000000098

Nakada, R. and Imaizumi, M. (2020). Adaptive approximation and generalization of deep neural network with intrinsic dimensionality. *J. Mach. Learn. Res.* **21** 1–38. MR4209460

Ohn, I. and Kim, Y. (2022). Nonconvex sparse regularization for deep neural networks and its optimality. *Neural Comput.* **34** 476–517. MR4381798 https://doi.org/10.1162/neco_a_01457

Schmidt-Hieber, J. (2019). Deep ReLU network approximation of functions on a manifold. https://doi.org/10.48550/arXiv.1908.00695.

Schmidt-Hieber, J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Statist.* **48** 1875–1897. MR4134774 https://doi.org/10.1214/19-AOS1875

Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*, 1st ed. Springer, Berlin. MR2796580

Suzuki, T. (2019). Adaptivity of deep ReLU network for learning in Besov and mixed smooth Besov spaces: optimal rate and curse of dimensionality. In *International Conference on Learning Representations*.

Suzuki, T. and Nitanda, A. (2021). Deep learning is adaptive to intrinsic dimensionality of model smoothness in anisotropic Besov space. In *Advances in Neural Information Processing Systems* **34** 3609–3621.

Tsybakov, A.B. (2004). Optimal aggregation of classifiers in statistical learning. *Ann. Statist.* **32** 135–166. MR2051002 https://doi.org/10.1214/aos/1079120131

Tsybakov, A.B. (2008). *Introduction to Nonparametric Estimation*, 1st ed. Springer, Berlin. MR2724359 https://doi.org/10.1007/b13794

van de Geer, S. (2000). *Empirical Processes in M-Estimation*. Cambridge Univ. Press, Cambridge.

Wong, W.H. and Shen, X. (1995). Probability inequalities for likelihood ratios and convergence rates of sieve MLEs. *Ann. Statist.* **23** 339–362. MR1332570 https://doi.org/10.1214/aos/1176324524

Yara, A. and Terada, Y. (2026). Supplement to "Nonparametric logistic regression with deep learning." https://doi.org/10.3150/25-BEJ1887SUPP

Zhou, T.-Y. and Huo, X. (2024). Classification of data generated by Gaussian mixture models using deep ReLU networks. *J. Mach. Learn. Res.* **25** 1–54. MR4777432

Bernoulli **32**(2), 2026, 978–995
https://doi.org/10.3150/25-BEJ1888

A fractional stochastic differential equation with discontinuous diffusion driven by fBm with Hurst parameter less than 1/2

JOHANNA GARZÓN[1,a], JORGE A. LEÓN[2,b], JORGE LOZADA[2,c] and SOLEDAD TORRES[3,4,d]

[1]*Departamento de Matemáticas, Universidad Nacional de Colombia, Bogotá, Colombia,* [a]*mjgarzonm@unal.edu.co*
[2]*Departamento de Control Automático, Cinvestav-IPN, Ciudad de México, Mexico,* [b]*joleon@cinvestav.mx,* [c]*jorge.lozada@cinvestav.mx*
[3]*Facultad de Ingeniería, CIMFAV Universidad de Valparaíso, Valparaiso, Chile,* [d]*soledad.torres@uv.cl*
[4]*Universidad Central de Chile, Santiago, Chile.*

The aim of this paper is to show the existence and uniqueness for the solution to a stochastic differential equation driven by fractional Brownian motion with Hurst parameter $H < 1/2$, with a discontinuous diffusion coefficient. The stochastic integral used in this paper is an extension of the Stratonovich integral introduced by León (*Bernoulli* **26** (2020) 2436–2462). In this way, we are able to complement previous results for SDEs driven by fBms with $H > 1/2$.

Keywords: Derivative operator in the Malliavin calculus sense; discontinuous diffusion; extensions of the divergence operator and the Stratonovich integral; fractional Brownian motion; fractional stochastic differential equation

References

[1] Alòs, E. and García Lorite, D. (2021). *Malliavin Calculus in Finance—Theory and Practice*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: CRC Press. With a foreword by Dariusz Gatarek. MR4701113 https://doi.org/10.1201/9781003018681

[2] Alòs, E., León, J.A. and Nualart, D. (2001). Stochastic Stratonovich calculus fBm for fractional Brownian motion with Hurst parameter less than 1/2. *Taiwanese J. Math.* **5** 609–632. MR1849782 https://doi.org/10.11650/twjm/1500574954

[3] Alòs, E. and Nualart, D. (2003). Stochastic integration with respect to the fractional Brownian motion. *Stoch. Stoch. Rep.* **75** 129–152. MR1978896 https://doi.org/10.1080/1045112031000078917

[4] Araya, H., Slaoui, M. and Torres, S. (2022). Bayesian inference for fractional oscillating Brownian motion. *Comput. Statist.* **37** 887–907. MR4400227 https://doi.org/10.1007/s00180-021-01146-8

[5] Cheridito, P. and Nualart, D. (2005). Stochastic integral of divergence type with respect to fractional Brownian motion with Hurst parameter $H \in (0, \frac{1}{2})$. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 1049–1081. MR2172209 https://doi.org/10.1016/j.anihpb.2004.09.004

[6] Föllmer, H., Protter, P. and Shiryayev, A.N. (1995). Quadratic covariation and an extension of Itô's formula. *Bernoulli* **1** 149–169. MR1354459 https://doi.org/10.2307/3318684

[7] Garrido-Atienza, M.J., Lu, K. and Schmalfuss, B. (2015). Local pathwise solutions to stochastic evolution equations driven by fractional Brownian motions with Hurst parameters $H \in (1/3, 1/2$. *Discrete Contin. Dyn. Syst. Ser. B* **20** 2553–2581. MR3423246 https://doi.org/10.3934/dcdsb.2015.20.2553

[8] Garzón, J., León, J.A. and Torres, S. (2017). Fractional stochastic differential equation with discontinuous diffusion. *Stoch. Anal. Appl.* **35** 1113–1123. MR3740753 https://doi.org/10.1080/07362994.2017.1358643

1350-7265

[9] Garzón, J., León, J.A. and Torres, S. (2023). Forward integration of bounded variation coefficients with respect to Hölder continuous processes. *Bernoulli* **29** 1877–1904. MR4580900 https://doi.org/10.3150/22-bej1524
[10] Garzón, J., León, J.A. and Torres, S. (2023). Representation of solutions to sticky stochastic differential equations. *Stoch. Dyn.* **23** Paper No. 2350005, 15. MR4566958 https://doi.org/10.1142/S0219493723500053
[11] Garzón, J., León, J.A., Lozada, J. and Torres, S. (2026). Supplement to "A fractional stochastic differential equation with discontinuous diffusion driven by fBm with Hurst parameter less than $1/2$." https://doi.org/10.3150/25-BEJ1888SUPP
[12] Harrison, J.M. and Shepp, L.A. (1981). On skew Brownian motion. *Ann. Probab.* **9** 309–313. MR0606993
[13] Ikeda, N. and Watanabe, S. (2014). *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library.
[14] Ilmonen, P., Torres, S. and Viitasaari, L. (2020). Oscillating Gaussian processes. *Stat. Inference Stoch. Process.* **23** 571–593. MR4136705 https://doi.org/10.1007/s11203-020-09212-6
[15] Itô, K. (1944). Stochastic integral. *Proc. Imp. Acad. (Tokyo)* **20** 519–524. MR0014633
[16] Jien, Y.-J. and Ma, J. (2009). Stochastic differential equations driven by fractional Brownian motions. *Bernoulli* **15** 846–870. MR2555202 https://doi.org/10.3150/08-BEJ169
[17] Keilson, J. and Wellner, J.A. (1978). Oscillating Brownian motion. *J. Appl. Probab.* **15** 300–310. MR474526 https://doi.org/10.2307/3213403
[18] Kohatsu-Higa, A., León, J.A. and Nualart, D. (1997). Stochastic differential equations with random coefficients. *Bernoulli* **3** 233–245. MR1466309 https://doi.org/10.2307/3318589
[19] Lejay, A. (2006). On the constructions of the skew Brownian motion. *Probab. Surv.* **3** 413–466. MR2280299 https://doi.org/10.1214/154957807000000013
[20] Lejay, A. and Pigato, P. (2018). Statistical estimation of the oscillating Brownian motion. *Bernoulli* **24** 3568–3602. MR3788182 https://doi.org/10.3150/17-BEJ969
[21] León, J.A. (2020). Stratonovich type integration with respect to fractional Brownian motion with Hurst parameter less than $1/2$. *Bernoulli* **26** 2436–2462. MR4091115 https://doi.org/10.3150/20-BEJ1202
[22] León, J.A. and Márquez-Carreras, D. (2021). Semilinear fractional stochastic differential equations driven by a γ-Hölder continuous signal with $\gamma > 2/3$. *Stoch. Dyn.* **21** Paper No. 2050039, 29. MR4192901 https://doi.org/10.1142/S0219493720500392
[23] León, J.A. and Nualart, D. (2005). An extension of the divergence operator for Gaussian processes. *Stochastic Process. Appl.* **115** 481–492. MR2118289 https://doi.org/10.1016/j.spa.2004.09.008
[24] León, J.A. and San Martín, J. (2007). Linear stochastic differential equations driven by a fractional Brownian motion with Hurst parameter less than $1/2$. *Stoch. Anal. Appl.* **25** 105–126. MR2284483 https://doi.org/10.1080/07362990601052052
[25] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR242239 https://doi.org/10.1137/1010093
[26] Nourdin, I. (2008). A simple theory for the study of SDEs driven by a fractional Brownian motion in dimension one. **1934** 181–197. MR2483731 https://doi.org/10.1007/978-3-540-77913-1_8
[27] Nualart, D. (2003). Stochastic integration with respect to fractional Brownian motion and applications. **336** 3–39. MR2037156 https://doi.org/10.1090/conm/336/06025
[28] Nualart, D. (2006). *The Malliavin Calculus and Related Topics*. *Probability and Its Applications (New York)*. New York: Springer-Verlag. MR1344217 https://doi.org/10.1007/978-1-4757-2437-0
[29] Nualart, D. and Răşcanu, A. (2002). Differential equations driven by fractional Brownian motion. *Collect. Math.* **53** 55–81. MR1893308
[30] Øksendal, B. (2003). An introduction with applications. In *Stochastic Differential Equations*, 6th ed. *Universitext*. Berlin: Springer-Verlag. MR2001996 https://doi.org/10.1007/978-3-642-14394-6
[31] Pipiras, V. and Taqqu, M.S. (2001). Are classes of deterministic integrands for fractional Brownian motion on an interval complete? *Bernoulli* **7** 873–897. MR1873833 https://doi.org/10.2307/3318624
[32] Royden, H. and Fitzpatrick, P.M. (2010). *Real Analysis*. China Machine Press.
[33] Russo, F. and Vallois, P. (1993). Forward, backward and symmetric stochastic integration. *Probab. Theory Related Fields* **97** 403–421. MR1245252 https://doi.org/10.1007/BF01195073

[34] Samko, S.G., Kilbas, A.A. and Marichev, O.I. (1993). *Fractional Integrals and Derivatives: Theory and Applications*. Gordon and Breach Science Publishers. Edited and with a foreword by S. M. Nikolskiĭ, Translated from the 1987 Russian original, Revised by the authors. MR1347689.

[35] Zähle, M. (1998). Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Related Fields* **111** 333–374. MR1640795 https://doi.org/10.1007/s004400050171

Bernoulli **32**(2), 2026, 996–1019
https://doi.org/10.3150/25-BEJ1889

A general maximal projection approach to uniformity testing on the hypersphere

JAROSLAV I. BORODAVKA[1,a] and BRUNO EBNER[2,b]

[1] *Scientific Computing Center, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany,* [a] *jaroslav.borodavka@kit.edu*
[2] *Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany,* [b] *Bruno.Ebner@kit.edu*

We propose a novel approach to uniformity testing on the d-dimensional unit hypersphere $\mathcal{S}^{d-1}$ based on maximal projections. This approach gives a unifying view on the classical uniformity tests of Rayleigh and Bingham, and it links to measures of multivariate skewness and kurtosis. We derive the limit distribution under the null hypothesis of the newly proposed test statistics using limit theorems for Banach-space-valued stochastic processes and we present strategies to simulate the limit processes by applying results on the theory of spherical harmonics. We examine the behaviour of the test statistics under contiguous and fixed alternatives and show the consistency of the testing procedure for some classes of alternatives. For the first time in uniformity testing on the hypersphere, we derive local Bahadur efficiency statements. Finally, we evaluate the theoretical findings and empirical power of the procedures in a broad competitive Monte Carlo simulation study.

Keywords: Bahadur efficiency; contiguous alternatives; directional data; maximal projections; uniformity tests

References

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

Araujo, A. and Giné, E. (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*. *Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons, Inc.

Atkinson, K.E. and Han, W. (2012). *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*. *Lecture Notes in Mathematics* **2044**. Springer. https://doi.org/10.1007/978-3-642-25983-8

Axler, S.J., Bourdon, P. and Ramey, W. (2001). *Harmonic Function Theory*. *Graduate Texts in Mathematics* **137**. Springer. https://doi.org/10.1007/978-1-4757-8137-3

Bahadur, R.R. (1967). Rates of convergence of estimates and test statistics. *Ann. Math. Stat.* **38** 303–324. https://doi.org/10.1214/aoms/1177698949

Bahadur, R.R. (1971). Some limit theorems in statistics. In *CBMS-NSF Regional Conference Series in Applied Mathematics*. *Industrial and Applied Mathematics*. **4**. Society for. https://doi.org/10.1137/1.9781611970630.ch1

Baker, J.A. (1997). Integration over spheres and the divergence theorem for balls. *Amer. Math. Monthly* **104** 36–47. https://doi.org/10.1080/00029890.1997.11990594

Baringhaus, L. and Henze, N. (1991). Limit distributions for measures of multivariate skewness and kurtosis based on projections. *J. Multivariate Anal.* **38** 51–69. https://doi.org/10.1016/0047-259X(91)90031-V

Beran, R.J. (1968). Testing for uniformity on a compact homogeneous space. *J. Appl. Probab.* **5** 177–195. https://doi.org/10.2307/3212085

Bingham, C. (1974). An antipodally symmetric distribution on the sphere. *Ann. Statist.* **2** 1201–1225. https://doi.org/10.1214/aos/1176342874

Bogdan, M., Bogdan, K. and Futschik, A. (2002). A data driven smooth test for circular uniformity. *Ann. Inst. Statist. Math.* **54** 29–44. https://doi.org/10.1023/A:1016109603897

Borodavka, J.I. (2024). Maximal Projection GoF. Available at https://gitlab.kit.edu/jaroslav.borodavka/maximal-projection-gof.git.

Borodavka, J.I. and Ebner, B. (2026). Supplement to "A general maximal projection approach to uniformity testing on the hypersphere." https://doi.org/10.3150/25-BEJ1889SUPP

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

Cutting, C., Paindaveine, D. and Verdebout, T. (2017). Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. *Ann. Statist.* **45** 1024–1058. https://doi.org/10.1214/16-AOS1473

Cutting, C., Paindaveine, D. and Verdebout, T. (2022). Testing uniformity on high-dimensional spheres: The non-null behaviour of the Bingham test. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 567–602. https://doi.org/10.1214/21-AIHP1168

Dai, F. and Xu, Y. (2013). *Approximation Theory and Harmonic Analysis on Spheres and Balls*. *Springer Monographs in Mathematics*. Springer. https://doi.org/10.1007/978-1-4614-6660-4

Diggle, P.J., Fisher, N.I. and Lee, A.J. (1985). A comparison of tests of uniformity for spherical data. *Aust. J. Stat.* **27** 53–59. https://doi.org/10.1111/j.1467-842X.1985.tb00547.x

Dutordoir, V., Durrande, N. and Hensman, J. (2020). Sparse Gaussian processes with spherical harmonic features. In *Proceedings of the 37th International Conference on Machine Learning* (H.D. III and A. Singh, eds.). *Proceedings of Machine Learning Research* **119** 2793–2802. PMLR.

Ebner, B., Henze, N. and Yukich, J.E. (2018). Multivariate goodness-of-fit on flat and curved spaces via nearest neighbor distances. *J. Multivariate Anal.* **165** 231–242. https://doi.org/10.1016/j.jmva.2017.12.009

Fernández-de Marcos, A. and García-Portugués, E. (2023). On new omnibus tests of uniformity on the hypersphere. *TEST* **32** 1508–1529. https://doi.org/10.1007/s11749-023-00882-x

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

García-Portugués, E., Navarro-Esteban, P. and Cuesta-Albertos, J.A. (2021). A Cramér-von Mises test of uniformity on the hypersphere. Available at https://arxiv.org/abs/2008.10767.

García-Portugués, E., Paindaveine, D. and Verdebout, T. (2024). On a class of Sobolev tests for symmetry of directions, their detection thresholds, and asymptotic powers. Available at https://arxiv.org/abs/2108.09874.

García-Portugués, E. and Verdebout, T. (2018). An overview of uniformity tests on the hypersphere. Available at https://arxiv.org/abs/1804.00286.

Giné, E.M. (1975). Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. *Ann. Statist.* **3** 1243–1266. https://doi.org/10.1214/aos/1176343283

Groemer, H. (1996). *Geometric Applications of Fourier Series and Spherical Harmonics*. *Encyclopedia of Mathematics and Its Applications* **61**. Cambridge University Press. https://doi.org/10.1017/CBO9780511530005

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

Jammalamadaka, S.R. and Sen Gupta, A. (2001). *Topics in Circular Statistics 5*. World Scientific, Singapore. https://doi.org/10.1142/9789812779267

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

Kent, J.T., Ganeiber, A.M. and Mardia, K.V. (2013). A new method to simulate the Bingham and related distributions in directional data analysis with applications. Available at https://arxiv.org/abs/1310.8110v1.

Kuiper, N.H. (1960). Tests concerning random points on a circle. *Proc. K. Ned. Akad. Wet., Ser. A, Indag. Math.* **63** 38–47. https://doi.org/10.1016/S1385-7258(60)50006-0

Ledoux, M. and Talagrand, M. (2002). *Probability in Banach Spaces – Isoperimetry and Processes*. *A Series of Modern Surveys in Mathematics* **23**. Springer. https://doi.org/10.1007/978-3-642-20212-4

Ley, C. and Verdebout, T. (2017). *Modern Directional Statistics*. *Chapmann & Hall/CRC Interdisciplinary Statistics Series*. CRC Press. https://doi.org/10.1201/9781315119472

Ley, C. and Verdebout, T., eds. (2019). *Applied Directional Statistics*. *Modern Methods and Case Studies*. Boca Raton, FL: CRC Press. https://doi.org/10.1201/9781315228570

Malkovich, J.F. and Afifi, A.A. (1973). On tests for multivariate normality. *J. Amer. Statist. Assoc.* **68** 176–179. https://doi.org/10.1080/01621459.1973.10481358

Manzotti, A. and Quiroz, A.J. (2001). Spherical harmonics in quadratic forms for testing multivariate normality. *TEST* **10** 87–104. https://doi.org/10.1007/BF02595825

Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*, 2nd ed. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470316979

Nikitin, J. (1995). *Asymptotic Efficiency of Nonparametric Tests*. Cambridge University Press. https://doi.org/10.1017/CBO9780511530081.001

R-Core-Team (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

Rao, J.S. (1972). Bahadur efficiencies of some tests for uniformity on the circle. *Ann. Math. Stat.* **43** 468–479. https://doi.org/10.1214/aoms/1177692627

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

Schölkopf, B. and Smola, A.J. (2010). *Learning with Kernels – Support Vector Machines, Regularization, Optimization and Beyond. Adaptive Computation and Machine Learning*. The MIT Press. https://doi.org/10.1198/jasa.2003.s269

Watson, G.S. (1961). Goodness-of-fit tests on a circle. *Biometrika* **48** 109–114. https://doi.org/10.2307/2333135

Wolfram Research, Inc. (2019). Mathematica, Version 12.0. Champaign, IL.

Bernoulli **32**(2), 2026, 1020–1044
https://doi.org/10.3150/25-BEJ1890

Nonparametric Bayesian intensity estimation for covariate-driven inhomogeneous point processes

MATTEO GIORDANO[1,a], ALISA KIRICHENKO[2,b] and JUDITH ROUSSEAU[3,c]

[1]*ESOMAS Department, University of Turin, Turin, Italy,* [a]*matteo.giordano@unito.it*
[2]*Department of Statistics, University of Warwick, Coventry, United Kingdom,* [b]*alice.kirichenko@warwick.ac.uk*
[3]*Department of Statistics, University of Oxford, Oxford, United Kingdom,* [c]*judith.rousseau@stats.ox.ac.uk*

This work studies nonparametric Bayesian estimation of the intensity function of an inhomogeneous Poisson point process in the important case where the intensity depends on covariates, based on the observation of a single realisation of the point pattern over a large area. It is shown how the presence of covariates allows to borrow information from far away locations in the observation window, enabling consistent inference in the growing domain asymptotics. In particular, optimal posterior contraction rates under both global and point-wise loss functions are derived. The rates in global loss are obtained under conditions on the prior distribution resembling those in the well established theory of Bayesian nonparametrics, combined with concentration inequalities for functionals of stationary processes to control certain random covariate-dependent loss functions appearing in the analysis. The local rates are derived with an ad-hoc study that builds on recent advances in the theory of Pólya tree priors, extended to the present multivariate setting with a novel construction that makes use of the random geometry induced by the covariates.

Keywords: Cox process; Gaussian priors; frequentist analysis of Bayesian procedures; mixture priors; Poisson process; Pólya tree priors

References

Adams, R.P., Murray, I. and MacKay, D.J.C. (2009). Tractable nonparametric Bayesian inference in Poisson processes with Gaussian process intensities. In *Proceedings of the 26th Annual International Conference on Machine Learning. ICML '09* 9–16. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/1553374.1553376

Agapiou, S. and Wang, S. (2024). Laplace priors and spatial inhomogeneity in Bayesian inverse problems. *Bernoulli* **30** 878–910.

Aida, S. and Stroock, D. (1994). Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1** 75–86. MR1258492 https://doi.org/10.4310/MRL.1994.v1.n1.a9

Baddeley, A., Chang, Y.-M., Song, Y. and Turner, R. (2012). Nonparametric estimation of the dependence of a spatial point process on spatial covariates. *Stat. Interface* **5** 221–236. MR2928072 https://doi.org/10.4310/SII.2012.v5.n2.a7

Bandyopadhyay, S. and Subba Rao, S. (2017). A test for stationarity for irregularly spaced spatial data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 95–123. MR3597966 https://doi.org/10.1111/rssb.12161

Belitser, E., Serra, P. and van Zanten, H. (2015). Rate-optimal Bayesian intensity smoothing for inhomogeneous Poisson processes. *J. Statist. Plann. Inference* **166** 24–35. MR3390131 https://doi.org/10.1016/j.jspi.2014.03.009

Berenfeld, C., Rosa, P. and Rousseau, J. (2024). Estimating a density near an unknown manifold: A Bayesian nonparametric approach. *Ann. Statist.* **52** 2081–2111. https://doi.org/10.1214/24-AOS2423

Berman, M. and Diggle, P. (1989). Estimating weighted integrals of the second-order intensity of a spatial point process. *J. Roy. Statist. Soc. Ser. B, Methodol.* **51** 81–92. MR0984995

1350-7265

Bobkov, S. and Ledoux, M. (1997). Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields* **107** 383–400. MR1440138 https://doi.org/10.1007/s004400050090

Borrajo, M.I., González-Manteiga, W. and Martínez-Miranda, M.D. (2020). Bootstrapping kernel intensity estimation for inhomogeneous point processes with spatial covariates. *Comput. Statist. Data Anal.* **144** 106875, 21 pp. MR4023908 https://doi.org/10.1016/j.csda.2019.106875

Brillinger, D.R. (1978). Comparative aspects of the study of ordinary time series and of point processes. In *Developments in Statistics* **1** 33–133. New York–London: Academic Press. MR0501668

Canale, A. and De Blasi, P. (2017). Posterior asymptotics of nonparametric location-scale mixtures for multivariate density estimation. *Bernoulli* **23** 379–404. MR3556776 https://doi.org/10.3150/15-BEJ746

Castillo, I. (2017). Pólya tree posterior distributions on densities. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 2074–2102. MR3729648 https://doi.org/10.1214/16-AIHP784

Castillo, I. and Mismer, R. (2021). Spike and slab Pólya tree posterior densities: Adaptive inference. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1521–1548. https://doi.org/10.1214/20-AIHP1132

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

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

Castillo, I. and Randrianarisoa, T. (2022). Optional Pólya trees: Posterior rates and uncertainty quantification. *Electron. J. Stat.* **16** 6267–6312.

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

Cox, D.R. (1955). Some statistical methods connected with series of events. *J. Roy. Statist. Soc. Ser. B, Methodol.* **17** 129–157; discussion, 157–164. MR92301

Cox, D.D. (1993). An analysis of Bayesian inference for nonparametric regression. *Ann. Statist.* **21** 903–923.

Cressie, N.A.C. (2015). *Statistics for Spatial Data*, revised ed. *Wiley Classics Library*. New York: John Wiley & Sons, Inc. MR3559472

Cronie, O. and van Lieshout, M.N.M. (2018). A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika* **105** 455–462. MR3804414 https://doi.org/10.1093/biomet/asy001

Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I*, 2nd ed. New York: Springer-Verlag. MR1950431

Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann. Statist.* **14** 1–26.

Diggle, P.J. (1990). A point process modelling approach to raised incidence of a rare phenomenon in the vicinity of a prespecified point. *J. R. Stat. Soc. Ser. A* **153** 349–362.

Diggle, P.J. (2014). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* **128**, 3rd ed. Boca Raton, FL: CRC Press. MR3113855

DiMatteo, I., Genovese, C.R. and Kass, R.E. (2001). Bayesian curve-fitting with free-knot splines. *Biometrika* **88** 1055–1071. MR1872219 https://doi.org/10.1093/biomet/88.4.1055

Donnet, S., Rivoirard, V., Rousseau, J. and Scricciolo, C. (2017). Posterior concentration rates for counting processes with Aalen multiplicative intensities. *Bayesian Anal.* **12** 53–87. MR3597567 https://doi.org/10.1214/15-BA986

Duerinckx, M. and Gloria, A. (2020a). Multiscale functional inequalities in probability: Constructive approach. *Ann. Henri Lebesgue* **3** 825–872. MR4149827 https://doi.org/10.5802/ahl.47

Duerinckx, M. and Gloria, A. (2020b). Multiscale functional inequalities in probability: Concentration properties. *ALEA Lat. Amer. J. Probab. Math. Stat.* **17** 133–157. MR4057186 https://doi.org/10.30757/alea.v17-06

Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531.

Ghosal, S. and van der Vaart, A.W. (2017). *Fundamentals of Nonparametric Bayesian Inference*. New York: Cambridge University Press.

Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. New York: Cambridge University Press. MR3588285 https://doi.org/10.1017/CBO9781107337862

Giordano, M. (2023). Besov-Laplace priors in density estimation: Optimal posterior contraction rates and adaptation. *Electron. J. Stat.* **17** 2210–2249. MR4649387 https://doi.org/10.1214/23-ejs2161

Giordano, M., Kirichenko, A. and Rousseau, J. (2026). Supplement to "Nonparametric Bayesian intensity estimation for covariate-driven inhomogeneous point processes." https://doi.org/10.3150/25-BEJ1890SUPP

Giordano, M. and Nickl, R. (2020). Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. *Inverse Probl.* **36** 085001, 35 pp. MR4151406 https://doi.org/10.1088/1361-6420/ab7d2a

Giordano, M., Ray, K. and Schmidt-Hieber, J. (2022). On the inability of Gaussian process regression to optimally learn compositional functions. In *Advances in Neural Information Processing Systems* (S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho and A. Oh, eds.) **35** 22341–22353. Curran Associates, Inc.

Gloria, A., Neukamm, S. and Otto, F. (2015). Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.* **199** 455–515. MR3302119 https://doi.org/10.1007/s00222-014-0518-z

Grant, J.A. and Leslie, D.S. (2019). Posterior contraction rates for Gaussian Cox processes with non-identically distributed data. Preprint. Available at arXiv:1906.08799.

Guan, Y. (2008). On consistent nonparametric intensity estimation for inhomogeneous spatial point processes. *J. Amer. Statist. Assoc.* **103** 1238–1247. MR2528839 https://doi.org/10.1198/016214508000000526

Gugushvili, S. and Spreij, P. (2013). A note on non-parametric Bayesian estimation for Poisson point processes. Preprint. Available at arXiv:1304.7353.

Gugushvili, S., Van Der Meulen, F., Schauer, M. and Spreij, P. (2018). Fast and scalable non-parametric Bayesian inference for Poisson point processes. Preprint. Available at arXiv:1804.03616.

Guyon, X. (1995). *Random Fields on a Network*. New York: Springer-Verlag. MR1344683

Heikkinen, J. and Arjas, E. (1998). Non-parametric Bayesian estimation of a spatial Poisson intensity. *Scand. J. Stat.* **25** 435–450.

Heinrich, L. and Schmidt, V. (1985). Normal convergence of multidimensional shot noise and rates of this convergence. *Adv. Appl. Probab.* **17** 709–730.

Illian, J.B., Møller, J. and Waagepetersen, R.P. (2009). Hierarchical spatial point process analysis for a plant community with high biodiversity. *Environ. Ecol. Stat.* **16** 389–405. MR2749847 https://doi.org/10.1007/s10651-007-0070-8

Illian, J.B., Sørbye, S.H. and Rue, H.V. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *Ann. Appl. Stat.* **6** 1499–1530. MR3058673 https://doi.org/10.1214/11-AOAS530

Jensen, J.L. (1993). Asymptotic normality of estimates in spatial point processes. *Scand. J. Stat.* **20** 97–109.

Jolivet, E. (1981). Central limit theorem and convergence of empirical processes for stationary point processes. In *Point Processes and Queuing Problems (Colloqium, Keszthely, 1978)* **24** 117–161.

Kirichenko, A. and van Zanten, H. (2015). Optimality of Poisson processes intensity learning with Gaussian processes. *J. Mach. Learn. Res.* **16** 2909–2919. MR3450529

Kottas, A. and Sansó, B. (2007). Bayesian mixture modeling for spatial Poisson process intensities, with applications to extreme value analysis. *J. Statist. Plann. Inference* **137** 3151–3163. MR2365118 https://doi.org/10.1016/j.jspi.2006.05.022

Kruijer, W., Rousseau, J. and van der Vaart, A. (2010). Adaptive Bayesian density estimation with location-scale mixtures. *Electron. J. Stat.* **4** 1225–1257. MR2735885 https://doi.org/10.1214/10-EJS584

Kuo, L. and Ghosh, S.K. (1997). Bayesian nonparametric inference for nonhomogeneous Poisson processes. Technical report, University of Connecticut, Department of Statistics.

Kutoyants, Y.A. (1998). *Statistical Inference for Spatial Poisson Processes*. *Lecture Notes in Statistics* **134**. New York: Springer-Verlag. MR1644620 https://doi.org/10.1007/978-1-4612-1706-0

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

Lember, J. and van der Vaart, A. (2007). On universal Bayesian adaptation. *Statist. Decisions* **25** 127–152. MR2388859 https://doi.org/10.1524/stnd.2007.25.2.127

Lions, J.L. and Magenes, E. (1972). *Non-Homogeneous Boundary Value Problems and Applications. Vol. I*. New York–Heidelberg: Springer-Verlag. MR0350177

Lo, A.Y. (1982). Bayesian nonparametric statistical inference for Poisson point processes. *Z. Wahrsch. Verw. Gebiete* **59** 55–66. MR643788 https://doi.org/10.1007/BF00575525

Ma, L. (2017). Adaptive shrinkage in Pólya tree type models. *Bayesian Anal.* **12** 779–805. https://doi.org/10.1214/16-BA1021

Møller, J. and Stoyan, D. (2007). Stochastic Geometry and Random Tessellations. Research Report Series R-2007-28, Department of Mathematical Sciences, Aalborg University.

