BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty Number Three August 2024 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 Number Three August 2024 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 30 Number 3 August 2024 Pages 1695–2545 ISI/BS

Volume 30 Number 3   August 2024   ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

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

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

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

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

---

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

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

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **30**(3), 2024, 1695–1722
https://doi.org/10.3150/23-BEJ1644

Noise covariance estimation in multi-task high-dimensional linear models

KAI TAN[1,a], GABRIEL ROMON[2,c] and PIERRE C. BELLEC[1,b]

[1]*Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA* , [a]*kai.tan@rutgers.edu*,
[b]*pierre.bellec@rutgers.edu*
[2]*CREST, ENSAE, IP Paris, Palaiseau 91120 Cedex, France* , [c]*gabriel.romon@ensae.fr*

This paper studies the multi-task high-dimensional linear regression models where the noise among different tasks is correlated, in the moderately high dimensional regime where sample size n and dimension p are of the same order. Our goal is to estimate the covariance matrix of the noise random vectors, or equivalently the correlation of the noise variables on any pair of two tasks. Treating the regression coefficients as a nuisance parameter, we leverage the multi-task elastic-net and multi-task lasso estimators to estimate the nuisance. By precisely understanding the bias of the squared residual matrix and by correcting this bias, we develop a novel estimator of the noise covariance that converges in Frobenius norm at the rate $n^{-1/2}$ when the covariates are Gaussian distributed with a known covariance matrix. This novel estimator is efficiently computable. Under suitable conditions, the proposed estimator of the noise covariance attains the same rate of convergence as the "oracle" estimator that knows in advance the regression coefficients of the multi-task model. The Frobenius error bounds obtained in this paper also illustrate the advantage of this new estimator compared to a method-of-moments estimator that does not attempt to estimate the nuisance. As byproducts of our techniques, we obtain estimates of the generalization error and out-of-sample error of the multi-task elastic-net and multi-task lasso estimators. Extensive simulation studies are carried out to illustrate the numerical performance of the proposed method.

Keywords: Elastic-net; high dimensional analysis; lasso; multi-task model; noise covariance

References

Bayati, M., Erdogdu, M.A. and Montanari, A. (2013). Estimating lasso risk and noise level. *Adv. Neural Inf. Process. Syst.* **26**.

Bayati, M. and Montanari, A. (2012). The LASSO risk for Gaussian matrices. *IEEE Trans. Inf. Theory* **58** 1997–2017. MR2951312 https://doi.org/10.1109/TIT.2011.2174612

Bellec, P.C. (2023). Out-of-sample error estimation for M-estimators with convex penalty. *Inf. Inference* **12** Paper No. iaad031, 36 pp. MR4660702 https://doi.org/10.1093/imaiai/iaad031

Bellec, P. and Kuchibhotla, A. (2019). First order expansion of convex regularized estimators. *Adv. Neural Inf. Process. Syst.* **32**.

Bellec, P.C. and Romon, G. (2021). Chi-square and normal inference in high-dimensional multi-task regression. Preprint. Available at arXiv:2107.07828.

Bellec, P.C. and Zhang, C.-H. (2021). Second-order Stein: SURE for SURE and other applications in high-dimensional inference. *Ann. Statist.* **49** 1864–1903. MR4319234 https://doi.org/10.1214/20-aos2005

Belloni, A., Chernozhukov, V. and Wang, L. (2011). Square-root lasso: Pivotal recovery of sparse signals via conic programming. *Biometrika* **98** 791–806. MR2860324 https://doi.org/10.1093/biomet/asr043

Belloni, A., Chernozhukov, V. and Wang, L. (2014). Pivotal estimation via square-root Lasso in nonparametric regression. *Ann. Statist.* **42** 757–788. MR3210986 https://doi.org/10.1214/14-AOS1204

Bertrand, Q., Massias, M., Gramfort, A. and Salmon, J. (2019). Handling correlated and repeated measurements with the smoothed multivariate square-root lasso. *Adv. Neural Inf. Process. Syst.* **32**.

Bickel, P.J. and Levina, E. (2008). Covariance regularization by thresholding. *Ann. Statist.* **36** 2577–2604. MR2485008 https://doi.org/10.1214/08-AOS600

Cai, T.T., Zhang, C.-H. and Zhou, H.H. (2010). Optimal rates of convergence for covariance matrix estimation. *Ann. Statist.* **38** 2118–2144. MR2676885 https://doi.org/10.1214/09-AOS752

Celentano, M. and Montanari, A. (2021). CAD: Debiasing the Lasso with inaccurate covariate model. Preprint. Available at arXiv:2107.14172.

Chen, S. and Banerjee, A. (2017). Alternating estimation for structured high-dimensional multi-response models. *Adv. Neural Inf. Process. Syst.* **30**.

Davidson, K.R. and Szarek, S.J. (2001). Local operator theory, random matrices and Banach spaces. In *Handbook of the Geometry of Banach Spaces, Vol. 1* 131. Amsterdam: North-Holland.

Dicker, L.H. (2014). Variance estimation in high-dimensional linear models. *Biometrika* **101** 269–284. MR3215347 https://doi.org/10.1093/biomet/ast065

Dicker, L.H. and Erdogdu, M.A. (2016). Maximum likelihood for variance estimation in high-dimensional linear models. In *Artificial Intelligence and Statistics* 159–167. PMLR.

Dobriban, E. and Wager, S. (2018). High-dimensional asymptotics of prediction: Ridge regression and classification. *Ann. Statist.* **46** 247–279. MR3766952 https://doi.org/10.1214/17-AOS1549

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

Fan, J., Guo, S. and Hao, N. (2012). Variance estimation using refitted cross-validation in ultrahigh dimensional regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 37–65. MR2885839 https://doi.org/10.1111/j.1467-9868.2011.01005.x

Fourdrinier, D., Haddouche, A.M. and Mezoued, F. (2021). Covariance matrix estimation under data-based loss. *Statist. Probab. Lett.* **177** Paper No. 109160, 7 pp. MR4271904 https://doi.org/10.1016/j.spl.2021.109160

Friedman, J.H., Hastie, T. and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33** 1–22.

Janson, L., Foygel Barber, R. and Candès, E. (2017). EigenPrism: Inference for high dimensional signal-to-noise ratios. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1037–1065. MR3689308 https://doi.org/10.1111/rssb.12203

Koltchinskii, V. and Lounici, K. (2017). Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli* **23** 110–133. MR3556768 https://doi.org/10.3150/15-BEJ730

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

Lecué, G. and Mendelson, S. (2017). Sparse recovery under weak moment assumptions. *J. Eur. Math. Soc. (JEMS)* **19** 881–904. MR3612870 https://doi.org/10.4171/JEMS/682

Liu, H. and Zhang, J. (2009). Estimation consistency of the group lasso and its applications. In *Artificial Intelligence and Statistics* 376–383. PMLR.

Lounici, K., Pontil, M., van de Geer, S. and Tsybakov, A.B. (2011). Oracle inequalities and optimal inference under group sparsity. *Ann. Statist.* **39** 2164–2204. MR2893865 https://doi.org/10.1214/11-AOS896

Miolane, L. and Montanari, A. (2021). The distribution of the Lasso: Uniform control over sparse balls and adaptive parameter tuning. *Ann. Statist.* **49** 2313–2335. MR4319252 https://doi.org/10.1214/20-aos2038

Molstad, A.J. (2022). New insights for the multivariate square-root lasso. *J. Mach. Learn. Res.* **23** Paper No. [66], 52 pp. MR4576651 https://doi.org/10.22405/2226-8383-2022-23-4-52-63

Obozinski, G., Wainwright, M.J. and Jordan, M.I. (2011). Support union recovery in high-dimensional multivariate regression. *Ann. Statist.* **39** 1–47. MR2797839 https://doi.org/10.1214/09-AOS776

Pedregosa, F., Varoquaux, G., Gramfort, A. et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12** 2825–2830. MR2854348

Simon, N., Friedman, J. and Hastie, T. (2013). A blockwise descent algorithm for group-penalized multiresponse and multinomial regression. Preprint. Available at arXiv:1311.6529.

Stein, C.M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9** 1135–1151. MR0630098

Sun, T. and Zhang, C.-H. (2012). Scaled sparse linear regression. *Biometrika* **99** 879–898. MR2999166 https://doi.org/10.1093/biomet/ass043

Tan, K., Romon, G. and Bellec, P.C. (2024). Supplement to "Noise covariance estimation in multi-task high-dimensional linear models." https://doi.org/10.3150/23-BEJ1644SUPPA, https://doi.org/10.3150/23-BEJ1644SUPPB van de Geer, S. and Stucky, B. (2016). χ^2-confidence sets in high-dimensional regression. In *Statistical Analysis for High-Dimensional Data*. *Abel Symp.* **11** 279–306. Cham: Springer. MR3616273

Yu, G. and Bien, J. (2019). Estimating the error variance in a high-dimensional linear model. *Biometrika* **106** 533–546. MR3992388 https://doi.org/10.1093/biomet/asz017

Bernoulli **30**(3), 2024, 1723–1745
https://doi.org/10.3150/23-BEJ1650

Asymptotics for isotropic Hilbert-valued spherical random fields

ALESSIA CAPONERA[1,2,a]

[1]*Institut de Mathématiques, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland*
[2]*DEMS, University of Milano-Bicocca, Milan, Italy,* [a]*alessia.caponera@unimib.it*

In this paper, we introduce the concept of isotropic Hilbert-valued spherical random field, thus extending the notion of isotropic spherical random field to an infinite-dimensional setting. We then establish a spectral representation theorem and a functional Schoenberg's theorem. Following some key results established for the real-valued case, we prove consistency and quantitative central limit theorem for the sample power spectrum operators in the high-frequency regime.

Keywords: High-frequency asymptotics; Hilbert spaces; isotropy; quantitative central limit theorem; spectral representation; spherical random fields

References

[1] Berg, C. and Porcu, E. (2017). From Schoenberg coefficients to Schoenberg functions. *Constr. Approx.* **45** 217–241. MR3619442 https://doi.org/10.1007/s00365-016-9323-9
[2] Bourguin, S. and Campese, S. (2020). Approximation of Hilbert-valued Gaussians on Dirichlet structures. *Electron. J. Probab.* **25** 1–30. MR4193891 https://doi.org/10.1214/20-ejp551
[3] Brockwell, P.J. and Davis, R.A. (1991). *Time Series: Theory and Methods*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1093459 https://doi.org/10.1007/978-1-4419-0320-4
[4] Cammarota, V. and Marinucci, D. (2018). A quantitative central limit theorem for the Euler-Poincaré characteristic of random spherical eigenfunctions. *Ann. Probab.* **46** 3188–3228. MR3857854 https://doi.org/10.1214/17-AOP1245
[5] Caponera, A. (2021). SPHARMA approximations for stationary functional time series on the sphere. *Stat. Inference Stoch. Process.* **24** 609–634. MR4321852 https://doi.org/10.1007/s11203-021-09244-6
[6] Caponera, A., Durastanti, C. and Vidotto, A. (2021). LASSO estimation for spherical autoregressive processes. *Stochastic Process. Appl.* **137** 167–199. MR4244190 https://doi.org/10.1016/j.spa.2021.03.009
[7] Caponera, A., Fageot, J., Simeoni, M. and Panaretos, V.M. (2022). Functional estimation of anisotropic covariance and autocovariance operators on the sphere. *Electron. J. Stat.* **16** 5080–5148. MR4492081 https://doi.org/10.1214/22-ejs2064
[8] Caponera, A. and Marinucci, D. (2021). Asymptotics for spherical functional autoregressions. *Ann. Statist.* **49** 346–369. MR4206681 https://doi.org/10.1214/20-AOS1959
[9] Caramellino, L., Giorgio, G. and Rossi, M. (2022). Convergence in Total Variation for nonlinear functionals of random hyperspherical harmonics. Available at arXiv:2206.02605.
[10] Dick, J., Remazeilles, M. and Delabrouille, J. (2010). Impact of calibration errors on CMB component separation using FastICA and ILC. *Mon. Not. R. Astron. Soc.* **401** 1602–1612. https://doi.org/10.1111/j.1365-2966.2009.15798.x
[11] Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. MR3102554 https://doi.org/10.3150/12-BEJSP06
[12] Hsing, T. and Eubank, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR3379106 https://doi.org/10.1002/9781118762547