Møller, J., Syversveen, A.R. and Waagepetersen, R.P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482. MR1650019 https://doi.org/10.1111/1467-9469.00115

Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. Boca Raton, FL: Chapman & Hall/CRC. MR2004226

Naulet, Z. and Rousseau, J. (2017). Posterior concentration rates for mixtures of normals in random design regression. *Electron. J. Stat.* **11** 4065–4102. https://doi.org/10.1214/17-EJS1344

Ng, T.L.J. and Murphy, T.B. (2019). Estimation of the intensity function of an inhomogeneous Poisson process with a change-point. *Canad. J. Statist.* **47** 604–618. MR4035791 https://doi.org/10.1002/cjs.11514

Nickl, R. and Pötscher, B.M. (2007). Bracketing metric entropy rates and empirical central limit theorems for function classes of Besov-and Sobolev-type. *J. Theor. Probab.* **20** 177–199.

Palacios, J.A. and Minin, V.N. (2013). Gaussian process-based Bayesian nonparametric inference of population size trajectories from gene genealogies. *Biometrics* **69** 8–18. MR3058047 https://doi.org/10.1111/biom.12003

Ročková, V. and Rousseau, J. (2024). Ideal Bayesian spatial adaptation. *J. Amer. Statist. Assoc.* **119** 2078–2091.

Rosenblatt, M. (2000). *Gaussian and Non-Gaussian Linear Time Series and Random Fields*. *Springer Series in Statistics*. New York: Springer-Verlag. MR1742357 https://doi.org/10.1007/978-1-4612-1262-1

Rousseau, J. and Szabo, B. (2020). Asymptotic frequentist coverage properties of Bayesian credible sets for sieve priors. *Ann. Statist.* **48** 2155–2179. https://doi.org/10.1214/19-AOS1881

Rue, H., Martino, S. and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B* **71** 319–392. MR2649602 https://doi.org/10.1111/j.1467-9868.2008.00700.x

Shen, W., Tokdar, S.T. and Ghosal, S. (2013). Adaptive Bayesian multivariate density estimation with Dirichlet mixtures. *Biometrika* **100** 623–640. MR3094441 https://doi.org/10.1093/biomet/ast015

Stein, M.L. (1999). *Interpolation of Spatial Data*. New York: Springer-Verlag. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

Sulem, D., Rivoirard, V. and Rousseau, J. (2024). Bayesian estimation of nonlinear Hawkes processes. *Bernoulli* **30** 1257–1286.

Szabo, B., van der Vaart, A. and van Zanten, J.H. (2015). Frequentist coverage of adaptive nonparametric Bayesian credible sets. *Ann. Statist.* **43** 1391–1428.

Torquato, S. (2002). *Random Heterogeneous Materials*. *Interdisciplinary Applied Mathematics* **16**. New York: Springer-Verlag. MR1862782 https://doi.org/10.1007/978-1-4757-6355-3

van der Vaart, A.W. and van Zanten, J.H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463.

van Waaij, J. and van Zanten, H. (2016). Gaussian process methods for one-dimensional diffusions: Optimal rates and adaptation. *Electron. J. Stat.* **10** 628–645. MR3471991 https://doi.org/10.1214/16-EJS1117

Waagepetersen, R.P. (2007). An estimating function approach to inference for inhomogeneous Neyman-Scott processes. *Biometrics* **63** 252–258, 315. MR2345595 https://doi.org/10.1111/j.1541-0420.2006.00667.x

Wong, W.H. and Ma, L. (2010). Optional Pólya tree and Bayesian inference. *Ann. Statist.* 1433–1459.

Yue, Y.R. and Loh, J.M. (2011). Bayesian semiparametric intensity estimation for inhomogeneous spatial point processes. *Biometrics* **67** 937–946. MR2829268 https://doi.org/10.1111/j.1541-0420.2010.01531.x

Zhang, H. and Zimmerman, D.L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika* **92** 921–936.

Bernoulli **32**(2), 2026, 1045–1072
https://doi.org/10.3150/25-BEJ1891

Simulating conditioned diffusions on manifolds

MARC CORSTANJE[1,a], FRANK VAN DER MEULEN[1,b], MORITZ SCHAUER[2,c] and STEFAN SOMMER[3,d]

[1] *Vrije Universiteit, Amsterdam, Netherlands,* [a]*m.a.corstanje@vu.nl,* [b]*f.h.van.der.meulen@vu.nl*
[2] *University of Gothenburg and Chalmers Technical University, Göteborg, Sweden,* [c]*smoritz@chalmers.se*
[3] *University of Copenhagen, Copenhagen, Denmark,* [d]*sommer@di.ku.dk*

To date, most methods for simulating conditioned diffusions are limited to the Euclidean setting. The conditioned process can be constructed using a change of measure known as Doob's h-transform. The specific type of conditioning depends on a function h which is typically unknown in closed form. To resolve this, we extend the notion of guided processes to a manifold M, where one replaces h by a function based on the heat kernel on M. We consider the case of a Brownian motion with drift, constructed using the frame bundle of M, conditioned to hit a point x_T at time T. We prove equivalence of the laws of the conditioned process and the guided process with a tractable Radon-Nikodym derivative. Subsequently, we show how one can obtain guided processes on any manifold N that is diffeomorphic to M without assuming knowledge of the heat kernel on N. We illustrate our results with numerical simulations of guided processes and Bayesian parameter estimation based on discrete-time observations. For this, we consider both the torus and the Poincaré disk.

Keywords: Bridge simulation; Doob's h-transform; geometric statistics; guided processes; Poincaré disk; Riemannian manifolds

References

Agarwal, R., Deng, S. and Zhang, W. (2005). Generalization of a retarded Gronwall-like inequality and its applications. *J. Appl. Math. Comput.* **165** 599–612. https://doi.org/10.1016/j.amc.2004.04.067

Arnaudon, A., Van der Meulen, F., Schauer, M. and Sommer, S. (2022). Diffusion bridges for stochastic Hamiltonian systems and shape evolutions. *SIAM J. Imaging Sci.* **15** 293–323. https://doi.org/10.1137/21M1406283

Axen, S.D., Baran, M., Bergmann, R. and Rzecki, K. (2021). Manifolds.jl: An Extensible Julia Framework for Data Analysis on Manifolds. Available at arXiv:2106.08777.

Baudoin, F. (2002). Conditioned stochastic differential equations: Theory, examples and application to finance. *Stochastic Process. Appl.* **100** 109–145. https://doi.org/10.1016/S0304-4149(02)00109-6

Baudoin, F. (2004). Conditioning and initial enlargement of filtration on a Riemannian manifold. *Ann. Appl. Probab.* **32** 2286–2303. https://doi.org/10.1214/009117904000000126

Beskos, A., Papaspiliopoulos, O. and Roberts, G.O. (2006). Retrospective exact simulation of diffusion sample paths with applications. *Bernoulli* **12** 1077–1098. MR2274855 https://doi.org/10.3150/bj/1165269151

Beskos, A., Papaspiliopoulos, O., Roberts, G.O. and Fearnhead, P. (2006). Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 333–382. MR2278331 https://doi.org/10.1111/j.1467-9868.2006.00552.x

Beskos, A., Roberts, G., Stuart, A. and Voss, J. (2008). MCMC methods for diffusion bridges. *Stoch. Dyn.* **08** 319–350. https://doi.org/10.1142/S0219493708002378

Bierkens, J., Van Der Meulen, F. and Schauer, M. (2020). Simulation of elliptic and hypo-elliptic conditional diffusions. *Adv. Appl. Probab.* **52** 173–212.

Bierkens, J., Grazzi, S., Van der Meulen, F. and Schauer, M. (2021). A piecewise deterministic Monte Carlo method for diffusion bridges. *Stat. Comput.* **31** 37. https://doi.org/10.1007/s11222-021-10008-8

Bladt, M., Finch, S. and Sørensen, M. (2016). Simulation of multivariate diffusion bridges. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 343–369. MR3454200 https://doi.org/10.1111/rssb.12118

Bui, M.N., Pokern, Y. and Dellaportas, P. (2023). Inference for partially observed Riemannian Ornstein–Uhlenbeck diffusions of covariance matrices. *Bernoulli* **29** 2961–2986.

1350-7265

Clark, J.M.C. (1990). The simulation of pinned diffusions. In *Proceedings of the 29th IEEE Conference on Decision and Control* 1418–1420. IEEE.

Corstanje, M. (2024). Code examples for simulations. Available at https://github.com/macorstanje/manifold_bridge.jl.

Corstanje, M., Van der Meulen, F. and Schauer, M. (2023). Conditioning continuous-time Markov processes by guiding. *Stochastics* **95** 963–996. https://doi.org/10.1080/17442508.2022.2150081

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. https://doi.org/10.1214/13-STS421

Delyon, B. and Hu, Y. (2006). Simulation of conditioned diffusions and application to parameter estimation. *Stochastic Process. Appl.* **116** 1660–1675.

Durham, G.B. and Gallant, A.R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *J. Bus. Econom. Statist.* **20** 297–338. MR1939904 https://doi.org/10.1198/073500102288618397

Elworthy, D. (1988). Geometric aspects of diffusions on manifolds. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87* 277–425. Berlin, Heidelberg: Springer.

Emery, M. (1989). *Stochastic Calculus in Manifolds*. Berlin, Heidelberg: Springer.

Fu, X., Gao, Y., Wei, Y., Sun, Q., Peng, H., Li, J. and Li, X. (2024). Hyperbolic geometric latent diffusion model for graph generation. In *Proceedings of the 41st International Conference on. Machine Learning. Proc. Mach. Learn. Res. (PMLR)* **235** 14102–14124.

Golightly, A. and Wilkinson, D.J. (2010). Learning and inference in computational systems biology. In *Markov Chain Monte Carlo Algorithms for SDE Parameter Estimation* 253–276. MIT Press.

Grigor'yan, A. and Noguchi, M. (1998). The heat kernel on hyperbolic space. *Bull. Lond. Math. Soc.* **30** 643–650. https://doi.org/10.1112/S0024609398004780

Hairer, M., Stuart, A.M. and Voss, J. (2009). Sampling conditioned diffusions. In *Trends in Stochastic Analysis. London Math. Soc. Lecture Note Ser.* **353** 159–186. Cambridge University Press. MR2562154

Hansen, P., Eltzner, B. and Sommer, S. (2021). Diffusion means and heat kernel on manifolds. In *Geometric Science of Information* 111–118. Cham: Springer International Publishing.

Hsu, E.P. (2002). *Stochastic Analysis on Manifolds* **38**. Grad. Stud. Math. https://doi.org/10.1090/gsm/038

Jensen, M.H., Joshi, S. and Sommer, S. (2022). Discrete-time observations of Brownian motion on Lie groups and homogeneous spaces: Sampling and metric estimation. *Algorithms* **15** 290.

Jensen, M.H., Mallasto, A. and Sommer, S. (2019). Simulation of conditioned diffusions on the flat torus. In *Geometric Science of Information* 685–694. Cham: Springer International Publishing.

Jensen, M.H. and Sommer, S. (2023). Simulation of Conditioned Semimartingales on Riemannian Manifolds. Available at arXiv:2105.13190.

Kühnel, L. and Sommer, S. (2017). Computational anatomy in theano. In *Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics* 164–176. Cham: Springer International Publishing.

Lin, M., Chen, R. and Mykland, P. (2010). On generating Monte Carlo samples of continuous diffusion bridges. *J. Amer. Statist. Assoc.* **105** 820–838. MR2724864 https://doi.org/10.1198/jasa.2010.tm09057

Lovett, S. (2010). *Differential Geometry of Manifolds*. AK Peters/CRC Press.

Manton, J.H. (2013). A primer on stochastic differential geometry for signal processing. *IEEE J. Sel. Top. Signal Process.* **7** 681–699.

Mider, M., Schauer, M. and Van der Meulen, F. (2021). Continuous-discrete smoothing of diffusions. *Electron. J. Stat.* **15** 4295–4342. https://doi.org/10.1214/21-EJS1894

Palmowski, Z. and Rolski, T. (2002). A technique for exponential change of measure for Markov processes. *Bernoulli* **8** 767–785.

Papaspiliopoulos, O., Roberts, G.O. and Stramer, O. (2013). Data augmentation for diffusions. *J. Comput. Graph. Statist.* **22** 665–688. MR3173736 https://doi.org/10.1080/10618600.2013.783484

Roberts, G.O. and Rosenthal, J.S. (2009). Examples of adaptive MCMC. *J. Comput. Graph. Statist.* **18** 349–367.

Schauer, M., Van der Meulen, F. and van Zanten, H. (2017). Guided proposals for simulating multi-dimensional diffusion bridges. *Bernoulli* **23** 2917–2950.

Sommer, S. (2020). Probabilistic approaches to geometric statistics: Stochastic processes, transition distributions, and fiber bundle geometry. In *Riemannian Geometric Statistics in Medical Image Analysis* (X. Pennec, S. Sommer and T. Fletcher, eds.) 377–416. Academic Press. https://doi.org/10.1016/B978-0-12-814725-2.00018-2

Sommer, S., Arnaudon, A., Kuhnel, L. and Joshi, S. (2017). Bridge simulation and metric estimation on landmark manifolds. In *Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics* 79–91. Cham: Springer International Publishing.

Stuart, A.M., Voss, J. and Wiberg, P. (2004). Fast communication conditional path sampling of SDEs and the Langevin MCMC method. *Commun. Math. Sci.* **2** 685–697. MR2119934

Sturm, K.T. (1993). Schrödinger semigroups on manifolds. *J. Funct. Anal.* **118** 309–350. https://doi.org/10.1006/jfan.1993.1147

Van der Meulen, F. and Schauer, M. (2017). Bayesian estimation of discretely observed multi-dimensional diffusion processes using guided proposals. *Electron. J. Stat.* **11** 2358–2396. https://doi.org/10.1214/17-EJS1290

Wen, L., Tang, X., Ouyang, M., Shen, X., Yang, J., Zhu, D., Chen, M. and Wei, X. (2024). Hyperbolic graph diffusion model. In *AAAI'24/IAAI'24/EAAI'24*. AAAI Press. https://doi.org/10.1609/aaai.v38i14.29512

Whitaker, G.A., Golightly, A., Boys, R.J. and Sherlock, C. (2017). Improved bridge constructs for stochastic differential equations. *Stat. Comput.* **27** 885–900. https://doi.org/10.1007/s11222-016-9660-3

Bernoulli **32**(2), 2026, 1073–1097
https://doi.org/10.3150/25-BEJ1892

Probabilistic cellular automata with local transition matrices: Synchronization, ergodicity, and inference

ERHAN BAYRAKTAR[1,a], FEI LU[2,b], MAURO MAGGIONI[2,3,c], RUOYU WU[4,e] and SICHEN YANG[3,d]

[1] *Department of Mathematics, University of Michigan, Ann Arbor, USA,* [a] *erhan@umich.edu*
[2] *Department of Mathematics, Johns Hopkins University, Baltimore, USA,* [b] *feilu@math.jhu.edu*
[3] *Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, USA,* [c] *mauromaggionijhu@icloud.com,* [d] *syang114@jhu.edu*
[4] *Department of Mathematics, Iowa State University, Ames, USA,* [e] *ruoyu@iastate.edu*

We introduce a class of probabilistic cellular automata that are capable of exhibiting rich dynamics such as synchronization and ergodicity, and can be easily inferred from data. The system is a finite-state locally interacting Markov chain on a circular graph. Each site's subsequent state is random, with a distribution determined by its neighborhood's empirical distribution multiplied by a local transition matrix. We establish sufficient and necessary conditions on the local transition matrix for synchronization and ergodicity. Also, we introduce novel least squares estimators for inferring the local transition matrix from various types of data, which may consist of either multiple trajectories, a long trajectory, or ensemble sequences without trajectory information. Under suitable identifiability conditions, we show the asymptotic normality of these estimators and provide non-asymptotic bounds for their accuracy.

Keywords: Ergodicity; inference; probabilistic cellular automata; synchronization

References

[1] Aldous, D. (2013). Interacting particle systems as stochastic social dynamics. *Bernoulli* **19** 1122–1149.
[2] Bayraktar, E., Chakraborty, S. and Wu, R. (2023). Graphon mean field systems. *Ann. Appl. Probab.* **33** 3587–3619.
[3] Bayraktar, E. and Zhou, H. (2025). Non-parametric estimates for graphon mean-field particle systems. *Bernoulli* **31** 2940–2961.
[4] Bayraktar, E., Lu, F., Maggioni, M., Wu, R. and Yang, S. (2026). Supplement to "Probabilistic cellular automata with local transition matrices: synchronization, ergodicity, and inference." https://doi.org/10.3150/25-BEJ1892SUPP
[5] Bérard, J. (2023). Coupling from the past for exponentially ergodic one-dimensional probabilistic cellular automata. *Electron. J. Probab.* **28** 1–17.
[6] Casse, J. (2023). Ergodicity of some probabilistic cellular automata with binary alphabet via random walks. *Electron. J. Probab.* **28** 1–17.
[7] Cattiaux, P., Delebecque, F. and Pédèches, L. (2018). Stochastic Cucker–Smale models: Old and new. *Ann. Appl. Probab.* **28** 3239–3286.
[8] Cho, G.E. and Meyer, C.D. (2001). Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra Appl.* **335** 137–150.
[9] Cucker, F. and Mordecki, E. (2008). Flocking in noisy environments. *J. Math. Pures Appl.* **89** 278–296.
[10] Dawson, D.A. (1975). Information flow in graphs. *Stochastic Process. Appl.* **3** 137–151.
[11] Della Maestra, L. and Hoffmann, M. (2022). Nonparametric estimation for interacting particle systems: McKean–Vlasov models. *Probab. Theory Related Fields* **182** 551–613.

[12] Della Maestra, L. and Hoffmann, M. (2023). The LAN property for McKean–Vlasov models in a mean-field regime. *Stochastic Process. Appl.* **155** 109–146.
[13] Durrett, R. (2007). *Random Graph Dynamics 200*. Cambridge University Press.
[14] Föllmer, H. and Horst, U. (2001). Convergence of locally and globally interacting Markov chains. *Stochastic Process. Appl.* **96** 99–121.
[15] Grimmett, G. (2018). *Probability on Graphs: Random Processes on Graphs and Lattices 8*. Cambridge University Press.
[16] Haviv, M. and Van der Heyden, L. (1984). Perturbation bounds for the stationary probabilities of a finite Markov chain. *Adv. Appl. Probab.* **16** 804–818.
[17] Lacker, D., Ramanan, K. and Wu, R. (2021). Locally interacting diffusions as Markov random fields on path space. *Stochastic Process. Appl.* **140** 81–114.
[18] Lang, Q., Wang, X., Lu, F. and Maggioni, M. (2024). Interacting particle systems on networks: Joint inference of the network and the interaction kernel. arXiv Preprint. Available at arXiv:2402.08412.
[19] Lebowitz, J.L., Maes, C. and Speer, E.R. (1990). Statistical mechanics of probabilistic cellular automata. *J. Stat. Phys.* **59** 117–170.
[20] Li, Z. and Lu, F. (2023). On the coercivity condition in the learning of interacting particle systems. *Stoch. Dyn.* **23** 2340003.
[21] Li, Z., Lu, F., Maggioni, M., Tang, S. and Zhang, C. (2021). On the identifiability of interaction functions in systems of interacting particles. *Stochastic Process. Appl.* **132** 135–163.
[22] Liggett, T.M. (1985). *Interacting Particle Systems 2*. Springer.
[23] Louis, P.Y. and Nardi, F.R. (2018). Probabilistic cellular automata. *Emerg. Complex. Comput.* **27**.
[24] Lu, F., Maggioni, M. and Tang, S. (2022). Learning interaction kernels in stochastic systems of interacting particles from multiple trajectories. *Found. Comput. Math.* **22** 1013–1067.
[25] Lu, F., Zhong, M., Tang, S. and Maggioni, M. (2019). Nonparametric inference of interaction laws in systems of agents from trajectory data. *Proc. Natl. Acad. Sci. USA* **116** 14424–14433.
[26] Moran, P.A.P. (1958). Random processes in genetics. In *Math. Proc. Camb. Philos. Soc* **54** 60–71.
[27] Norris, J.R. (1998). *Markov Chains 2*. Cambridge University Press.
[28] Schweitzer, P.J. (1968). Perturbation theory and finite Markov chains. *J. Appl. Probab.* **5** 401–413.
[29] Seneta, E. (1988). Perturbation of the stationary distribution measured by ergodicity coefficients. *Adv. Appl. Probab.* **20** 228–230.
[30] Toom, A. (1994). On critical phenomena in interacting growth systems. Part I: General. *J. Stat. Phys.* **74** 91–109.

Bernoulli **32**(2), 2026, 1098–1121
https://doi.org/10.3150/25-BEJ1893

Moderate deviation principles for a reaction diffusion model in non-equilibrium

LINJIE ZHAO[a]

School of Mathematics and Statistics, and Hubei Key Laboratory of Engineering Modeling and Scientific Computing, Huazhong University of Science and Technology, Wuhan, China, [a]*linjie_zhao@hust.edu.cn*

We study moderate deviations from hydrodynamic limits of a reaction diffusion model. The process is defined as the superposition of the symmetric exclusion process with a Glauber dynamics. When the process starts from a product measure with a constant density, which is a non-equilibrium measure for the process, we prove that the re-scaled density fluctuation field satisfies the moderate deviation principle. Our proof relies on the so-called main lemma developed in (Jara and Menezes (2018); *Markov Process. Related Fields* **26** (2020) 95–124).

Keywords: Hydrodynamic limits; moderate deviations; non-equilibrium; reaction diffusion model

References

De Masi, A., Ferrari, P.A. and Lebowitz, J.L. (1986). Reaction-diffusion equations for interacting particle systems. *J. Stat. Phys.* **44** 589–644.

Farfan, J., Landim, C. and Tsunoda, K. (2019). Static large deviations for a reaction–diffusion model. *Probab. Theory Related Fields* **174** 49–101.

Gao, F. and Quastel, J. (2003). Moderate deviations from the hydrodynamic limit of the symmetric exclusion process. *Sci. China Ser. A* **46** 577–592.

Gonçalves, P., Jara, M., Marinho, R. and Menezes, O. (2024). CLT for NESS of a reaction-diffusion model. *Probab. Theory Related Fields* **190** 337–377.

Jara, M. and Menezes, O. (2018). Non-equilibrium fluctuations of interacting particle systems. arXiv preprint. Available at arXiv:1810.09526.

Jara, M. and Menezes, O. (2020). Non-equilibrium fluctuations for a reaction-diffusion model via relative entropy. *Markov Process. Related Fields* **26** 95–124.

Jona-Lasinio, G., Landim, C. and Vares, M. (1993). Large deviations for a reaction diffusion model. *Probab. Theory Related Fields* **97** 339–361.

Kipnis, C. and Landim, C. (2013). *Scaling Limits of Interacting Particle Systems 320*. Springer Science & Business Media.

Landim, C. and Tsunoda, K. (2018). Hydrostatics and dynamical large deviations for a reaction-diffusion model. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 51–74.

Wang, X. and Gao, F. (2006). Moderate deviations from hydrodynamic limit of a Ginzburg-Landau model. *Acta Math. Sci.* **26** 691–701.

Xu, L. and Zhao, L. (2023). Equilibrium perturbations for stochastic interacting systems. *Electron. J. Probab.* **28** 1–30.

Xue, X. (2024). Nonequilibrium moderate deviations from hydrodynamics of the simple symmetric exclusion process. *Electron. Commun. Probab.* **29** 1–16.

Xue, X. and Zhao, L. (2021). Moderate deviations for the SSEP with a slow bond. *J. Stat. Phys.* **182** 1–26.

Zhao, L. (2024). Moderate deviation principles for the WASEP. arXiv preprint. Available at arXiv:2405.16151.

Bernoulli **32**(2), 2026, 1122–1141
https://doi.org/10.3150/25-BEJ1894

Principles of statistical inference in online problems

MAN FUNG LEUNG[1,2,a] and KIN WAI CHAN[2,b]

[1]*Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA,* [a]*mfleung2@illinois.edu*
[2]*Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong SAR,* [b]*kinwaichan@sta.cuhk.edu.hk*

To investigate a dilemma of statistical and computational efficiency faced by long-run variance estimators, we propose a decomposition of kernel weights in a quadratic form and some online inference principles. These proposals allow us to characterize efficient online long-run variance estimators. Our asymptotic theory and simulations show that this principle-driven approach leads to online estimators with a uniformly lower mean squared error than all existing works. We also discuss practical enhancements such as mini-batch and automatic updates to handle fast streaming data and optimal parameters tuning. Beyond variance estimation, we consider the proposals in the context of online quantile regression, online change point detection, Markov chain Monte Carlo convergence diagnosis, and stochastic approximation. Substantial improvements in computational cost and finite-sample statistical properties are observed when we apply our principle-driven variance estimator to original and modified inference procedures.

Keywords: Long-run variance; nonparametric estimation; online learning; recursive estimation; streaming data

References

Andrews, D.W.K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59** 817–858. MR1106513 https://doi.org/10.2307/2938229

Bäßler, D., Kortus, T. and Gühring, G. (2022). Unsupervised anomaly detection in multivariate time series with online evolving spiking neural networks. *Mach. Learn.* **111** 1377–1408. MR4423163 https://doi.org/10.1007/s10994-022-06129-4

Chan, K.W. (2022). Optimal difference-based variance estimators in time series: A general framework. *Ann. Statist.* **50** 1376–1400. MR4441124 https://doi.org/10.1214/21-aos2154

Chan, K.W. and Yau, C.Y. (2017a). Automatic optimal batch size selection for recursive estimators of time-average covariance matrix. *J. Amer. Statist. Assoc.* **112** 1076–1089. MR3735361 https://doi.org/10.1080/01621459.2016.1189337

Chan, K.W. and Yau, C.Y. (2017b). High-order corrected estimator of asymptotic variance with optimal bandwidth. *Scand. J. Stat.* **44** 866–898. MR3730019 https://doi.org/10.1111/sjos.12279

Chen, X., Liu, W. and Zhang, Y. (2019). Quantile regression under memory constraint. *Ann. Statist.* **47** 3244–3273. MR4025741 https://doi.org/10.1214/18-AOS1777

Chen, X., Lee, J.D., Tong, X.T. and Zhang, Y. (2020). Statistical inference for model parameters in stochastic gradient descent. *Ann. Statist.* **48** 251–273. MR4065161 https://doi.org/10.1214/18-AOS1801

Flegal, J.M. and Jones, G.L. (2010). Batch means and spectral variance estimators in Markov chain Monte Carlo. *Ann. Statist.* **38** 1034–1070. MR2604704 https://doi.org/10.1214/09-AOS735

Flegal, J.M., Hughes, J., Vats, D., Dai, N., Gupta, K. and Maji, U. (2021). mcmcse: Monte Carlo standard errors for MCMC, Riverside, CA, and Kanpur, India R package version 1.5-0.

Gösmann, J., Kley, T. and Dette, H. (2021). A new approach for open-end sequential change point monitoring. *J. Time Series Anal.* **42** 63–84. MR4192928 https://doi.org/10.1111/jtsa.12555

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109. MR3363437 https://doi.org/10.1093/biomet/57.1.97

Huang, Y., Chen, X. and Wu, W.B. (2014). Recursive nonparametric estimation for time series. *IEEE Trans. Inf. Theory* **60** 1301–1312. MR3164976 https://doi.org/10.1109/TIT.2013.2292813

Jentsch, C. and Politis, D.N. (2015). Covariance matrix estimation and linear process bootstrap for multivariate time series of possibly increasing dimension. *Ann. Statist.* **43** 1117–1140. MR3346699 https://doi.org/10.1214/14-AOS1301

Jones, G.L., Haran, M., Caffo, B.S. and Neath, R. (2006). Fixed-width output analysis for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **101** 1537–1547. MR2279478 https://doi.org/10.1198/016214506000000492

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

Lee, S., Liao, Y., Seo, M.H. and Shin, Y. (2024). Fast inference for quantile regression with tens of millions of observations. *J. Econometrics* **249** 105673. https://doi.org/10.1016/j.jeconom.2024.105673

Leung, M.F. and Chan, K.W. (2026). Supplement to "Principles of statistical inference in online problems." https://doi.org/10.3150/25-BEJ1894SUPPA, https://doi.org/10.3150/25-BEJ1894SUPPB

Ljung, L. and Söderström, T. (1983). *Theory and Practice of Recursive Identification. MIT Press Series in Signal Processing, Optimization, and Control* **4**. Cambridge, Mass.–London: MIT Press. MR0719192

Maleki, S., Maleki, S. and Jennings, N.R. (2021). Unsupervised anomaly detection with LSTM autoencoders using statistical data-filtering. *Appl. Soft Comput.* **108** 107443. https://doi.org/10.1016/j.asoc.2021.107443

Meketon, M.S. and Schmeiser, B. (1984). Overlapping batch means: Something for nothing? In *Proceedings of the Winter Simulation Conference. WSC'84* 226–230.

Montiel, J., Halford, M., Mastelini, S.M., Bolmier, G., Sourty, R., Vaysse, R., Zouitine, A., Gomes, H.M., Read, J., Abdessalem, T. and Bifet, A. (2021). River: Machine learning for streaming data in Python. *J. Mach. Learn. Res.* **22** 1–8.

Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.* **28** 329–348. MR88833 https://doi.org/10.1214/aoms/1177706962

Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22** 400–407. MR42668 https://doi.org/10.1214/aoms/1177729586

Vats, D. and Flegal, J.M. (2022). Lugsail lag windows for estimating time-average covariance matrices. *Biometrika* **109** 735–750. MR4472845 https://doi.org/10.1093/biomet/asab049

Welford, B.P. (1962). Note on a method for calculating corrected sums of squares and products. *Technometrics* **4** 419–420. MR143307 https://doi.org/10.2307/1266577

Wickham, H. (2019). *Advanced R*. Chapman and Hall/CRC. https://doi.org/10.1201/9781351201315

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

Wu, W.B. (2009). Recursive estimation of time-average variance constants. *Ann. Appl. Probab.* **19** 1529–1552. MR2538079 https://doi.org/10.1214/08-AAP587

Wu, W.B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4** 207–226. MR2812816 https://doi.org/10.4310/SII.2011.v4.n2.a15

Xiao, H. and Wu, W.B. (2011). A single-pass algorithm for spectrum estimation with fast convergence. *IEEE Trans. Inf. Theory* **57** 4720–4731. MR2840487 https://doi.org/10.1109/TIT.2011.2145610

Yau, C.Y. and Chan, K.W. (2016). New recursive estimators of the time-average variance constant. *Stat. Comput.* **26** 609–627. MR3489860 https://doi.org/10.1007/s11222-015-9548-7

Yeh, Y. and Schmeiser, B. (2000). Simulation output analysis via dynamic batch means. In *In Proceedings of the Winter Simulation Conference* **1** 637–645. IEEE. https://doi.org/10.1109/wsc.2000.899775

Zeileis, A., Köll, S. and Graham, N. (2020). Various versatile variances: An object-oriented implementation of clustered covariances in R. *J. Stat. Softw.* **95** 1–36. https://doi.org/10.18637/jss.v095.i01

Zhu, W., Chen, X. and Wu, W.B. (2023). Online covariance matrix estimation in stochastic gradient descent. *J. Amer. Statist. Assoc.* **118** 393–404. MR4571129 https://doi.org/10.1080/01621459.2021.1933498

Bernoulli **32**(2), 2026, 1142–1159
https://doi.org/10.3150/25-BEJ1895

Diffusion processes as Wasserstein gradient flows via stochastic control of the volatility matrix

BERTRAM TSCHIDERER[a]

Faculty of Mathematics, University of Vienna, Vienna, Austria, [a]*bertram.tschiderer@univie.ac.at*

We study a class of time-homogeneous diffusion processes on $\mathbb{R}^n$ that share a common invariant measure but differ in their volatility matrices. In the Euclidean setting, we show that when the volatility matrix is the identity, the time-marginal distributions evolve as an entropic gradient flow in the quadratic Wasserstein space. This result recovers the gradient flow formulation of the Fokker–Planck equation, as established by Jordan, Kinderlehrer, and Otto. When $\mathbb{R}^n$ is equipped with a Riemannian metric, we prove that the diffusion process becomes a gradient flow in the Wasserstein space induced by the metric. This characterization holds when the volatility matrix is the inverse of the metric tensor. Our approach combines stochastic control of the diffusion coefficient and time-reversal techniques. These findings align with results by Lisini, which build on the metric theory of Ambrosio, Gigli, and Savaré, and connect to Fathi's work on large deviations for diffusion processes via gradient flows.

Keywords: Diffusion process; gradient flow; relative entropy dissipation; Riemannian metric; stochastic control; volatility matrix; Wasserstein distance

References

[1] Adams, S., Dirr, N., Peletier, M. and Zimmer, J. (2013). Large deviations and gradient flows. *Philos. Trans. R. Soc. A* **371**.

[2] Ambrosio, L. and Gigli, N. (2013). A user's guide to optimal transport. In *Modelling and Optimisation of Flows on Networks*. *Lecture Notes in Math.* **2062** 1–155. Berlin: Springer.

[3] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. *Lect. Math. ETH Zürich.* Basel: Birkhäuser.

[4] Arnold, A., Markowich, P., Toscani, G. and Unterreiter, A. (2001). On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations. *Comm. Partial Differential Equations* **26** 43–100.

[5] Backhoff-Veraguas, J., Conforti, G., Gentil, I. and Léonard, C. (2020). The mean field Schrödinger problem: Ergodic behavior, entropy estimates and functional inequalities. *Probab. Theory Related Fields* **178** 475–530.

[6] Bakry, D. and Émery, M. (1985). Diffusions hypercontractives. In *Sémin. Probab. XIX 1983/84* (J. Azéma and M. Yor, eds.). *Lecture Notes in Math.* **1123** 177–206. Berlin: Springer.

[7] Bayraktar, E., Feng, Q. and Li, W. (2024). Exponential entropy dissipation for weakly self-consistent Vlasov–Fokker–Planck equations. *J. Nonlinear Sci.* **34**.

[8] Carlen, E.A. and Gangbo, W. (2003). Constrained steepest descent in the 2-Wasserstein metric. *Ann. Math.* **157** 807–846.

[9] Carrillo, J.A., McCann, R.J. and Villani, C. (2003). Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoam.* **19** 971–1018.

[10] Carrillo, J.A., McCann, R.J. and Villani, C. (2006). Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.* **179** 217–263.

[11] Casella, G. and Lehmann, E.L. (1998). *Theory of Point Estimation*, 2nd ed. *Springer Texts in Statistics*. New York: Springer.

[12] Conforti, G. and Léonard, C. (2022). Time reversal of Markov processes with jumps under a finite entropy condition. *Stochastic Process. Appl.* **144** 85–124.

[13] Cover, T.M. and Thomas, J.A. (2006). *Elements of Information Theory*, 2nd ed. *Wiley Ser. Telecommun. Signal Process.* Hoboken, NJ: John Wiley & Sons.

[14] Dai Pra, P. and Pavon, M. (1989). Variational path-integral representations for the density of a diffusion process. *Stochastics* **26** 205–226.

[15] De Giorgi, E. (1993). New problems on minimizing movements. In *Boundary Value Problems for Partial Differential Equations and Applications — Dedicated to E. Magenes* (C. Baiocchi and J.L. Lions, eds.). *Research Notes in Applied Mathematics* **29** 81–98. Paris: Masson.

[16] De Giorgi, E., Marino, A. and Tosques, M. (1980). Problemi di evoluzione in spazi metrici e curve di massima pendenza. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Ser. 8* **68** 180–187.

[17] Degiovanni, M., Marino, A. and Tosques, M. (1985). Evolution equations with lack of convexity. *Nonlinear Anal., Theory Methods Appl.* **9** 1401–1443.

[18] Erbar, M. (2010). The heat equation on manifolds as a gradient flow in the Wasserstein space. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 1–23.

[19] Fathi, M. (2016). A gradient flow approach to large deviations for diffusion processes. *J. Math. Pures Appl.* **106** 957–993.

[20] Feng, Q. and Li, W. (2023). Entropy dissipation for degenerate stochastic differential equations via sub-Riemannian density manifold. *Entropy* **25** 786.

[21] Feng, Q. and Li, W. (2025). Hypoelliptic entropy dissipation for stochastic differential equations. Available at arXiv:2102.00544.

[22] Föllmer, H. (1985). An entropy approach to the time reversal of diffusion processes. In *Stochastic Differential Systems — Filtering and Control* (M. Métivier and É. Pardoux, eds.). *Lect. Notes Control Inf. Sci.* **69** 156–163. Berlin: Springer.

[23] Föllmer, H. (1986). Time reversal on Wiener space. In *Stochastic Processes — Mathematics and Physics* (S.A. Albeverio, P. Blanchard and L. Streit, eds.). *Lecture Notes in Math.* **1158** 119–129. Berlin: Springer.

[24] Fontbona, J. and Jourdain, B. (2016). A trajectorial interpretation of the dissipations of entropy and Fisher information for stochastic differential equations. *Ann. Probab.* **44** 131–170.

[25] Friedman, A. (1975). *Stochastic Differential Equations and Applications — Volume 1*. *Probab. Math. Statist. Ser. Monogr. Textb.* **28**. New York: Academic Press.

[26] Gardiner, C. (2009). *Stochastic Methods — a Handbook for the Natural and Social Sciences*, 4th ed. *Springer Ser. Synergetics* **13**. Berlin: Springer.

[27] Haussmann, U.G. and Pardoux, É. (1986). Time reversal of diffusions. *Ann. Probab.* **14** 1188–1205.

[28] 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.

[29] Jahnel, B. and Köppl, J. (2023). Trajectorial dissipation of Φ-entropies for interacting particle systems. *J. Stat. Phys.* **190** 119.

[30] Jordan, R., Kinderlehrer, D. and Otto, F. (1998). The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29** 1–17.

[31] Karatzas, I., Maas, J. and Schachermayer, W. (2021). Trajectorial dissipation and gradient flow for the relative entropy in Markov chains. *Commun. Inf. Syst.* **21** 481–536.

[32] Karatzas, I., Schachermayer, W. and Tschiderer, B. (2020). Trajectorial Otto calculus. Available at arXiv: 1811.08686.

[33] Karatzas, I., Schachermayer, W. and Tschiderer, B. (2022). A trajectorial approach to the gradient flow properties of Langevin–Smoluchowski diffusions. *Theory Probab. Appl.* **66** 668–707.

[34] Karatzas, I. and Tschiderer, B. (2022). A variational characterization of Langevin–Smoluchowski diffusions. In *Stochastic Analysis, Filtering, and Stochastic Optimization — a Commemorative Volume to Honor Mark H.A. Davis's Contributions* (G. Yin and T. Zariphopoulou, eds.) 239–265. Cham: Springer.

[35] Kim, D. and Yeung, L.C. (2024). A trajectorial approach to entropy dissipation for degenerate parabolic equations. *Bernoulli* **30** 2253–2274.

[36] Kolmogorov, A.N. (1931). Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* **104** 415–458.

[37] Lee, J.M. (2018). *Introduction to Riemannian Manifolds*, 2nd ed. *Grad. Texts in Math.* **176**. Switzerland: Springer International Publishing.

[38] Léonard, C. (2014). Some properties of path measures. In *Sémin. Probab. XLVI* (C. Donati-Martin, A. Lejay and A. Rouault, eds.). *Lecture Notes in Math.* **2123** 207–230. Switzerland: Springer International Publishing.

[39] Lisini, S. (2009). Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. *ESAIM Control Optim. Calc. Var.* **15** 712–740.

[40] Lott, J. (2008). Some geometric calculations on Wasserstein space. *Comm. Math. Phys.* **277** 423–437.

[41] Lott, J. and Villani, C. (2009). Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math.* **169** 903–991.

[42] McCann, R.J. (2001). Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11** 589–608.

[43] Meyer, P.A. (1994). Sur une transformation du mouvement brownien due à Jeulin et Yor. In *Sémin. Probab. XXVIII* (J. Azéma, M. Yor and P.A. Meyer, eds.). *Lecture Notes in Math.* **1583** 98–101. Berlin: Springer.

[44] Nelson, E. (2001). *Dynamical Theories of Brownian Motion*, 2nd ed. *Math. Notes*. Princeton, NJ: Princeton University Press.

[45] Ohta, S. (2009). Gradient flows on Wasserstein spaces over compact Alexandrov spaces. *Amer. J. Math.* **131** 475–516.

[46] Otto, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations* **26** 101–174.

[47] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400.

[48] Pardoux, É. (1986). Grossissement d'une filtration et retournement du temps d'une diffusion. In *Sémin. Probab. XX 1984/85* (J. Azéma and M. Yor, eds.). *Lecture Notes in Math.* **1204** 48–55. Berlin: Springer.

[49] Pavon, M. (1989). Stochastic control and nonequilibrium thermodynamical systems. *Appl. Math. Optim.* **19** 187–202.

[50] Risken, H. (1996). *The Fokker–Planck Equation — Methods of Solution and Applications*, 2nd ed. *Springer Ser. Synergetics* **18**. Berlin: Springer.

[51] Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians — Calculus of Variations, PDEs, and Modeling*. *Progress in Nonlinear Differential Equations and Their Applications* **87**. Switzerland: Birkhäuser.

[52] Schuss, Z. (1980). Singular perturbation methods in stochastic differential equations of mathematical physics. *SIAM Rev.* **22** 119–155.

[53] Sturm, K.T. (2006). On the geometry of metric measure spaces I. *Acta Math.* **196** 65–131.

[54] Sturm, K.T. (2006). On the geometry of metric measure spaces II. *Acta Math.* **196** 133–177.

[55] Tschiderer, B. and Yeung, L.C. (2023). A trajectorial approach to relative entropy dissipation of McKean–Vlasov diffusions: Gradient flows and HWBI inequalities. *Bernoulli* **29** 725–756.

[56] Villani, C. (2003). *Topics in Optimal Transportation*. *Grad. Stud. Math.* **58**. Providence, RI: American Mathematical Society.

[57] Villani, C. (2009). *Optimal Transport — Old and New*. *Grundlehren Math. Wiss.* **338**. Berlin: Springer.

Bernoulli **32**(2), 2026, 1160–1185
https://doi.org/10.3150/25-BEJ1897

Scaling limit for the cover time of the λ-biased random walk on a binary tree with $\lambda < 1$

DAVID A. CROYDON[a]

Research Institute for Mathematical Sciences, Kyoto University, Kyoto, Japan, [a]*croydon@kurims.kyoto-u.ac.jp*

The λ-biased random walk on a binary tree of depth n is the continuous-time Markov chain that has unit mean holding times and, when at a vertex other than the root or a leaf of the tree in question, has a probability of jumping to the parent vertex that is λ times the probability of jumping to a particular child. (From the root, it chooses one of the two children with equal probability.) For this process, when $\lambda < 1$, we derive an $n \to \infty$ scaling limit for the cover time, that is, the time taken to visit every vertex. The distributional limit is described in terms of a jump process on a Cantor set that can be seen as the asymptotic boundary of the tree. This conclusion complements previous results obtained when $\lambda \geq 1$.

Keywords: Binary tree; Cantor set; cover time; jump process; random walk

References

[1] Abe, Y. (2014). Cover times for sequences of reversible Markov chains on random graphs. *Kyoto J. Math.* **54** 555–576. MR3263552 https://doi.org/10.1215/21562261-2693442

[2] Aldous, D.J. (1991). Random walk covering of some special trees. *J. Math. Anal. Appl.* **157** 271–283. MR1109456 https://doi.org/10.1016/0022-247X(91)90149-T

[3] Aldous, D.J. (1991). Threshold limits for cover times. *J. Theoret. Probab.* **4** 197–211. MR1088401 https://doi.org/10.1007/BF01047002

[4] Andriopoulos, G., Croydon, D.A., Margarint, V. and Menard, L. (2024). On the cover time of Brownian motion on the Brownian continuum random tree. Available at arXiv:2410.03922.

[5] Athreya, S., Eckhoff, M. and Winter, A. (2013). Brownian motion on $\mathbb{R}$-trees. *Trans. Amer. Math. Soc.* **365** 3115–3150. MR3034461 https://doi.org/10.1090/S0002-9947-2012-05752-7

[6] Athreya, S., Löhr, W. and Winter, A. (2017). Invariance principle for variable speed random walks on trees. *Ann. Probab.* **45** 625–667. MR3630284 https://doi.org/10.1214/15-AOP1071

[7] Bai, T. (2020). On the cover time of λ-biased walk on supercritical Galton-Watson trees. *Stoch. Process. Appl.* **130** 6863–6879. MR4158805 https://doi.org/10.1016/j.spa.2020.07.001

[8] Barlow, M.T. (2017). *Random Walks and Heat Kernels on Graphs*. *London Mathematical Society Lecture Note Series* **438**. Cambridge: Cambridge University Press. MR3616731 https://doi.org/10.1017/9781107415690

[9] Baxter, M. (1992). Markov processes on the boundary of the binary tree. In *Séminaire de Probabilités, XXVI*. *Lecture Notes in Math.* **1526** 210–224. Berlin: Springer. MR1231996 https://doi.org/10.1007/BFb0084323

[10] Belius, D., Rosen, J. and Zeitouni, O. (2019). Barrier estimates for a critical Galton-Watson process and the cover time of the binary tree. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 127–154. MR3901643 https://doi.org/10.1214/17-aihp878

[11] Bramson, M. and Zeitouni, O. (2009). Tightness for a family of recursion equations. *Ann. Probab.* **37** 615–653. MR2510018 https://doi.org/10.1214/08-AOP414

[12] Cortines, A., Louidor, O. and Saglietti, S. (2021). A scaling limit for the cover time of the binary tree. *Adv. Math.* **391** Paper No. 107974. MR4303734 https://doi.org/10.1016/j.aim.2021.107974

[13] Croydon, D.A. (2008). Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 987–1019. MR2469332 https://doi.org/10.1214/07-AIHP153

1350-7265

[14] Croydon, D.A. (2010). Scaling limits for simple random walks on random ordered graph trees. *Adv. Appl. Probab.* **42** 528–558. MR2675115 https://doi.org/10.1239/aap/1275055241
[15] Croydon, D.A. (2015). Moduli of continuity of local times of random walks on graphs in terms of the resistance metric. *Trans. Lond. Math. Soc.* **2** 57–79. MR3355578 https://doi.org/10.1112/tlms/tlv003
[16] Croydon, D.A. (2017). An introduction to stochastic processes associated with resistance forms and their scaling limits. RIMS Kôkyûroku 2030 paper no. 1.
[17] Croydon, D.A. (2018). Scaling limits of stochastic processes associated with resistance forms. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 1939–1968. MR3865663 https://doi.org/10.1214/17-AIHP861
[18] Croydon, D.A., Hambly, B.M. and Kumagai, T. (2017). Time-changes of stochastic processes associated with resistance forms. *Electron. J. Probab.* **22** Paper No. 82. MR3718710 https://doi.org/10.1214/17-EJP99
[19] Ding, J. (2014). Asymptotics of cover times via Gaussian free fields: Bounded-degree graphs and general trees. *Ann. Probab.* **42** 464–496. MR3178464 https://doi.org/10.1214/12-AOP822
[20] Ding, J., Lee, J.R. and Peres, Y. (2012). Cover times, blanket times, and majorizing measures. *Ann. Math.* **175** 1409–1471. MR2912708 https://doi.org/10.4007/annals.2012.175.3.8
[21] Ding, J. and Zeitouni, O. (2012). A sharp estimate for cover times on binary trees. *Stoch. Process. Appl.* **122** 2117–2133. MR2921974 https://doi.org/10.1016/j.spa.2012.03.008
[22] Fukushima, M., Oshima, Y. and Takeda, M. (2011). *Dirichlet Forms and Symmetric Markov Processes*, extended ed. *De Gruyter Studies in Mathematics* **19**. Berlin: Walter de Gruyter & Co. MR2778606
[23] Kallenberg, O. (2002). *Foundations of Modern Probability*, second ed. *Probability and Its Applications (New York)*. New York: Springer-Verlag. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[24] Kigami, J. (1995). Harmonic calculus on limits of networks and its application to dendrites. *J. Funct. Anal.* **128** 48–86. MR1317710 https://doi.org/10.1006/jfan.1995.1023
[25] Kigami, J. (2001). *Analysis on Fractals*. *Cambridge Tracts in Mathematics* **143**. Cambridge: Cambridge University Press. MR1840042 https://doi.org/10.1017/CBO9780511470943
[26] Kigami, J. (2010). Dirichlet forms and associated heat kernels on the Cantor set induced by random walks on trees. *Adv. Math.* **225** 2674–2730. MR2680180 https://doi.org/10.1016/j.aim.2010.04.029
[27] Kigami, J. (2012). Resistance forms, quasisymmetric maps and heat kernel estimates. *Mem. Amer. Math. Soc.* **216 vi+132**. MR2919892 https://doi.org/10.1090/S0065-9266-2011-00632-5
[28] Kigami, J. (2013). Transitions on a noncompact Cantor set and random walks on its defining tree. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1090–1129. MR3127915 https://doi.org/10.1214/12-AIHP496
[29] Krebs, W.B. (1995). Brownian motion on the continuum tree. *Probab. Theory Related Fields* **101** 421–433. MR1324094 https://doi.org/10.1007/BF01200505
[30] Levin, D.A., Peres, Y. and Wilmer, E.L. (2009). *Markov Chains and Mixing Times*. Providence, RI: American Mathematical Society. MR2466937 https://doi.org/10.1090/mbk/058
[31] Noda, R. (2023). Convergence of local times of stochastic processes associated with resistance forms. Available at arXiv:2305.13224.
[32] Talagrand, M. (1987). Regularity of Gaussian processes. *Acta Math.* **159** 99–149. https://doi.org/10.1007/BF02392556 MR906527
[33] Talagrand, M. (2001). Majorizing measures without measures. *Ann. Probab.* **29** 411–417. MR1825156 https://doi.org/10.1214/aop/1008956336
[34] Tokushige, Y. (2020). Jump processes on the boundaries of random trees. *Stoch. Process. Appl.* **130** 584–604. MR4046511 https://doi.org/10.1016/j.spa.2019.02.004
[35] Zhai, A. (2018). Exponential concentration of cover times. *Electron. J. Probab.* **23**. Paper No. 32. MR3785402 https://doi.org/10.1214/18-EJP149

Bernoulli **32**(2), 2026, 1186–1212
https://doi.org/10.3150/25-BEJ1898

Stein's method of moments on the sphere

ADRIAN FISCHER[1,a], ROBERT E. GAUNT[2,b] and YVIK SWAN[3,c]

[1] *Department of Statistics, University of Oxford, Oxford, UK,* [a]*adrian.fischer@stats.ox.ac.uk*
[2] *Department of Mathematics, The University of Manchester, Manchester, UK,* [b]*robert.gaunt@manchester.ac.uk*
[3] *Département de Mathématique, Université libre de Bruxelles, Brussels, Belgium,* [c]*yvik.swan@ulb.be*

We use Stein characterizations to obtain new moment-type estimators for the parameters of three classical spherical distributions (namely the Fisher-Bingham, the von Mises-Fisher, and the Watson distributions) in the i.i.d. case. This leads to explicit estimators which have good asymptotic properties (close to efficiency) and therefore provide interesting alternatives to classical maximum likelihood methods or more recent score matching estimators. We perform competitive simulation studies to assess the quality of the new estimators. Finally, the practical relevance of our estimators is illustrated on a real data application in spherical latent representations of handwritten numbers.

Keywords: Autoencoder; Fisher-Bingham distribution; point estimation; spherical distributions; Stein's method

References

Ameijeiras-Alonso, J. and Ley, C. (2022). Sine-skewed toroidal distributions and their application in protein bioinformatics. *Biostatistics* **23** 685–704.

Ameijeiras-Alonso, J., Benali, A., Crujeiras, R.M., Rodríguez-Casal, A. and Pereira, J.M. (2019). Fire seasonality identification with multimodality tests. *Ann. Appl. Stat.* **13** 2120–2139.

Bee, M., Benedetti, R. and Espa, G. (2017). Approximate maximum likelihood estimation of the Bingham distribution. *Comput. Statist. Data Anal.* **108** 84–96.

Chen, Y. and Tanaka, K. (2021). Maximum likelihood estimation of the Fisher–Bingham distribution via efficient calculation of its normalizing constant. *Stat. Comput.* **31** 40.

Christie, D. (2015). Efficient von Mises–Fisher concentration parameter estimation using Taylor series. *J. Stat. Comput. Simul.* **85** 3259–3265.

Cinelli, L.P., Marins, M.A., Da Silva, E.A.B. and Netto, S.L. (2021). *Variational Methods for Machine Learning with Applications to Deep Networks 15*. Springer.

Davidson, T.R., Falorsi, L., De Cao, N., Kipf, T. and Tomczak, J.M. (2018). Hyperspherical variational autoencoders. In *34th Conf. Uncertain. Artif. Intell. 2018, UAI 2018* 856–865. Association For Uncertainty in Artificial Intelligence (AUAI).

Dey, S. and Jana, N. (2022). Inference on parameters of Watson distributions and application to classification of observations. *J. Comput. Appl. Math.* **403** 113847.

Dortet-Bernadet, J.-L. and Wicker, N. (2008). Model-based clustering on the unit sphere with an illustration using gene expression profiles. *Biostatistics* **9** 66–80.

Ebner, B., Fischer, A., Gaunt, R.E., Picker, B. and Swan, Y. (2025). Stein's method of moments. *Scand. J. Stat.* **52** 1594–1624.

Fischer, A., Gaunt, R.E. and Swan, Y. (2025a). Stein's method of moments for truncated multivariate distributions. *Electron. J. Stat.* **19** 1784–1808.

Fischer, A., Gaunt, R.E. and Swan, Y. (2026b). Supplement to "Stein's method of moments on the sphere." https://doi.org/10.3150/25-BEJ1898SUPP

Fisher, N.I., Lewis, T. and Embleton, B.J. (1993). *Statistical Analysis of Spherical Data*. Cambridge University Press.

García-Portugués, E., Navarro-Esteban, P. and Cuesta-Albertos, J.A. (2023). On a projection-based class of uniformity tests on the hypersphere. *Bernoulli* **29** 181–204.

1350-7265

García-Portugués, E., Paindaveine, D. and Verdebout, T. (2020). On optimal tests for rotational symmetry against new classes of hyperspherical distributions. *J. Amer. Statist. Assoc.* **115** 1873–1887.

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

Ghosh, P., Chatterjee, D. and Banerjee, A. (2024). On the directional nature of celestial object's fall on the Earth (Part 1: Distribution of fireball shower, meteor fall, and crater on Earth's surface). *Mon. Not. R. Astron. Soc.* **531** 1294–1307.

Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* 249–256. JMLR Workshop and Conference Proceedings.

Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. MIT Press.

Gopal, S. and Yang, Y. (2014). von Mises-Fisher clustering models. In *Int. Conf. Mach. Learn.* 154–162. PMLR.

Hendriks, H. and Landsman, Z. (1996). Asymptotic behavior of sample mean location for manifolds. *Statist. Probab. Lett.* **26** 169–178.

Hendriks, H., Landsman, Z. and Ruymgaart, F. (1996). Asymptotic behavior of sample mean direction for spheres. *J. Multivariate Anal.* **59** 141–152.

Hillen, T., Painter, K.J., Swan, A.C. and Murtha, A.D. (2017). Moments of von Mises and Fisher distributions and applications. *Math. Biosci. Eng.* **14** 673–694.

Horn, R.A. and Johnson, C.R. (2013). *Matrix Analysis*, 2nd ed. Cambridge: Cambridge University Press.

Hornik, K. and Grün, B. (2014a). movMF: An R package for fitting mixtures of von Mises-Fisher distributions. *J. Stat. Softw.* **58** 1–31.

Hornik, K. and Grün, B. (2014b). On maximum likelihood estimation of the concentration parameter of von Mises–Fisher distributions. *Comput. Statist.* **29** 945–957.

Hyvärinen, A. (2007). Some extensions of score matching. *Comput. Statist. Data Anal.* **51** 2499–2512.

Hyvärinen, A. and Dayan, P. (2005). Estimation of non-normalized statistical models by score matching. *J. Mach. Learn. Res.* **6** 695–709.

Jupp, P. (1995). Some applications of directional statistics to astronomy. In *New Trends in Probability and Statistics* **3** 123–133. De Gruyter.

Kent, J.T., Ganeiber, A.M. and Mardia, K.V. (2018). A new unified approach for the simulation of a wide class of directional distributions. *J. Comput. Graph. Statist.* **27** 291–301.

Kent, J.T. and Hamelryck, T. (2005). Using the Fisher-Bingham distribution in stochastic models for protein structure. *Quantitative Biology, Shape Analysis, and Wavelets* **24** 57–60.

Kingma, D.P. and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv:1412.6980.

Kingma, D.P. and Welling, M. (2019). An introduction to variational autoencoders. *Found. Trends Mach. Learn.* **12** 307–392.

Kollo, T. (2005). *Advanced Multivariate Statistics with Matrices*. Springer.

Koyama, T., Nakayama, H., Nishiyama, K. and Takayama, N. (2014). Holonomic gradient descent for the Fisher–Bingham distribution on the d-dimensional sphere. *Comput. Statist.* **29** 661–683.

Kume, A. and Sei, T. (2018). On the exact maximum likelihood inference of Fisher–Bingham distributions using an adjusted holonomic gradient method. *Stat. Comput.* **28** 835–847.

Kume, A. and Wood, A.T. (2005). Saddlepoint approximations for the Bingham and Fisher–Bingham normalising constants. *Biometrika* **92** 465–476.

Landler, L., Ruxton, G.D. and Malkemper, E.P. (2018). Circular data in biology: Advice for effectively implementing statistical procedures. *Behav. Ecol. Sociobiol.* **72** 1–10.

Le, H., Lewis, A., Bharath, K. and Fallaize, C. (2024). A diffusion approach to Stein's method on Riemannian manifolds. *Bernoulli* **30** 1079–1104.

Lee, J.M. (2013). *Introduction to Smooth Manifolds*. Springer.

Ley, C., Reinert, G. and Swan, Y. (2017). Stein's method for comparison of univariate distributions. *Probab. Surv.* **14** 1–52.

Ley, C. and Swan, Y. (2013). Stein's density approach and information inequalities. *Electron. Commun. Probab.* **18** 1–14.

Ley, C. and Verdebout, T. (2017). *Modern Directional Statistics*. CRC Press.

Magnus, J.R. and Neudecker, H. (2019). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons.

Mardia, K.V. (1975). Distribution theory for the von Mises-Fisher distribution and its application. *Statistical Distributions for Scientific Work* **1** 113–30.

Mardia, K.V. (2024). Fisher's legacy of directional statistics, and beyond to statistics on manifolds. arXiv:2405.17919.

Mardia, K.V., Foldager, J.I. and Frellsen, J. (2018). Directional statistics in protein bioinformatics. In *Applied Directional Statistics 17–40* Chapman and Hall/CRC.

Mardia, K.V. and Jupp, P.E. (2000). *Directional Statistics*. John Wiley & Sons.

Mardia, K.V., Kent, J.T. and Laha, A.K. (2016). Score matching estimators for directional distributions. arXiv:1604.08470.

Merrifield, A., Myerscough, M.R. and Weber, N. (2006). Statistical tests for analysing directed movement of self-organising animal groups. *Math. Biosci.* **203** 64–78.

Mijoule, G., Raič, M., Reinert, G. and Swan, Y. (2023). Stein's density method for multivariate continuous distributions. *Electron. J. Probab.* **28** 1–40.

Nakayama, H., Nishiyama, K., Noro, M., Ohara, K., Sei, T., Takayama, N. and Takemura, A. (2011). Holonomic gradient descent and its application to the Fisher–Bingham integral. *Adv. in Appl. Math.* **47** 639–658.

Pewsey, A. and García-Portugués, E. (2021). Recent advances in directional statistics. *TEST* **30** 1–58.

Pewsey, A., Neuhäuser, M. and Ruxton, G.D. (2013). *Circular Statistics in R*. Oxford: Oxford University Press.

Sablica, L., Hornik, K. and Leydold, J. (2023). watson: fitting and simulating mixtures of Watson distributions R package version 0.3.

Sra, S. (2012). A short note on parameter approximation for von Mises-Fisher distributions: And a fast implementation of $I_s(x)$. *Comput. Statist.* **27** 177–190.

Sra, S. (2018). Directional statistics in machine learning: A brief review. In *Applied Directional Statistics: Modern Methods and Case Studies* **225**. New York: Chapman & Hall/CRC Press.

Sra, S. and Karp, D. (2013). The multivariate Watson distribution: Maximum-likelihood estimation and other aspects. *J. Multivariate Anal.* **114** 256–269.

Takasu, Y., Yano, K. and Komaki, F. (2018). Scoring rules for statistical models on spheres. *Statist. Probab. Lett.* **138** 111–115.

Tanabe, A., Fukumizu, K., Oba, S., Takenouchi, T. and Ishii, S. (2007). Parameter estimation for von Mises–Fisher distributions. *Comput. Statist.* **22** 145–157.

Wang, P., Wu, D., Chen, C., Liu, K., Fu, Y., Huang, J., Zhou, Y., Zhan, J. and Hua, X. (2023). Deep adaptive graph clustering via von Mises-Fisher distributions. *ACM Trans. Web* **18** 1–21.

Williams, D.J. and Liu, S. (2022). Score matching for truncated density estimation on a manifold. In *Topological, Algebraic and Geometric Learning Workshops 2022* 312–321. PMLR.

Zhao, D., Zhu, J. and Zhang, B. (2019). Latent variables on spheres for autoencoders in high dimensions. arXiv:1912.10233.

Bernoulli **32**(2), 2026, 1213–1242
https://doi.org/10.3150/25-BEJ1899

Monotone measure-transportation maps in Hilbert spaces, with statistical applications

ALBERTO GONZÁLEZ-SANZ[1,a], MARC HALLIN[2,3,b] and BODHISATTVA SEN[4,c]

To Juan Antonio Cuesta Albertos and Carlos Matrán Bea, on the occasion of their 70th birthday.