1350-7265

[13] Lang, A. and Schwab, C. (2015). Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations. *Ann. Appl. Probab.* **25** 3047–3094. MR3404631 https://doi.org/10.1214/14-AAP1067

[14] Marinucci, D. and Peccati, G. (2010). Ergodicity and Gaussianity for spherical random fields. *J. Math. Phys.* **51** 043301. MR2662485 https://doi.org/10.1063/1.3329423

[15] Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere: Representation, limit theorems and cosmological applications*. *London Mathematical Society Lecture Note Series* **389**. Cambridge: Cambridge Univ. Press. MR2840154 https://doi.org/10.1017/CBO9780511751677

[16] Marinucci, D. and Peccati, G. (2013). Mean-square continuity on homogeneous spaces of compact groups. *Electron. Commun. Probab.* **18** 1–10. MR3064996 https://doi.org/10.1214/ECP.v18-2400

[17] Marinucci, D. and Rossi, M. (2015). Stein-Malliavin approximations for nonlinear functionals of random eigenfunctions on $\mathbb{S}^d$. *J. Funct. Anal.* **268** 2379–2420. MR3318653 https://doi.org/10.1016/j.jfa.2015.02.004

[18] Marinucci, D., Rossi, M. and Wigman, I. (2020). The asymptotic equivalence of the sample trispectrum and the nodal length for random spherical harmonics. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 374–390. MR4058991 https://doi.org/10.1214/19-AIHP964

[19] Marinucci, D. and Wigman, I. (2011). The defect variance of random spherical harmonics. *J. Phys. A* **44** 355206. https://doi.org/10.1088/1751-8113/44/35/355206

[20] Marinucci, D. and Wigman, I. (2014). On nonlinear functionals of random spherical eigenfunctions. *Comm. Math. Phys.* **327** 849–872. MR3192051 https://doi.org/10.1007/s00220-014-1939-7

[21] Nourdin, I. and Peccati, G. (2012). *Normal Approximations with Malliavin Calculus: From Stein's method to universality*. *Cambridge Tracts in Mathematics* **192**. Cambridge: Cambridge Univ. Press. MR2962301 https://doi.org/10.1017/CBO9781139084659

[22] Panaretos, V.M. and Tavakoli, S. (2013). Fourier analysis of stationary time series in function space. *Ann. Statist.* **41** 568–603. MR3099114 https://doi.org/10.1214/13-AOS1086

[23] Panaretos, V.M. and Tavakoli, S. (2013). Cramér-Karhunen-Loève representation and harmonic principal component analysis of functional time series. *Stochastic Process. Appl.* **123** 2779–2807. MR3054545 https://doi.org/10.1016/j.spa.2013.03.015

[24] Porcu, E., Furrer, R. and Nychka, D. (2020). 30 years of space-time covariance functions. *Wiley Interdiscip. Rev.: Comput. Stat.* e1512. MR4218945 https://doi.org/10.1002/wics.1512

[25] Rossi, M. (2019). The defect of random hyperspherical harmonics. *J. Theoret. Probab.* **32** 2135–2165. MR4020703 https://doi.org/10.1007/s10959-018-0849-6

[26] Schoenberg, I.J. (1942). Positive definite functions on spheres. *Duke Math. J.* **9** 96–108. MR0005922 https://doi.org/10.1215/S0012-7094-42-00908-6

[27] Szegö, G. (1975). *Orthogonal Polynomials*, 4th ed. *American Mathematical Society Colloquium Publications* **23**. Providence: Amer. Math. Soc. MR0106295 https://doi.org/10.1090/coll/023

[28] Todino, A.P. (2019). A quantitative central limit theorem for the excursion area of random spherical harmonics over subdomains of $\mathbb{S}^2$. *J. Math. Phys.* **60** 023505. MR3916834 https://doi.org/10.1063/1.5048976

[29] Todino, A.P. (2020). Nodal lengths in shrinking domains for random eigenfunctions on S^2. *Bernoulli* **26** 3081–3110. MR4140538 https://doi.org/10.3150/20-BEJ1216

[30] Trübner, M. and Ziegel, J.F. (2017). Derivatives of isotropic positive definite functions on spheres. *Proc. Amer. Math. Soc.* **145** 3017–3031. MR3637950 https://doi.org/10.1090/proc/13561

[31] Ziegel, J. (2014). Convolution roots and differentiability of isotropic positive definite functions on spheres. *Proc. Amer. Math. Soc.* **142** 2063–2077. MR3182025 https://doi.org/10.1090/S0002-9939-2014-11989-7

Bernoulli **30**(3), 2024, 1746–1773
https://doi.org/10.3150/23-BEJ1651

Signal detection in degree corrected ERGMs

YUANZHE XU[a] and SUMIT MUKHERJEE[b]

Columbia University, NY, USA, [a]*yuanzhe.xu@columbia.edu,* [b]*sm3949@columbia.edu*

In this paper, we study sparse signal detection problems in "degree corrected" Exponential Random Graph Models (ERGMs). We study the performance of two tests based on conditionally centered sum of degrees/maximum of degrees, for a wide class of such ERGMs. The performance of these tests match the performance of corresponding uncentered tests in the $\boldsymbol{\beta}$ model (*Ann. Statist.* **46** (2018) 1288–1317). Focusing on the degree corrected two star ERGM, we show that improved detection is possible at "criticality" using a test based on (unconditional) sum of degrees. In this setting we provide matching lower bounds in all parameter regimes, which is based on correlations estimates between degrees under the alternative, and is of possible independent interest.

Keywords: Asymptotic efficiency; auxiliary variables; ERGM; phase transition; signal detection; two star

References

Anderson, C.J., Wasserman, S. and Crouch, B. (1999). A p^* primer: Logit models for social networks. *Soc. Netw.* **21** 37–66.

Bhamidi, S., Steele, J.M. and Zaman, T. (2015). Twitter event networks and the superstar model. *Ann. Appl. Probab.* **25** 2462–2502. MR3375881 https://doi.org/10.1214/14-AAP1053

Blitzstein, J. and Diaconis, P. (2011). A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet Math.* **6** 489–522.

Burnašev, M.V. (1979). Minimax detection of an imperfectly known signal against a background of Gaussian white noise. *Teor. Veroyatn. Primen.* **24** 106–118. MR0522240

Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072 https://doi.org/10.1007/s00440-006-0029-y

Chatterjee, S. and Diaconis, P. (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41** 2428–2461. MR3127871 https://doi.org/10.1214/13-AOS1155

Chatterjee, S., Diaconis, P. and Sly, A. (2011). Random graphs with a given degree sequence. *Ann. Appl. Probab.* **21** 1400–1435. MR2857452 https://doi.org/10.1214/10-AAP728

Chatterjee, S. and Mukherjee, S. (2019). Estimation in tournaments and graphs under monotonicity constraints. *IEEE Trans. Inf. Theory* **65** 3525–3539. MR3959003 https://doi.org/10.1109/TIT.2019.2893911

Deb, N. and Mukherjee, S. (2023). Fluctuations in mean-field Ising models. *Ann. Appl. Probab.* **33** 1961–2003. MR4583662 https://doi.org/10.1214/22-aap1857

Deb, N., Mukherjee, R., Mukherjee, S. and Yuan, M. (to appear). Detecting structured signals in Ising models. *Ann. Appl. Probab.*

Dembo, A. and Montanari, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563 https://doi.org/10.1214/09-BJPS027

Frank, O. and Strauss, D. (1986). Markov graphs. *J. Amer. Statist. Assoc.* **81** 832–842. MR0860518

Götze, F., Sambale, H. and Sinulis, A. (2021). Concentration inequalities for polynomials in α-sub-exponential random variables. *Electron. J. Probab.* **26** 48. MR4247973 https://doi.org/10.1214/21-ejp606

Holland, P.W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *J. Amer. Statist. Assoc.* **76** 33–65. MR0608176

Ingster, Y.I. (1994). Minimax detection of a signal in l_p metrics. *J. Math. Sci.* **68** 503–515.

Ingster, Y., Ingster, J.I. and Suslina, I. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models* **169**. Berlin: Springer.

Lebowitz, J.L. (1974). GHS and other inequalities. *Comm. Math. Phys.* **35** 87–92. MR0339738

1350-7265

Mukherjee, R., Mukherjee, S. and Sen, S. (2018). Detection thresholds for the β-model on sparse graphs. *Ann. Statist.* **46** 1288–1317. MR3798004 https://doi.org/10.1214/17-AOS1585

Mukherjee, R., Mukherjee, S. and Yuan, M. (2018). Global testing against sparse alternatives under Ising models. *Ann. Statist.* **46** 2062–2093. MR3845011 https://doi.org/10.1214/17-AOS1612

Mukherjee, S. and Xu, Y. (2023). Statistics of the two star ERGM. *Bernoulli* **29** 24–51. MR4497238 https://doi.org/10.3150/21-bej1448

Park, J. and Newman, M.E.J. (2004). Solution of the two-star model of a network. *Phys. Rev. E (3)* **70** 066146. MR2133810 https://doi.org/10.1103/PhysRevE.70.066146

Rinaldo, A., Petrović, S. and Fienberg, S.E. (2013). Maximum likelihood estimation in the β-model. *Ann. Statist.* **41** 1085–1110. MR3113804 https://doi.org/10.1214/12-AOS1078

Robins, G., Pattison, P., Kalish, Y. and Lusher, D. (2007). An introduction to exponential random graph (p^*) models for social networks. *Soc. Netw.* **29** 173–191.

Schweinberger, M. and Stewart, J. (2020). Concentration and consistency results for canonical and curved exponential-family models of random graphs. *Ann. Statist.* **48** 374–396. MR4065166 https://doi.org/10.1214/19-AOS1810

Shalizi, C.R. and Rinaldo, A. (2013). Consistency under sampling of exponential random graph models. *Ann. Statist.* **41** 508–535. MR3099112 https://doi.org/10.1214/12-AOS1044

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications* **8**. Cambridge: Cambridge University Press.

Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks. I. An introduction to Markov graphs and *p*. *Psychometrika* **61** 401–425. MR1424909 https://doi.org/10.1007/BF02294547

Xu, Y. and Mukherjee, S. (2024). Supplement to "Signal detection in degree corrected ERGMs." https://doi.org/10.3150/23-BEJ1651SUPP

Bernoulli **30**(3), 2024, 1774–1799
https://doi.org/10.3150/23-BEJ1652

Stochastic fractional diffusion equations with Gaussian noise rough in space

YUHUI GUO[1,a], JIAN SONG[2,b] and XIAOMING SONG[3,c]

[1]*School of Mathematics, Shandong University, Jinan, Shandong, 250100, China,* [a]*guoyuhui@mail.sdu.edu.cn*
[2]*Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, Shandong, 266237, China,* [b]*txjsong@sdu.edu.cn*
[3]*Department of Mathematics, Drexel University, Philadelphia, PA 19104, USA,* [c]*xs73@drexel.edu*

In this article, we consider the stochastic fractional diffusion equation

$$\left(\partial^{\beta} + \frac{\nu}{2}(-\Delta)^{\alpha/2}\right) u(t,x) = \lambda\, I_{0_+}^{\gamma} \left[u(t,x)\dot{W}(t,x)\right], \quad t>0,\ x\in\mathbb{R},$$

where $\alpha > 0$, $\beta \in (0,2]$, $\gamma \geq 0$, $\lambda \neq 0$, $\nu > 0$, and $\dot{W}$ is a Gaussian noise which is white or fractional in time and rough in space. We prove the existence and uniqueness of the solution in the Itô-Skorohod sense and obtain the lower and upper bounds for the p-th moment. The Hölder regularity of the solution is also studied.