[1]*Department of Statistics, Columbia University, New York, USA,* [a]*ag4855@columbia.edu*
[2]*ECARES and Département de Mathématique Université libre de Bruxelles, Brussels, Belgium*
[3]*Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic,* [b]*mhallin@ulb.ac.be*
[4]*Department of Statistics, Columbia University, New York, USA,* [c]*bodhi@stat.columbia.edu*

The contribution of this work is twofold. The first part deals with a Hilbert-space version of McCann's celebrated result on the existence and uniqueness of monotone measure-preserving maps: given two probability measures P and Q on a separable Hilbert space $\mathcal{H}$ where P does not give mass to "small" sets (namely, *Lipschitz hypersurfaces*), we show, without imposing any moment assumptions, that there exists a gradient of convex function $\nabla\psi$ pushing P forward to Q. In case $\mathcal{H}$ is infinite-dimensional, P-a.e. uniqueness is not guaranteed, though; we show that uniqueness holds among all gradients of convex functions $\nabla\psi$ pushing P forward to Q for which the boundary of the domain of ψ has P-probability zero. This condition (hence the uniqueness of the gradient of convex function $\nabla\psi$ pushing P forward to Q) is automatically satisfied in the finite-dimensional case or when Q is boundedly supported (a natural assumption in several statistical applications). Furthermore, we establish stability results for transport maps in the sense of uniform convergence over compact "regularity sets." As a consequence, we obtain a central limit theorem for the fluctuations of the optimal quadratic transport cost in a separable Hilbert space. In the second part of the paper we consider several important statistical applications of our results—center-outward ranks and quantiles for Hilbert-space-valued data, nonparametric distribution-free testing, and the construction of quantile regions. We show that the measure-transportation-based ranks are distribution-free and maximal ancillary, while the corresponding quantile functions fully characterize the underlying probability measures. These are the first notions of ranks and quantiles in non-locally compact spaces satisfying these properties.

Keywords: Central limit theorem; Lipschitz hypersurfaces; McCann's theorem; measure transportation; nonparametric distribution-free testing; quantiles for functional data; stability and uniqueness of transport maps; Wasserstein distance

References

Adly, S., Attouch, H. and Rockafellar, R.T. (2023). Preservation or not of the maximally monotone property by graph-convergence. *J. Convex Anal.* **30** 413–440.
Ambrosio, L., Gigli, N. and Savare, G. (2005). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Basel: Birkhäuser. https://doi.org/10.1007/978-3-7643-8722-8
Anderson, R.D. and Klee, V.L. Jr. (1952). Convex functions and upper semi-continuous collections. *Duke Math. J.* **19** 349–357. https://doi.org/10.1215/S0012-7094-52-01935-2
Avella-Medina, M. and González-Sanz, A. (2025). On the breakdown point of transport-based quantiles. *Bernoulli*. To appear.
Bauschke, H.H. and Combettes, P.L. (2011). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Cham: Springer. https://doi.org/10.1007/978-1-4419-9467-7
Beer, G. (1993). *Topologies on Closed and Closed Convex Sets*. Dordrecht: Springer.
Bogachev, V. (2008). *Gaussian Measures*. Providence: American Mathematical Society.

1350-7265 This research was funded, in whole or in part, by Czech Science Foundation, 100785 and Fonds Thelam of the Fondation Roi Baudouin.

Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417.

Brezis, H. (2010). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. New York: Springer.

Chakraborty, A. and Chaudhuri, P. (2014a). On data depth in infinite-dimensional spaces. *Ann. Inst. Statist. Math.* **66** 303–324. MR3171407 https://doi.org/10.1007/s10463-013-0416-y

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

Chaudhuri, P. (1996). On a geometric notion of quantiles for multivariate data. *J. Amer. Statist. Assoc.* **91** 862–872.

Chernozhukov, V., Galichon, A., Hallin, M. and Henry, M. (2017). Monge-Kantorovich depth, quantiles, ranks and signs. *Ann. Statist.* **45** 223–256.

Chowdhury, J. and Chaudhuri, P. (2019). Nonparametric depth and quantile regression for functional data. *Bernoulli* **25** 395–423.

Cordero-Erausquin, D. and Figalli, A. (2019). Regularity of monotone transport maps between unbounded domains. *Discrete Contin. Dyn. Syst.* **39** 7101–7112.

Csörnyei, M. (1999). Aronszajn null and Gaussian null sets coincide. *Israel J. Math.* **111** 191–201.

Cuesta-Albertos, J.A., Fraiman, R. and Ransford, T. (2006). Random projections and goodness-of-fit tests in infinite-dimensional spaces. *Bull. Braz. Math. Soc. (N.S.)* **37** 477–501. MR2284883 https://doi.org/10.1007/s00574-006-0023-0

Cuesta-Albertos, J.A. and Matrán, C. (1989). Notes on the Wasserstein metric in Hilbert spaces. *Ann. Probab.* **17** 1264–1276.

Cuesta-Albertos, J.A., Matrán, C. and Tuero-Diáz, A. (1996). On lower bounds for the L_2-Wasserstein metric in a Hilbert space. *J. Theoret. Probab.* **9** 263–283.

Deb, N., Bhattacharya, B.B. and Sen, B. (2021). Pitman efficiency lower bounds for multivariate distribution-free tests based on optimal transport. arXiv preprint. arXiv:2104.01986.

Deb, N. and Sen, B. (2023). Multivariate rank-based distribution-free nonparametric testing using measure transportation. *J. Amer. Statist. Assoc.* **118** 192–207. MR4571116 https://doi.org/10.1080/01621459.2021.1923508

del Barrio, E. and González-Sanz, A. (2024). Regularity of center-outward distribution functions in non-convex domains. *Adv. Nonlinear Stud.* **24** 880–894.

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

del Barrio, E., González-Sanz, A. and Hallin, M. (2025). Nonparametric multiple-output center-outward quantile regression. *J. Amer. Statist. Assoc.* **120** 818–832. https://doi.org/10.1080/01621459.2024.2366029

del Barrio, E., González-Sanz, A. and Loubes, J.-M. (2024a). A central limit theorem for semidiscrete Wasserstein distances. *Bernoulli* **30** 554–580.

del Barrio, E., González-Sanz, A. and Loubes, J.-M. (2024b). Central limit theorems for general transportation costs. *Ann. Inst. Henri Poincaré Probab. Stat.* **60**.

del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** 926–951. https://doi.org/10.1214/18-AOP1275

Dutta, S., Ghosh, A.K. and Chaudhuri, P. (2011). Some intriguing properties of Tukey's half-space depth. *Bernoulli* **17**. https://doi.org/10.3150/10-bej322

Edgar, G.A. (1977). Measurability in a Banach space. *Indiana Univ. Math. J.* **26** 663–677.

Feyel, D. and Üstünel, A.S. (2003). Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theory Related Fields* **128** 347–385. https://doi.org/10.1007/s00440-003-0307-x

Figalli, A. (2017). *The Monge-Ampère Equation and its Applications*. *Zurich Lectures in Advanced Mathematics*. Zurich: European Mathematical Society (EMS).

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

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

Ghiglietti, A., Ieva, F. and Paganoni, A.M. (2017). Statistical inference for stochastic processes: Two-sample hypothesis tests. *J. Statist. Plann. Inference* **180** 49–68. MR3554193 https://doi.org/10.1016/j.jspi.2016.08.004

Ghosal, P. and Sen, B. (2022). Multivariate ranks and quantiles using optimal transport: Consistency, rates and nonparametric testing. *Ann. Statist.* **50** 1012–1037. https://doi.org/10.1214/21-AOS2136

Gijbels, I. and Nagy, S. (2017). On a general definition of depth for functional data. *Statist. Sci.* **32** 630–639. https://doi.org/10.1214/17-sts625

Girard, S. and Stupfler, G. (2017). Intriguing properties of extreme geometric quantiles. *REVSTAT* **15** 107–139.

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.* 1547–1586.

González-Sanz, A., Hallin, M. and Sen, B. (2026). Supplement to "Monotone measure-transportation maps in Hilbert spaces, with statistical applications." https://doi.org/10.3150/25-BEJ1899SUPP

González-Sanz, A., Hallin, M. and Yao, Y. (2025). Nonparametric vector quantile autoregression. arXiv:2510.03166.

Gretton, A., Fukumizu, K. and Sriperumbudur, B.K. (2009). Discussion of Brownian distance covariance. *Ann. Appl. Stat.* **3** 1285–1294. https://doi.org/10.1214/13-AOS1140

Hallin, M. (2022). Measure transportation and statistical decision theory. *Annu. Rev. Stat. Appl.* **9** 401–424.

Hallin, M., Hlubinka, D. and Hudecová, Š. (2022). Fully distribution-free center-outward rank tests for multiple-output regression and MANOVA. *J. Amer. Statist. Assoc.* **118** 1923–1939.

Hallin, M. and Konen, D. (2024). Multivariate quantiles: Geometric and measure-transportation-based contours. In *Applications of Optimal Transport to Economics and Related Topics* (V. Kreinovich, W. Yamaka and S. Leurcharusmee, eds.). *Proceedings of the 17th International Conference of the Econometrics Society of Thailand* 61–78. Springer.

Hallin, M., La Vecchia, D. and Liu, H. (2022a). Center-outward R-estimation for semiparametric VARMA models. *J. Amer. Statist. Assoc.* **117** 925–938.

Hallin, M., La Vecchia, D. and Liu, H. (2022b). Rank-based testing for semiparametric VAR models: A measure transportation approach. *Bernoulli* **29** 229–273.

Hallin, M. and Liu, H. (2022). Center-outward rank- and sign-based VARMA portmanteau tests: Chitturi, Hosking, and Li–McLeod revisited. *Econom. Stat.* Special issue in honor of Elvezio Ronchetti and Peter Rousseeuw, to appear. Available at arXiv. arXiv:2208.12143. https://doi.org/10.48550/ARXIV.2208.12143

Hallin, M. and Liu, H. (2024). Quantiles and quantile regression on Riemannian manifolds: a measure-transportation-based approach. arXiv:2410.15711.

Hallin, M., Liu, H. and Verdebout, T. (2024). Nonparametric measure-transportation-based methods for directional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **86** 1172–1196.

Hallin, M. and Mordant, G. (2025). Multiple-attribute Lorenz functions and Gini indices: A measure transportation approach. *J. Bus. Econom. Statist.* **43** 1092–1104. https://doi.org/10.1080/07350015.2025.2475964

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. https://doi.org/10.1214/20-AOS1996

Helmberg, G. (2006). Curiosities concerning weak topology in Hilbert space. *Amer. Math. Monthly* **113** 447–452.

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

Horváth, L., Kokoszka, P. and Reeder, R. (2013). Estimation of the mean of functional time series and a two-sample problem. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 103–122. MR3008273 https://doi.org/10.1111/j.1467-9868.2012.01032.x

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: John Wiley & Sons, Ltd. MR3379106 https://doi.org/10.1002/9781118762547

Huang, Z. and Sen, B. (2023). *Multivariate Symmetry: Distribution-Free Testing via Optimal Transport*. arXiv:2305.01839.

Hundrieser, S., Staudt, T. and Munk, A. (2022). Empirical optimal transport between different measures adapts to lower complexity. *Ann. Inst. Henri Poincaré Probab. Stat.* **60** 824–846.

Jayasumana, S., Salzmann, M., Li, H. and Harandi, M. (2013). A framework for shape analysis via Hilbert space embedding. In *Proceedings of the IEEE International Conference on Computer Vision* 1249–1256.

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. *Texts in Statistical Science Series*. Boca Raton, FL: CRC Press. MR3793167

Konen, D. and Paindaveine, D. (2022). Multivariate ρ-quantiles: a spatial approach. *Bernoulli* **28** 1912–1934.
Lai, T., Zhang, Z., Wang, Y. and Kong, L. (2021). Testing independence of functional variables by angle covariance. *J. Multivariate Anal.* **182**. Paper No. 104711, 15. MR4187269 https://doi.org/10.1016/j.jmva.2020.104711
Lyons, R. (2013). Distance covariance in metric spaces. *Ann. Probab.* **41** 3284–3305.
Marron, J.S. and Alonso, A.M. (2014). Overview of object oriented data analysis. *Biom. J.* **56** 732–753. MR3258083 https://doi.org/10.1002/bimj.201300072
McCann, R.J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80** 309–323. https://doi.org/10.1215/S0012-7094-95-08013-2
Menafoglio, A. and Petris, G. (2016). Kriging for Hilbert-space valued random fields: The operatorial point of view. *J. Multivariate Anal.* **146** 84–94. MR3477651 https://doi.org/10.1016/j.jmva.2015.06.012
Nagy, S. (2021). Halfspace depth does not characterize probability distributions. *Stat. Pap.* **62** 1135–1139.
Paindaveine, D. and Passeggeri, R. (2025). On the robustness of semi-discrete optimal transport. *Ann. Appl. Probab.* To appear.
Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993
Rizzo, M.L. and Székely, G.J. (2016). Energy distance. *Wiley Interdiscip. Rev.: Comput. Stat.* **8** 27–38. https://doi.org/10.1002/wics.1375
Rockafellar, R.T. (1970). On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.* **33** 209–216. https://doi.org/pjm/1102977253
Rockafellar, R.T. and Wets, R.J.-B. (2009). *Variational Analysis*. Springer Science & Business Media.
Rudin, W. (1953). *Principles of Mathematical Analysis*. New York-Toronto-London: McGraw-Hill Book Company, Inc. (14, 1070c). MR0055409
Rudin, W. (1990). *Funct. Anal.* New York: McGraw-Hill.
Rüschendorf, L. and Rachev, S.T. (1990). A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.* **32** 48–54.
Segers, J. (2022). Graphical and uniform consistency of estimated optimal transport plans. arXiv:2208.02508.
Sejdinovic, D., Sriperumbudur, B., Gretton, A. and Fukumizu, K. (2013). Equivalence of distance-based and RKHS-based statistics in hypothesis testing. *Ann. Statist.* **41** 2263–2291. https://doi.org/10.1214/13-AOS1140
Shapiro, A. (1990). On concepts of directional differentiability. *J. Optim. Theory Appl.* **66** 477–487.
Shi, H., Drton, M., Hallin, M. and Han, F. (2022). On universally consistent and fully distribution-free rank tests of vector independence. *Ann. Statist.* **50** 1933–1959.
Shi, H., Drton, M., Hallin, M. and Han, F. (2025). Distribution-free tests of multivariate independence based on center-outward quadrant, Spearman, Kendall, and van der Waerden statistics. *Bernoulli* **31** 106–129.
Small, C.G. and McLeish, D.L. (1994). *Hilbert Space Methods in Probability and Statistical Inference*. *Wiley Series in Probability and Mathematical Statistics*. New York: John Wiley & Sons, Inc. MR1269321 https://doi.org/10.1002/9781118165522
Székely, G.J. and Rizzo, M.L. (2004). Testing for equal distributions in high dimensions. *InterStat* **5** 1249–1272.
Székely, G.J. and Rizzo, M.L. (2013). Energy statistics: A class of statistics based on distances. *J. Statist. Plann. Inference* **143** 1249–1272.
Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians, Vancouver, 1975* **2** 23–531.
Villani, C. (2003). *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Providence, RI: American Mathematical Society. MR1964483 https://doi.org/10.1090/gsm/058
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
Wynne, G. and Duncan, A.B. (2022). A kernel two-sample test for functional data. *J. Mach. Learn. Res.* **23**. Paper No. [73], 51. MR4576658
Zajíček, L. (1978). On the points of multivaluedness of metric projections in separable Banach spaces. *Comment. Math. Univ. Carolin.* **19** 513–523.
Zajíček, L. (1979). On the differentiation of convex functions in finite- and infinite-dimensional spaces. *Czechoslovak Math. J.* **29** 340–348.
Zajíček, L. (1983). Differentiability of the distance function and points of multi-valuedness of the metric projection in Banach space. *Czechoslovak Math. J.* **33** 292–308.

Zhang, J.-T., Liang, X. and Xiao, S. (2010). On the two-sample Behrens-Fisher problem for functional data. *J. Stat. Theory Pract.* **4** 571–587. MR2758746 https://doi.org/10.1080/15598608.2010.10412005

Bernoulli **32**(2), 2026, 1243–1266
https://doi.org/10.3150/25-BEJ1900

Detecting spectral breaks in spiked covariance models

NINA DÖRNEMANN[1,a] and DEBASHIS PAUL[2,3,b]

[1]*Department of Mathematics, Aarhus University, Aarhus, Denmark,* [a]*ndoernemann@math.au.dk*
[2]*Department of Statistics, University of California, Davis, USA,* [b]*debpaul@ucdavis.edu*
[3]*Indian Statistical Institute, Kolkata, India*

In this paper, the key objects of interest are the sequential covariance matrices $\mathbf{S}_{n,t}$ and their largest eigenvalues. Here, the matrix $\mathbf{S}_{n,t}$ is computed as the empirical covariance associated with observations $\{\mathbf{x}_1, \dots, \mathbf{x}_{\lfloor nt \rfloor}\}$, for $t \in [0,1]$. The observations $\mathbf{x}_1, \dots, \mathbf{x}_n$ are assumed to be i.i.d. p-dimensional vectors with zero mean, and a covariance matrix that is a fixed-rank perturbation of the identity matrix. Treating $\{\mathbf{S}_{n,t}\}_{t\in[0,1]}$ as a matrix-valued stochastic process indexed by t, we study the behavior of the largest eigenvalues of $\mathbf{S}_{n,t}$, as t varies, with n and p increasing simultaneously, so that $p/n \to y \in (0,1)$. As a key contribution of this work, we establish the weak convergence of the stochastic process corresponding to the sample spiked eigenvalues, if their population counterparts exceed the critical phase-transition threshold. Our analysis of the limiting process is fully comprehensive revealing, in general, non-Gaussian limiting processes. As an application, we consider a class of change-point problems, where the interest is in detecting structural breaks in the covariance caused by a change in magnitude of the spiked eigenvalues. For this purpose, we propose two different maximal statistics corresponding to centered spiked eigenvalues of the sequential covariances. We show the existence of limiting null distributions for these statistics, and prove consistency of the test under fixed alternatives. Moreover, we compare the behavior of the proposed tests through a simulation study.

Keywords: Change-point problems; high-dimensional statistics; hypothesis testing; spiked covariance model

References

Aue, A. and Horvath, L. (2013). Structural breaks in time series. *J. Time Series Anal.* **34** 1–16.
Aue, A., Hörmann, S., Horváth, L. and Reimherr, M. (2009). Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** 4046–4087.
Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices 20*. Springer.
Bai, Z. and Yao, J. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474.
Bai, Z. and Yao, J. (2012). On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.* **106** 167–177.
Bai, Z.-D. and Yin, Y.-Q. (2008). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. In *In Advances in Statistics* 108–127. World Sci. Publ.
Baik, J., Ben Arous, G. and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697.
Baik, J. and Silverstein, J.W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408.
Bao, Z., Pan, G. and Zhou, W. (2015). Universality for the largest eigenvalue of sample covariance matrices with general population. *Ann. Statist.* **43** 382–421.
Bao, Z., Ding, X., Wang, J. and Wang, K. (2022). Statistical inference for principal components of spiked covariance matrices. *Ann. Statist.* **50** 1144–1169.
Beisson, R., Vallet, P., Giremus, A. and Ginolhac, G. (2024). A new statistic for testing covariance equality in high-dimensional Gaussian low-rank models. *IEEE Trans. Signal Process.*

1350-7265

Billingsley, P. (1968). *Convergence of Probability Measures*, second ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: John Wiley & Sons Inc.

Cai, T.T., Han, X. and Pan, G. (2020). Limiting laws for divergent spiked eigenvalues and largest nonspiked eigenvalue of sample covariance matrices. *Ann. Statist.* **48** 1255–1280.

Chen, L., Wang, W. and Wu, W.B. (2022). Inference of breakpoints in high-dimensional time series. *J. Amer. Statist. Assoc.* **117** 1951–1963.

Dette, H. and Gösmann, J. (2020). A likelihood ratio approach to sequential change point detection for a general class of parameters. *J. Amer. Statist. Assoc.* **115** 1361–1377.

Dette, H., Pan, G. and Yang, Q. (2022). Estimating a change point in a sequence of very high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **117** 444–454.

Dette, H. and Wied, D. (2016). Detecting relevant changes in time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 371–394.

Ding, X. and Yang, F. (2018). A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices. *Ann. Appl. Probab.* **28** 1679–1738.

Dörnemann, N. and Dette, H. (2024a). Linear spectral statistics of sequential sample covariance matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **60** 946–970.

Dörnemann, N. and Dette, H. (2024b). Detecting Change Points of Covariance Matrices in High Dimensions. arXiv preprint, arXiv:2409.15588.

Dörnemann, N. and Paul, D. (2026). Supplement to "Detecting spectral breaks in spiked covariance models." https://doi.org/10.3150/25-BEJ1900SUPP

El Karoui, N. (2007). Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.* **35** 663–714.

Fabozzi, F.J., Kolm, P.N., Pachamanova, D.A. and Focardi, S.M. (2007). *Robust Portfolio Optimization and Management*. John Wiley & Sons.

Feng, G.-C., Yuen, P.C. and Dai, D.-Q. (2000). Human face recognition using PCA on wavelet subband. *J. Electron. Imaging* **9** 226–233.

Galeano, P. and Peña, D. (2007). Covariance changes detection in multivariate time series. *J. Statist. Plann. Inference* **137** 194–211.

Horvath, L. and Rice, G. (2019). Asymptotics for empirical eigenvalue processes in high-dimensional linear factor models. *J. Multivariate Anal.* **169** 138–165.

Jiang, D. and Bai, Z. (2021a). Generalized four moment theorem and an application to CLT for spiked eigenvalues of high-dimensional covariance matrices. *Bernoulli* **27** 274–294.

Jiang, D. and Bai, Z. (2021b). Partial generalized four moment theorem revisited. *Bernoulli* **27** 2337–2352.

Jirak, M. (2015). Uniform change point tests in high dimension. *Ann. Statist.* **43** 2451–2483.

Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327.

Johnstone, I.M. and Paul, D. (2018). PCA in high dimensions: An orientation. *Proc. IEEE* **108** 1277–1292.

Kao, C., Trapani, L. and Urga, G. (2018). Testing for instability in covariance structures. *Bernoulli* **24** 740–771.

Ke, Z.T., Ma, Y. and Lin, X. (2023). Estimation of the number of spiked eigenvalues in a covariance matrix by bulk eigenvalue matching analysis. *J. Amer. Statist. Assoc.* **118** 374–392.

Knowles, A. and Yin, J. (2017). Anisotropic local laws for random matrices. *Probab. Theory Related Fields* **169** 257–352.

Lee, J.O. and Schnelli, K. (2016). Tracy-Widom distribution for the largest eigenvalue of real sample covariance matrices with general population. *Ann. Appl. Probab.* **26** 3786–3839.

Li, Z., Han, F. and Yao, J. (2020). Asymptotic joint distribution of extreme eigenvalues and trace of large sample covariance matrix in a generalized spiked population model. *Ann. Statist.* **48** 3138–3160.

Lorenz, E.N. (1956). Empirical Orthogonal Functions and Statistical Weather Prediction 1. Massachusetts Institute of Technology. Department of Meteorology Cambridge.

Onatski, A. (2008). The Tracy-Widom limit for the largest eigenvalues of singular complex Wishart matrices. *Ann. Appl. Probab.* **18** 470–490.

Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica* **77** 1447–1479.

Passemier, D., Li, Z. and Yao, J. (2017). On estimation of the noise variance in high dimensional probabilistic principal component analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* 51–67.

Passemier, D. and Yao, J. (2014). Estimation of the number of spikes, possibly equal, in the high-dimensional case. *J. Multivariate Anal.* **127** 173–183.

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

Paul, D. and Aue, A. (2014). Random matrix theory in statistics: A review. *J. Statist. Plann. Inference* **150** 1–29.

Ruppert, D. and Matteson, D.S. (2011). *Statistics and Data Analysis for Financial Engineering 13*. Springer.

Ryan, S. and Killick, R. (2023). Detecting changes in covariance via random matrix theory. *Technometrics* **65** 480–491.

Schnelli, K. and Xu, Y. (2023). Convergence rate to the Tracy-Widom laws for the largest eigenvalue of sample covariance matrices. *Ann. Appl. Probab.* **33** 677–725.

Tao, T. (2012). *Topics in Random Matrix Theory 132*. Am. Math. Soc.

Van Der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence*. Springer.

Wang, Q., Su, Z. and Yao, J. (2014). Joint CLT for several random sesquilinear forms with applications to large-dimensional spiked population models. *Electron. J. Probab.* **19** 1–28.

Wang, D., Yu, Y. and Rinaldo, A. (2021). Optimal covariance change point localization in high dimensions. *Bernoulli* **27** 554–574.

Wang, R., Zhu, C., Volgushev, S. and Shao, X. (2022). Inference for change points in high-dimensional data via selfnormalization. *Ann. Statist.* **50** 781–806.

Yao, J., Zheng, S. and Bai, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press.

Zhang, Y., Wang, R. and Shao, X. (2022). Adaptive inference for change points in high-dimensional data. *J. Amer. Statist. Assoc.* **117** 1751–1762.

Zhang, Z., Zheng, S., Pan, G. and Zhong, P.-S. (2022). Asymptotic independence of spiked eigenvalues and linear spectral statistics for large sample covariance matrices. *Ann. Statist.* **50** 2205–2230.

Zheng, S. (2012). Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 444–476.

Bernoulli **32**(2), 2026, 1267–1284
https://doi.org/10.3150/25-BEJ1901

Bernstein-type inequalities for Markov chains and Markov processes: A simple and robust proof

DE HUANG[a] and XIANGYUAN LI[b]

School of Mathematical Sciences, Peking University, Beijing, China, [a]*dhuang@math.pku.edu.cn,* [b]*lixiangyuan23@stu.pku.edu.cn*

We establish a new Bernstein-type deviation inequality for general (non-reversible) discrete-time Markov chains via an elementary approach. More robust than existing works in the literature, our result only requires the Markov chain to satisfy an iterated Poincaré inequality. Moreover, our method can be readily generalized to continuous-time Markov processes.

Keywords: Concentration inequality; Markov chain; Markov process; spectral gap

References

[1] Chatterjee, S. (2023). Spectral gap of nonreversible Markov chains. arXiv preprint. Available at arXiv:2310.10876. MR4945088 https://doi.org/10.1214/25-AAP2183
[2] Chung, K.-M., Lam, H., Liu, Z. and Mitzenmacher, M. (2012). Chernoff–Hoeffding bounds for Markov chains: Generalized and simplified. arXiv preprint. Available at arXiv:1201.0559. MR2909308
[3] Fan, J., Jiang, B. and Sun, Q. (2021). Hoeffding's inequality for general Markov chains and its applications to statistical learning. *J. Mach. Learn. Res.* **22** 1–35. MR4318495
[4] Fill, J.A. (1991). Eigenvalue bounds on convergence to stationarity for nonreversible Markov chains, with an application to the exclusion process. *Ann. Appl. Probab.* **1** 62–87. MR1097464
[5] Jerison, D. (2013). General mixing time bounds for finite Markov chains via the absolute spectral gap. arXiv preprint. Available at arXiv:1310.8021.
[6] Jiang, B., Sun, Q. and Fan, J. (2018). Bernstein's inequality for general Markov chains. arXiv preprint. Available at arXiv:1805.10721.
[7] Kontoyiannis, I. and Meyn, S.P. (2012). Geometric ergodicity and the spectral gap of non-reversible Markov chains. *Probab. Theory Related Fields* **154** 327–339. MR2981426 https://doi.org/10.1007/s00440-011-0373-4
[8] León, C.A. and Perron, F. (2004). Optimal Hoeffding bounds for discrete reversible Markov chains. *Ann. Appl. Probab.* **14** 958–970. MR2052909 https://doi.org/10.1214/105051604000000170
[9] Levin, D.A. and Peres, Y. (2017). *Markov Chains and Mixing Times 107*. American Mathematical Soc. MR3726904 https://doi.org/10.1090/mbk/107
[10] Lezaud, P. (1998). Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.* **8** 849–867. MR1627795 https://doi.org/10.1214/aoap/1028903453
[11] Miasojedow, B. (2014). Hoeffding's inequalities for geometrically ergodic Markov chains on general state space. *Statist. Probab. Lett.* **87** 115–120. MR3168944 https://doi.org/10.1016/j.spl.2014.01.013
[12] Montenegro, R. and Tetali, P. (2006). Mathematical aspects of mixing times in Markov chains. *Found. Trends Theor. Comput. Sci.* **1** 237–354. MR2341319 https://doi.org/10.1561/0400000003
[13] Paulin, D. (2015). Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electron. J. Probab.* **20** 1–32. MR3383563 https://doi.org/10.1214/EJP.v20-4039
[14] Pearcy, C. (1966). An elementary proof of the power inequality for the numerical radius. *Michigan Math. J.* **13** 289–291. MR0201976

[15] Rao, S. (2019). A Hoeffding inequality for Markov chains. *Electron. Commun. Probab.* **24** 1–11. MR3933038 https://doi.org/10.1214/19-ECP219

[16] Wolfer, G. and Kontorovich, A. (2019). Estimating the mixing time of ergodic Markov chains. In *Conference on Learning Theory 3120–3159*. PMLR. MR3932874

[17] Wolfer, G. and Kontorovich, A. (2024). Improved estimation of relaxation time in nonreversible Markov chains. *Ann. Appl. Probab.* **34** 249–276. MR4696277 https://doi.org/10.1214/23-aap1963

Bernoulli **32**(2), 2026, 1285–1302
https://doi.org/10.3150/25-BEJ1904

Empirical Bayes large-scale multiple testing for high-dimensional binary outcome data

YU-CHIEN BO NING[a]

Harvard T.H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02155, USA,
[a]*bycning@hsph.harvard.edu*

This paper explores the multiple testing problem for sparse high-dimensional data with binary outcomes. We propose novel empirical Bayes multiple testing procedures based on a spike-and-slab posterior and then evaluate their performance in controlling the false discovery rate (FDR). A surprising finding is that the procedure using the default conjugate prior (namely, the ℓ-value procedure) can be overly conservative in estimating the FDR. To address this, we introduce two new procedures that provide accurate FDR control. Sharp frequentist theoretical results are established for these procedures, and numerical experiments are conducted to validate our theory in finite samples. To the best of our knowledge, we obtain the first *uniform* FDR control result in multiple testing for high-dimensional data with binary outcomes under the sparsity assumption.

Keywords: Binomial distribution; empirical Bayes; false discovery rate; multiple testing; sparse binary data; spike-and-slab posterior

References

Abraham, K., Castillo, I. and Roquain, E. (2024). Sharp multiple testing boundary for sparse sequences. *Ann. Statist.* **52** 1564–1591. https://doi.org/10.1214/24-AOS2404

Bai, R., Moran, G.E., Antonelli, J.L., Chen, Y. and Boland, M.R. (2022). Spike-and-slab group lassos for grouped regression and sparse generalized additive models. *J. Amer. Statist. Assoc.* **117** 184–197. https://doi.org/10.1080/01621459.2020.1765784

Belitser, E. and Ghosal, S. (2020). Empirical Bayes oracle uncertainty quantification for regression. *Ann. Statist.* **48** 3113–3137. https://doi.org/10.1214/19-AOS1845

Butucea, C., Ndaoud, M., Stepanova, N.A. and Tsybakov, A.B. (2018). Variable selection with Hamming loss. *Ann. Statist.* **46** 1837–1875. https://doi.org/10.1214/17-AOS1572

Castillo, I. and Roquain, E. (2020). On spike and slab empirical Bayes multiple testing. *Ann. Statist.* **48** 2548–2574. https://doi.org/10.1214/19-AOS1897

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

Castillo, I. and Szabó, B. (2020). Spike and slab empirical Bayes sparse credible sets. *Bernoulli* **26** 127–158. https://doi.org/10.3150/19-BEJ1119

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

Efron, B. (2004). Large-scale simultaneous hypothesis testing: The choice of a null hypothesis. *J. Amer. Statist. Assoc.* **99** 96–104. https://doi.org/10.1198/016214504000000089

Efron, B. (2010). *Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction*. Cambridge University Press. https://doi.org/10.1017/CBO9780511761362

Gao, C., Lu, Y. and Zhou, D. (2016). Exact exponent in optimal rates for crowdsourcing. In *Proceedings of the 33rd International Conference on Machine Learning* **48**. New York, NY: JMLR: W&CP.

Greenshtein, E. and Ritov, Y. (2009). Asymptotic efficiency of simple decisions for the compound decision problem. *IMS Lecture Notes Monogr. Ser.* **57** 266–275. https://doi.org/10.1214/09-LNMS5716

1350-7265

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

Marshall, P.J., Verma, A., More, A., Davis, C.P., More, S., Kapadia, A., Parrish, M., Snyder, C., Wilcox, J., Baeten, E., Macmillan, C., Cornen, C., Baumer, M., Simpson, E., Lintott, C.J., Miller, D., Paget, E., Simpson, R., Smith, A.M., Küng, R., Saha, P. and Collett, T.E. (2015). Space warps – I. Crowdsourcing the discovery of gravitational lenses. *Mon. Not. R. Astron. Soc.* **455** 1171–1190. https://doi.org/10.1093/mnras/stv2009

Martin, R., Mess, R. and Walker, S.G. (2017). Empirical Bayes posterior concentration in sparse high-dimensional linear models. *Bernoulli* **23** 1822–1847. https://doi.org/10.3150/15-BEJ797

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

Mukherjee, R., Pillai, N.S. and Lin, X. (2015). Hypothesis testing for high-dimensional sparse binary regression. *Ann. Statist.* **43** 352–381. https://doi.org/10.1214/14-AOS1279

Ning, Y.-C.B. (2026). Supplement to "Empirical Bayes large-scale multiple testing for high-dimensional binary outcome data." https://doi.org/10.3150/25-BEJ1904SUPP

Ning, B., Joeng, S. and Ghosal, S. (2020). Bayesian linear regression for multivariate responses under group sparsity. *Bernoulli* **26** 2353–2382. https://doi.org/10.3150/20-BEJ1198

Ray, K. and Szabó, B. (2022). Variational Bayes for high-dimensional linear regression with sparse priors. *J. Amer. Statist. Assoc.* **117** 1270–1281. https://doi.org/10.1080/01621459.2020.1847121

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

Storey, J.D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Ann. Statist.* **31** 2013–2035. https://doi.org/10.1214/aos/1074290335

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

Wang, R. and Ramdas, A. (2022). False discovery rate control with e-values. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 822–852. https://doi.org/10.1111/rssb.12489

Bernoulli **32**(2), 2026, 1303–1327
https://doi.org/10.3150/25-BEJ1907

Addressing both variable selection and misclassified responses with parametric and semiparametric methods

HUI GUO[1,a], GRACE Y. YI[1,2,b] and BOYU WANG[1,c]

[1]*Department of Computer Science, University of Western Ontario, Ontario, Canada,* [a]*hguo288@uwo.ca,* [b]*gyi5@uwo.ca,* [c]*bwang@csd.uwo.ca*
[2]*Department of Statistical and Actuarial Sciences, University of Western Ontario, Ontario, Canada*

While variable selection has received extensive attention in the literature, its exploration in the presence of response measurement error remains underexplored. In this paper, we investigate this important problem within the context of binary classification with error-prone responses. We present valid variable selection procedures to address the complexities of response errors. Leveraging validation data, we introduce both parametric and semiparametric methodologies to accommodate the mismeasurement effects. By rigorously establishing theoretical results, we offer insights and justifications of the validity of the proposed methods. By properly choosing the penalty function and regularization parameter, we demonstrate that the resulting estimators possess the oracle property. To assess the finite sample properties of the proposed methods, we conduct numerical studies that confirm the effectiveness of our proposed methods.

Keywords: Errors in response; generalized linear models; logistic regression; measurement error; parametric method; penalty function; semiparametric method; variable selection

References

[1] Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika* **60** 255–265.
[2] Bach, F., Jenatton, R., Mairal, J. and Obozinski, G. (2012). Optimization with sparsity-inducing penalties. *Found. Trends Mach. Learn.* **4** 1–106.
[3] Brown, B., Weaver, T. and Wolfson, J. (2019). MEBoost: Variable selection in the presence of measurement error. *Stat. Med.* **38** 2705–2718.
[4] Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective*. Chapman and Hall/CRC.
[5] Chen, L.-P. and Yi, G.Y. (2021). Analysis of noisy survival data with graphical proportional hazards measurement error models. *Biometrics* **77** 956–969.
[6] Datta, A. and Zou, H. (2017). Cocolasso for high-dimensional error-in-variables regression. *Ann. Statist.* **45** 2400–2426.
[7] Eckstein, S., Iske, A. and Trabs, M. (2023). Dimensionality reduction and Wasserstein stability for kernel regression. *J. Mach. Learn. Res.* **24** 1–35.
[8] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360.
[9] Foster, D.P. and George, E.I. (1994). The risk inflation criterion for multiple regression. *Ann. Statist.* **22** 1947–1975.
[10] Ge, J., Li, X., Jiang, H., Liu, H., Zhang, T., Wang, M. and Zhao, T. (2019). Picasso: A sparse learning library for high dimensional data analysis in R and Python. *J. Mach. Learn. Res.* **20** 44–1.
[11] Guo, H., Wang, B. and Yi, G. (2023). Label correction of crowdsourced noisy annotations with an instance-dependent noise transition model. In *Adv. Neural Inf. Process. Syst* **36** 347–386.

[12] Guo, H., Yi, G. and Wang, B. (2024). Learning from noisy labels via conditional distributionally robust optimization. In *Adv. Neural Inf. Process. Syst* **37** 82627–82672.

[13] Guo, H., Yi, G.Y. and Wang, B. (2026). Supplement to "Addressing both variable selection and misclassified responses with parametric and semiparametric methods." https://doi.org/10.3150/25-BEJ1907SUPPA, https://doi.org/10.3150/25-BEJ1907SUPPB

[14] Hannan, E.J. and Quinn, B.G. (1979). The determination of the order of an autoregression. *J. Roy. Statist. Soc. Ser. B, Methodol.* **41** 190–195.

[15] Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.

[16] Liang, H. and Li, R. (2009). Variable selection for partially linear models with measurement errors. *J. Amer. Statist. Assoc.* **104** 234–248.

[17] Loh, P.-L. and Wainwright, M.J. (2011). High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. In *Adv. Neural Inf. Process. Syst* **24**.

[18] Ma, Y. and Li, R. (2010). Variable selection in measurement error models. *Bernoulli* **16** 274–300.

[19] Nesterov, Y. (2013). Gradient methods for minimizing composite functions. *Math. Program.* **140** 125–161.

[20] Racine, J. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *J. Econometrics* **119** 99–130.

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

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

[23] Van der Vaart, A.W. (2000). *Asymptotic Statistics*. Cambridge University Press.

[24] Wang, H., Li, R. and Tsai, C.-L. (2007). Tuning parameter selectors for the smoothly clipped absolute deviation method. *Biometrika* **94** 553–568.

[25] Wang, Z., Liu, H. and Zhang, T. (2014). Optimal computational and statistical rates of convergence for sparse nonconvex learning problems. *Ann. Statist.* **42** 2164–2201.

[26] Yi, G.Y. (2017). *Statistical Analysis with Measurement Error or Misclassification: Strategy, Method and Application*. Springer.

[27] Yi, G.Y. and Chen, L.-P. (2023). Estimation of the average treatment effect with variable selection and measurement error simultaneously addressed for potential confounders. *Stat. Methods Med. Res.* **32** 691–711.

[28] Yi, G.Y., Delaigle, A. and Gustafson, P. (2021). *Handbook of Measurement Error Models*. CRC Press.

[29] Yi, G.Y., Tan, X. and Li, R. (2015). Variable selection and inference procedures for marginal analysis of longitudinal data with missing observations and covariate measurement error. *Canad. J. Statist.* **43** 498–518.

[30] Yi, G.Y., Yan, Y., Liao, X. and Spiegelman, D. (2018). Parametric regression analysis with covariate misclassification in main study/validation study designs. *Int. J. Biostat.* **15** 20170002.

[31] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942.

Bernoulli **32**(2), 2026, 1328–1355
https://doi.org/10.3150/25-BEJ1908

Multitype branching processes in random environments with not strictly positive expectation matrices

VILMA ORGOVÁNYI[1,a] and KÁROLY SIMON[2,b]

[1]*Department of Stochastics, Institute of Mathematics, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary,* [a]*orgovanyi.vilma@gmail.com*
[2]*HUN-REN-BME Stochastics Research Group, Műegyetem rkp. 3., H-1111 Budapest, Hungary,* [b]*karoly.simon51@gmail.com*

It is well known that under some conditions the almost sure survival probability of a multitype branching process in random environment is positive if the Lyapunov exponent corresponding to the expectation matrices is positive, and zero if the Lyapunov exponent is negative. The goal of this note is to establish similar results when certain positivity conditions on the expectation matrices are not met. One application of such a result is to classify the positivity of Lebesgue measure of certain overlapping random self-similar sets.

Keywords: Multitype branching process in random environment; random fractals

References

Athreya, K.B. and Karlin, S. (1971). On branching processes with random environments: I: Extinction probabilities. *Ann. Math. Statist.* **42** 1499–1520. https://doi.org/10.1214/aoms/1177693150

Bárány, B., Simon, K. and Solomyak, B. (2023). *Self-Similar and Self-Affine Sets and Measures*. *Math. Surv. Monogr.* Providence, RI: American Mathematical Society (AMS).

Dekking, F.M. and Grimmet, G.R. (1988). Superbranching processes and projections of random Cantor sets. *Probab. Theory Related Fields* **78** 335–355. https://doi.org/10.1007/BF00334199

Dekking, F.M. and Meester, R.W.J. (1990). On the structure of Mandelbrot's percolation process and other random Cantor sets. *J. Stat. Phys.* **58** 1109–1126. https://doi.org/10.1007/BF01026566

Dekking, F.M. and Simon, K. (2008). On the size of the algebraic difference of two random Cantor sets. *Random Structures Algorithms* **32** 205–222. https://doi.org/10.1002/rsa.20178

Falconer, K.J. (1989). Projections of random Cantor sets. *J. Theoret. Probab.* **2** 65–70. https://doi.org/10.1007/BF01048269

Falconer, K.J. and Grimmett, G.R. (1992). On the geometry of random Cantor sets and fractal percolation. *J. Theoret. Probab.* **5** 465–485. https://doi.org/10.1007/BF01060430

Hennion, H. (1997). Limit theorems for products of positive random matrices. *Ann. Probab.* **25** 1545–1587. https://doi.org/10.1214/aop/1023481103

Karlin, S. and Taylor, H.M. (1975). *A First Course in Stochastic Processes*.

Kersting, G. and Vatutin, V. (2017) In *Discrete Time Branching Processes in Random Environment*. *Math. Stat. Ser.* Hoboken, NJ: John Wiley & Sons. https://doi.org/10.1002/9781119452898

Móra, P., Simon, K. and Solomyak, B. (2009). The Lebesgue measure of the algebraic difference of two random Cantor sets. *Indag. Math.* **20** 131–149. https://doi.org/10.1016/S0019-3577(09)80007-4

Orgoványi, V. and Simon, K. (2023). Projections of the random Menger sponge. *Asian J. Math.* **27** 893–936. https://doi.org/10.4310/AJM.2023.v27.n6.a4

Orgoványi, V. and Simon, K. (2024). Interior points and Lebesgue measure of overlapping Mandelbrot percolation sets. Preprint.

1350-7265

Pollicott, M. (2023). In Lectures on Equilibrium states, mixing and dimension. Lecture Notes from Banach. Center, Warsaw.

Tanny, D. (1981). On multitype branching processes in a random environment. *Adv. Appl. Probab.* **13** 464–497. https://doi.org/10.2307/1426781

Vatutin, V.A. and Dyakonova, E.E. (2021). Multitype branching processes in random environment. *Russian Math. Surveys* **76** 1019–1063. https://doi.org/10.1070/RM10012

Walters, P. (2000). *An Introduction to Ergodic Theory*. *Graduate Texts in Mathematics*. New York: Springer.

Weissner, E.W. (1971). Multitype branching processes in random environments. *J. Appl. Probab.* **8** 17–31. https://doi.org/10.2307/3211834

Bernoulli **32**(2), 2026, 1356–1382
https://doi.org/10.3150/25-BEJ1909

Causal inference on process graphs: Causal structure and effect identification

NICOLAS-DOMENIC REITER[1,a], JONAS WAHL[2,b], ANDREAS GERHARDUS[3,c] and JAKOB RUNGE[4,d]

[1] *TUM School of Computation, Information and Technology, Technical University of Munich, Munich, Germany,* [a]*nicolas-domenic.reiter@tum.de*
[2] *Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI), Saarbrücken, Germany,* [b]*jonas.wahl@dfki.de*
[3] *German Aerospace Center (DLR), Institute of Data Science, Jena, Germany,* [c]*andreas.gerhardus@dlr.de*
[4] *Department of Computer Science, University of Potsdam, Potsdam, Germany,* [d]*jakob.runge@uni-potsdam.de*

A structural vector autoregressive (SVAR) process is a linear causal model for variables that evolve over a discrete set of time points and between which there may be lagged and instantaneous effects. The qualitative causal structure of an SVAR process can be represented by its finite and directed process graph, where a directed link connects two processes whenever there is a lagged or instantaneous effect between them. At the process graph level, the causal structure of SVAR processes is compactly parameterized in the frequency domain. In this paper, we consider the problem of causal discovery and causal effect estimation from the spectral density, the frequency domain analogue of the autocovariance, of the SVAR process. Causal discovery concerns the recovery of the process graph. Causal effect estimation concerns the identification and estimation of frequency domain causal effects. We show that information about the process graph, in terms of d- and t-separation, is generically characterized by algebraic constraints on the spectral density. Furthermore, we introduce a notion of rational identifiability for frequency domain causal effects between processes that may be confounded by exogenous latent processes, and show that the recent graphical latent factor half-trek criterion can be used on the process graph to assess whether a given (confounded) effect can be identified by rational operations on the entries of the spectral density.

Keywords: Causal inference; graphical models; identifiability; spectral density; SVAR processes

References

[1] Améndola, C., Drton, M., Grosdos, A., Homs, R. and Robeva, E. (2023). Third-order moment varieties of linear non-Gaussian graphical models. *Inf. Inference* **12** 1405–1436. https://doi.org/10.1093/imaiai/iaad007
[2] Assaad, C.K., Devijver, E. and Gaussier, E. (2022). Survey and evaluation of causal discovery methods for time series. *J. Artificial Intelligence Res.* **73** 767–819. https://doi.org/10.1613/jair.1.13428
[3] Barber, R.F., Drton, M., Sturma, N. and Weihs, L. (2022). Half-trek criterion for identifiability of latent variable models. *Ann. Statist.* **50** 3174–3196. https://doi.org/10.1214/22-aos2221
[4] Bernanke, B.S. and Gertler, M. (1995). Inside the black box: The credit channel of monetary policy transmission. *J. Econ. Perspect.* **9** 27–48. https://doi.org/10.1257/jep.9.4.27
[5] Bollen, K.A. (1989). *Structural Equations with Latent Variables* **210**. John Wiley & Sons.
[6] Brillinger, D.R. (2001). Time series. *Soc. Ind. Appl. Math.*. https://doi.org/10.1137/1.9780898719246
[7] Brockwell, P.J. and Davis, R.A. (2009). *Time Series: Theory and Methods*, 2nd ed. New York, NY: Springer. https://doi.org/10.1007/978-1-4419-0320-4
[8] Dahlhaus, R. and Eichler, M. (2003). In *Causality and Graphical Models in Time Series Analysis*. *Oxford Statistical Science Series* 115–137. https://doi.org/10.1093/oso/9780198510550.003.0011
[9] Di, Y. (2009). t-separation and d-separation for Directed Acyclic Graphs. Preprint.
[10] Drton, M. (2018). Algebraic problems in structural equation modeling. In *The 50th Anniversary of Gröbner Bases* **77** 35–87. Tokyo: Mathematical Society of Japan. https://doi.org/10.2969/aspm/07710035

1350-7265

[11] Drton, M., Foygel, R. and Sullivant, S. (2011). Global identifiability of linear structural equation models. *Ann. Statist.* **39** 865–886. https://doi.org/10.1214/10-AOS859

[12] Drton, M., Sturmfels, B. and Sullivant, S. (2009). *Lectures on Algebraic Statistics*. *Oberwolfach Seminars* **39**. Springer. https://doi.org/10.1007/978-3-7643-8905-5

[13] Ebert-Uphoff, I. and Deng, Y. (2012). Causal discovery for climate research using graphical models. *J. Climate* **25** 5648–5665. https://doi.org/10.1175/JCLI-D-11-00387.1

[14] Eichler, M. (2007). Granger causality and path diagrams for multivariate time series. *J. Econometrics* **137** 334–353. https://doi.org/10.1016/j.jeconom.2005.06.032

[15] Eichler, M. and Didelez, V. (2010). On Granger causality and the effect of interventions in time series. *Lifetime Data Anal.* **16** 3–32. https://doi.org/10.1007/s10985-009-9143-3

[16] Etesami, J., Kiyavash, N., Zhang, K. and Singhal, K. (2016). Learning network of multivariate Hawkes processes: A time series approach. In *32nd Conference on Uncertainty in Artificial Intelligence 2016, UAI 2016* (D. Janzing and A. Ihler, eds.). 162–171.

[17] Foygel, R., Draisma, J. and Drton, M. (2012). Half-trek criterion for generic identifiability of linear structural equation models. *Ann. Statist.* **40** 1682–1713. https://doi.org/10.1214/12-AOS1012

[18] Friston, K., Moran, R. and Seth, A.K. (2013). Analysing connectivity with Granger causality and dynamic causal modelling. *Curr. Opin. Neurobiol.* **23** 172–178. Macrocircuits. https://doi.org/10.1016/j.conb.2012.11.010

[19] Geiger, D., Verma, T. and Pearl, J. (1990). Identifying independence in Bayesian networks. *Networks* **20** 507–534. https://doi.org/10.1002/net.3230200504

[20] Gelfand, I.M., Kapranov, M.M. and Zelevinsky, A.V. (1994). *Discriminants, Resultants, and Multidimensional Determinants*. Boston, MA: Birkhäuser Boston. https://doi.org/10.1007/978-0-8176-4771-1_13

[21] Gerhardus, A. (2024). Characterization of causal ancestral graphs for time series with latent confounders. *Ann. Statist.* **52** 103–130. https://doi.org/10.1214/23-AOS2325

[22] Gerhardus, A. and Runge, J. (2020). High-recall causal discovery for autocorrelated time series with latent confounders. *Adv. Neural Inf. Process. Syst.* **33** 12615–12625.

[23] Gerhardus, A., Wahl, J., Faltenbacher, S., Ninad, U. and Runge, J. (2023). Projecting infinite time series graphs to finite marginal graphs using number theory. arXiv preprint. Available at arXiv:2310.05526.

[24] Gessel, I. and Viennot, G. (1985). Binomial determinants, paths, and hook length formulae. *Adv. Math.* **58** 300–321. https://doi.org/10.1016/0001-8708(85)90121-5

[25] González-Pérez, I. (2025). Causality for VARMA processes with instantaneous effects: The global Markov property, faithfulness and instrumental variables. arXiv preprint. Available at arXiv:2501.13037.

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

[27] Herwartz, H., Lange, A. and Maxand, S. (2022). Data-driven identification in SVARs—when and how can statistical characteristics be used to unravel causal relationships? *Econ. Inq.* **60** 668–693. https://doi.org/10.1111/ecin.13035

[28] Hochsprung, T., Runge, J. and Gerhardus, A. (2024). A global Markov property for solutions of stochastic difference equations and the corresponding full time graphs. In *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence* (N. Kiyavash and J.M. Mooij, eds.). *Proceedings of Machine Learning Research* **244** 1698–1726. PMLR.

[29] Imbens, G.W. (2024). Causal inference in the social sciences. *Annu. Rev. Stat. Appl.* **11** 123–152. https://doi.org/10.1146/annurev-statistics-033121-114601

[30] Lindström, B. (1973). On the vector representations of induced matroids. *Bull. Lond. Math. Soc.* **5** 85–90. https://doi.org/10.1112/blms/5.1.85

[31] Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin, Heidelberg: Springer.

[32] McGraw, M.C. and Barnes, E.A. (2018). Memory matters: A case for Granger causality in climate variability studies. *J. Climate* **31** 3289–3300. https://doi.org/10.1175/JCLI-D-17-0334.1

[33] Mogensen, S.W. (2023). Instrumental processes using integrated covariances. In *Proceedings of the Second Conference on Causal Learning and Reasoning*. *Proceedings of Machine Learning Research* **213** 620–641.

[34] Niemiro, W. and Rajkowski, Ł. (2023). Local dependence graphs for discrete time processes. In *Proceedings of the Second Conference on Causal Learning and Reasoning* (M. van der Schaar, C. Zhang and D. Janzing, eds.). *Proceedings of Machine Learning Research* **213** 772–790. PMLR.

[35] Nikias, C.L. (1993). Higher-order spectral analysis. In *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Societ* 319–319. https://doi.org/10.1109/IEMBS.1993.978564.
[36] Nikias, C.L. and Mendel, J.M. (1993). Signal processing with higher-order spectra. *IEEE Signal Process. Mag.* **10** 10–37.
[37] Pearl, J. (2009). *Causality*, 2nd ed. Cambridge University Press.
[38] Peters, J., Janzing, D. and Schlkopf, B. (2017). *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press.
[39] Petropulu, A.P. (1994). Higher-order spectra in biomedical signal processing. In *IFAC Proceedings. IFAC Symposium on Modelling and Control in Biomedical Systems* **27** 47–52. Galveston, TX, USA. 27–30 March 1994. https://doi.org/10.1016/S1474-6670(17)46158-1
[40] Reiter, N., Gerhardus, A., Wahl, J. and Runge, J. (2023). Causal inference on process graphs, part I: the structural equation process representation. arXiv preprint. Available at arXiv:2305.11561.
[41] Reiter, N., Wahl, J., Hegerl, G. and Runge, J. (2024). Asymptotic uncertainty in the estimation of frequency domain causal effects for linear processes. arXiv preprint. Available at arXiv:2406.18191.
[42] Reiter, N.-D., Wahl, J., Gerhardus, A. and Runge, J. (2026). Supplement to "Causal inference on process graphs: causal structure and effect identification." https://doi.org/10.3150/25-BEJ1909SUPP
[43] Robeva, E. and Seby, J.-B. (2021). Multi-trek separation in linear structural equation models. *SIAM J. Appl. Algebra Geom.* **5** 278–303. https://doi.org/10.1137/20M1316470
[44] Runge, J. (2020). Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In *Conference on Uncertainty in Artificial Intelligence* 1388–1397. PMLR.
[45] Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M.D., Muñoz-Marí, J. et al. (2019). Inferring causation from time series in Earth system sciences. *Nat. Commun.* **10** 1–13. https://doi.org/10.1038/s41467-019-10105-3
[46] Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci. Adv.* **5** eaau4996. https://doi.org/10.1126/sciadv.aau4996
[47] Runge, J., Gerhardus, A., Varando, G., Eyring, V. and Camps-Valls, G. (2023). Causal inference for time series. *Nat. Rev., Earth Environ.* **4** 487–505. https://doi.org/10.1038/s43017-023-00431-y
[48] Seth, A.K., Barrett, A.B. and Barnett, L. (2015). Granger causality analysis in neuroscience and neuroimaging. *J. Neurosci.* **35** 3293–3297. https://doi.org/10.1523/JNEUROSCI.4399-14.2015
[49] Siddiqi, S.H., Kording, K.P., Parvizi, J. and Fox, M.D. (2022). Causal mapping of human brain function. *Nat. Rev. Neurosci.* **23** 361–375.
[50] Spirtes, P., Glymour, C.N., Scheines, R. and Heckerman, D. (2001). *Causation, Prediction, and Search*. The MIT Press. https://doi.org/10.7551/mitpress/1754.003.0007
[51] Sullivant, S. (2008). Algebraic geometry of Gaussian Bayesian networks. *Adv. in Appl. Math.* **40** 482–513. https://doi.org/10.1016/j.aam.2007.04.004
[52] Sullivant, S., Talaska, K. and Draisma, J. (2010). Trek separation for Gaussian graphical models. *Ann. Statist.* **38** 1665–1685. https://doi.org/10.1214/09-AOS760
[53] Thams, N., Søndergaard, R., Weichwald, S. and Peters, J. (2024). Identifying causal effects using instrumental time series: Nuisance IV and correcting for the past. *J. Mach. Learn. Res.* **25** 1–51.
[54] Tramontano, D., Monod, A. and Drton, M. (2022). Learning linear non-Gaussian polytree models. In *Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence* (J. Cussens and K. Zhang, eds.). *Proceedings of Machine Learning Research* **180** 1960–1969. PMLR.
[55] Weihs, L., Robinson, B., Dufresne, E., Kenkel, J., McGee, K.K.R. II, Reginald, M.I., Nguyen, N., Robeva, E. and Drton, M. (2018). Determinantal generalizations of instrumental variables. *J. Causal Inference* **6**. https://doi.org/10.1515/jci-2017-0009
[56] Wright, S. (1921). Correlation and causation. *J. Agric. Res.* **20** 557–585.
[57] Wright, S. (1934). The method of path coefficients. *Ann. Math. Statist.* **5** 161–215.

Bernoulli **32**(2), 2026, 1383–1403
https://doi.org/10.3150/25-BEJ1911

Locally sharp goodness-of-fit testing in sup norm for high-dimensional counts

SUBHODH KOTEKAL[1,a], JULIEN CHHOR[2,c] and CHAO GAO[1,b]

[1] *Department of Statistics, University of Chicago, Chicago, United States of America,* [a] *skotekal@uchicago.edu,* [b] *chaogao@uchicago.edu*
[2] *Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France,* [c] *julien.chhor@tse-fr.eu*

We consider testing the goodness-of-fit of a distribution against alternatives separated in sup norm. We study the twin settings of Poisson-generated count data with a large number of categories and high-dimensional multinomials. In previous studies of different separation metrics, it has been found that the local minimax separation rate exhibits substantial heterogeneity and is a complicated function of the null distribution; the rate-optimal test requires careful tailoring to the null. In the setting of sup norm, this remains the case and we establish that the local minimax separation rate is determined by the finer decay behavior of the category rates. The upper bound is obtained by a test involving the sample maximum, and the lower bound argument involves reducing the original heteroskedastic null to an auxiliary homoskedastic null determined by the decay of the rates. Further, in a particular asymptotic setup, the sharp constants are identified.

Keywords: Goodness-of-fit tests; multinomial data; Poisson distribution; sharp constant; sup norm

References

Arias-Castro, E. and Wang, M. (2015). The sparse Poisson means model. *Electron. J. Stat.* **9** 2170–2201. MR3406276 https://doi.org/10.1214/15-EJS1066

Balakrishnan, S. and Wasserman, L. (2018). Hypothesis testing for high-dimensional multinomials: A selective review. *Ann. Appl. Stat.* **12** 727–749. MR3834283 https://doi.org/10.1214/18-AOAS1155SF

Balakrishnan, S. and Wasserman, L. (2019). Hypothesis testing for densities and high-dimensional multinomials: Sharp local minimax rates. *Ann. Statist.* **47** 1893–1927. MR3953439 https://doi.org/10.1214/18-AOS1729

Batu, T., Fortnow, L., Rubinfeld, R., Smith, W.D. and White, P. (2000). Testing that distributions are close. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science (FOCS)* 259–269. https://doi.org/10.1109/SFCS.2000.892113

Berrett, T. and Butucea, C. (2020). Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms. In *Advances in Neural Information Processing Systems (NeurIPS)* **33** 3164–3173.

Bhattacharya, B.B. and Mukherjee, R. (2024). Sparse uniformity testing. *IEEE Trans. Inf. Theory* **70** 6371–6390. MR4792653 https://doi.org/10.1109/tit.2024.3424679

Blais, E., Canonne, C.L. and Gur, T. (2019). Distribution testing lower bounds via reductions from communication complexity. *ACM Trans. Comput. Theory* **11**. https://doi.org/10.1145/3305270

Blanchard, M., Cohen, D. and Kontorovich, A. (2024). Correlated binomial process. In *Proceedings of the Thirty Seventh Conference on Learning Theory (COLT)* 551–595.

Blanchard, M. and Voracek, V. (2024). Tight bounds for local Glivenko-Cantelli. In *Proceedings of the 35th International Conference on Algorithmic Learning Theory (ALT)* 179–220.

Cai, T.T., Jeng, X.J. and Jin, J. (2011). Optimal detection of heterogeneous and heteroscedastic mixtures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 629–662. MR2867452 https://doi.org/10.1111/j.1467-9868.2011.00778.x

Cai, T.T. and Wu, Y. (2014). Optimal detection of sparse mixtures against a given null distribution. *IEEE Trans. Inf. Theory* **60** 2217–2232. MR3181520 https://doi.org/10.1109/TIT.2014.2304295

Canonne, C.L. (2022). Topics and techniques in distribution testing: A biased but representative sample. *CIT* **19** 1032–1198. https://doi.org/10.1561/0100000114

Chan, S.-O., Diakonikolas, I., Valiant, P. and Valiant, G. (2014). Optimal algorithms for testing closeness of discrete distributions. In *Proceedings of the 2014 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)* 1193–1203. https://doi.org/10.1137/1.9781611973402.88

Chhor, J. and Carpentier, A. (2022). Sharp local minimax rates for goodness-of-fit testing in multivariate binomial and Poisson families and in multinomials. *Math. Stat. Learn.* **5** 1–54. MR4510329 https://doi.org/10.4171/msl/32

Chhor, J. and Carpentier, A. (2023). Goodness-of-fit testing for Hölder-continuous densities: sharp local minimax rates. arXiv preprint. Available at arXiv:2109.04346. https://doi.org/10.48550/arXiv

Chhor, J., Mukherjee, R. and Sen, S. (2024). Sparse signal detection in heteroscedastic Gaussian sequence models: Sharp minimax rates. *Bernoulli* **30** 2127–2153. MR4746602 https://doi.org/10.3150/23-bej1667

Chow, Y.S. and Teicher, H. (1997). *Probab. Theory Related Fields*, third ed. *Springer Texts in Statistics*. New York: Springer-Verlag. MR1476912 https://doi.org/10.1007/978-1-4612-1950-7

Collier, O., Comminges, L. and Tsybakov, A.B. (2017). Minimax estimation of linear and quadratic functionals on sparsity classes. *Ann. Statist.* **45** 923–958. MR3662444 https://doi.org/10.1214/15-AOS1432

Collier, O. and Dalalyan, A.S. (2018). Estimating linear functionals of a sparse family of Poisson means. *Stat. Inference Stoch. Process.* **21** 331–344. MR3824971 https://doi.org/10.1007/s11203-018-9173-0

Diakonikolas, I. and Kane, D.M. (2016). A new approach for testing properties of discrete distributions. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)* 685–694. https://doi.org/10.1109/FOCS.2016.78

Diakonikolas, I., Kane, D. and Nikishkin, V. (2017). Near-optimal closeness testing of discrete histogram distributions. In *44th International Colloquium on Automata, Languages, and Programming (ICALP)* 1–15. https://doi.org/10.4230/LIPIcs.ICALP.2017.8

Ditzhaus, M. (2019). Signal detection via phi-divergences for general mixtures. *Bernoulli* **25** 3041–3068. MR4003573 https://doi.org/10.3150/18-BEJ1079

Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195 https://doi.org/10.1214/009053604000000265

Donoho, D. and Jin, J. (2015). Higher criticism for large-scale inference, especially for rare and weak effects. *Statist. Sci.* **30** 1–25. MR3317751 https://doi.org/10.1214/14-STS506

Donoho, D.L. and Kipnis, A. (2022). Higher criticism to compare two large frequency tables, with sensitivity to possible rare and weak differences. *Ann. Statist.* **50** 1447–1472. MR4441127 https://doi.org/10.1214/21-aos2158

Dubois, A., Berrett, T.B. and Butucea, C. (2023). Goodness-of-fit testing for Hölder continuous densities under local differential privacy. In *Foundations of Modern Statistics* 53–119.