Keywords: Fractional Brownian field; Hölder continuity; Malliavin calculus; Mittag-Leffler function; moment estimates

References

Balan, R.M., Chen, L. and Chen, X. (2022). Exact asymptotics of the stochastic wave equation with time-independent noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 1590–1620. MR4452644 https://doi.org/10.1214/21-aihp1207

Balan, R.M. and Conus, D. (2016). Intermittency for the wave and heat equations with fractional noise in time. *Ann. Probab.* **44** 1488–1534. MR3474476 https://doi.org/10.1214/15-AOP1005

Balan, R.M., Jolis, M. and Quer-Sardanyons, L. (2015). SPDEs with affine multiplicative fractional noise in space with index $\frac{1}{4} < H < \frac{1}{2}$. *Electron. J. Probab.* **20** no. 54, 36. MR3354614 https://doi.org/10.1214/EJP.v20-3719

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

Balan, R.M., Quer-Sardanyons, L. and Song, J. (2019). Hölder continuity for the parabolic Anderson model with space-time homogeneous Gaussian noise. *Acta Math. Sci. Ser. B Engl. Ed.* **39** 717–730. MR4066501 https://doi.org/10.1007/s10473-019-0306-3

Balan, R.M. and Song, J. (2017). Hyperbolic Anderson model with space-time homogeneous Gaussian noise. *ALEA Lat. Am. J. Probab. Math. Stat.* **14** 799–849. MR3731795

Balan, R.M. and Song, J. (2019). Second order Lyapunov exponents for parabolic and hyperbolic Anderson models. *Bernoulli* **25** 3069–3089. MR4003574 https://doi.org/10.3150/18-BEJ1080

Bertini, L. and Cancrini, N. (1995). The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Stat. Phys.* **78** 1377–1401. MR1316109 https://doi.org/10.1007/BF02180136

Chen, L. (2017). Nonlinear stochastic time-fractional diffusion equations on $\mathbb{R}$: Moments, Hölder regularity and intermittency. *Trans. Amer. Math. Soc.* **369** 8497–8535. MR3710633 https://doi.org/10.1090/tran/6951

Chen, X. (2019). Parabolic Anderson model with rough or critical Gaussian noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** 941–976. MR3949959 https://doi.org/10.1214/18-aihp904

Chen, X. (2020). Parabolic Anderson model with a fractional Gaussian noise that is rough in time. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 792–825. MR4076766 https://doi.org/10.1214/19-AIHP983

1350-7265

Chen, L. and Eisenberg, N. (2023). Interpolating the stochastic heat and wave equations with time-independent noise: Solvability and exact asymptotics. *Stoch. Partial Differ. Equ. Anal. Comput.* **11** 1203–1253. MR4624137 https://doi.org/10.1007/s40072-022-00258-6

Chen, L., Guo, Y. and Song, J. (2022). Moments and asymptotics for a class of SPDEs with space-time white noise. arXiv preprint. Available at arXiv:2206.10069.

Chen, Z.-Q. and Hu, Y. (2022a). Solvability of parabolic Anderson equation with fractional Gaussian noise. *Commun. Math. Stat.* **1**. https://doi.org/10.1007/s40304-021-00264-5

Chen, L. and Hu, G. (2022b). Hölder regularity for the nonlinear stochastic time-fractional slow & fast diffusion equations on $\mathbb{R}^d$. *Fract. Calc. Appl. Anal.* **25** 608–629. MR4437294 https://doi.org/10.1007/s13540-022-00033-3

Chen, L., Hu, Y. and Nualart, D. (2019). Nonlinear stochastic time-fractional slow and fast diffusion equations on $\mathbb{R}^d$. *Stochastic Process. Appl.* **129** 5073–5112. MR4025700 https://doi.org/10.1016/j.spa.2019.01.003

Chen, Z.-Q., Kim, K.-H. and Kim, P. (2015). Fractional time stochastic partial differential equations. *Stochastic Process. Appl.* **125** 1470–1499. MR3310354 https://doi.org/10.1016/j.spa.2014.11.005

Chen, X., Hu, Y., Song, J. and Xing, F. (2015). Exponential asymptotics for time-space Hamiltonians. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1529–1561. MR3414457 https://doi.org/10.1214/13-AIHP588

Chen, L., Hu, G., Hu, Y. and Huang, J. (2017). Space-time fractional diffusions in Gaussian noisy environment. *Stochastics* **89** 171–206. MR3574699 https://doi.org/10.1080/17442508.2016.1146282

Chen, X., Hu, Y., Song, J. and Song, X. (2018). Temporal asymptotics for fractional parabolic Anderson model. *Electron. J. Probab.* **23** Paper No. 14, 39. MR3771751 https://doi.org/10.1214/18-EJP139

Gorenflo, R., Kilbas, A.A., Mainardi, F. and Rogosin, S.V. (2014). *Mittag-Leffler Functions, Related Topics and Applications*. *Springer Monographs in Mathematics*. Heidelberg: Springer. MR3244285 https://doi.org/10.1007/978-3-662-43930-2

Hu, Y. and Lê, K. (2019). Joint Hölder continuity of parabolic Anderson model. *Acta Math. Sci. Ser. B Engl. Ed.* **39** 764–780. MR4066504 https://doi.org/10.1007/s10473-019-0309-0

Hu, Y. and Nualart, D. (2009). Stochastic heat equation driven by fractional noise and local time. *Probab. Theory Related Fields* **143** 285–328. MR2449130 https://doi.org/10.1007/s00440-007-0127-5

Hu, Y., Nualart, D. and Song, J. (2011). Feynman-Kac formula for heat equation driven by fractional white noise. *Ann. Probab.* **39** 291–326. MR2778803 https://doi.org/10.1214/10-AOP547

Hu, Y. and Wang, X (2022). Matching upper and lower moment bounds for a large class of stochastic PDEs driven by general space-time Gaussian noises. *Stoch. Partial Differ. Equ. Anal. Comput.* 1–52. https://doi.org/10.1007/s40072-022-00278-2

Hu, Y., Huang, J., Nualart, D. and Tindel, S. (2015). Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. *Electron. J. Probab.* **20** no. 55, 50. MR3354615 https://doi.org/10.1214/EJP.v20-3316

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

Hu, Y., Huang, J., Lê, K., Nualart, D. and Tindel, S. (2018). Parabolic Anderson model with rough dependence in space. In *Computation and Combinatorics in Dynamics, Stochastics and Control*. *Abel Symp.* **13** 477–498. Cham: Springer. MR3967394

Huang, J., Lê, K. and Nualart, D. (2017a). Large time asymptotics for the parabolic Anderson model driven by space and time correlated noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **5** 614–651. MR3736656 https://doi.org/10.1007/s40072-017-0099-0

Huang, J., Lê, K. and Nualart, D. (2017b). Large time asymptotics for the parabolic Anderson model driven by spatially correlated noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1305–1340. MR3689969 https://doi.org/10.1214/16-AIHP756

Khoshnevisan, D. (2014). *Analysis of Stochastic Partial Differential Equations*. *CBMS Regional Conference Series in Mathematics* **119**. Providence, RI: Amer. Math. Soc.. Published for the Conference Board of the Mathematical Sciences, Washington, DC. MR3222416 https://doi.org/10.1090/cbms/119

Kilbas, A.A., Srivastava, H.M. and Trujillo, J.J. (2006). *Theory and Applications of Fractional Differential Equations*. *North-Holland Mathematics Studies* **204**. Amsterdam: Elsevier Science B.V. MR2218073

Liu, S., Hu, Y. and Wang, X. (2022). Nonlinear stochastic wave equation driven by rough noise. *J. Differ. Equ.* **331** 99–161. MR4435914 https://doi.org/10.1016/j.jde.2022.05.016

Mijena, J.B. and Nane, E. (2015). Space-time fractional stochastic partial differential equations. *Stochastic Process. Appl.* **125** 3301–3326. MR3357610 https://doi.org/10.1016/j.spa.2015.04.008

Nualart, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. *Probability and Its Applications (New York)*. Berlin: Springer. MR2200233

Podlubny, I. (1999). *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*. *Mathematics in Science and Engineering* **198**. San Diego, CA: Academic Press. MR1658022

Song, J. (2012). Asymptotic behavior of the solution of heat equation driven by fractional white noise. *Statist. Probab. Lett.* **82** 614–620. MR2887479 https://doi.org/10.1016/j.spl.2011.11.017

Song, J. (2017). On a class of stochastic partial differential equations. *Stochastic Process. Appl.* **127** 37–79. MR3575535 https://doi.org/10.1016/j.spa.2016.05.008

Song, J., Song, X. and Xu, F. (2020). Fractional stochastic wave equation driven by a Gaussian noise rough in space. *Bernoulli* **26** 2699–2726. MR4140526 https://doi.org/10.3150/20-BEJ1204

Bernoulli **30**(3), 2024, 1800–1824
https://doi.org/10.3150/23-BEJ1653

Functional linear quantile regression on a two-dimensional domain

NAN ZHANG[1,a], PENG LIU[2,b], LINGLONG KONG[3,c], BEI JIANG[3,d] and JIANHUA Z. HUANG[4,e]

[1]*School of Data Science, Fudan University, Shanghai, China,* [a]*zhangnan@fudan.edu.cn*
[2]*School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, UK,* [b]*p.liu@kent.ac.uk*
[3]*Department of Mathematical and Statistical Sciences University of Alberta, Edmonton, Canada,* [c]*lkong@ualberta.ca,* [d]*bei1@ualberta.ca*
[4]*School of Data Science, The Chinese University of Hong Kong, Shenzhen, Shenzhen, China,* [e]*jhuang@cuhk.edu.cn*

This article considers the functional linear quantile regression which models the conditional quantile of a scalar response given a functional predictor over a two-dimensional domain. We propose an estimator for the slope function by minimizing the penalized empirical check loss function. Under the framework of reproducing kernel Hilbert space, the minimax rate of convergence for the regularized estimator is established. Using the theory of interpolation spaces on a two- or multi-dimensional domain, we develop a novel result on simultaneous diagonalization of the reproducing and covariance kernels, revealing the interaction of the two kernels in determining the optimal convergence rate of the estimator. Sufficient conditions are provided to show that our analysis applies to many situations, for example, when the covariance kernel is from the Matérn class, and the slope function belongs to a Sobolev space. We implement the interior point method to compute the regularized estimator and illustrate the proposed method by applying it to the hippocampus surface data in the ADNI study.

Keywords: Functional linear regression; multi-dimensional domain; quantile regression; rate of convergence; reproducing kernel Hilbert space; simultaneous diagonalization

References

Adams, R.A. and Fournier, J.J.F. (2003). *Sobolev Spaces*, 2nd ed. *Pure and Applied Mathematics (Amsterdam)* **140**. Amsterdam: Elsevier/Academic Press. MR2424078

Arnone, E., Azzimonti, L., Nobile, F. and Sangalli, L.M. (2019). Modeling spatially dependent functional data via regression with differential regularization. *J. Multivariate Anal.* **170** 275–295. MR3913041 https://doi.org/10.1016/j.jmva.2018.09.006

Barnes, J. and Fox, N.C. (2014). The search for early markers of AD: Hippocampal atrophy and memory deficits. *Int. Psychogeriatr.* **26** 1065–1066. https://doi.org/10.1017/S1041610214000623

Beckett, L.A., Donohue, M.C., Wang, C., Aisen, P., Harvey, D.J., Saito, N. and Initiative, A.D.N. (2015). The Alzheimer's Disease Neuroimaging Initiative phase 2: Increasing the length, breadth, and depth of our understanding. *Alzheimer's Dement.* **11** 823–831.

Bertsimas, D. and Tsitsiklis, J.N. (1997). *Introduction to Linear Optimization*. Athena Scientific.

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3** 1–122.

Cai, T.T. and Hall, P. (2006). Prediction in functional linear regression. *Ann. Statist.* **34** 2159–2179. MR2291496 https://doi.org/10.1214/009053606000000830

Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *J. Amer. Statist. Assoc.* **107** 1201–1216. MR3010906 https://doi.org/10.1080/01621459.2012.716337

Cardot, H., Ferraty, F. and Sarda, P. (2003). Spline estimators for the functional linear model. *Statist. Sinica* **13** 571–591. MR1997162

Carl, B. and Stephani, I. (1990). *Entropy, Compactness and the Approximation of Operators*. *Cambridge Tracts in Mathematics* **98**. Cambridge: Cambridge Univ. Press. MR1098497 https://doi.org/10.1017/CBO9780511897467

Changyong, F., Hongyue, W., Naiji, L., Tian, C., Hua, H., Ying, L. and TU, X.M. (2014). Log-transformation and its implications for data analysis. *Shanghai Archives of Psychiatry* **26** 105–109.