Fienberg, S.E. (1979). The use of chi-squared statistics for categorical data problems. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **41** 54–64. MR0535545

Hall, P. and Jin, J. (2010). Innovated higher criticism for detecting sparse signals in correlated noise. *Ann. Statist.* **38** 1686–1732. MR2662357 https://doi.org/10.1214/09-AOS764

Hoeffding, W. (1965). Asymptotically optimal tests for multinomial distributions. *Ann. Math. Statist.* **36** 369–408. https://doi.org/10.1214/aoms/1177700150 MR173322

Ingster, Y.I. (1997). Some problems of hypothesis testing leading to infinitely divisible distributions. *Math. Methods Statist.* **6** 47–69.

Ingster, Y.I. and Kutoyants, Y.A. (2007). Nonparametric hypothesis testing for intensity of the Poisson process. *Math. Methods Statist.* **16** 217–245. https://doi.org/10.3103/S1066530707030039

Ingster, Y. and Lepski, O. (2003). Multichannel nonparametric signal detection. *Math. Methods Statist.* **12** 247–275.

Ingster, Y.I. and Suslina, I.A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. New York: Springer-Verlag. MR1991446 https://doi.org/10.1007/978-0-387-21580-8

Ingster, Y.I., Tsybakov, A.B. and Verzelen, N. (2010). Detection boundary in sparse regression. *Electron. J. Stat.* **4** 1476–1526. MR2747131 https://doi.org/10.1214/10-EJS589

Kipnis, A. (2023). The minimax risk in testing the histogram of discrete distributions for uniformity under missing ball alternatives. In *2023 59th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* 1–8. https://doi.org/10.1109/Allerton58177.2023.10313383

Kotekal, S. (2022). Statistical limits of sparse mixture detection. *Electron. J. Stat.* **16** 4982–5018. MR4490413 https://doi.org/10.1214/22-ejs2053

Kotekal, S., Chhor, J. and Gao, C. (2026). Supplement to "Locally sharp goodness-of-fit testing in sup norm for high-dimensional counts." https://doi.org/10.3150/25-BEJ1911SUPP

Lam-Weil, J., Carpentier, A. and Sriperumbudur, B.K. (2022). Local minimax rates for closeness testing of discrete distributions. *Bernoulli* **28** 1179–1197. MR4388934 https://doi.org/10.3150/21-bej1382

Laurent, B., Loubes, J.-M. and Marteau, C. (2012). Non asymptotic minimax rates of testing in signal detection with heterogeneous variances. *Electron. J. Stat.* **6** 91–122. MR2879673 https://doi.org/10.1214/12-EJS667

Paninski, L. (2008). A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Trans. Inf. Theory* **54** 4750–4755. MR2591136 https://doi.org/10.1109/TIT.2008.928987

Rosalsky, A. (1983). A remark on the fluctuation behavior of i.i.d. Poisson random variables. *Statist. Probab. Lett.* **1** 181–182. MR0709245 https://doi.org/10.1016/0167-7152(83)90027-5

Rosalsky, A. (1984). Acknowledgement of priority to: "A remark on the fluctuation behavior of i.i.d. Poisson random variables". *Statist. Probab. Lett.* **2** 117. MR0735244 https://doi.org/10.1016/0167-7152(84)90060-9

Spokoiny, V.G. (1996). Adaptive hypothesis testing using wavelets. *Ann. Statist.* **24** 2477–2498. MR1425962 https://doi.org/10.1214/aos/1032181163

Talagrand, M. (2021). *Upper and Lower Bounds for Stochastic Processes: Decomposition Theorems*. *A Series of Modern Surveys in Mathematics* **60**. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-82595-9

Valiant, G. and Valiant, P. (2017). An automatic inequality prover and instance optimal identity testing. *SIAM J. Comput.* **46** 429–455. MR3614697 https://doi.org/10.1137/151002526

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

Waggoner, B. (2015). L_p testing and learning of discrete distributions. In *Proceedings of the 2015 Conference on Innovations in Theoretical Computer Science (ITCS)* 347–356. https://doi.org/10.1145/2688073.2688095

Bernoulli **32**(2), 2026, 1404–1431
https://doi.org/10.3150/25-BEJ1912

Dimension-free relaxation times of informed MCMC samplers on discrete spaces

HYUNWOONG CHANG[1,a] and QUAN ZHOU[2,b]

[1]*Department of Mathematical Sciences, The University of Texas at Dallas, Dallas, TX 75080, USA,* [a]*hwchang@utdallas.edu*
[2]*Department of Statistics, Texas A&M University, College Station, TX 77843, USA,* [b]*quan@stat.tamu.edu*

Convergence analysis of Markov chain Monte Carlo methods in high-dimensional statistical applications is increasingly recognized. In this paper, we develop general mixing time bounds for Metropolis-Hastings algorithms on discrete spaces by building upon and refining some recent theoretical advancements in Bayesian model selection problems. We establish sufficient conditions for a class of informed Metropolis-Hastings algorithms to attain relaxation times that are independent of the problem dimension. These conditions are grounded in the high-dimensional statistical theory and allow for possibly multimodal posterior distributions. We obtain our results through two independent techniques: the multicommodity flow method and single-element drift condition analysis; we find that the latter yields a slightly tighter mixing time bound. Our results are readily applicable to a broad spectrum of statistical problems with discrete parameter spaces, as we demonstrate using both theoretical and numerical examples.

Keywords: Drift condition; finite Markov chains; informed Metropolis-Hastings; mixing time; model selection; multicommodity flow; random walk Metropolis-Hastings; restricted spectral gap

References

Anil Kumar, V. and Ramesh, H. (2001). Coupling vs. conductance for the Jerrum–Sinclair chain. *Random Structures Algorithms* **18** 1–17. MR1799799 https://doi.org/10.1002/1098-2418(200101)18:1<1::AID-RSA1>3.0.CO;2-7

Ascolani, F., Roberts, G.O. and Zanella, G. (2024). Scalability of Metropolis-within-Gibbs schemes for high-dimensional Bayesian models. arXiv preprint. https://doi.org/10.48550/arXiv.2403.09416.

Ascolani, F. and Zanella, G. (2024). Dimension-free mixing times of Gibbs samplers for Bayesian hierarchical models. *Ann. Statist.* **52** 869–894. MR4784062 https://doi.org/10.1214/24-AOS2367

Atchadé, Y.F. (2021). Approximate spectral gaps for Markov chain mixing times in high dimensions. *SIAM J. Math. Data Sci.* **3** 854–872. MR4303259 https://doi.org/10.1137/19M1283082

Belloni, A. and Chernozhukov, V. (2009). On the computational complexity of MCMC-based estimators in large samples. *Ann. Statist.* **37** 2011–2055. MR2533478 https://doi.org/10.1214/08-AOS634

Bhattacharya, R. and Jones, G. (2025). Explicit constraints on the geometric rate of convergence of random walk Metropolis-Hastings. *Bernoulli* **31** 2042–2076. https://doi.org/10.3150/24-bej1796

Castillo, I. and Ročková, V. (2021). Uncertainty quantification for Bayesian CART. *Ann. Statist.* **49** 3482–3509. MR4352538 https://doi.org/10.1214/21-aos2093

Chang, H., Cai, J. and Zhou, Q. (2023). Order-based structure learning without score equivalence. *Biometrika*. MR4745581 https://doi.org/10.1093/biomet/asad052.

Chang, H. and Zhou, Q. (2026). Supplement to "Dimension-free relaxation times of informed MCMC samplers on discrete spaces." https://doi.org/10.3150/25-BEJ1912SUPP

Chang, H., Lee, C., Luo, Z.T., Sang, H. and Zhou, Q. (2022). Rapidly mixing multiple-try Metropolis algorithms for model selection problems. *Adv. Neural Inf. Process. Syst.* **35** 25842–25855. https://doi.org/10.48550/arXiv.2207.00689

Cheng, X., Chatterji, N.S., Bartlett, P.L. and Jordan, M.I. (2018). Underdamped Langevin MCMC: A non-asymptotic analysis. In *Conference on Learning Theory* 300–323. PMLR. https://doi.org/10.48550/arXiv.1707.03663.

1350-7265

Chewi, S., Lu, C., Ahn, K., Cheng, X., Le Gouic, T. and Rigollet, P. (2021). Optimal dimension dependence of the Metropolis-adjusted Langevin algorithm. In *Conference on Learning Theory* 1260–1300. PMLR. https://doi.org/10.48550/arXiv.2012.12810.

Cowles, M.K. and Carlin, B.P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *J. Amer. Statist. Assoc.* **91** 883–904. MR1395755 https://doi.org/10.2307/2291683

Dalalyan, A.S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183

Diaconis, P. and Shahshahani, M. (1981). Generating a random permutation with random transpositions. *Z. Wahrsch. Verw. Gebiete* **57** 159–179. MR626813 https://doi.org/10.1007/BF00535487

Donoho, D.L. (1997). CART and best-ortho-basis: A connection. *Ann. Statist.* **25** 1870–1911. MR1474073 https://doi.org/10.1214/aos/1069362377

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

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

Dwivedi, R., Chen, Y., Wainwright, M.J. and Yu, B. (2019). Log-concave sampling: Metropolis-Hastings algorithms are fast. *J. Mach. Learn. Res.* **20** 1–42. MR4048994 https://doi.org/10.48550/arXiv.1801.02309

Fort, G., Moulines, E., Roberts, G.O. and Rosenthal, J. (2003). On the geometric ergodicity of hybrid samplers. *J. Appl. Probab.* **40** 123–146. MR1953771 https://doi.org/10.1239/jap/1044476831

Gagnon, P., Maire, F. and Zanella, G. (2023). Improving multiple-try Metropolis with local balancing. *J. Mach. Learn. Res.* **24** 1–59. MR4633975 https://doi.org/10.48550/arXiv.2211.11613

Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–472. https://doi.org/10.1214/ss/1177011136

George, E.I. and McCulloch, R.E. (1997). Approaches for Bayesian variable selection. *Statist. Sinica* 339–373. https://www.jstor.org/stable/pdf/24306083.pdf.

Giné, E., Grimmett, G.R. and Saloff-Coste, L. (1996). *Lectures on Probability Theory and Statistics: École d'été de Probabilités de Saint-Flour XXVI-1996*. Springer. MR1490042 https://doi.org/10.1007/BFb0092617

Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 123–214. MR2814492 https://doi.org/10.1111/j.1467-9868.2010.00765.x

Gong, L. and Flegal, J.M. (2016). A practical sequential stopping rule for high-dimensional Markov chain Monte Carlo. *J. Comput. Graph. Statist.* **25** 684–700. MR3533633 https://doi.org/10.1080/10618600.2015.1044092

Grathwohl, W., Swersky, K., Hashemi, M., Duvenaud, D. and Maddison, C. (2021). Oops I took a gradient: Scalable sampling for discrete distributions. In *International Conference on Machine Learning* 3831–3841. PMLR. https://doi.org/10.48550/arXiv.2102.04509

Griffin, J.E., Łatuszyński, K. and Steel, M.F. (2021). In search of lost mixing time: Adaptive Markov chain Monte Carlo schemes for Bayesian variable selection with very large p. *Biometrika* **108** 53–69. MR4226189 https://doi.org/10.1093/biomet/asaa055

Guan, Y. and Krone, S.M. (2007). Small-world MCMC and convergence to multi-modal distributions: From slow mixing to fast mixing. *Ann. Appl. Probab.* **17** 284–304. MR2292588 https://doi.org/10.1214/105051606000000772

Guan, Y. and Stephens, M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *Ann. Appl. Stat.* **5** 1780–1815. MR2884922 https://doi.org/10.1214/11-AOAS455

Guruswami, V. (2000). Rapidly mixing Markov chains: A comparison of techniques. arXiv preprint. https://doi.org/10.48550/arXiv.1603.01512

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109. MR3363437 https://doi.org/10.1093/biomet/57.1.97

Jarner, S.F. and Hansen, E. (2000). Geometric ergodicity of Metropolis algorithms. *Stoch. Process. Appl.* **85** 341–361. MR4305824 https://doi.org/10.1080/03610926.2020.1719418

Jerison, D. (2016). The drift and minorization method for reversible Markov chains. Thesis (Ph.D.), Stanford University. ProQuest LLC, Ann Arbor, MI. MR4172228.

Jerrum, M. and Sinclair, A. (1996). The Markov chain Monte Carlo method: An approach to approximate counting and integration. In *Approximation Algorithms for NP-Hard Problems* (D.S. Hochbaum, ed.) 482–520. PWS Publishing. https://doi.org/10.1007/978-3-662-12788-9_4

Jerrum, M., Son, J.-B., Tetali, P. and Vigoda, E. (2004). Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Probab.* 1741–1765. MR2099650 https://doi.org/10.1214/105051604000000639.

Jin, Z. and Hobert, J.P. (2022). Dimension free convergence rates for Gibbs samplers for Bayesian linear mixed models. *Stoch. Process. Appl.* **148** 25–67. MR4393342 https://doi.org/10.1016/j.spa.2022.02.003

Johndrow, J., Orenstein, P. and Bhattacharya, A. (2020). Scalable approximate MCMC algorithms for the horseshoe prior. *J. Mach. Learn. Res.* **21** 1–61. MR4095352

Kim, J. and Rockova, V. (2023). On mixing rates for Bayesian CART. https://doi.org/10.48550/arXiv.2306.00126

Lamnisos, D., Griffin, J.E. and Steel, M.F. (2009). Transdimensional sampling algorithms for Bayesian variable selection in classification problems with many more variables than observations. *J. Comput. Graph. Statist.* **18** 592–612. MR2751643 https://doi.org/10.1198/jcgs.2009.08027

Levin, D.A. and Peres, Y. (2017). *Markov Chains and Mixing Times* **107**. *American Mathematical Soc.* MR3726904 https://doi.org/10.1090/mbk/107

Li, G., Smith, A. and Zhou, Q. (2023). Importance is Important: Generalized Markov Chain Importance Sampling Methods. arXiv preprint. https://doi.org/10.48550/arXiv.2304.06251

Liang, X., Livingstone, S. and Griffin, J. (2022). Adaptive random neighbourhood informed Markov chain Monte Carlo for high-dimensional Bayesian variable selection. *Stat. Comput.* **32**. MR4491497 https://doi.org/10.1007/s11222-022-10137-8

Liu, J.S., Liang, F. and Wong, W.H. (2000). The multiple-try method and local optimization in Metropolis sampling. *J. Amer. Statist. Assoc.* **95** 121–134. MR1803145 https://doi.org/10.2307/2669532

Madigan, D., York, J. and Allard, D. (1995). Bayesian graphical models for discrete data. In *International Statistical Review/Revue Internationale de Statistique* 215–232. https://doi.org/10.2307/1403615.

Madras, N. and Randall, D. (2002). Markov chain decomposition for convergence rate analysis. *Ann. Appl. Probab.* 581–606. MR1910641 https://doi.org/10.1214/aoap/1026915617

Mira, A. (2001). Ordering and improving the performance of Monte Carlo Markov chains. *Statist. Sci.* 340–350. MR1888449 https://doi.org/10.1214/ss/1015346319.

Peres, Y. and Sousi, P. (2015). Mixing times are hitting times of large sets. *J. Theoret. Probab.* **28** 488–519. MR3370663 https://doi.org/10.1007/s10959-013-0497-9

Pollard, D. and Yang, D. (2019). Rapid mixing of a Markov chain for an exponentially weighted aggregation estimator. arXiv preprint. https://doi.org/10.48550/arXiv.1909.11773.

Qin, Q. and Hobert, J.P. (2019). Convergence complexity analysis of Albert and Chib's algorithm for Bayesian probit regression. *Ann. Statist.* **47** 2320–2347. MR3953453 https://doi.org/10.1214/18-AOS1749

Qin, Q. and Hobert, J.P. (2022). Wasserstein-based methods for convergence complexity analysis of MCMC with applications. *Ann. Appl. Probab.* **32** 124–166. MR4386523 https://doi.org/10.1214/21-aap1673

Robert, C.P., Casella, G. and Casella, G. (1999). *Monte Carlo Statistical Methods* **2**. Springer. MR1707311 https://doi.org/10.1007/978-1-4757-3071-5

Roberts, G.O. and Stramer, O. (2002). Langevin diffusions and Metropolis-Hastings algorithms. *Methodol. Comput. Appl. Probab.* **4** 337–357. https://doi.org/10.1023/A:1023562417138

Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika* **83** 95–110. MR1399158 https://doi.org/10.1093/biomet/83.1.95

Robinson, R.W. (1977). Counting unlabeled acyclic digraphs. In *Combinatorial Mathematics V* (C.H.C. Little, ed.) 28–43. Berlin, Heidelberg: Springer. MR476569 https://doi.org/10.1007/BFb0069178

Rosenthal, J.S. (1995). Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **90** 558–566. MR1340509 https://doi.org/10.1080/01621459.1995.10476548

Rosenthal, J.S. (2011). Optimal proposal distributions and adaptive MCMC. In *Handbook of Bayesian Variable Selection* 93–111. Boca Raton, FL: Chapman & Hall/CRC. MR2858446 https://doi.org/10.1201/9781003089018

Roy, V. and Hobert, J.P. (2007). Convergence rates and asymptotic standard errors for Markov chain Monte Carlo algorithms for Bayesian probit regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 607–623. MR2370071 https://doi.org/10.1111/j.1467-9868.2007.00602.x

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

Tadesse, M.G. and Vannucci, M. (2021). *Handbook of Bayesian Variable Selection*. CRC Press. https://doi.org/10.1201/9781003089018

Tang, R. and Yang, Y. (2024). On the computational complexity of Metropolis-adjusted Langevin algorithms for Bayesian posterior sampling. *J. Mach. Learn. Res.* **25** 1–79. MR4758128 https://doi.org/10.48550/arXiv.2206.06491

Titsias, M.K. and Yau, C. (2017). The Hamming ball sampler. *J. Amer. Statist. Assoc.* **112** 1598–1611. MR3750884 https://doi.org/10.1080/01621459.2016.1222288

Woodard, D.B., Schmidler, S.C. and Huber, M. (2009). Conditions for rapid mixing of parallel and simulated tempering on multimodal distributions. *Ann. Appl. Probab.* **19** 617–640. MR2521882 https://doi.org/10.1214/08-AAP555

Yang, J. and Rosenthal, J.S. (2023). Complexity results for MCMC derived from quantitative bounds. *Ann. Appl. Probab.* **33** 1459–1500. MR4564431 https://doi.org/10.1214/22-aap1846

Yang, Y., Wainwright, M.J. and Jordan, M.I. (2016). On the computational complexity of high-dimensional Bayesian variable selection. *Ann. Statist.* **44** 2497–2532. MR3576552 https://doi.org/10.1214/15-AOS1417

Zanella, G. (2020). Informed proposals for local MCMC in discrete spaces. *J. Amer. Statist. Assoc.* **115** 852–865. MR4107684 https://doi.org/10.1080/01621459.2019.1585255

Zanella, G. and Roberts, G. (2019). Scalable importance tempering and Bayesian variable selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 489–517. MR3961496 https://doi.org/10.1111/rssb.12316

Zhang, R., Liu, X. and Liu, Q. (2022). A Langevin-like sampler for discrete distributions. In *International Conference on Machine Learning* 26375–26396. PMLR. https://doi.org/10.48550/arXiv.2206.09914.

Zhou, S. (2010). Thresholded Lasso for high dimensional variable selection and statistical estimation. arXiv preprint. https://doi.org/10.48550/arXiv.1002.1583.

Zhou, Q. and Chang, H. (2023). Complexity analysis of Bayesian learning of high-dimensional DAG models and their equivalence classes. *Ann. Statist.* **51** 1058–1085. MR4630940 https://doi.org/10.1214/23-aos2280

Zhou, Q. and Guan, Y. (2019). Fast model-fitting of Bayesian variable selection regression using the iterative complex factorization algorithm. *Bayesian Anal.* **14** 573–594. MR3934098 https://doi.org/10.1214/18-BA1120

Zhou, Q. and Smith, A. (2022). Rapid convergence of informed importance tempering. In *International Conference on Artificial Intelligence and Statistics* 10939–10965. PMLR. https://doi.org/10.48550/arXiv.2107.10827

Zhou, Q., Yang, J., Vats, D., Roberts, G.O. and Rosenthal, J.S. (2022). Dimension-free mixing for high-dimensional Bayesian variable selection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1751–1784. MR4515557 https://doi.org/10.1111/rssb.12546

Zhuo, B. and Gao, C. (2021). Mixing time of Metropolis-Hastings for Bayesian community detection. *J. Mach. Learn. Res.* **22** 1–89. MR4253703 https://doi.org/10.48550/arXiv.1811.02612

Bernoulli **32**(2), 2026, 1432–1455
https://doi.org/10.3150/25-BEJ1913

Long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos for mean field interacting particle system

XING HUANG[1,a], FEN-FEN YANG[2,3,b] and CHENGGUI YUAN[4,c]

[1]*Center for Applied Mathematics, Tianjin University, Tianjin 300072, China,* [a]*xinghuang@tju.edu.cn*
[2]*Department of Mathematics, Shanghai University, Shanghai 200444, China,* [b]*yangfenfen@shu.edu.cn*
[3]*Newtouch Center for Mathematics of Shanghai University, Shanghai, 200444 China*
[4]*Department of Mathematics, Swansea University, Bay campus, SA1 8EN, UK,* [c]*c.yuan@swansea.ac.uk*

In this paper, a general result on the long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos (PoC), one type of the PoC with regularization effect, is derived for mean field interacting particle system driven by Lévy noise, where TV is the total variation distance and $\mathbb{W}_{\ell_1}$ is the L^1-Wasserstein distance. By using the method of coupling, the general result is applied to mean field interacting particle system driven by Brownian motion and $\alpha(\alpha > 1)$-stable noise respectively, where the non-interacting drift is assumed to be dissipative in long distance.

Keywords: α-stable noise; McKean-Vlasov SDEs; mean field interacting particle system; quantitative PoC; reflection coupling; total variation distance

References

Bao, J. and Wang, J. (2025). Long time behavior of one-dimensional McKean-Vlasov SDEs with common noise. *J. Math. Anal. Appl.* **552** 129819. MR4925916 https://doi.org/10.1016/j.jmaa.2025.129819

Bresch, D., Jabin, P.-E. and Wang, Z. (2023). Mean-field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.* **172** 2591–2641. MR4658923 https://doi.org/10.1215/00127094-2022-0088

Chen, F., Lin, Y., Ren, Z. and Wang, S. (2024). Uniform-in-time propagation of chaos for kinetic mean field Langevin dynamics. *Electron. J. Probab.* **29** 1–43. MR4701825 https://doi.org/10.1214/24-EJP1079

Durmus, A., Eberle, A., Guillin, A. and Zimmer, R. (2020). An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** 5387–5398. MR4163850 https://doi.org/10.1090/proc/14612

Huang, X., Yang, F.-F. and Yuan, C. (2026). Supplement to "Long time TV-$\mathbb{W}_{\ell_1}$ type propagation of chaos for mean field interacting particle system." https://doi.org/10.3150/25-BEJ1913SUPP

Jabin, P.-E. and Wang, Z. (2016). Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *J. Funct. Anal.* **271** 3588–3627. MR3558251 https://doi.org/10.1016/j.jfa.2016.09.014

Jabin, P.-E. and Wang, Z. (2018). Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.* **214** 523–591. MR3858403 https://doi.org/10.1007/s00222-018-0808-y

Kac, M. (1954–1955). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. III* 171–197. University of California Press. MR0084985

Lacker, D. (2023). Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.* **4** 377–432. MR4595391 https://doi.org/10.2140/pmp.2023.4.377

Lacker, D. and Flem, L.L. (2023). Sharp uniform-in-time propagation of chaos. *Probab. Theory Related Fields* **187** 443–480. MR4634344 https://doi.org/10.1007/s00440-023-01192-x

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

Liu, W., Wu, L. and Zhang, C. (2021). Long-time behaviors of mean-field interacting particle systems related to McKean-Vlasov equations. *Comm. Math. Phys.* **387** 179–214. MR4312362 https://doi.org/10.1007/s00220-021-04198-5

Lukkarinen, J. (2023). Generation and propagation of chaos in the stochastic Kac model. Talk at Rutgers.

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

McKean, H.P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907

McKean, H.P. and Singer, I.M. (1967). Curvature and the eigenvalues of the Laplacian. *J. Differential Geom.* **1** 43–69. MR0217739

Monmarché, P., Ren, Z. and Wang, S. (2024). Time-uniform log-Sobolev inequalities and applications to propagation of chaos. *Electron. J. Probab.* **29** 1–38. MR4816013

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

Ren, P. and Wang, F.-Y. (2025). Bi-coupling method and applications. *Probab. Theory Related Fields*. https://doi.org/10.1007/s00440-025-01394-5

Rosenzweig, M. and Serfaty, S. (2025). Modulated logarithmic Sobolev inequalities and generation of chaos. *Ann. Fac. Sci. Toulouse Math. (6)* **34** 107–134. https://doi.org/10.5802/afst.1807

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

Sznitman, A.-S. (1991). Topics in propagation of chaos. In *"École d'Été de Probabilités de Sain-Flour XIX-1989"*. *Lecture Notes in Mathematics* **1464** 165–251. Berlin: Springer. MR1108185 https://doi.org/10.1007/BFb0085169

Wang, F.-Y. (2011). Harnack inequality for SDE with multiplicative noise and extension to Neumann semigroup on nonconvex manifolds. *Ann. Probab.* **39** 1449–1467. MR2857246 https://doi.org/10.1214/10-AOP600

Wang, F.-Y. (2015). Asymptotic couplings by reflection and applications for nonlinear monotone SPDES. *Nonlinear Anal.* **117** 169–188. MR3316613 https://doi.org/10.1016/j.na.2015.01.012

Wang, F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stoch. Process. Appl.* **128** 595–621. MR3739509 https://doi.org/10.1016/j.spa.2017.05.006

Wang, F.-Y. (2023). Exponential ergodicity for non-dissipative McKean-Vlasov SDEs. *Bernoulli* **29** 1035–1062. MR4550214 https://doi.org/10.3150/22-bej1489

Wang, F.-Y. and Wang, J. (2014). Harnack inequalities for stochastic equations driven by Lévy noise. *J. Math. Anal. Appl.* **410** 513–523. MR3109860 https://doi.org/10.1016/j.jmaa.2013.08.013

Bernoulli **32**(2), 2026, 1456–1480
https://doi.org/10.3150/25-BEJ1914

Covariance change point localisation and inference in fragmented functional data

GENGYU XUE[a], HAOTIAN XU[b] and YI YU[c]

Department of Statistics, University of Warwick, Coventry, UK, [a]*gengyu.xue@warwick.ac.uk,* [b]*haotian.xu.1@warwick.ac.uk,* [c]*yi.yu.2@warwick.ac.uk*

We study the problem of covariance change point localisation and inference for sequentially collected fragmented functional data, where each curve is observed only over discrete grids randomly sampled over a short fragment. The sequence of underlying covariance functions is assumed to be piecewise constant, with changes happening at unknown time points. To localise the change points, we propose a computationally efficient fragmented functional dynamic programming (FFDP) algorithm with consistent change point localisation rates. With an extra step of local refinement, we derive the limiting distributions for the refined change point estimators in two different regimes where the minimal jump size vanishes and where it remains constant as the sample size diverges. Such results are the first time seen in the fragmented functional data literature. As a byproduct of independent interest, we also present a non-asymptotic result on the estimation error of the covariance function estimators over intervals with change points. Our result accounts for the effects of the sampling grid size within each fragment under novel identifiability conditions. Extensive numerical studies are also provided to support our theoretical results.

Keywords: Change point analysis; covariance function estimation; fragmented functional data; inference

References

Aue, A., Rice, G. and Sönmez, O. (2018). Detecting and dating structural breaks in functional data without dimension reduction. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 509–529.

Aue, A., Rice, G. and Sönmez, O. (2020). Structural break analysis for spectrum and trace of covariance operators. *Environmetrics* **31** e2617.

Berkes, I., Gabrys, R., Horváth, L. and Kokoszka, P. (2009). Detecting changes in the mean of functional observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 927–946.

Cai, T.T., Kim, D. and Pu, H. (2024). Transfer learning for functional mean estimation: phase transition and adaptive algorithms. arXiv preprint. Available at arXiv:2401.12331.

Cai, T. and Yuan, M. (2010). Nonparametric covariance function estimation for functional and longitudinal data Technical report, University of Pennsylvania, Philadelphia, PA.

Cai, T.T. and Yuan, M. (2011). Optimal estimation of the mean function based on discretely sampled functional data: Phase transition. *Ann. Statist.* **39** 2330–2355.

Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *J. Amer. Statist. Assoc.* **107** 1201–1216.

Canuto, C., Hussaini, M.Y., Quarteroni, A. and Zang, T.A. (2007). *Spectral Methods: Fundamentals in Single Domains*. Springer Science & Business Media.

Chan, N.H., Ng, W.L., Yau, C.Y. and Yu, H. (2021). Optimal change-point estimation in time series. *Ann. Statist.* **49** 2336–2355.

Cho, H. and Kirch, C. (2022). Bootstrap confidence intervals for multiple change points based on moving sum procedures. *Comput. Statist. Data Anal.* **175** 107552.

Cho, H. and Owens, D. (2022). High-dimensional data segmentation in regression settings permitting heavy tails and temporal dependence. arXiv preprint. Available at arXiv:2209.08892.

Crawford, L., Monod, A., Chen, A.X., Mukherjee, S. and Rabadán, R. (2020). Predicting clinical outcomes in glioblastoma: An application of topological and functional data analysis. *J. Amer. Statist. Assoc.* **115** 1139–1150.

Dai, X., Müller, H.-G., Wang, J.-L. and Deoni, S.C. (2019). Age-dynamic networks and functional correlation for early white matter myelination. *Brain Struct. Funct.* **224** 535–551.

Delaigle, A. and Hall, P. (2012). Achieving near perfect classification for functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 267–286.

Delaigle, A. and Hall, P. (2016). Approximating fragmented functional data by segments of Markov chains. *Biometrika* **103** 779–799.

Delaigle, A., Hall, P., Huang, W. and Kneip, A. (2021). Estimating the covariance of fragmented and other related types of functional data. *J. Amer. Statist. Assoc.* **116** 1383–1401.

Descary, M.-H. and Panaretos, V.M. (2019). Recovering covariance from functional fragments. *Biometrika* **106** 145–160.

Dette, H. and Kokot, K. (2022). Detecting relevant differences in the covariance operators of functional time series: A sup-norm approach. *Ann. Inst. Statist. Math.* **74** 195–231.

Dette, H. and Kutta, T. (2021). Detecting structural breaks in eigensystems of functional time series. *Electron. J. Stat.* **15** 944–983.

Ember (2024). European wholesale electricity price data. https://ember-climate.org/data-catalogue/european-wholesale-electricity-price-data/.

European Network of Transmission System Operators for Electricity (2023). Monthly Hourly Load Values. https://www.entsoe.eu/data/power-stats/.

Fan, J., Huang, T. and Li, R. (2007). Analysis of longitudinal data with semiparametric estimation of covariance function. *J. Amer. Statist. Assoc.* **102** 632–641.

Fraiman, R., Justel, A., Liu, R. and Llop, P. (2014). Detecting trends in time series of functional data: A study of Antarctic climate change. *Canad. J. Statist.* **42** 597–609.

Friedrich, F., Kempe, A., Liebscher, V. and Winkler, G. (2008). Complexity penalized M-estimation: Fast computation. *J. Comput. Graph. Statist.* **17** 201–224.

Gromenko, O., Kokoszka, P. and Sojka, J. (2017). Evaluation of the cooling trend in the ionosphere using functional regression with incomplete curves. *Ann. Appl. Stat.* **11** 898–918.

Guo, S., Li, D., Qiao, X. and Wang, Y. (2023). From sparse to dense functional data in high dimensions: Revisiting phase transitions from a non-asymptotic perspective. arXiv preprint. Available at arXiv:2306.00476.

Hasenstab, K., Scheffler, A., Telesca, D., Sugar, C.A., Jeste, S., DiStefano, C. and Şentürk, D. (2017). A multi-dimensional functional principal components analysis of EEG data. *Biometrics* **73** 999–1009.

Horváth, L., Rice, G. and Zhao, Y. (2022). Change point analysis of covariance functions: A weighted cumulative sum approach. *J. Multivariate Anal.* **189** 104877.

International Trade Administration (2023). Germany Country Country Commercial Guide: Energy. https://www.trade.gov/country-commercial-guides/germany-energy.

Jarušková, D. (2013). Testing for a change in covariance operator. *J. Statist. Plann. Inference* **143** 1500–1511.

Jiao, S., Frostig, R.D. and Ombao, H. (2023). Break point detection for functional covariance. *Scand. J. Stat.* **50** 477–512.

Kneip, A. and Liebl, D. (2020). On the optimal reconstruction of partially observed functional data. *Ann. Statist.* **48** 1692–1717.

Kovács, S., Bühlmann, P., Li, H. and Munk, A. (2023). Seeded binary segmentation: A general methodology for fast and optimal changepoint detection. *Biometrika* **110** 249–256.

Krantz, S.G. and Parks, H.R. (2002). *A Primer of Real Analytic Functions*. Springer Science & Business Media.

Kraus, D. (2015). Components and completion of partially observed functional data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 777–801.

Lei, E., Yao, F., Heckman, N. and Meyer, K. (2015). Functional data model for genetically related individuals with application to cow growth. *J. Comput. Graph. Statist.* **24** 756–770.

Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. *Ann. Appl. Stat.* 1562–1592.

Liebl, D. (2019). Nonparametric testing for differences in electricity prices: The case of the Fukushima nuclear accident. *Ann. Appl. Stat.* **13** 1128–1146.

Liebl, D. and Rameseder, S. (2019). Partially observed functional data: The case of systematically missing parts. *Comput. Statist. Data Anal.* **131** 104–115.

Lin, Z. and Wang, J.-L. (2022). Mean and covariance estimation for functional snippets. *J. Amer. Statist. Assoc.* **117** 348–360.

Lin, Z., Wang, J.-L. and Zhong, Q. (2021). Basis expansions for functional snippets. *Biometrika* **108** 709–726.

Ling, S. (2016). Estimation of change-points in linear and nonlinear time series models. *Econometric Theory* **32** 402–430.

Ng, W.L., Pan, S. and Yau, C.Y. (2022). Bootstrap inference for multiple change-points in time series. *Econometric Theory* **38** 752–792.

Padilla, O.H.M., Yu, Y. and Priebe, C.E. (2022). Change point localization in dependent dynamic nonparametric random dot product graphs. *J. Mach. Learn. Res.* **23** 10661–10719.

Padilla, O.H.M., Yu, Y., Wang, D. and Rinaldo, A. (2021). Optimal nonparametric multivariate change point detection and localization. *IEEE Trans. Inf. Theory* **68** 1922–1944.

Padilla, C.M.M., Wang, D., Zhao, Z. and Yu, Y. (2022). Change-point detection for sparse and dense functional data in general dimensions. *Adv. Neural Inf. Process. Syst.* **35** 37121–37133.

Padilla, C.M.M., Xu, H., Wang, D., Padilla, O.H.M. and Yu, Y. (2024). Change point detection and inference in multivariate non-parametric models under mixing conditions. *Adv. Neural Inf. Process. Syst.* **36**.

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.

Rice, G. and Zhang, C. (2022). Consistency of binary segmentation for multiple change-point estimation with functional data. *Statist. Probab. Lett.* **180** 109228.

Stoehr, C., Aston, J.A. and Kirch, C. (2021). Detecting changes in the covariance structure of functional time series with application to fMRI data. *Econom. Stat.* **18** 44–62.

van der Vaart, A.W. and Wellner, J.A. (2023). *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York, New York, NY: Springer.

Vershynin, R. (2018). *High-Dimensional Probability: An Introduction with Applications in Data Science* **47**. Cambridge university press.

Verzelen, N., Fromont, M., Lerasle, M. and Reynaud-Bouret, P. (2023). Optimal change-point detection and localization. *Ann. Statist.* **51** 1586–1610.

VoxEU, Centre for Economic Policy Research (2023). Reasons behind the 2022 energy price increases and prospects for next year. https://cepr.org/voxeu/columns/reasons-behind-2022-energy-price-increases-and-prospects-next-year.

Waghmare, K.G. and Panaretos, V.M. (2022). The completion of covariance kernels. *Ann. Statist.* **50** 3281–3306.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Wang, D., Yu, Y. and Rinaldo, A. (2020). Univariate mean change point detection: Penalization, CUSUM and optimality. *Electron. J. Stat.* **14** 1917–1961.

Wang, D., Yu, Y. and Rinaldo, A. (2021). Optimal change point detection and localization in sparse dynamic networks. *Ann. Statist.* **49** 203–232.

Xiao, L., Li, Y. and Ruppert, D. (2013). Fast bivariate P-splines: The sandwich smoother. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 577–599.

Xu, H., Wang, D., Zhao, Z. and Yu, Y. (2022). Change point inference in high-dimensional regression models under temporal dependence. arXiv preprint. Available at arXiv:2207.12453.

Xue, G., Xu, H. and Yu, Y. (2026). Supplement to "Covariance change point localisation and inference in fragmented functional data." https://doi.org/10.3150/25-BEJ1914SUPPA, https://doi.org/10.3150/25-BEJ1914SUPPB

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590.

Yau, C.Y. and Zhao, Z. (2016). Inference for multiple change points in time series via likelihood ratio scan statistics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 895–916.

Zhang, A.R. and Chen, K. (2022). Nonparametric covariance estimation for mixed longitudinal studies, with applications in midlife women's health. *Statist. Sinica* **32** 345–365.

Zhang, X. and Wang, J.-L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321.

Zhou, H., Yao, F. and Zhang, H. (2023). Functional linear regression for discretely observed data: From ideal to reality. *Biometrika* **110** 381–393.

Bernoulli **32**(2), 2026, 1481–1502
https://doi.org/10.3150/25-BEJ1916

Decompounding under general mixing distributions

DENIS BELOMESTNY[1,2,a,b], EKATERINA MOROZOVA[1,c] and VLADIMIR PANOV[2,d]

[1]*University of Duisburg-Essen, Essen, Germany,* [a]*denis.belomestny@uni-due.de,* [b]*https://denbel.github.io,* [c]*ekaterina.morozova@uni-due.de*
[2]*HSE University, Moscow, Russia,* [d]*vpanov@hse.ru*

This study focuses on statistical inference for compound models of the form $X = \xi_1 + \ldots + \xi_N$, where N is a random variable denoting the count of summands, which are independent and identically distributed (i.i.d.) random variables $\xi_1, \xi_2, \ldots$. The paper addresses the problem of reconstructing the distribution of ξ from observed samples of X's distribution, a process referred to as decompounding, with the assumption that N's distribution is known. This work diverges from the conventional scope by not limiting N's distribution to the Poisson type, thus embracing a broader context. We propose a nonparametric estimate for the density of ξ, derive its rates of convergence and prove that these rates are minimax optimal for suitable classes of distributions for ξ and N. Finally, we illustrate the numerical performance of the algorithm on simulated examples.

Keywords: Compound models; decompounding; inverse problem; minimax rates

References

Asmussen, S. and Albrecher, H. (2010). Ruin Probabilities 14. World Scientific.

Belomestny, D., Morozova, E. and Panov, V. (2024). Nonparametric statistical inference for compound models. *Statistics* **58** 943–960.

Belomestny, D., Morozova, E. and Panov, V. (2026). Supplement to "Decompounding under general mixing distributions." https://doi.org/10.3150/25-BEJ1916SUPP

Belomestny, D. and Reiss, M. (2015). Estimation and calibration of Lévy models via Fourier methods. In *Lévy Matters IV. Lecture Notes in Math.* **2128**. Springer.

Bøgsted, M. and Pitts, S.M. (2010). Decompounding random sums: A nonparametric approach. *Ann. Inst. Statist. Math.* **62** 855–872.

Brenner Miguel, S., Comte, F. and Johannes, J. (2023). Linear functional estimation under multiplicative measurement error. *Bernoulli* **29** 2247–2271.

Buchmann, B. and Grübel, R. (2003). Decompounding: An estimation problem for Poisson random sums. *Ann. Statist.* **31** 1054–1074.

Buchmann, B. and Grübel, R. (2004). Decompounding Poisson random sums: Recursively truncated estimates in the discrete case. *Ann. Inst. Statist. Math.* **56** 743–756.

Chydzinski, A. (2020). Queues with the dropping function and non-Poisson arrivals. *IEEE Access* **8** 39819–39829.

Coca, A.J. (2018). Efficient nonparametric inference for discretely observed compound Poisson processes. *Probab. Theory Related Fields* **170** 475–523.

Comte, F., Duval, C. and Genon-Catalot, V. (2014). Nonparametric density estimation in compound Poisson processes using convolution power estimators. *Metrika* **77** 163–183.

Comte, F., Johannes, J. and Neubert, B. (2025). Quadratic functional estimation from observations with multiplicative measurement error. *Ann. Inst. Statist. Math.* In press.

Den Boer, A. and Mandjes, M. (2017). Convergence rates of Laplace-transform based estimators. *Bernoulli* **23** 2533–2557.

1350-7265

Goldenshluger, A. and Kim, T. (2021). Density deconvolution with non–standard error distributions: Rates of convergence and adaptive estimation. *Electron. J. Stat.* **15** 3394–3427.

Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *Ann. Statist.* **39** 1608–1632.

Gorenflo, R. and Hofmann, B. (1994). On autoconvolution and regularization. *Inverse Probl.* **10** 353.

Gugushvili, S., van der Meulen, F. and Spreij, P. (2018). A non-parametric Bayesian approach to decompounding from high frequency data. *Stat. Inference Stoch. Process.* **21** 53–79.

Hansen, M.B. and Pitts, S.M. (2006). Nonparametric inference from the M/G/1 workload. *Bernoulli* **12** 737–759.

Lin, Z. and Bai, Z. (2011). *Probability Inequalities*. Springer.

Meister, A. (2009). Deconvolution problems in nonparametric statistics. *Lecture Notes in Statist.* 5–138.

Sundt, B. and Vernic, R. (2009). *Recursions for Convolutions and Compound Distributions with Insurance Applications*. Springer.

Van Es, B., Gugushvili, S. and Spreij, P. (2007). A kernel type nonparametric density estimator for decompounding. *Bernoulli* **13** 672–694.

Bernoulli **32**(2), 2026, 1503–1522
https://doi.org/10.3150/25-BEJ1917

Parametrization, prior independence, and the semiparametric Bernstein-von Mises theorem for the partially linear model

CHRISTOPHER D. WALKER[a]

Department of Economics, Duke University, Durham, NC 27708, U.S.A , [a]christopher.walker@duke.edu

I prove a semiparametric Bernstein-von Mises theorem for a partially linear regression model with independent priors for the low-dimensional parameter of interest and the infinite-dimensional nuisance parameters. My result avoids a challenging prior invariance condition that arises from a loss of information associated with not knowing the nuisance parameter. The key idea is to employ a feasible reparametrization of the partially linear regression model that reflects the semiparametric structure of the model. This allows a researcher to assume independent priors for the model parameters while automatically accounting for the loss of information associated with not knowing the nuisance parameters. The theorem is verified for uniform wavelet series priors and Matérn Gaussian process priors.

Keywords: Adaptive Parametrization; Bernstein-von Mises; Partially linear model; Prior Invariance

References

Ackerberg, D.A., Caves, K. and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica* **83** 2411–2451.

Ahn, H. and Powell, J.L. (1993). Semiparametric estimation of censored selection models with a nonparametric selection mechanism. *J. Econometrics* **58** 3–29.

Andrews, D.W.K. (1994). Asymptotics for semiparametric econometric models via stochastic equicontinuity. *Econometrica* 43–72.

Bickel, P.J. (1982). On adaptive estimation. *Ann. Statist.* **10** 647–671. https://doi.org/10.1214/aos/1176345863

Bickel, P.J. and Kleijn, B.J.K. (2012). The semiparametric Bernstein–von Mises theorem. *Ann. Statist.* **40** 206–237. https://doi.org/10.1214/11-AOS921

Breunig, C., Liu, R. and Yu, Z. (2022). Double Robust Bayesian Inference on Average Treatment Effects. arXiv preprint. Available at arXiv:2211.16298.

Castillo, I. (2008). Lower bounds for posterior rates with Gaussian process priors. *Electron. J. Stat.* **2** 1281–1299. https://doi.org/10.1214/08-EJS273

Castillo, I. (2012a). A semiparametric Bernstein–von Mises theorem for Gaussian process priors. *Probab. Theory Related Fields* **152** 53–99.

Castillo, I. (2012b). Semiparametric Bernstein–von Mises theorem and bias, illustrated with Gaussian process priors. *Sankhya A* **74** 194–221.

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

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

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

Chae, M., Kim, Y. and Kleijn, B.J. (2019). The semi-parametric Bernstein-von Mises theorem for regression models with symmetric errors. *Statist. Sinica* **29** 1465–1487.

1350-7265

Chamberlain, G. (1992). Efficiency bounds for semiparametric regression. *Econometrica* 567–596.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econom. J.*

Cohen, A., Daubechies, I. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.*

Cox, D.R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *J. Roy. Statist. Soc. Ser. B, Methodol.* **49** 1–18. https://doi.org/10.1111/j.2517-6161.1987.tb01422.x

Das, M., Newey, W.K. and Vella, F. (2003). Nonparametric estimation of sample selection models. *Rev. Econ. Stud.* **70** 33–58.

de Jonge, R. and van Zanten, J.H. (2012). Adaptive estimation of multivariate functions using conditionally Gaussian tensor-product spline priors. *Electron. J. Stat.* **6** 1984–2001. https://doi.org/10.1214/12-EJS735

de Jonge, R. and van Zanten, H. (2013). Semiparametric Bernstein–von Mises for the error standard deviation. *Electron. J. Stat.* **7** 217–243. https://doi.org/10.1214/13-EJS768

Donald, S.G. and Newey, W.K. (1994). Series estimation of semilinear models. *J. Multivariate Anal.* **50** 30–40. https://doi.org/10.1006/jmva.1994.1032

Engle, R.F., Granger, C.W., Rice, J. and Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *J. Amer. Statist. Assoc.* **81** 310–320.

Ghosal, S., Ghosh, J.K. and Van Der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* 500–531.

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

Ghosal, S. and Van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference* **44**. Cambridge University Press.

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

Giné, E. and Nickl, R. (2015). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge University Press.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica* **52** 681–700.

Hahn, P.R., Carvalho, C.M., Puelz, D. and He, J. (2018). Regularization and confounding in linear regression for treatment effect estimation. *Bayesian Anal.* **13** 163–182. https://doi.org/10.1214/16-BA1044

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

Kleijn, B.J.K. and van der Vaart, A.W. (2006). Misspecification in infinite-dimensional Bayesian statistics. *Ann. Statist.* **34** 837–877. https://doi.org/10.1214/009053606000000029

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles. *J. Econometrics* **128** 137–164. https://doi.org/10.1016/j.jeconom.2004.08.010

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables. *Rev. Econ. Stud.* **70** 317–341.

L'Huillier, A., Travis, L., Castillo, I. and Ray, K. (2023). Semiparametric inference using fractional posteriors. *J. Mach. Learn. Res.* **24** 1–61.

Monard, F., Nickl, R. and Paternain, G.P. (2021). Statistical guarantees for Bayesian uncertainty quantification in nonlinear inverse problems with Gaussian process priors. *Ann. Statist.* **49** 3255–3298. https://doi.org/10.1214/21-AOS2082

Murphy, S.A. and Van der Vaart, A.W. (2000). On profile likelihood. *J. Amer. Statist. Assoc.* **95** 449–465.

Newey, W.K. (1990). Semiparametric efficiency bounds. *J. Appl. Econometrics* **5** 99–135. https://doi.org/10.1002/jae.3950050202

Newey, W.K. (1994). The asymptotic variance of semiparametric estimators. *Econometrica* **62** 1349–1382.

Nickl, R. (2020). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc.* **22** 2697–2750.

Nickl, R. and Söhl, J. (2019). Bernstein–von Mises theorems for statistical inverse problems II: Compound Poisson processes. *Electron. J. Stat.* **13** 3513–3571. https://doi.org/10.1214/19-EJS1609

Norets, A. (2015). Bayesian regression with nonparametric heteroskedasticity. *J. Econometrics* **185** 409–419.

Olley, G.S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica* **64** 1263–1297.

Rasmussen, C.E. and Williams, C.K.I. (2005). *Gaussian Processes for Machine Learning*. The MIT Press. https://doi.org/10.7551/mitpress/3206.001.0001

Ray, K. and Szabó, B. (2019). Debiased Bayesian inference for average treatment effects. *Adv. Neural Inf. Process. Syst.* **32**.

Ray, K. and van der Vaart, A. (2020). Semiparametric Bayesian causal inference. *Ann. Statist.* **48** 2999–3020. https://doi.org/10.1214/19-AOS1919

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

Robinson, P.M. (1988). Root-N-consistent semiparametric regression. *Econometrica* 931–954.

Severini, T.A. and Wong, W.H. (1992). Profile likelihood and conditionally parametric models. *Ann. Statist.* 1768–1802.

Shen, X. (2002). Asymptotic normality of semiparametric and nonparametric posterior distributions. *J. Amer. Statist. Assoc.* **97** 222–235.

Shen, W. and Ghosal, S. (2015). Adaptive Bayesian procedures using random series priors. *Scand. J. Stat.* **42** 1194–1213.

Shen, X. and Wasserman, L. (2001). Rates of convergence of posterior distributions. *Ann. Statist.* **29** 687–714. https://doi.org/10.1214/aos/1009210686

van de Geer, S. (2000). *Empirical Processes in M-Estimation 6*. Cambridge university press.

van der Vaart, A.W. and van Zanten, J.H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463. https://doi.org/10.1214/009053607000000613

van der Vaart, A.W. and van Zanten, J.H. (2009). Adaptive Bayesian estimation using a Gaussian random field with inverse Gamma bandwidth. *Ann. Statist.* **37** 2655–2675. https://doi.org/10.1214/08-AOS678

van der Vaart, A.W. and van Zanten, J.H. (2011). Information rates of nonparametric Gaussian process methods. *J. Mach. Learn. Res.* **12**.

van der Vaart, A.W. and Wellner, J.A. (2023). *Weak Convergence and Empirical Process: With Applications to Statistics*. Springer.

Walker, C.D. (2026). Supplement to "Parametrization, Prior Independence, and the Semiparametric Bernstein-von Mises Theorem for the Partially Linear Model." https://doi.org/10.3150/25-BEJ1917SUPP

White, H. (1994). *Estimation, Inference and Specification Analysis* **22**. Cambridge university press.

Wooldridge, J.M. (2009). On estimating firm-level production functions using proxy variables to control for unobservables. *Econom. Lett.* **104** 112–114.

Xie, F. and Xu, Y. (2020). Adaptive Bayesian nonparametric regression using a kernel mixture of polynomials with application to partial linear models. *Bayesian Anal.* **15** 159–186. https://doi.org/10.1214/19-BA1148

Yang, D. (2019). Posterior asymptotic normality for an individual coordinate in high-dimensional linear regression. *Electron. J. Stat.* **13** 3082–3094. https://doi.org/10.1214/19-EJS1605

Yang, Y., Cheng, G. and Dunson, D.B. (2015). Semiparametric Bernstein-von Mises theorem: Second order studies. arXiv preprint. Available at arXiv:1503.04493.

Yiu, A., Fong, E., Holmes, C. and Rousseau, J. (2025). Semiparametric posterior corrections. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **qkaf005**. https://doi.org/10.1093/jrsssb/qkaf005

Bernoulli **32**(2), 2026, 1523–1546
https://doi.org/10.3150/25-BEJ1918

Estimating invertible processes in Hilbert spaces, with applications to functional ARMA processes

SEBASTIAN KÜHNERT[1,a], GREGORY RICE[2,b] and ALEXANDER AUE[3,c]

[1]*Department of Mathematics, Ruhr University Bochum, Bochum, Germany,* [a]*sebastian.kuehnert@ruhr-uni-bochum.de*
[2]*Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada,* [b]*grice@uwtaerloo.ca*
[3]*Department of Statistics, University of California, Davis, United States of America,* [c]*aaue@ucdavis.edu*

Invertible processes are central to functional time series analysis, making the estimation of their defining operators a key problem. While asymptotic error bounds have been established for specific ARMA models on $L^2[0,1]$, a general theoretical framework has not yet been considered. This paper fills in this gap by deriving consistent estimators for the operators characterizing the invertible representation of a functional time series with white noise innovations in a general separable Hilbert space. Under mild conditions covering a broad class of functional time series, we establish explicit asymptotic error bounds, with rates determined by operator smoothness and eigenvalue decay. These results further provide consistency-rate estimates for operators in Hilbert space-valued causal linear processes, including functional MA, AR, and ARMA models of arbitrary order.

Keywords: ARMA; functional time series; invertible processes; linear processes; operator estimation

References

Aue, A., Horváth, L. and Pellatt, D. (2017). Functional generalized autoregressive conditional heteroskedasticity. *J. Time Series Anal.* **38** 3–21.

Aue, A. and Klepsch, J. (2017). Estimating functional time series by moving average model fitting. arXiv:1701.00770. https://arxiv.org/abs/1701.00770.

Aue, A. and van Delft, A. (2020). Testing for stationarity of functional time series in the frequency domain. *Ann. Statist.* **48** 2505–2547.

Bosq, D. (2000). *Linear Processes in Function Spaces*. *Lecture Notes in Statistics*. New York: Springer.

Caponera, A. and Panaretos, V. (2022). On the rate of convergence for the autocorrelation operator in functional autoregression. *Statist. Probab. Lett.* **189**.

Cerovecki, C., Francq, C., Hörmann, S. and Zakoïan, J.-M. (2019). Functional GARCH models: The quasi-likelihood approach and its applications. *J. Econometrics* **209** 353–375.

Düker, M.-C. (2018). Limit theorems for Hilbert space-valued linear processes under long-range dependence. *Stoch. Process. Appl.* **128** 1439–1465.

Granger, C.W.J. and Andersen, A. (1978). On the invertibility of time series models. *Stoch. Process. Appl.* **8** 87–92.

Hall, P. and Horowitz, J. (2005). Nonparametric methods for inference in the presence of instrumental variables. *Ann. Statist.* **41** 2904–2929.

Hall, P. and Horowitz, J. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91.

Hall, P. and Meister, A. (2007). A ridge-parameter approach to deconvolution. *Ann. Statist.* **35** 1535–1558.

Hallin, M. (1980). Invertibility and generalized invertibility of time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **42** 210–212.

1350-7265

Hallin, M. (1981). Addendum to: “Invertibility and generalized invertibility of time series models”. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **43** 103.

Hannan, E.J. and Rissanen, J. (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika* **69** 81–94.

Hörmann, S., Horváth, L. and Reeder, R. (2013). A functional version of the ARCH model. *Econometric Theory* **29** 267–288.

Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884.

Horváth, L. and Kokoszka, P. (2012). *Inference for Functional Data with Applications*. New York: Springer.

Horváth, L., Kokoszka, P. and Rice, G. (2014). Testing stationarity of functional time series. *J. Econometrics* **179** 66–82.

Jaimez, R. and Bonnet, M. (1987). On the Karhunen-Loeve expansion for transformed processes. *Trabajos de Estadistica* **2** 81–90.

Kara Terki, N. and Mourid, T. (2024). Exponential bounds and convergence rates of sieve estimators for functional autoregressive processes. *Sankhya A* **86** 364–391.

Klepsch, J. and Klüppelberg, C. (2017). An innovations algorithm for the prediction of functional linear processes. *J. Multivariate Anal.* **155** 252–271.

Klepsch, J., Klüppelberg, C. and Wei, T. (2017). Prediction of functional ARMA processes with an application to traffic data. *Econom. Stat.* **1** 128–149.

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. New York: Chapman and Hall/CRC.

Kuenzer, T. (2024). Estimation of functional ARMA models. *Bernoulli* **30** 117–142.

Kühnert, S. (2020). Functional ARCH and GARCH models: A Yule-Walker approach. *Electron. J. Stat.* **14** 4321–4360.

Kühnert, S. (2025). Estimating lagged (cross-)covariance operators of L^p-m-approximable processes in Cartesian product Hilbert spaces. *J. Time Series Anal.* **46** 582–595.

Kutta, T. (2025). Approximately mixing time series. *Statist. Probab. Lett.* **220**.

Liebl, D. (2013). Modeling and forecasting electricity spot prices: A functional data perspective. *Ann. Appl. Stat.* **7** 1562–1592.

Mas, A. and Pumo, B. (2009). Linear processes for functional data. arXiv:0901.2503. https://arxiv.org/abs/0901.2503.

Merlevède, F. (1996). Central limit theorem for linear processes with values in a Hilbert space. *Stoch. Process. Appl.* **65** 103–114.

Merlevède, F., Peligrad, M. and Utev, S. (1997). Sharp conditions for the CLT of linear processes in a Hilbert space. *J. Theoret. Probab.* **10** 681–693.

Račkauskas, A. and Suquet, C. (2010). On limit theorems for Banach-space-valued linear processes. *Lith. Math. J.* **50** 71–87.

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

Reimherr, M. (2015). Functional regression with repeated eigenvalues. *Statist. Probab. Lett.* **107** 62–70.

Spangenberg, F. (2013). Strictly stationary solutions of ARMA equations in Banach spaces. *J. Multivariate Anal.* **121** 127–138.

Turbillon, C., Bosq, D., Marion, J.-M. and Pumo, B. (2008). Estimation du paramètre des moyennes mobiles Hilbertiennes. *C. R. Math. Acad. Sci. Paris* **346** 347–350.

van Delft, A., Characiejus, V. and Dette, H. (2021). A nonparametric test for stationarity in functional time series. *Statist. Sinica* **31** 1375–1395.

Weidmann, J. (1980). *Linear Operators in Hilbert Spaces*. *Graduate Texts in Mathematics* **68**. New York-Berlin: Springer.

Bernoulli **32**(2), 2026, 1547–1569
https://doi.org/10.3150/25-BEJ1919

Optimal matching problem on the Boolean cube

SHI FENG[a]

Department of Mathematics, Cornell University, Ithaca, USA, [a]*sf599@cornell.edu*

We establish upper and lower bounds for the expected Wasserstein-1 distance between the random empirical measure and the uniform measure on the Boolean cube. Our analysis leverages techniques from Fourier analysis, following the framework introduced in (*Ann. Appl. Probab.* **31** (2021) 2567–2584), as well as methods from large deviations theory.

Keywords: Boolean cube; optimal matching problem

References

Ajtai, M., Komlós, J. and Tusnády, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264.
Ambrosio, L., Stra, F. and Trevisan, D. (2019). A PDE approach to a 2-dimensional matching problem. *Probab. Theory Related Fields* **173** 433–477.
Ash, R.B. (2012). *Information Theory*. Courier Corporation.
Barthe, F. and Bordenave, C. (2013). Combinatorial optimization over two random point sets. In *Séminaire de Probabilités XLV* 483–535. Springer.
Bobkov, S.G. and Ledoux, M. (2021). A simple Fourier analytic proof of the AKT optimal matching theorem. *Ann. Appl. Probab.* **31** 2567–2584.
Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford Univ. Press.
Del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 896–926.
Dereich, S., Scheutzow, M. and Schottstedt, R. (2013). Constructive quantization: Approximation by empirical measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1183–1203.
Dobrić, V. and Yukich, J.E. (1995). Asymptotics for transportation cost in high dimensions. *J. Theoret. Probab.* **8** 97–118.
Efron, B. and Stein, C. (1981). The jackknife estimate of variance. *Ann. Statist.* **9** 586–596.
Fournier, N. (2023). Convergence of the empirical measure in expected Wasserstein distance: Non-asymptotic explicit bounds in $\mathbb{R}^d$. *ESAIM Probab. Stat.* **27** 749–775.
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.
O'Donnell, R. (2014). *Analysis of Boolean Functions*. Cambridge Univ. Press.
Talagrand, M. (1992). Matching random samples in many dimensions. *Ann. Appl. Probab.* **2** 846–856.
Talagrand, M. (2014). *Upper and Lower Bounds for Stochastic Processes* **60**. Springer.
Villani, C. (2009). *Optimal Transport: Old and New 338*. 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.

Bernoulli **32**(2), 2026, 1570–1593
https://doi.org/10.3150/25-BEJ1920

Robust functional data analysis: From sparse to dense designs

LINGXUAN SHAO[1,a] and FANG YAO[2,b]

[1]*Department of Statistics and Data Science, School of Management, Fudan University, Shanghai, China,* [a]*shao_lingxuan@fudan.edu.cn*
[2]*Department of Probability and Statistics, Center for Statistical Science, School of Mathematical Sciences, Peking University, Beijing, China,* [b]*fyao@math.pku.edu.cn*

We propose a new perspective to conduct robust functional data analysis of discretely observed functional data ranging from sparse to dense sampling designs. This analysis caters to processes with various distributions, including heavy-tailed, skewed, or contaminated distributions. We study the robust functional mean (M-location) and introduce a robust dimension reduction method via principal component analysis. Theoretical outcomes for the robust functional mean and eigenfunction estimates, derived from pooled discretely observed data, are elucidated, matching their non-robust counterparts. The established convergence rates for these estimated eigenfunctions, with indices increasing with sample size, pave the way for further modeling and analysis.

Keywords: Perturbation analysis; robust dimension reduction; sparse functional data

References

Bali, J.L., Boente, G., Tyler, D.E. and Wang, J.-L. (2011). Robust functional principal components: A projection-pursuit approach. *Ann. Statist.* **39** 852–882.
Boente, G. and Salibián-Barrera, M. (2021). Robust functional principal components for sparse longitudinal data. *Metron* **79** 159–188.
Bosq, D. (2000). *Linear Process in Function Spaces*. *Lecture Notes in Statistics*. Springer.
Cardot, H. and Godichon-Baggioni, A. (2017). Fast estimation of the median covariation matrix with application to online robust principal components analysis. *TEST* **26** 461–480.
Dai, W., Mrkvicka, T., Sun, Y. and Genton, M.G. (2020). Functional outlier detection and taxonomy by sequential transformations. *Comput. Statist. Data Anal.* **149** 106960.
Dou, W.W., Pollard, D. and Zhou, H. (2012). Estimation in functional regression for general exponential families. *Ann. Statist.* **40** 2421–2451.
Dubey, P. and Müller, H.-G. (2022). Modeling time-varying random objects and dynamic networks. *J. Amer. Statist. Assoc.* **117** 2252–2267.
Fletcher, T. and Joshi, S. (2007). Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Process.* **87** 250–262.
Hall, P. (1984). Integrated square error properties of kernel estimators of regression functions. *Ann. Statist.* **12** 241–260.
Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91.
Hall, P. and Hosseini-Nasab, M. (2006). On properties of functional principal components analysis. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 109–126.
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.
Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley.
Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

1350-7265

Huber, P.J. (2011). Robust statistics. In *International Encyclopedia of Statistical Science 1248–1251*. Berlin, Heidelberg: Springer.

Jiang, J. and Mack, Y. (2001). Robust local polynomial regression for dependent data. *Statist. Sinica* **11** 705–722.