Crambes, C., Kneip, A. and Sarda, P. (2009). Smoothing splines estimators for functional linear regression. *Ann. Statist.* **37** 35–72. MR2488344 https://doi.org/10.1214/07-AOS563

Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables (Proc. Conf., Math. Res. Inst., Oberwolfach, 1976)*. (W. Schemp and K. Zeller, eds.). *Lecture Notes in Math., Vol. 571* 85–100. Berlin: Springer. MR0493110

Edmunds, D.E. and Triebel, H. (1996). *Function Spaces, Entropy Numbers, Differential Operators*. *Cambridge Tracts in Mathematics* **120**. Cambridge: Cambridge Univ. Press. MR1410258 https://doi.org/10.1017/CBO9780511662201

Fasiolo, M., Wood, S.N., Zaffran, M., Nedellec, R. and Goude, Y. (2021). Fast calibrated additive quantile regression. *J. Amer. Statist. Assoc.* **116** 1402–1412. MR4309281 https://doi.org/10.1080/01621459.2020.1725521

Galea, M. and Woodward, M. (2005). Mini-mental state examination (MMSE). *Aust. J. Physiother.* **51** 198.

Geraci, M. (2019). Additive quantile regression for clustered data with an application to children's physical activity. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **68** 1071–1089. MR4002384

Hall, P. and Horowitz, J.L. (2007). Methodology and convergence rates for functional linear regression. *Ann. Statist.* **35** 70–91. MR2332269 https://doi.org/10.1214/009053606000000957

Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming. In *Proceedings of the Sixteenth Annual ACM Symposium on Theory of Computing* 302–311.

Kato, K. (2012). Estimation in functional linear quantile regression. *Ann. Statist.* **40** 3108–3136. MR3097971 https://doi.org/10.1214/12-AOS1066

Kimeldorf, G. and Wahba, G. (1971). Some results on Tchebycheffian spline functions. *J. Math. Anal. Appl.* **33** 82–95. MR0290013 https://doi.org/10.1016/0022-247X(71)90184-3

Knight, K. (1998). Limiting distributions for L_1 regression estimators under general conditions. *Ann. Statist.* **26** 755–770. MR1626024 https://doi.org/10.1214/aos/1028144858

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

Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50. MR0474644 https://doi.org/10.2307/1913643

Koenker, R. and Machado, J.A.F. (1999). Goodness of fit and related inference processes for quantile regression. *J. Amer. Statist. Assoc.* **94** 1296–1310. MR1731491 https://doi.org/10.2307/2669943

Koenker, R., Ng, P. and Portnoy, S. (1994). Quantile smoothing splines. *Biometrika* **81** 673–680. MR1326417 https://doi.org/10.1093/biomet/81.4.673

Koenker, R., Chernozhukov, V., He, X. and Peng, L. (2017). *Handbook of Quantile Regression*. Boca Raton: CRC Press.

Li, Y., Liu, Y. and Zhu, J. (2007). Quantile regression in reproducing kernel Hilbert spaces. *J. Amer. Statist. Assoc.* **102** 255–268. MR2293307 https://doi.org/10.1198/016214506000000979

Li, R., Lu, W., Zhu, Z. and Lian, H. (2021). Optimal prediction of quantile functional linear regression in reproducing kernel Hilbert spaces. *J. Statist. Plann. Inference* **211** 162–170. MR4122078 https://doi.org/10.1016/j.jspi.2020.06.010

Li, M., Wang, K., Maity, A. and Staicu, A.-M. (2022). Inference in functional linear quantile regression. *J. Multivariate Anal.* **190** Paper No. 104985. MR4402003 https://doi.org/10.1016/j.jmva.2022.104985

Liu, Y., Li, M. and Morris, J.S. (2020). Function-on-scalar quantile regression with application to mass spectrometry proteomics data. *Ann. Appl. Stat.* **14** 521–541. MR4117818 https://doi.org/10.1214/19-AOAS1319

Lustig, I.J., Marsten, R.E. and Shanno, D.F. (1994). Interior point methods for linear programming: Computational state of the art. *ORSA J. Comput.* **6** 1–14. MR1261376 https://doi.org/10.1287/ijoc.6.1.1

Lv, S., Lin, H., Lian, H. and Huang, J. (2018). Oracle inequalities for sparse additive quantile regression in reproducing kernel Hilbert space. *Ann. Statist.* **46** 781–813. MR3782384 https://doi.org/10.1214/17-AOS1567

Ma, H., Li, T., Zhu, H. and Zhu, Z. (2019). Quantile regression for functional partially linear model in ultra-high dimensions. *Comput. Statist. Data Anal.* **129** 135–147. MR3855203 https://doi.org/10.1016/j.csda.2018.06.005

Morris, J.S. (2015). Functional regression. *Annu. Rev. Stat. Appl.* **2** 321–359.

Morris, J.S., Baladandayuthapani, V., Herrick, R.C., Sanna, P. and Gutstein, H. (2011). Automated analysis of quantitative image data using isomorphic functional mixed models, with application to proteomics data. *Ann. Appl. Stat.* **5** 894–923. MR2840180 https://doi.org/10.1214/10-AOAS407

Muggeo, V.M.R., Torretta, F., Eilers, P.H.C., Sciandra, M. and Attanasio, M. (2021). Multiple smoothing parameters selection in additive regression quantiles. *Stat. Model.* **21** 428–448. MR4307778 https://doi.org/10.1177/1471082X20929802

Riekkinen, P. Jr., Soininen, H., Helkala, E.L., Partanen, K., Laakso, M., Vanhanen, M. and Riekkinen, P. (1995). Hippocampal atrophy, acute THA treatment and memory in Alzheimer's disease. *NeuroReport* **6** 1297–1300. https://doi.org/10.1097/00001756-199506090-00017

Pearson, J.W. and Gondzio, J. (2017). Fast interior point solution of quadratic programming problems arising from PDE-constrained optimization. *Numer. Math.* **137** 959–999. MR3719049 https://doi.org/10.1007/s00211-017-0892-8

Portnoy, S. and Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Statist. Sci.* **12** 279–300. MR1619189 https://doi.org/10.1214/ss/1030037960

Potra, F.A. and Wright, S.J. (2000). Interior-point methods. *J. Comput. Appl. Math.* **124** 281–302.

Qiu, A., Brown, T., Fischl, B., Ma, J. and Miller, M.I. (2010). Atlas generation for subcortical and ventricular structures with its applications in shape analysis. *IEEE Trans. Image Process.* **19** 1539–1547. MR2814625 https://doi.org/10.1109/TIP.2010.2042099

Ramsay, J.O. and Dalzell, C.J. (1991). Some tools for functional data analysis. *J. Roy. Statist. Soc. Ser. B* **53** 539–572. MR1125714

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Rasmussen, C.E. and Williams, C.K.I. (2006). *Gaussian Processes for Machine Learning*. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR2514435

Reiss, P.T. and Huang, L. (2012). Smoothness selection for penalized quantile regression splines. *Int. J. Biostat.* **8** Art. 10. MR2997664 https://doi.org/10.1515/1557-4679.1381

Ritter, K., Wasilkowski, G.W. and Woźniakowski, H. (1995). Multivariate integration and approximation for random fields satisfying Sacks-Ylvisaker conditions. *Ann. Appl. Probab.* **5** 518–540. MR1336881

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

Santacruz, K.S. and Swagerty, D. (2001). Early diagnosis of dementia. *Amer. Fam. Phys.* **63** 703–718.

Schölkopf, B., Herbrich, R. and Smola, A.J. (2001). A generalized representer theorem. In *Computational Learning Theory (Amsterdam, 2001)*. (D. Helmbold and B. Williamson, eds.). *Lecture Notes in Computer Science* **2111** 416–426. Berlin: Springer. MR2042050 https://doi.org/10.1007/3-540-44581-1_27

Sherwood, B. and Wang, L. (2016). Partially linear additive quantile regression in ultra-high dimension. *Ann. Statist.* **44** 288–317. MR3449769 https://doi.org/10.1214/15-AOS1367

Shi, H., Yang, Y., Wang, L., Ma, D., Beg, M.F., Pei, J. and Cao, J. (2022). Two-dimensional functional principal component analysis for image feature extraction. *J. Comput. Graph. Statist.* **31** 1127–1140. MR4513375 https://doi.org/10.1080/10618600.2022.2035738

Stein, M.L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. *Springer Series in Statistics*. New York: Springer. MR1697409 https://doi.org/10.1007/978-1-4612-1494-6

Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*. Springer Science & Business Media.

Steinwart, I. and Scovel, C. (2012). Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs. *Constr. Approx.* **35** 363–417. MR2914365 https://doi.org/10.1007/s00365-012-9153-3

Sun, X., Du, P., Wang, X. and Ma, P. (2018). Optimal penalized function-on-function regression under a reproducing kernel Hilbert space framework. *J. Amer. Statist. Assoc.* **113** 1601–1611. MR3902232 https://doi.org/10.1080/01621459.2017.1356320

Thompson, P.M., Hayashi, K.M., Zubicaray, G.I.D., Janke, A.L., Rose, S.E., Semple, J., Hong, M.S., Herman, D.H., Gravano, D., Doddrell, D.M. and Toga, A.W. (2004). Mapping hippocampal and ventricular change in Alzheimer disease. *NeuroImage* **22** 1754–1766. https://doi.org/10.1016/j.neuroimage.2004.03.040

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

Volgushev, S., Chao, S.-K. and Cheng, G. (2019). Distributed inference for quantile regression processes. *Ann. Statist.* **47** 1634–1662. MR3911125 https://doi.org/10.1214/18-AOS1730

Wahba, G. (1990). *Spline Models for Observational Data*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. Philadelphia, PA: SIAM. MR1045442 https://doi.org/10.1137/1.9781611970128

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

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Wang, J., Wong, R.K.W. and Zhang, X. (2022). Low-rank covariance function estimation for multidimensional functional data. *J. Amer. Statist. Assoc.* **117** 809–822. MR4436314 https://doi.org/10.1080/01621459.2020.1820344

Wang, Y., Lui, L.M., Gu, X., Hayashi, K.M., Chan, T.F., Toga, A.W., Thompson, P.M. and Yau, S.-T. (2007). Brain surface conformal parameterization using Riemann surface structure. *IEEE Trans. Med. Imag.* **26** 853–865.

Wang, Y., Chan, T.F., Toga, A.W. and Thompson, P.M. (2009). Multivariate tensor-based brain anatomical surface morphometry via holomorphic one-forms. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 337–344. Springer.

Wendland, H. (2005). *Scattered Data Approximation*. *Cambridge Monographs on Applied and Computational Mathematics* **17**. Cambridge: Cambridge Univ. Press. MR2131724

Yang, Y. and Barron, A. (1999). Information-theoretic determination of minimax rates of convergence. *Ann. Statist.* **27** 1564–1599. MR1742500 https://doi.org/10.1214/aos/1017939142

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. MR2253106 https://doi.org/10.1214/009053605000000660

Yao, F., Sue-Chee, S. and Wang, F. (2017). Regularized partially functional quantile regression. *J. Multivariate Anal.* **156** 39–56. MR3624684 https://doi.org/10.1016/j.jmva.2017.02.001

Yuan, M. (2006). GACV for quantile smoothing splines. *Comput. Statist. Data Anal.* **50** 813–829. MR2207010 https://doi.org/10.1016/j.csda.2004.10.008

Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. MR2766857 https://doi.org/10.1214/09-AOS772

Zhang, P. (1993). Model selection via multifold cross validation. *Ann. Statist.* **21** 299–313. MR1212178 https://doi.org/10.1214/aos/1176349027

Zhang, Z., Wang, X., Kong, L. and Zhu, H. (2022). High-dimensional spatial quantile function-on-scalar regression. *J. Amer. Statist. Assoc.* **117** 1563–1578. MR4480732 https://doi.org/10.1080/01621459.2020.1870984

Zhou, L. and Pan, H. (2014). Principal component analysis of two-dimensional functional data. *J. Comput. Graph. Statist.* **23** 779–801. MR3224656 https://doi.org/10.1080/10618600.2013.827986

Bernoulli **30**(3), 2024, 1825–1844
https://doi.org/10.3150/23-BEJ1654

Finitely additive mass transportation

PIETRO RIGO[a]

Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, via delle Belle Arti 41, 40126 Bologna, Italy, [a]*pietro.rigo@unibo.it*

Some classical mass transportation problems are investigated in a finitely additive setting. Let $\Omega = \prod_{i=1}^n \Omega_i$ and $\mathcal{A} = \otimes_{i=1}^n \mathcal{A}_i$, where $(\Omega_i, \mathcal{A}_i, \mu_i)$ is a (σ-additive) probability space for $i = 1, \dots, n$. Let $c : \Omega \to [0, \infty]$ be an $\mathcal{A}$-measurable cost function. Let M be the collection of finitely additive probabilities on $\mathcal{A}$ with marginals $\mu_1, \dots, \mu_n$. If couplings are meant as elements of M, most classical results of mass transportation theory, including duality and attainability of the Kantorovich inf, are valid without any further assumptions. Special attention is devoted to martingale transport. Let $(\Omega_i, \mathcal{A}_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ for all i and

$$M_1 = \{P \in M : P \ll P^* \text{ and } (\pi_1, \dots, \pi_n) \text{ is a } P\text{-martingale}\}$$

where P^* is a reference probability on $\mathcal{A}$ and $\pi_1, \dots, \pi_n$ are the canonical projections on $\Omega = \mathbb{R}^n$. If $M_1 \neq \emptyset$, the Kantorovich inf over M_1 is attained, in the sense that $\int c\, dP = \inf_{Q \in M_1} \int c\, dQ$ for some $P \in M_1$. Conditions for $M_1 \neq \emptyset$ are given as well.

Keywords: Coupling; duality theorem; finitely additive probability; martingale; mass transportation

References

[1] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows*, 2nd ed. Basel: Birkhäuser.
[2] Backhoff-Veraguas, J. and Pammer, G. (2022). Stability of martingale optimal transport and weak optimal transport. *Ann. Appl. Probab.* **32** 721–752. MR4386541 https://doi.org/10.1214/21-aap1694
[3] Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985 https://doi.org/10.1007/s00780-013-0205-8
[4] Beiglböck, M., Jourdain, B., Margheriti, W. and Pammer, G. (2022). Approximation of martingale couplings on the line in the adapted weak topology. *Probab. Theory Related Fields* **183** 359–413. MR4421177 https://doi.org/10.1007/s00440-021-01103-y
[5] Beiglböck, M., Nutz, M. and Touzi, N. (2017). Complete duality for martingale optimal transport on the line. *Ann. Probab.* **45** 3038–3074. MR3706738 https://doi.org/10.1214/16-AOP1131
[6] Beiglböck, M. and Schachermayer, W. (2011). Duality for Borel measurable cost functions. *Trans. Amer. Math. Soc.* **363** 4203–4224. MR2792985 https://doi.org/10.1090/S0002-9947-2011-05174-3
[7] Berti, P., Pratelli, L. and Rigo, P. (2013). Finitely additive equivalent martingale measures. *J. Theoret. Probab.* **26** 46–57. MR3023834 https://doi.org/10.1007/s10959-010-0337-0
[8] Berti, P., Pratelli, L. and Rigo, P. (2015). Two versions of the fundamental theorem of asset pricing. *Electron. J. Probab.* **20** Paper No. 34, 21. MR3335825 https://doi.org/10.1214/EJP.v20-3321
[9] Berti, P., Pratelli, L., Rigo, P. and Spizzichino, F. (2015). Equivalent or absolutely continuous probability measures with given marginals. *Depend. Model.* **3** 47–58. MR3418656 https://doi.org/10.1515/demo-2015-0004
[10] Berti, P. and Rigo, P. (2021). Finitely additive mixtures of probability measures. *J. Math. Anal. Appl.* **500** Paper No. 125114, 16. MR4227103 https://doi.org/10.1016/j.jmaa.2021.125114
[11] Bhaskara Rao, K.P.S. and Bhaskara Rao, M. (1983). *Theory of Charges: A Study of Finitely Additive Measures*. *Pure and Applied Mathematics* **109**. New York: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0751777

[12] Brückerhoff, M. and Juillet, N. (2022). Instability of martingale optimal transport in dimension $d \geq 2$. *Electron. Commun. Probab.* **27** Paper No. 24, 10. MR4416823 https://doi.org/10.1134/s1560354722010051

[13] Dalang, R.C., Morton, A. and Willinger, W. (1990). Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stoch. Stoch. Rep.* **29** 185–201. MR1041035 https://doi.org/10.1080/17442509008833613

[14] Ekren, I. and Soner, H.M. (2018). Constrained optimal transport. *Arch. Ration. Mech. Anal.* **227** 929–965. MR3744379 https://doi.org/10.1007/s00205-017-1178-0

[15] Galichon, A., Henry-Labordère, P. and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24** 312–336. MR3161649 https://doi.org/10.1214/13-AAP925

[16] Kellerer, H.G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** 399–432. MR0761565 https://doi.org/10.1007/BF00532047

[17] Korman, J. and McCann, R.J. (2015). Optimal transportation with capacity constraints. *Trans. Amer. Math. Soc.* **367** 1501–1521. MR3286490 https://doi.org/10.1090/S0002-9947-2014-06032-7

[18] Luschgy, H. and Thomsen, W. (1983). Extreme points in the Hahn-Banach-Kantorovič setting. *Pacific J. Math.* **105** 387–398. MR0691610

[19] Rachev, S.T. and Rüschendorf, L. (1988). *Mass Transportation Problems, Volume I: Theory*. New York: Springer.

[20] Ramachandran, D. (1979). *Perfect Measures. Part I: Basic Theory*. *ISI Lecture Notes* **5**. New Delhi: Macmillan Co. of India, Ltd. MR0553600

[21] Ramachandran, D. (1996). The marginal problem in arbitrary product spaces. In *Distributions with Fixed Marginals and Related Topics (Seattle, WA, 1993)*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **28** 260–272. Hayward, CA: IMS. MR1485537 https://doi.org/10.1214/lnms/1215452624

[22] Ramachandran, D. and Rüschendorf, L. (1995). A general duality theorem for marginal problems. *Probab. Theory Related Fields* **101** 311–319. MR1324088 https://doi.org/10.1007/BF01200499

[23] Rigo, P. (2020). A note on duality theorems in mass transportation. *J. Theoret. Probab.* **33** 2337–2350. MR4166202 https://doi.org/10.1007/s10959-019-00932-x

[24] Rüschendorf, L. (1991). Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals (Rome, 1990)* (G. Dall'Aglio, S. Kotz and G. Salinetti, eds.). *Math. Appl.* **67** 151–187. Dordrecht: Kluwer Academic. MR1215951

[25] Rüschendorf, L. (1996). On c-optimal random variables. *Statist. Probab. Lett.* **27** 267–270. MR1395577 https://doi.org/10.1016/0167-7152(95)00078-X

[26] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36** 423–439. MR0177430 https://doi.org/10.1214/aoms/1177700153

[27] Villani, C. (2009). *Optimal Transport, Old and New*. New York: Springer.

[28] Zaev, D.A. (2015). On the Monge-Kantorovich problem with additional linear constraints. *Math. Notes* **98** 725–741.

Bernoulli **30**(3), 2024, 1845–1877
https://doi.org/10.3150/23-BEJ1655

Local convergence rates of the nonparametric least squares estimator with applications to transfer learning

JOHANNES SCHMIDT-HIEBER[a] and PETR ZAMOLODTCHIKOV[b]

Department of Applied Mathematics, University of Twente, Enschede, The Netherlands,
[a]*a.j.schmidt-hieber@utwente.nl,* [b]*p.zamolodtchikov@utwente.nl*

Convergence properties of empirical risk minimizers can be conveniently expressed in terms of the associated population risk. To derive bounds for the performance of the estimator under covariate shift, however, pointwise convergence rates are required. Under weak assumptions on the design distribution, it is shown that least squares estimators (LSE) over 1-Lipschitz functions are also minimax rate optimal with respect to a weighted uniform norm, where the weighting accounts in a natural way for the non-uniformity of the design distribution. This implies that although least squares is a global criterion, the LSE adapts locally to the size of the design density. We develop a new indirect proof technique that establishes the local convergence behavior based on a carefully chosen local perturbation of the LSE. The obtained local rates are then applied to analyze the LSE for transfer learning under covariate shift.

Keywords: Covariate shift; domain adaptation; local rates; mean squared error; minimax estimation; nonparametric least squares; nonparametric regression; transfer learning

References

[1] 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
[2] Baxter, J. (1997). A Bayesian/information theoretic model of learning to learn via multiple task sampling. *Mach. Learn.* **28** 7–39. https://doi.org/10.1023/A:1007327622663
[3] Beliakov, G. (2007). Smoothing Lipschitz functions. *Optim. Methods Softw.* **22** 901–916. MR2360803 https://doi.org/10.1080/10556780701393591
[4] Ben-David, S. and Schuller, R. (2003). Exploiting task relatedness for multiple task learning. In *Learning Theory and Kernel Machines* (B. Schölkopf and M.K. Warmuth, eds.) 567–580. Berlin, Heidelberg: Springer Berlin Heidelberg.
[5] Birgé, L. and Massart, P. (1993). Rates of convergence for minimum contrast estimators. *Probab. Theory Related Fields* **97** 113–150. MR1240719 https://doi.org/10.1007/BF01199316
[6] Brunk, H.D. (1955). Maximum likelihood estimates of monotone parameters. *Ann. Math. Stat.* **26** 607–616. MR0073894 https://doi.org/10.1214/aoms/1177728420
[7] Brunk, H.D. (1958). On the estimation of parameters restricted by inequalities. *Ann. Math. Stat.* **29** 437–454. MR0132632 https://doi.org/10.1214/aoms/1177706621
[8] Buckley, S.M. and MacManus, P. (2000). Singular measures and the key of *G*. *Publ. Mat.* **44** 483–489. MR1800819 https://doi.org/10.5565/PUBLMAT_44200_07
[9] Cai, T.T. and Wei, H. (2021). Transfer learning for nonparametric classification: Minimax rate and adaptive classifier. *Ann. Statist.* **49** 100–128. MR4206671 https://doi.org/10.1214/20-AOS1949
[10] Caruana, R. (1997). Multitask learning. *Mach. Learn.* **28** 41–75. https://doi.org/10.1023/A:1007379606734
[11] Chinot, G., Löffler, M. and van de Geer, S. (2022). On the robustness of minimum norm interpolators and regularized empirical risk minimizers. *Ann. Statist.* **50** 2306–2333. MR4474492 https://doi.org/10.1214/22-aos2190

1350-7265

[12] Dümbgen, L. and Rufibach, K. (2009). Maximum likelihood estimation of a log-concave density and its distribution function: Basic properties and uniform consistency. *Bernoulli* **15** 40–68. MR2546798 https://doi.org/10.3150/08-BEJ141

[13] Gaïffas, S. (2009). Uniform estimation of a signal based on inhomogeneous data. *Statist. Sinica* **19** 427–447. MR2514170

[14] Groeneboom, P., Jongbloed, G. and Wellner, J.A. (2001). Estimation of a convex function: Characterizations and asymptotic theory. *Ann. Statist.* **29** 1653–1698. MR1891742 https://doi.org/10.1214/aos/1015345958

[15] Guntuboyina, A. and Sen, B. (2018). Nonparametric shape-restricted regression. *Statist. Sci.* **33** 568–594. MR3881209 https://doi.org/10.1214/18-STS665

[16] Györfi, L., Kohler, M., Krzyżak, A. and Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. *Springer Series in Statistics*. New York: Springer. MR1920390 https://doi.org/10.1007/b97848

[17] Han, Q. (2021). Set structured global empirical risk minimizers are rate optimal in general dimensions. *Ann. Statist.* **49** 2642–2671. MR4338378 https://doi.org/10.1214/21-aos2049