Kai, B., Li, R. and Zou, H. (2010). Local composite quantile regression smoothing: An efficient and safe alternative to local polynomial regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 49–69.

Kalogridis, I. and Van Aelst, S. (2022). Robust optimal estimation of location from discretely sampled functional data. *Scand. J. Stat.* 1–41.

Kokoszka, P. and Reimherr, M. (2017). *Introduction to Functional Data Analysis*. New York: Chapman and Hall/CRC.

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.

Lima, I.R., Cao, G. and Billor, N. (2018). M-based simultaneous inference for the mean function of functional data. *Ann. Inst. Statist. Math.* **71** 577–598.

Lindberg, O., Walterfang, M., Looi, J.C.L., Malykhin, N., Östberg, P., Zandbelt, B., Styner, M., Velakoulis, D., Örndahl, E., Cavallin, L. and Wahlund, L.-O. (2012). Shape analysis of the hippocampus in Alzheimer's disease and subtypes of frontotemporal lobar degeneration. *J. Alzheimer's Dis.* **30** 355–365.

Lugosi, G. and Mendelson, S. (2019). Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.* **10** 1145–1190.

Nieto-Reyes, A. and Battey, H. (2016). A topologically valid definition of depth for functional data. *Statist. Sci.* **31** 61–79.

Pennec, X., Fillard, P. and Ayache, N. (2006). A Riemannian framework for tensor computing. *Int. J. Comput. Vis.* **66** 41–66.

Petersen, A. and Müller, H.-G. (2019). Fréchet regression for random objects with Euclidean predictors. *Ann. Statist.* **47** 691–719.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*. New York: Springer.

Ren, H., Chen, N. and Zou, C. (2017). Projection-based outlier detection in functional data. *Biometrika* **104** 411–423.

Shao, L., Lin, Z. and Yao, F. (2022). Intrinsic Riemannian functional data analysis for sparse longitudinal observations. *Ann. Statist.* **50** 1696–1721.

Shao, L. and Yao, F. (2026). Supplement to "Robust functional data analysis: from sparse to dense designs." https://doi.org/10.3150/25-BEJ1920SUPPA, https://doi.org/10.3150/25-BEJ1920SUPPB

Sinova, B., González-Rodríguez, G. and Van Aelst, S. (2018). M-estimators of location for functional data. *Bernoulli* **24** 2328–2357.

Sun, Y. and Genton, M.G. (2011). Functional boxplots. *J. Comput. Graph. Statist.* **20** 316–334.

Wahl, M. (2022). Lower bounds for invariant statistical models with applications to principal component analysis. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 1565–1589.

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Review of functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005a). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005b). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590.

Zhang, X. and Wang, J.L. (2016). From sparse to dense functional data and beyond. *Ann. Statist.* **44** 2281–2321.

Zhou, H., Wei, D. and Yao, F. (2022). Theory of functional principal components analysis for discretely observed data. Preprint. Available at arXiv:2209.08768v1.

Zhou, H., Yao, F. and Zhang, H. (2022). Functional linear regression for discretely observed data: From ideal to reality. *Biometrika* **110** 381–393.

Bernoulli **32**(2), 2026, 1594–1619
https://doi.org/10.3150/25-BEJ1922

First contact percolation

BENEDIKT JAHNEL[1,2,a], LUKAS LÜCHTRATH[2,b] and ANH DUC VU[2,c]

[1] *Technische Universität Braunschweig, Braunschweig, Germany,* [a]*benedikt.jahnel@tu-braunschweig.de*
[2] *Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany,* [b]*lukas.luechtrath@wias-berlin.de,*
[c]*anhduc.vu@wias-berlin.de*

We study a version of first passage percolation on $\mathbb{Z}^d$ where the random passage times on the edges are replaced by contact times represented by random closed sets on $\mathbb{R}$. Similarly to the contact process without recovery, an infection can spread into the system along increasing sequences of contact times. In case of stationary contact times, we can identify associated first passage percolation models, which in turn establish shape theorems also for first contact percolation. In case of periodic contact times that reflect some reoccurring daily pattern, we also present shape theorems with limiting shapes that are universal with respect to the within-one-day contact distribution. In this case, we also prove a Poisson approximation for increasing numbers of within-one-day contacts. Finally, we present a comparison of the limiting speeds of three models – all calibrated to have one expected contact per day – that suggests that less randomness is beneficial for the speed of the infection. The proofs rest on coupling and subergodicity arguments.

Keywords: Contact process; first-passage percolation; pure-growth process; shape theorem

References

[1] Ahlberg, D., Deijfen, M. and Sfragara, M. (2024). Chaos, concentration and multiple valleys in first-passage percolation. arXiv preprint.arXiv:2302.11367.
[2] Auffinger, A., Damron, M. and Hanson, J. (2017). *50 Years of First-Passage Percolation*. *University Lecture Series* **68**. Providence, RI: American Mathematical Society. https://doi.org/10.1090/ulect/068
[3] Basu, R., Ganguly, S. and Sly, A. (2021). Upper tail large deviations in first passage percolation. *Comm. Pure Appl. Math.* **74** 1577–1640. https://doi.org/10.1002/cpa.22010
[4] Beutler, F.J. and Leneman, O.A.Z. (1966). The theory of stationary point processes. *Acta Math.* **116** 159–190. https://doi.org/10.1007/BF02392816
[5] Cardona-Tobón, N., Ortgiese, M., Seiler, M. and Sturm, A. (2024). The contact process on dynamical random trees with degree dependence. arXiv preprint. arXiv:2406.12689.
[6] Chow, Y. and Zhang, Y. (2003). Large deviations in first-passage percolation. *Ann. Appl. Probab.* **13** 1601–1614. https://doi.org/10.1214/aoap/1069786513
[7] Concannon, R.J. and Blythe, R.A. (2014). Spatiotemporally complete condensation in a non-Poissonian exclusion process. *Phys. Rev. Lett.* **112**. https://doi.org/10.1103/PhysRevLett.112.050603
[8] Cox, J.T. and Kesten, H. (1981). On the continuity of the time constant of first-passage percolation. *J. Appl. Probab.* **18** 809–819. https://doi.org/10.1017/s0021900200034161
[9] Damron, M. and Hochman, M. (2013). Examples of nonpolygonal limit shapes in iid first-passage percolation and infinite coexistence in spatial growth models. *Ann. Appl. Probab.* **23** 1074–1085. https://doi.org/10.1214/12-AAP864
[10] Deijfen, M. (2003). Asymptotic shape in a continuum growth model. *Adv. Appl. Probab.* **35** 303–318. https://doi.org/10.1239/aap/1051201647
[11] Dembin, B. and Garban, C. (2024). Superconcentration for minimal surfaces in first passage percolation and disordered Ising ferromagnets. *Probab. Theory Related Fields* 1–28. https://doi.org/10.1007/s00440-023-01252-2
[12] Durrett, R. (1980). On the growth of one dimensional contact processes. *Ann. Probab.* **8** 890–907. https://doi.org/10.1214/aop/1176994619

1350-7265

[13] Durrett, R. (1991). The contact process, 1974–1989. In *Mathematics of Random Media (Blacksburg, VA, 1989)*. *Lectures in Appl. Math.* **27** 1–18. Providence, RI: Amer. Math. Soc.

[14] Durrett, R. and Griffeath, D. (1982). Contact processes in several dimensions. *Probab. Theory Related Fields* **59** 535–552. https://doi.org/10.1007/BF00532808

[15] Fontes, L.R., Gomes, P.A. and Sanchis, R. (2021). Contact process under heavy-tailed renewals on finite graphs. *Bernoulli* **27** 1745–1763. https://doi.org/10.3150/20-BEJ1290

[16] Fontes, L.R., Mountford, T.S. and Vares, M.E. (2020). Contact process under renewals II. *Stoch. Process. Appl.* **130** 1103–1118. https://doi.org/10.1016/j.spa.2019.04.008

[17] Fontes, L.R., Marchetti, D.H., Mountford, T.S. and Vares, M.E. (2019). Contact process under renewals I. *Stoch. Process. Appl.* **129** 2903–2911. https://doi.org/10.1016/j.spa.2018.08.007

[18] Fontes, L.R., Mountford, T.S., Ungaretti, D. and Vares, M.E. (2023). Renewal contact processes: Phase transition and survival. *Stoch. Process. Appl.* **161** 102–136. https://doi.org/10.1016/j.spa.2023.03.005

[19] Ganguly, A. and Ramanan, K. (2024). Hydrodynamic limits of non-Markovian interacting particle systems on sparse graphs. *Electron. J. Probab.* **29** 1–63. https://doi.org/10.1214/24-EJP1243

[20] Garet, O. and Marchand, R. (2004). Asymptotic shape for the chemical distance and first-passage percolation on the infinite Bernoulli cluster. *ESAIM Probab. Stat.* **8** 169–199. https://doi.org/10.1051/ps:2004009

[21] Garet, O. and Marchand, R. (2005). Coexistence in two-type first-passage percolation models. *Ann. Appl. Probab.* **15** 298–330. https://doi.org/10.1214/105051604000000503

[22] Hilário, M., Ungaretti, D., Valesin, D. and Vares, M.E. (2022). Results on the contact process with dynamic edges or under renewals. *Electron. J. Probab.* **27** 1–31. https://doi.org/10.1214/22-EJP811

[23] Hoffman, C. (2008). Geodesics in first passage percolation. *Ann. Appl. Probab.* **18** 1944–1969. https://doi.org/10.1214/07-AAP510

[24] Kesten, H. (1986). Aspects of first passage percolation. In *École d'Été de Probabilités de Saint Flour XIV - 1984* (P.L. Hennequin, ed.) 125–264. Berlin: Springer. https://doi.org/10.1007/BFb0074919

[25] Kesten, H. (1987). Percolation theory and first-passage percolation. *Ann. Probab.* **15** 1231–1271. https://doi.org/10.1214/aop/1176991975

[26] Kesten, H. (2003). First-passage percolation. In *From Classical to Modern Probability*. *Progr. Probab.* **54** 93–143. Basel: Birkhäuser. https://doi.org/10.1007/978-3-0348-8053-4_4

[27] Khoromskaia, D., Harris, R.J. and Grosskinsky, S. (2014). Dynamics of non-Markovian exclusion processes. *J. Stat. Mech. Theory Exp.* **2014** P12013. https://doi.org/10.1088/1742-5468/2014/12/P12013

[28] Liggett, T.M. (2005). *Interacting Particle Systems*. *Classics in Mathematics*. Berlin: Springer. (Reprint of the 1985 original). https://doi.org/10.1007/b138374

[29] Linker, A. and Remenik, D. (2020). The contact process with dynamic edges on $\mathbb{Z}$. *Electron. J. Probab.* **25** 1–21. https://doi.org/10.1214/20-EJP480

[30] Marchand, R. (2002). Strict inequalities for the time constant in first passage percolation. *Ann. Appl. Probab.* **12** 1001–1038. https://doi.org/10.1214/aoap/1031863179

[31] Molchanov, I. (2017). *Theory of Random Sets*. *Probability Theory and Stochastic Modelling* **87**, 2nd ed. London: Springer. https://doi.org/10.1007/978-1-4471-7349-6

[32] Seiler, M. and Sturm, A. (2023). Contact process on a dynamical long range percolation. *Electron. J. Probab.* **28** 1–36. https://doi.org/10.1214/23-EJP1042

[33] Smythe, R.T. and Wierman, J.C. (1978). *First-Passage Percolation on the Square Lattice*. *Lecture Notes in Mathematics* **671**. Berlin: Springer. https://doi.org/10.1007/BFb0063306

[34] van den Berg, J. and Kesten, H. (1993). Inequalities for the time constant in first-passage percolation. *Ann. Appl. Probab.* **3** 56–80. https://doi.org/10.1214/aoap/1177005507

Bernoulli **32**(2), 2026, 1620–1644
https://doi.org/10.3150/25-BEJ1924

The asymptotic properties of the extreme eigenvectors of high-dimensional generalized spiked covariance models

ZHANGNI PU[1,a], XIAOZHUO ZHANG[1,b], JIANG HU[1,c] and ZHIDONG BAI[1,2,d]

[1] *KLASMOE and School of Mathematics & Statistics, Northeast Normal University, No. 5268 People's Street, Changchun 130024, China,* [a]*puzn687@nenu.edu.cn,* [b]*zhangxz722@nenu.edu.cn,* [c]*huj156@nenu.edu.cn*
[2] *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China,* [d]*baizd@nenu.edu.cn*

In this paper, we investigate the asymptotic behaviors of the extreme eigenvectors in a general spiked covariance matrix, where the dimension and sample size increase proportionally. We eliminate the restrictive assumption of the block diagonal structure in the population covariance matrix. Moreover, there is no requirement for the spiked eigenvalues and the 4th moment to be bounded. Specifically, we apply random matrix theory to derive the convergence and limiting distributions of certain projections of the extreme eigenvectors in a large sample covariance matrix within a generalized spiked population model. Furthermore, our techniques are robust and effective, even when spiked eigenvalues differ significantly in magnitude from nonspiked ones. Finally, we propose a powerful test for the eigenspaces of covariance matrices.

Keywords: Eigenvector; high-dimensional covariance matrix; random matrix theory; spiked model

References

[1] Bai, Z. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. *Statist. Sinica* **9** 611–662. MR1711663
[2] Bai, Z. and Ding, X. (2012). Estimation of spiked eigenvalues in spiked models. *Random Matrices Theory Appl.* **1** 1150011. MR2934717 https://doi.org/10.1142/S2010326311500110
[3] Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica* **70** 191–221. MR1926259 https://doi.org/10.1111/1468-0262.00273
[4] Bai, Z. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845
[5] 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
[6] Bai, Z. and Yao, J. (2008). Central limit theorems for eigenvalues in a spiked population model. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 447–474. MR2451053 https://doi.org/10.1214/07-AIHP118
[7] Bai, Z. and Yao, J. (2012). On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.* **106** 167–177. MR2887686 https://doi.org/10.1016/j.jmva.2011.10.009
[8] Baik, J., Ben Arous, G. and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. MR2165575 https://doi.org/10.1214/009117905000000233
[9] Baik, J. and Silverstein, J.W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.* **97** 1382–1408. MR2279680 https://doi.org/10.1016/j.jmva.2005.08.003
[10] Bao, Z., Ding, X. and Wang, K. (2021). Singular vector and singular subspace distribution for the matrix denoising model. *Ann. Statist.* **49** 370–392. MR4206682 https://doi.org/10.1214/20-AOS1960
[11] Bao, Z., Ding, X., Wang, J. and Wang, K. (2022). Statistical inference for principal components of spiked covariance matrices. *Ann. Statist.* **50** 1144–1169. MR4404931 https://doi.org/10.1214/21-AOS2143

[12] Benaych-Georges, F. and Nadakuditi, R.R. (2011). The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. Math.* **227** 494–521. MR2782201 https://doi.org/10.1016/j.aim.2011.02.007
[13] Bianchi, P., Najim, J., Maida, M. and Debbah, M. (2009). Performance analysis of some eigen-based hypothesis tests for collaborative sensing. In *2009 IEEE/SP 15th Workshop on Statistical Signal Processing* 5–8. https://doi.org/10.1109/SSP.2009.5278654
[14] Bloemendal, A., Knowles, A., Yau, H.-T. and Yin, J. (2014). On the principal components of sample covariance matrices. *Probab. Theory Related Fields* **164** 459–552. MR3449395 https://doi.org/10.1007/s00440-015-0616-x
[15] Capitaine, M. and Donati-Martin, C. (2021). Non universality of fluctuations of outlier eigenvectors for block diagonal deformations of Wigner matrices. *ALEA Lat. Amer. J. Probab. Math. Stat.* **18** 129–165. MR4198872 https://doi.org/10.30757/ALEA.v18-07
[16] Capitaine, M., Donati-Martin, C. and Féral, D. (2009). The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *Ann. Probab.* **37** 1–47. https://doi.org/10.1214/08-AOP394
[17] Couillet, R. and Debbah, M. (2011). *Random Matrix Methods for Wireless Communications*. Cambridge: Cambridge University Press. MR2884783 https://doi.org/10.1017/CBO9780511994746
[18] Couillet, R. and Hachem, W. (2013). Fluctuations of spiked random matrix models and failure diagnosis in sensor networks. *IEEE Trans. Inf. Theory* **59** 509–525. MR3008163 https://doi.org/10.1109/TIT.2012.2218572
[19] Ding, X. (2021). Spiked sample covariance matrices with possibly multiple bulk components. *Random Matrices Theory Appl.* **10** 2150014. https://doi.org/10.1142/S2010326321500143
[20] Ding, X. and Yang, F. (2018). A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices. *Ann. Appl. Probab.* **28** 1679–1738. https://doi.org/10.1214/17-AAP1341
[21] Fan, J., Fan, Y., Han, X. and Lv, J. (2022). Asymptotic theory of eigenvectors for random matrices with diverging spikes. *J. Amer. Statist. Assoc.* **117** 996–1009. MR4436328 https://doi.org/10.1080/01621459.2020.1840990
[22] Hoyle, D.C. and Rattray, M. (2004). Principal-component-analysis eigenvalue spectra from data with symmetry-breaking structure. *Phys. Rev. E* **69** 026124. https://doi.org/10.1103/PhysRevE.69.026124
[23] Hu, J. and Bai, Z. (2014). Strong representation of weak convergence. *Sci. China Math.* **57** 2399–2406. MR3266500 https://doi.org/10.1007/s11425-014-4855-6
[24] Jiang, D. and Bai, Z. (2021). Generalized four moment theorem and an application to CLT for spiked eigenvalues of high-dimensional covariance matrices. *Bernoulli* **27** 274–294. MR4177370 https://doi.org/10.3150/20-BEJ1237
[25] Jiang, D. and Bai, Z. (2021). Partial generalized four moment theorem revisited. *Bernoulli* **27** 2337–2352. MR4303885 https://doi.org/10.3150/20-BEJ1310
[26] 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
[27] Johnstone, I.M. and Yang, J. (2018). Notes on asymptotics of sample eigenstructure for spiked covariance models with non-Gaussian data. arXiv:1810.10427.
[28] Knowles, A. and Yin, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1749. MR3103909 https://doi.org/10.1002/cpa.21450
[29] Knowles, A. and Yin, J. (2014). The outliers of a deformed Wigner matrix. *Ann. Probab.* **42** 1980–2031. MR3262497 https://doi.org/10.1214/13-AOP855
[30] Marchenko, V.A. and Pastur, L.A. (1967). Distribution of eigenvalues for some sets of random matrices. *Sb. Math.* **1** 457–483. https://doi.org/10.1070/SM1967v001n04ABEH001994
[31] Mehta, M.L. (1991). *Random Matrices*. Academic Press.
[32] Morales-Jimenez, D., Johnstone, I.M., McKay, M.R. and Yang, J. (2021). Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. *Statist. Sinica* **31** 571–601. MR4286186 https://doi.org/10.5705/ss.202019.0052
[33] Onatski, A. (2009). Testing hypotheses about the number of factors in large factor models. *Econometrica* **77** 1447–1479. MR2561070 https://doi.org/10.3982/ECTA6964

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

[35] Pu, Z., Zhang, X., Hu, J. and Bai, Z. (2026). Supplement to "The asymptotic properties of the extreme eigenvectors of high-dimensional generalized spiked covariance models." https://doi.org/10.3150/25-BEJ1924SUPP

[36] Silverstein, J.W. (1995). Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices. *J. Multivariate Anal.* **55** 331–339. https://doi.org/10.1006/jmva.1995.1083

[37] Skorokhod, A.V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261–290. https://doi.org/10.1137/1101022

[38] Yao, J., Zheng, S. and Bai, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics* **39**. New York: Cambridge University Press. MR3468554 https://doi.org/10.1017/CBO9781107588080

Bernoulli **32**(2), 2026, 1645–1664
https://doi.org/10.3150/25-BEJ1925

Contraction rates and projection subspace estimation with Gaussian process priors in high dimension

ELIE ODIN[1,a], FRANÇOIS BACHOC[1,2,4,b] and AGNÈS LAGNOUX[3,4,c]

[1] *Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT3, F-31062 Toulouse, France,* [a] *elie.odin@math.univ-toulouse.fr*
[2] *Institut Universitaire de France (IUF), France,* [b] *francois.bachoc@math.univ-toulouse.fr*
[3] *Institut de Mathématiques de Toulouse; UMR5219. Université de Toulouse; CNRS. UT2J, F-31058 Toulouse, France,* [c] *lagnoux@univ-tlse2.fr*
[4] *Artificial and Natural Intelligence Toulouse Institute (ANITI), Toulouse, France*

This work explores the dimension reduction problem for Bayesian nonparametric regression and density estimation. More precisely, we are interested in estimating a functional parameter f over the unit ball in $\mathbb{R}^d$, which depends only on a d^*-dimensional subspace of $\mathbb{R}^d$, with $d^* < d$. It is well-known that rescaled Gaussian process priors over a given function space achieve smoothness adaptation and posterior contraction with near minimax-optimal rates. Furthermore, hierarchical extensions of this approach, equipped with subspace projection, can also adapt to the intrinsic dimension d^* (Tokdar (2011)). When the ambient dimension d does not vary with n, the minimax rate remains of the order $n^{-\beta/(2\beta+d^*)}$, where β denotes the smoothness of f. However, this is up to multiplicative constants that can become prohibitively large as d increases. The dependence between the contraction rate and the ambient dimension has not been fully explored yet and this work provides a first insight: we let the dimensions d^* and d grow with n, and by combining the arguments of Tokdar (2011) and Castillo and Randrianarisoa (2024), we derive growth rates for them that still lead to posterior consistency with minimax rate. We also discuss the optimality of the growth rate for d. Additionally, we provide a set of assumptions under which a consistent estimation of f leads to correct estimation of the subspace projection, assuming that d^* is known.

Keywords: Bayesian nonparametrics; contraction rates; density estimation; Gaussian process prior; nonparametric regression; reproducing kernel Hilbert space; sufficient dimension reduction

References

Bachoc, F. and Lagnoux, A. (2025). Posterior contraction rates for constrained deep Gaussian processes in density estimation and classification. *Comm. Statist. Theory Methods* **54** 774–811.

Berenfeld, C., Rosa, P. and Rousseau, J. (2024). Estimating a density near an unknown manifold: A Bayesian nonparametric approach. *Ann. Statist.* **52** 2081–2111.

Bhatia, R. (1997). *Matrix Analysis*. *Grad. Texts in Math* **169**. New York, NY: Springer.

Birgé, L. (1986). On estimating a density using Hellinger distance and some other strange facts. *Probab. Theory Related Fields* **71** 271–291.

Cai, T.T., Ma, Z. and Wu, Y. (2013). Sparse PCA: Optimal rates and adaptive estimation. *Ann. Statist.* **41** 3074–3110.

Castillo, I. and Egels, P. (2024). Posterior and variational inference for deep neural networks with heavy-tailed weights. Available at https://arxiv.org/abs/2406.03369.

Castillo, I. and Randrianarisoa, T. (2024). Deep Horseshoe Gaussian processes. Available at https://arxiv.org/abs/2403.01737.

Comminges, L. and Dalalyan, A.S. (2012). Tight conditions for consistency of variable selection in the context of high dimensionality. *Ann. Statist.* **40** 2667–2696.

1350-7265

Cook, R.D. (1998). *Regression Graphics: Ideas for Studying Regressions Through Graphics*. John Wiley & Sons.
Dunson, D.B. and Wu, N. (2024). Inferring manifolds from noisy data using Gaussian Processes. Available at https://arxiv.org/abs/2110.07478.
Dunson, D.B., Wu, H.-T. and Wu, N. (2022). Graph based Gaussian processes on restricted domains. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 414–439.
Finocchio, G. and Schmidt-Hieber, J. (2023). Posterior contraction for deep Gaussian process priors. *J. Mach. Learn. Res.* **24** 1–49.
Ghosal, S., Ghosh, J.K. and Van Der Vaart, A. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531.
Giné, E. and Nickl, R. (2011). Rates of contraction for posterior distributions in L^r-metrics, $1 \le r \le \infty$. *Ann. Statist.* **39** 2883–2911.
Jiang, S. and Tokdar, S.T. (2021). Variable selection consistency of Gaussian process regression. *Ann. Statist.* **49** 2491–2505.
Johnstone, I.M. and Lu, A.Y. (2004). Sparse principal components analysis Technical Report. Available at arXiv:0901.4392.
Lafferty, J. and Wasserman, L. (2008). Rodeo: Sparse, greedy nonparametric regression. *Ann. Statist.* **36** 28–63.
Li, K.-C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86** 316–327.
Li, D., Mukhopadhyay, M. and Dunson, D.B. (2022). Efficient manifold approximation with spherelets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1129–1149.
Lin, Q., Zhao, Z. and Liu, J.S. (2018). On consistency and sparsity for sliced inverse regression in high dimensions. *Ann. Statist.* **46** 580–610.
Lin, Q., Zhao, Z. and Liu, J.S. (2019). Sparse sliced inverse regression via lasso. *J. Amer. Statist. Assoc.* **114** 1726–1739.
Lin, Q., Li, X., Huang, D. and Liu, J.S. (2021). On the optimality of sliced inverse regression in high dimensions. *Ann. Statist.* **49** 1–20.
Nakada, R. and Imaizumi, M. (2020). Adaptive approximation and generalization of deep neural network with intrinsic dimensionality. *J. Mach. Learn. Res.* **21** 1–38.
Odin, E., Bachoc, F. and Lagnoux, A. (2026). Supplement to "Contraction rates and projection subspace estimation with Gaussian process priors in high dimension." https://doi.org/10.3150/25-BEJ1925SUPP
Rosa, P., Borovitskiy, S., Terenin, A. and Rousseau, J. (2023). Posterior contraction rates for Matérn Gaussian processes on Riemannian manifolds. In *Advances in Neural Information Processing Systems* **36** 34087–34121. Curran Associates, Inc.
Rousseau, J. and Mengersen, K. (2011). Asymptotic behaviour of the posterior distribution in overfitted mixture models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 689–710.
Schmidt-Hieber, A.J. (2020). Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Statist.* **48** 1875–1897.
Shen, W., Tokdar, S.T. and Ghosal, S. (2013). Adaptive Bayesian multivariate density estimation with Dirichlet mixtures. *Biometrika* **100** 623–640.
Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053.
Tan, K., Shi, L. and Yu, Z. (2020). Sparse SIR: Optimal rates and adaptive estimation. *Ann. Statist.* **48** 64–85.
Tang, T., Wu, N., Cheng, X. and Dunson, D. (2024). Adaptive Bayesian regression on data with low intrinsic dimensionality. Available at https://arxiv.org/abs/2407.09286.
Tokdar, S.T. (2011). Dimension adaptability of Gaussian process models with variable selection and projection Technical Report. https://arxiv.org/pdf/1112.0716.
Tokdar, S.T., Zhu, Y.M. and Ghosh, J.K. (2010). Bayesian density regression with logistic Gaussian process and subspace projection. *Bayesian Anal.* **5** 319–344.
Van Der Vaart, A. and Van Zanten, H. (2008a). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463.
Van Der Vaart, A. and Van Zanten, H. (2008b). Reproducing kernel Hilbert spaces of Gaussian priors. In *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh. Inst. Math. Stat. Monogr.* **3** 200–223.
Van Der Vaart, A. and Van Zanten, H. (2009). Adaptative Bayesian estimation using a Gaussian random field with inverse gamma bandwidth. *Ann. Statist.* **37** 2655–2675.

Verzelen, N. (2012). Minimax risks for sparse regressions: Ultra-high dimensional phenomenons. *Electron. J. Stat.* **6** 38–90.

Wainwright, M.J. (2009). Sharp thresholds for high-dimensional and noisy sparsity recovery using ℓ_1-constrained quadratic programming (Lasso). *IEEE Trans. Inf. Theory* **55** 2183–2202.

Yang, Y. and Dunson, D.B. (2016). Bayesian manifold regression. *Ann. Statist.* **44** 876–905.

Yang, Y. and Tokdar, S.T. (2015). Minimax-optimal nonparametric regression in high dimensions. *Ann. Statist.* **43** 652–674.

Zeng, J., Mai, Q. and Zhang, X. (2024). Subspace estimation with automatic dimension and variable selection in sufficient dimension reduction. *J. Amer. Statist. Assoc.* **119** 343–355.

Zhu, L., Miao, B. and Peng, H. (2006). On sliced inverse regression with high-dimensional covariates. *J. Amer. Statist. Assoc.* **101** 630–643.

Bernoulli **32**(2), 2026, 1665–1685
https://doi.org/10.3150/25-BEJ1926

Tame sparse exponential random graphs

SUMAN CHAKRABORTY[1,a], REMCO VAN DER HOFSTAD[2,b] and FRANK DEN HOLLANDER[3,c]

[1]*Pijnacker, Netherlands,* [a]*contact@sumanc.com*
[2]*Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands,* [b]*rhofstad@win.tue.nl*
[3]*Mathematical Institute, Leiden University, Leiden, Netherlands,* [c]*denholla@math.leidenuniv.nl*

In this paper we obtain a precise estimate of the probability that the sparse binomial random graph contains a large number of vertices in a triangle. We compute the logarithm of this probability up to second order, which enables us to propose an exponential random graph model based on the number of vertices in a triangle. Specifically, by tuning a single parameter, we can with high probability induce any given fraction of vertices in a triangle. Moreover, in the proposed exponential random graph model we derive a large deviation principle for the number of edges. As a byproduct, we propose a consistent estimator of the tuning parameter.

Keywords: Consistent estimation; exponential random graph; nonlinear large deviations; random graphs

References

Bhamidi, S., Bresler, G. and Sly, A. (2011). Mixing time of exponential random graphs. *Ann. Appl. Probab.* **21** 2146–2170.

Bhamidi, S., Chakraborty, S., Cranmer, S. and Desmarais, B. (2018). Weighted exponential random graph models: Scope and large network limits. *J. Stat. Phys.* **173** 704–735.

Chakraborty, S., van der Hofstad, R. and den Hollander, F. (2021). Sparse random graphs with many triangles. Preprint, arXiv:2112.06526 [math.PR].

Chatterjee, S. and Dembo, A. (2016). Nonlinear large deviations. *Adv. Math.* **299** 396–450.

Chatterjee, S. and Diaconis, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41** 2428–2461.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, second ed. *Applications of Mathematics (New York)* **38**. New York: Springer-Verlag.

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

Häggström, O. and Jonasson, J. (1999). Phase transition in the random triangle model. *J. Appl. Probab.* **36** 1101–1115.

Handcock, M.S. (2003). Assessing degeneracy in statistical models of social networks. *J. Amer. Statist. Assoc.* **76** 33–50.

Harris, T.E. (1960). A lower bound for the critical probability in a certainvpercolation process. *Proc. Camb. Philos. Soc.* **56** 13–20.

Holland, P.W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *J. Amer. Statist. Assoc.* **76** 33–50.

Hunter, D.R. and Handcock, M.S. (2006). Inference in curved exponential family models for networks. *J. Comput. Graph. Statist.* **15** 565–583.

Jonasson, J. (1999). The random triangle model. *J. Appl. Probab.* **36** 852–867.

Mukherjee, S. (2020). Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics. *Bernoulli* **26** 1016–1043.

Newman, M.E. (2009). Random graphs with clustering. *Phys. Rev. Lett.* **103** 058701.

Rapoport, A. (1948). Cycle distributions in random nets. *Bull. Math. Biophys.* **10** 145–157.

Serrano, M.Á. and Boguña, M. (2006). Clustering in complex networks. I. General formalism. *Phys. Rev. E* **74** 056114.

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

Watts, D.J. and Strogatz, S.H. (1998). Collective dynamics of 'small-world' networks. *Nature* **393** 440–442.