[18] Han, Q., Wang, T., Chatterjee, S. and Samworth, R.J. (2019). Isotonic regression in general dimensions. *Ann. Statist.* **47** 2440–2471. MR3988762 https://doi.org/10.1214/18-AOS1753

[19] Han, Q. and Wellner, J.A. (2019). Convergence rates of least squares regression estimators with heavy-tailed errors. *Ann. Statist.* **47** 2286–2319. MR3953452 https://doi.org/10.1214/18-AOS1748

[20] Hanson, D.L., Pledger, G. and Wright, F.T. (1973). On consistency in monotonic regression. *Ann. Statist.* **1** 401–421. MR0353540

[21] Kahane, J.-P. (1969). Trois notes sur les ensembles parfaits linéaires. *Enseign. Math. (2)* **15** 185–192. MR0245734

[22] 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

[23] Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *Ann. Statist.* **34** 2593–2656. MR2329442 https://doi.org/10.1214/009053606000001019

[24] Koltchinskii, V. and Mendelson, S. (2015). Bounding the smallest singular value of a random matrix without concentration. *Int. Math. Res. Not. IMRN* **2015** 12991–13008. MR3431642 https://doi.org/10.1093/imrn/rnv096

[25] Kpotufe, S. and Martinet, G. (2021). Marginal singularity and the benefits of labels in covariate-shift. *Ann. Statist.* **49** 3299–3323. MR4352531 https://doi.org/10.1214/21-aos2084

[26] Kuchibhotla, A.K. and Patra, R.K. (2022). On least squares estimation under heteroscedastic and heavy-tailed errors. *Ann. Statist.* **50** 277–302. MR4382017 https://doi.org/10.1214/21-aos2105

[27] Kur, G., Gao, F., Guntuboyina, A. and Sen, B. (2020). Convex regression in multidimensions: suboptimality of least squares estimators. Preprint. Available at arXiv:2006.02044v1.

[28] Lecué, G. and Mendelson, S. (2017). Regularization and the small-ball method II: Complexity dependent error rates. *J. Mach. Learn. Res.* **18** Paper No. 146, 48 pp. MR3763780

[29] Lecué, G. and Mendelson, S. (2018). Regularization and the small-ball method I: Sparse recovery. *Ann. Statist.* **46** 611–641. MR3782379 https://doi.org/10.1214/17-AOS1562

[30] Mammen, E. and van de Geer, S. (1997). Locally adaptive regression splines. *Ann. Statist.* **25** 387–413. MR1429931 https://doi.org/10.1214/aos/1034276635

[31] Mendelson, S. (2014). Learning without concentration. In *Proceedings of the 27th Conference on Learning Theory* (M.F. Balcan, V. Feldman and C. Szepesvári, eds.). *Proceedings of Machine Learning Research* **35** 25–39. Barcelona, Spain: PMLR.

[32] Micchelli, C. and Pontil, M. (2004). Kernels for multi-task learning. In *Advances in Neural Information Processing Systems* (L. Saul, Y. Weiss and L. Bottou, eds.) **17**. Cambridge, MA: MIT Press.

[33] Pathak, R., Ma, C. and Wainwright, M.J. (2022). A new similarity measure for covariate shift with applications to nonparametric regression. Preprint. Available at arXiv:2202.02837.

[34] Patschkowski, T. and Rohde, A. (2016). Adaptation to lowest density regions with application to support recovery. *Ann. Statist.* **44** 255–287. MR3449768 https://doi.org/10.1214/15-AOS1366

[35] Ray, K. and Schmidt-Hieber, J. (2017). A regularity class for the roots of nonnegative functions. *Ann. Mat. Pura Appl. (4)* **196** 2091–2103. MR3714756 https://doi.org/10.1007/s10231-017-0655-2

[36] Reeve, H.W.J., Cannings, T.I. and Samworth, R.J. (2021). Adaptive transfer learning. *Ann. Statist.* **49** 3618–3649. MR4352543 https://doi.org/10.1214/21-aos2102

[37] Saumard, A. (2010). Convergence in sup-norm of least-squares estimators in regression with random design and nonparametric heteroscedastic noise. HAL Id: hal-00528539.

[38] 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

[39] Schmidt-Hieber, J. and Zamolodtchikov, P. (2024). Supplement to "Local convergence rates of the nonparametric least squares estimator with applications to transfer learning." https://doi.org/10.3150/23-BEJ1655SUPP

[40] Shimodaira, H. (2000). Improving predictive inference under covariate shift by weighting the log-likelihood function. *J. Statist. Plann. Inference* **90** 227–244. MR1795598 https://doi.org/10.1016/S0378-3758(00)00115-4

[41] Soloff, J.A., Guntuboyina, A. and Pitman, J. (2019). Distribution-free properties of isotonic regression. *Electron. J. Stat.* **13** 3243–3253. MR4010598 https://doi.org/10.1214/19-ejs1594

[42] Stone, C.J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.* **8** 1348–1360. MR0594650

[43] Sugiyama, M., Nakajima, S., Kashima, H., Buenau, P. and Kawanabe, M. (2007). Direct importance estimation with model selection and its application to covariate shift adaptation. In *Advances in Neural Information Processing Systems* (J. Platt, D. Koller, Y. Singer and S. Roweis, eds.) **20**. Curran Associates, Inc.

[44] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794

[45] van de Geer, S. (1990). Estimating a regression function. *Ann. Statist.* **18** 907–924. MR1056343 https://doi.org/10.1214/aos/1176347632

[46] van de Geer, S.A. (2000). *Applications of Empirical Process Theory*. *Cambridge Series in Statistical and Probabilistic Mathematics* **6**. Cambridge: Cambridge Univ. Press. https://doi.org/1739079

[47] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

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

[49] Wright, F.T. (1981). The asymptotic behavior of monotone regression estimates. *Ann. Statist.* **9** 443–448. MR0606630

Bernoulli **30**(3), 2024, 1878–1893
https://doi.org/10.3150/23-BEJ1656

Representation of random variables as Lebesgue integrals

SARA BIAGINI[1,a] and GORDAN ŽITKOVIĆ[2,b]

[1]*Department of Economics and Finance, LUISS University, Rome, Italy,* [a]*sbiagini@luiss.it*
[2]*Department of Mathematics, University of Texas at Austin, Texas, US,* [b]*gordanz@math.utexas.edu*

We study representations of a random variable ξ as an integral of an adapted process with respect to the Lebesgue measure. The existence of such representations in two different regularity classes is characterized in terms of the quadratic variation of (local) martingales closed by ξ.

Keywords: Absolutely continuous representation; Girsanov theorem; martingale representation; quadratic variation

References

Aïd, R. and Biagini, S. (2023). Optimal dynamic regulation of carbon emissions market. *Math. Finance* **33** 80–115. MR4543017

Biagini, S., Gozzi, F. and Zanella, M. (2022). Robust portfolio choice with sticky wages. *SIAM J. Financial Math.* **13** 1004–1039. MR4468616 https://doi.org/10.1137/21M1429722

Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge: Cambridge Univ. Press. MR1417491 https://doi.org/10.1017/CBO9780511662829

Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*, 2nd ed. *Encyclopedia of Mathematics and Its Applications* **152**. Cambridge: Cambridge Univ. Press. MR3236753 https://doi.org/10.1017/CBO9781107295513

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300** 463–520. MR1304434 https://doi.org/10.1007/BF01450498

Fabbri, G., Gozzi, F. and Świech, A. (2017). *Stochastic Optimal Control in Infinite Dimension*. *Probability Theory and Stochastic Modelling* **82**. Cham: Springer. Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore. MR3674558 https://doi.org/10.1007/978-3-319-53067-3

He, S.W., Wang, J.G. and Yan, J.A. (1992). *Semimartingale Theory and Stochastic Calculus*. Beijing: Kexue Chubanshe (Science Press). MR1219534

Marinucci, D. and Robinson, P.M. (1999). Alternative forms of fractional Brownian motion. *J. Statist. Plann. Inference* **80** 111–122. MR1713794 https://doi.org/10.1016/S0378-3758(98)00245-6

Samko, S.G., Kilbas, A.A. and Marichev, O.I. (1993). *Fractional Integrals and Derivatives*. Yverdon: Gordon and Breach Science Publishers. MR1347689

Veraar, M. (2012). The stochastic Fubini theorem revisited. *Stochastics* **84** 543–551. MR2966093 https://doi.org/10.1080/17442508.2011.618883

Bernoulli **30**(3), 2024, 1894–1920
https://doi.org/10.3150/23-BEJ1657

A large-sample theory for infinitesimal gradient boosting

CLÉMENT DOMBRY[a] and JEAN-JIL DUCHAMPS[b]

Université de Franche-Comté, CNRS, LmB, F-25000 Besançon, France, [a]*clement.dombry@univ-fcomte.fr,* [b]*jean-jil.duchamps@univ-fcomte.fr*

Infinitesimal gradient boosting (Dombry and Duchamps (2021)) is defined as the vanishing-learning-rate limit of the popular tree-based gradient boosting algorithm from machine learning. It is characterized as the solution of a nonlinear ordinary differential equation in an infinite-dimensional function space where the infinitesimal boosting operator driving the dynamics depends on the training sample. We consider the asymptotic behavior of the model in the large sample limit and prove its convergence to a deterministic process. This population limit is again characterized by a differential equation that depends on the population distribution. We explore some properties of this population limit: we prove that the dynamics makes the test error decrease and we consider its long time behavior.

Keywords: Gradient boosting; large sample theory; softmax gradient tree

References

Blanchard, G., Lugosi, G. and Vayatis, N. (2004). On the rate of convergence of regularized boosting classifiers. *J. Mach. Learn. Res.* **4** 861–894. MR2076000 https://doi.org/10.1162/1532443041424319

Breiman, L. (2004). Population theory for boosting ensembles. *Ann. Statist.* **32** 1–11. MR2050998 https://doi.org/10.1214/aos/1079120126

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J. (1984). *Classification and Regression Trees*. *Wadsworth Statistics/Probability Series*. Belmont, CA: Wadsworth Advanced Books and Software. MR0726392

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794. San Francisco California USA: ACM.

Dieuleveut, A., Durmus, A. and Bach, F. (2020). Bridging the gap between constant step size stochastic gradient descent and Markov chains. *Ann. Statist.* **48** 1348–1382. MR4124326 https://doi.org/10.1214/19-AOS1850

Dombry, C. and Duchamps, J.-J. (2021). Infinitesimal gradient boosting. Preprint. Available at https://arxiv.org/abs/2104.13208.

Dombry, C. and Duchamps, J.-J. (2024). Supplement to "A large-sample theory for infinitesimal gradient boosting." https://doi.org/10.3150/23-BEJ1657SUPP

Durrett, R. (2010). *Probability: Theory and Examples*, 4th ed. *Cambridge Series in Statistical and Probabilistic Mathematics* **31**. Cambridge: Cambridge Univ. Press. MR2722836 https://doi.org/10.1017/CBO9780511779398

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Ann. Statist.* **29** 1189–1232. MR1873328 https://doi.org/10.1214/aos/1013203451

Geurts, P., Ernst, D. and Wehenkel, L. (2006). Extremely randomized trees. *Mach. Learn.* **63** 3–42.

Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2722294 https://doi.org/10.1007/978-0-387-84858-7

Jiang, W. (2004). Process consistency for AdaBoost. *Ann. Statist.* **32** 13–29. MR2050999 https://doi.org/10.1214/aos/1079120128

Latz, J. (2021). Analysis of stochastic gradient descent in continuous time. *Stat. Comput.* **31** Paper No. 39. MR4257846 https://doi.org/10.1007/s11222-021-10016-8

1350-7265

Lugosi, G. and Vayatis, N. (2004). On the Bayes-risk consistency of regularized boosting methods. *Ann. Statist.* **32** 30–55. MR2051000 https://doi.org/10.1214/aos/1079120129

Lykov, A., Muzychka, S. and Vaninsky, K. (2015). The AdaBoost flow. *Comm. Pure Appl. Math.* **68** 865–886. MR3333843 https://doi.org/10.1002/cpa.21555

Scornet, E., Biau, G. and Vert, J.-P. (2015). Consistency of random forests. *Ann. Statist.* **43** 1716–1741. MR3357876 https://doi.org/10.1214/15-AOS1321

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

Zhang, T. and Yu, B. (2005). Boosting with early stopping: Convergence and consistency. *Ann. Statist.* **33** 1538–1579. MR2166555 https://doi.org/10.1214/009053605000000255

Bernoulli **30**(3), 2024, 1921–1944
https://doi.org/10.3150/23-BEJ1658

Kernel-weighted specification testing under general distributions

SID KANKANALA[1,a] and VICTORIA ZINDE-WALSH[2,b]

[1]*Department of Economics, Yale University, New Haven, USA,* [a]*sid.kankanala@yale.edu*
[2]*Department of Economics, McGill University, Montreal, Canada,* [b]*victoria.zinde-walsh@mcgill.ca*

Kernel-weighted test statistics have been widely used in a variety of settings including non-stationary regression, survival analysis, propensity score and panel data models. We develop the limit theory for a kernel-weighted specification test of a parametric conditional mean when the law of the regressors may not be absolutely continuous to the Lebesgue measure and admits non-trivial singular components. In the special case of absolutely continuous measures, our approach weakens the usual regularity conditions. This result is of independent interest and may be useful in other applications that utilize kernel smoothed statistics. Simulations illustrate the non-trivial impact of the distribution of the conditioning variables on the power properties of the test statistic.

Keywords: Fractal; goodness-of-fit; kernel smoothing; singular distribution; small ball probability

References

[1] Blundell, R., Chen, X. and Kristensen, D. (2007). Semi-nonparametric IV estimation of shape-invariant Engel curves. *Econometrica* **75** 1613–1669. MR2351452 https://doi.org/10.1111/j.1468-0262.2007.00808.x
[2] Burrough, P.A. (1981). Fractal dimensions of landscapes and other environmental data. *Nature* **294** 240–242.
[3] Chen, S.X. and Van Keilegom, I. (2009). A goodness-of-fit test for parametric and semi-parametric models in multiresponse regression. *Bernoulli* **15** 955–976. MR2597579 https://doi.org/10.3150/09-BEJ208
[4] Davies, S. and Hall, P. (1999). Fractal analysis of surface roughness by using spatial data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 3–37. MR1664088 https://doi.org/10.1111/1467-9868.00160
[5] De Lellis, C. (2008). *Rectifiable Sets, Densities and Tangent Measures*. *Zurich Lectures in Advanced Mathematics*. Zürich: European Mathematical Society (EMS). MR2388959 https://doi.org/10.4171/044
[6] Dette, H. (1999). A consistent test for the functional form of a regression based on a difference of variance estimators. *Ann. Statist.* **27** 1012–1040. MR1724039 https://doi.org/10.1214/aos/1018031266
[7] Gao, J., King, M., Lu, Z. and Tjøstheim, D. (2009). Specification testing in nonlinear and nonstationary time series autoregression. *Ann. Statist.* **37** 3893–3928. MR2572447 https://doi.org/10.1214/09-AOS698
[8] González-Manteiga, W. and Crujeiras, R.M. (2013). An updated review of goodness-of-fit tests for regression models. *TEST* **22** 361–411. MR3093195 https://doi.org/10.1007/s11749-013-0327-5
[9] Hall, P. (1984). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Multivariate Anal.* **14** 1–16. MR0734096 https://doi.org/10.1016/0047-259X(84)90044-7
[10] Hall, P. and Heyde, C.C. (2014). *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0624435
[11] Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21** 1926–1947. MR1245774 https://doi.org/10.1214/aos/1176349403
[12] Härdle, W., Marron, J.S. and Wand, M.P. (1990). Bandwidth choice for density derivatives. *J. Roy. Statist. Soc. Ser. B* **52** 223–232. MR1049312
[13] Horowitz, J.L. and Spokoiny, V.G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica* **69** 599–631. MR1828537 https://doi.org/10.1111/1468-0262.00207
[14] Hsiao, C., Li, Q. and Racine, J.S. (2007). A consistent model specification test with mixed discrete and continuous data. *J. Econometrics* **140** 802–826. MR2408927 https://doi.org/10.1016/j.jeconom.2006.07.015

[15] Hutchinson, J.E. (1981). Fractals and self-similarity. *Indiana Univ. Math. J.* **30** 713–747. MR0625600 https://doi.org/10.1512/iumj.1981.30.30055

[16] Kankanala, S. and Zinde-Walsh, V. (2024). Supplement to "Kernel-weighted specification testing under general distributions." https://doi.org/10.3150/23-BEJ1658SUPP

[17] Koltchinskii, V. and Lounici, K. (2017). Normal approximation and concentration of spectral projectors of sample covariance. *Ann. Statist.* **45** 121–157. MR3611488 https://doi.org/10.1214/16-AOS1437

[18] Kunze, H., La Torre, D., Mendivil, F. and Vrscay, E.R. (2012). *Fractal-Based Methods in Analysis*. New York: Springer. MR3014680 https://doi.org/10.1007/978-1-4614-1891-7

[19] Lin, Z., Li, Q. and Sun, Y. (2014). A consistent nonparametric test of parametric regression functional form in fixed effects panel data models. *J. Econometrics* **178** 167–179. MR3137900 https://doi.org/10.1016/j.jeconom.2013.08.014

[20] Luukkainen, J. and Saksman, E. (1998). Every complete doubling metric space carries a doubling measure. *Proc. Amer. Math. Soc.* **126** 531–534. MR1443161 https://doi.org/10.1090/S0002-9939-98-04201-4

[21] Mackay, J.M. and Tyson, J.T. (2010). *Conformal Dimension: Theory and Application*. *University Lecture Series* **54**. Providence, RI: Amer. Math. Soc. MR2662522 https://doi.org/10.1090/ulect/054

[22] Mammen, E., Van Keilegom, I. and Yu, K. (2019). Expansion for moments of regression quantiles with applications to nonparametric testing. *Bernoulli* **25** 793–827. MR3920357 https://doi.org/10.3150/17-bej986

[23] Marron, J.S. and Wand, M.P. (1992). Exact mean integrated squared error. *Ann. Statist.* **20** 712–736. MR1165589 https://doi.org/10.1214/aos/1176348653

[24] Meilán-Vila, A., Opsomer, J.D., Francisco-Fernández, M. and Crujeiras, R.M. (2020). A goodness-of-fit test for regression models with spatially correlated errors. *TEST* **29** 728–749. MR4140781 https://doi.org/10.1007/s11749-019-00678-y

[25] Moore, E.F. (1950). Density ratios and $(\phi, 1)$ rectifiability in n-space. *Trans. Amer. Math. Soc.* **69** 324–334. MR0037894 https://doi.org/10.2307/1990362

[26] Müller, U.U. and Van Keilegom, I. (2019). Goodness-of-fit tests for the cure rate in a mixture cure model. *Biometrika* **106** 211–227. MR3912392 https://doi.org/10.1093/biomet/asy058

[27] Preiss, D. (1987). Geometry of measures in $\mathbf{R}^n$: Distribution, rectifiability, and densities. *Ann. of Math. (2)* **125** 537–643. MR0890162 https://doi.org/10.2307/1971410

[28] Rudin, W. (1987). *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill. MR0924157

[29] Shaikh, A.M., Simonsen, M., Vytlacil, E.J. and Yildiz, N. (2009). A specification test for the propensity score using its distribution conditional on participation. *J. Econometrics* **151** 33–46. MR2538270 https://doi.org/10.1016/j.jeconom.2009.01.014

[30] van der Vaart, A.W. (2000). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

[31] Volberg, A.L. and Konyagin, S.V. (1987). On measures with the doubling condition. *Math. USSR* **51** 666–675. MR0903629

[32] Wang, Q. and Phillips, P.C.B. (2012). A specification test for nonlinear nonstationary models. *Ann. Statist.* **40** 727–758. MR2933664 https://doi.org/10.1214/12-AOS975

[33] Zheng, J.X. (1996). A consistent test of functional form via nonparametric estimation techniques. *J. Econometrics* **75** 263–289. MR1413644 https://doi.org/10.1016/0304-4076(95)01760-7

[34] Zheng, J.X. (1998). A consistent nonparametric test of parametric regression models under conditional quantile restrictions. *Econometric Theory* **14** 123–138. MR1613710 https://doi.org/10.1017/S0266466698141051

Bernoulli **30**(3), 2024, 1945–1970
https://doi.org/10.3150/23-BEJ1659

Asymptotics of discrete Schrödinger bridges via chaos decomposition

ZAID HARCHAOUI[1,a], LANG LIU[1,b] and SOUMIK PAL[2,c]

[1]*Department of Statistics, University of Washington, Seattle, United States,* [a]*zaid@uw.edu,* [b]*liu16@uw.edu*
[2]*Department of Mathematics, University of Washington, Seattle, United States,* [c]*soumik@uw.edu*

Consider the problem of matching two independent i.i.d. samples of size N from two distributions P and Q in $\mathbb{R}^d$. For an arbitrary continuous cost function, the optimal assignment problem looks for the matching that minimizes the total cost. We consider instead in this paper the problem where each matching is endowed with a Gibbs probability weight proportional to the exponential of the negative total cost of that matching. Viewing each matching as a joint distribution with N atoms, we then take a convex combination with respect to the above Gibbs probability measure. We show that this resulting random joint distribution converges, as $N \to \infty$, to the solution of a variational problem, introduced by Föllmer, called the Schrödinger problem. We also prove a limiting Gaussian fluctuation for this convergence in the form of central limit theorems for integrated test functions. This establishes a novel passage for the transition from discrete to continuum in Schrödinger's lazy gas experiment.

Keywords: Chaos decomposition; contiguity; entropy regularization; Hoeffding decomposition; infinite-order U-statistics; optimal matching; optimal transport; Schrödinger bridge

References

[1] Adams, S., Bru, J.-B. and König, W. (2006). Large deviations for trapped interacting Brownian particles and paths. *Ann. Probab.* **34** 1370–1422. MR2257650 https://doi.org/10.1214/009117906000000214

[2] Adams, S. and Dorlas, T. (2008). Asymptotic Feynman-Kac formulae for large symmetrised systems of random walks. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 837–875. MR2453847 https://doi.org/10.1214/07-AIHP132

[3] Adams, S. and König, W. (2008). Large deviations for many Brownian bridges with symmetrised initial-terminal condition. *Probab. Theory Related Fields* **142** 79–124. MR2413267 https://doi.org/10.1007/s00440-007-0099-5

[4] Ajtai, M., Komlós, J. and Tusnády, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264. MR0779885 https://doi.org/10.1007/BF02579135

[5] Arratia, R., Barbour, A.D. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach*. *EMS Monographs in Mathematics*. Zürich: European Mathematical Society (EMS). MR2032426 https://doi.org/10.4171/000

[6] Barvinok, A. (2002). *A Course in Convexity*. *Graduate Studies in Mathematics* **54**. Providence, RI: Amer. Math. Soc. MR1940576 https://doi.org/10.1090/gsm/054

[7] Beichl, I. and Sullivan, F. (1999). Approximating the permanent via importance sampling with application to the dimer covering problem. *J. Comput. Phys.* **149** 128–147. MR1669817 https://doi.org/10.1006/jcph.1998.6149

[8] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1998). *Efficient and Adaptive Estimation for Semiparametric Models*. New York: Springer. Reprint of the 1993 original. MR1623559

[9] Bigot, J., Cazelles, E. and Papadakis, N. (2019). Central limit theorems for entropy-regularized optimal transport on finite spaces and statistical applications. *Electron. J. Stat.* **13** 5120–5150. MR4041704 https://doi.org/10.1214/19-EJS1637

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

1350-7265

[11] Chen, Y., Georgiou, T.T. and Pavon, M. (2021). Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge. *SIAM Rev.* **63** 249–313. MR4253788 https://doi.org/10.1137/20M1339982

[12] Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **3** 146–158. MR0365798 https://doi.org/10.1214/aop/1176996454

[13] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Proceedings of Advances in Neural Information Processing Systems*.

[14] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999 https://doi.org/10.1214/aop/1022677394

[15] del Barrio, E., Giné, E. and Utzet, F. (2005). Asymptotics for L_2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* **11** 131–189. MR2121458 https://doi.org/10.3150/bj/1110228245

[16] del Barrio, E. and Loubes, J.-M. (2019). Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** 926–951. MR3916938 https://doi.org/10.1214/18-AOP1275

[17] Dobrić, V. and Yukich, J.E. (1995). Asymptotics for transportation cost in high dimensions. *J. Theoret. Probab.* **8** 97–118. MR1308672 https://doi.org/10.1007/BF02213456

[18] Dynkin, E.B. and Mandelbaum, A. (1983). Symmetric statistics, Poisson point processes, and multiple Wiener integrals. *Ann. Statist.* **11** 739–745. MR0707925 https://doi.org/10.1214/aos/1176346241

[19] Ferradans, S., Papadakis, N., Peyré, G. and Aujol, J.-F. (2014). Regularized discrete optimal transport. *SIAM J. Imaging Sci.* **7** 1853–1882. MR3264566 https://doi.org/10.1137/130929886

[20] Feynman, R.P. (1953). Atomic theory of the λ transition in helium. *Phys. Rev.* **91** 1291–1301.

[21] Foata, D. (1981). Some Hermite polynomial identities and their combinatorics. *Adv. in Appl. Math.* **2** 250–259. MR0626861 https://doi.org/10.1016/0196-8858(81)90006-3

[22] Föllmer, H. (1988). Random fields and diffusion processes. In *École D'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*. *Lecture Notes in Math.* **1362** 101–203. Berlin: Springer. MR0983373 https://doi.org/10.1007/BFb0086180

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

[24] Galichon, A. and Salanié, B. (2010). Matching with trade-offs: Revealed preferences over competing characteristics. CEPR Discussion Paper No. DP7858.

[25] Genevay, A., Cuturi, M., Peyré, G. and Bach, F. (2016). Stochastic optimization for large-scale optimal transport. In *Proceedings of Advances in Neural Information Processing Systems*.

[26] Gohberg, I., Goldberg, S. and Kaashoek, M.A. (1990). *Classes of Linear Operators Vol. I*. Basel: Birkhäuser.

[27] González-Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2022). Weak limits of entropy regularized optimal transport; potentials, plans and divergences. arXiv preprint.

[28] Halász, G. and Székely, G.J. (1976). On the elementary symmetric polynomials of independent random variables. *Acta Math. Acad. Sci. Hung.* **28** 397–400. MR0423491 https://doi.org/10.1007/BF01896806

[29] Halmos, P.R. (1946). The theory of unbiased estimation. *Ann. Math. Stat.* **17** 34–43. MR0015746 https://doi.org/10.1214/aoms/1177731020

[30] Harchaoui, Z., Liu, L. and Pal, S. (2024). Supplement to "Asymptotics of discrete Schrödinger bridges via chaos decomposition." https://doi.org/10.3150/23-BEJ1659SUPP

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

[32] Hundrieser, S., Klatt, M., Staudt, T. and Munk, A. (2022). A unifying approach to distributional limits for empirical optimal transport. arXiv preprint.

[33] Kenyon, R., Kráľ, D., Radin, C. and Winkler, P. (2020). Permutations with fixed pattern densities. *Random Structures Algorithms* **56** 220–250. MR4052852 https://doi.org/10.1002/rsa.20882

[34] Klatt, M., Munk, A. and Zemel, Y. (2022). Limit laws for empirical optimal solutions in random linear programs. *Ann. Oper. Res.* **315** 251–278. MR4458612 https://doi.org/10.1007/s10479-022-04698-0

[35] Klatt, M., Tameling, C. and Munk, A. (2020). Empirical regularized optimal transport: Statistical theory and applications. *SIAM J. Math. Data Sci.* **2** 419–443. MR4105566 https://doi.org/10.1137/19M1278788

[36] Kosowsky, J.J. and Yuille, A.L. (1994). The invisible hand algorithm: Solving the assignment problem with statistical physics. *Neural Netw.* **7** 477–490.

[37] Lei, J. (2020). Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces. *Bernoulli* **26** 767–798. MR4036051 https://doi.org/10.3150/19-BEJ1151
[38] Léonard, C. (2012). From the Schrödinger problem to the Monge-Kantorovich problem. *J. Funct. Anal.* **262** 1879–1920. MR2873864 https://doi.org/10.1016/j.jfa.2011.11.026
[39] Léonard, C. (2014). A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin. Dyn. Syst.* **34** 1533–1574. MR3121631 https://doi.org/10.3934/dcds.2014.34.1533
[40] Liu, L. (2022). Statistical Divergences for Learning and Inference: Limit Laws and Non-Asymptotic Bounds. Ph.D. thesis, Univ. Washington.
[41] Luise, G., Rudi, A., Pontil, M. and Ciliberto, C. (2018). Differential properties of Sinkhorn approximation for learning with Wasserstein distance. In *Proceedings of Advances in Neural Information Processing Systems*.
[42] Major, P. (1999). The limit behavior of elementary symmetric polynomials of i.i.d. random variables when their order tends to infinity. *Ann. Probab.* **27** 1980–2010. MR1742897 https://doi.org/10.1214/aop/1022677557
[43] Mallows, C.L. (1957). Non-null ranking models. I. *Biometrika* **44** 114–130. MR0087267 https://doi.org/10.1093/biomet/44.1-2.114
[44] Mena, G. and Niles-Weed, J. (2019). Statistical bounds for entropic optimal transport: Sample complexity and the central limit theorem. In *Proceedings of Advances in Neural Information Processing Systems*.
[45] Mikami, T. (2004). Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. *Probab. Theory Related Fields* **129** 245–260. MR2063377 https://doi.org/10.1007/s00440-004-0340-4
[46] Móri, T.F. and Székely, G.J. (1982). Asymptotic behaviour of symmetric polynomial statistics. *Ann. Probab.* **10** 124–131. MR0637380
[47] Mukherjee, S. (2016). Estimation in exponential families on permutations. *Ann. Statist.* **44** 853–875. MR3476619 https://doi.org/10.1214/15-AOS1389
[48] Munk, A. and Czado, C. (1998). Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 223–241. MR1625620 https://doi.org/10.1111/1467-9868.00121
[49] Pal, S. and Wong, T.-K.L. (2020). Multiplicative Schrödinger problem and the Dirichlet transport. *Probab. Theory Related Fields* **178** 613–654. MR4146546 https://doi.org/10.1007/s00440-020-00987-6
[50] Peyré, G. and Cuturi, M. (2019). *Computational Optimal Transport: With Applications to Data Science*. Found. Trends Mach. Learn.
[51] Ramdas, A., García Trillos, N. and Cuturi, M. (2017). On Wasserstein two-sample testing and related families of nonparametric tests. *Entropy* **19** Paper No. 47. MR3608466 https://doi.org/10.3390/e19020047
[52] Rempała, G.A. and Wesołowski, J. (1999). Limiting behavior of random permanents. *Statist. Probab. Lett.* **45** 149–158. MR1718443 https://doi.org/10.1016/S0167-7152(99)00054-1
[53] Rempała, G.A. and Wesołowski, J. (2005). Approximation theorems for random permanents and associated stochastic processes. *Probab. Theory Related Fields* **131** 442–458. MR2123252 https://doi.org/10.1007/s00440-004-0380-9
[54] Rempała, G.A. and Wesołowski, J. (2008). *Symmetric Functionals on Random Matrices and Random Matchings Problems*. *The IMA Volumes in Mathematics and Its Applications* **147**. New York: Springer. MR2368286
[55] Rigollet, P. and Weed, J. (2018). Entropic optimal transport is maximum-likelihood deconvolution. *C. R. Math. Acad. Sci. Paris* **356** 1228–1235. MR3907589 https://doi.org/10.1016/j.crma.2018.10.010
[56] Rippl, T., Munk, A. and Sturm, A. (2016). Limit laws of the empirical Wasserstein distance: Gaussian distributions. *J. Multivariate Anal.* **151** 90–109. MR3545279 https://doi.org/10.1016/j.jmva.2016.06.005
[57] Rugh, H.H. (2010). Cones and gauges in complex spaces: Spectral gaps and complex Perron-Frobenius theory. *Ann. of Math.* **171** 1707–1752. MR2680397 https://doi.org/10.4007/annals.2010.171.1707
[58] Rüschendorf, L. and Thomsen, W. (1993). Note on the Schrödinger equation and I-projections. *Statist. Probab. Lett.* **17** 369–375. MR1237783 https://doi.org/10.1016/0167-7152(93)90257-J
[59] Schrödinger, E. (1932). Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Ann. Inst. Henri Poincaré* **2** 269–310. MR1508000
[60] Sommerfeld, M. and Munk, A. (2018). Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 219–238. MR3744719 https://doi.org/10.1111/rssb.12236
[61] Talagrand, M. (1992). Matching random samples in many dimensions. *Ann. Appl. Probab.* **2** 846–856. MR1189420

[62] Tameling, C., Sommerfeld, M. and Munk, A. (2019). Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications. *Ann. Appl. Probab.* **29** 2744–2781. MR4019874 https://doi.org/10.1214/19-AAP1463

[63] Trashorras, J. (2008). Large deviations for symmetrised empirical measures. *J. Theoret. Probab.* **21** 397–412. MR2391251 https://doi.org/10.1007/s10959-007-0121-y

[64] van Es, A.J. and Helmers, R. (1988). Elementary symmetric polynomials of increasing order. *Probab. Theory Related Fields* **80** 21–35. MR0970469 https://doi.org/10.1007/BF00348750

[65] van Es, B. (1986). On the weak limits of elementary symmetric polynomials. *Ann. Probab.* **14** 677–695. MR0832030

[66] van der Vaart, A.W. (2000). *Asymptotic Statistics*. Cambridge: Cambridge Univ. Press.

[67] Varadarajan, V.S. (1958). Weak convergence of measures on separable metric spaces. *Sankhyā* **19** 15–22. MR0094838

[68] 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

Bernoulli **30**(3), 2024, 1971–1997
https://doi.org/10.3150/23-BEJ1660

Parametric inference for ergodic McKean-Vlasov stochastic differential equations

VALENTINE GENON-CATALOT[1,a] and CATHERINE LARÉDO[2,b]

[1]*Université Paris Cité, MAP5, UMR 8145 CNRS, F-75006, France,*
[a]*valentine.genon-catalot@mi.parisdescartes.fr*
[2]*Université Paris Cité, LPSM, UMR 8001 CNRS, F-75006, France,* [b]*catherine.laredo@inrae.fr*

We consider a one-dimensional McKean-Vlasov stochastic differential equation with potential and interaction terms depending on unknown parameters. The sample path is continuously observed on a time interval $[0,2T]$. We assume that the process is in the stationary regime. As this distribution is not explicit, the exact likelihood does not lead to computable estimators. To overcome this difficulty, we consider a kernel estimator of the invariant density based on the sample path on $[0,T]$ and obtain new properties for this estimator. Then, we derive an explicit approximate likelihood using the sample path on $[T,2T]$, including the kernel estimator of the invariant density and study the associated estimators of the unknown parameters. We prove their consistency and asymptotic normality with rate $\sqrt{T}$ as T grows to infinity. Several classes of models illustrate the theory.

Keywords: Approximate likelihood; asymptotic properties of estimators; continuous observations; invariant distribution; long time asymptotics; McKean-Vlasov stochastic differential equations; parametric and nonparametric inference

References

Amorino, C. and Gloter, A. (2020). Invariant density adaptive estimation for ergodic jump diffusion processes over anisotropic classes. *J. Statist. Plann. Inference* **213** 106–129.

Amorino, C., Heidari, A., Pilipauskaitė, V. and Podolskij, M. (2023). Parameter estimation of discretely observed interacting particle systems. *Stochastic Process. Appl.* **163** 350–386. MR4612302 https://doi.org/10.1016/j.spa.2023.06.011

Baladron, J., Fasoli, D., Faugeras, O. and Touboul, J. (2012). Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons. *J. Math. Neurosci.* **2** 10. MR2974499 https://doi.org/10.1186/2190-8567-2-10

Ball, F. and Sirl, D. (2019). Stochastic SIR in structured populations. In *Stochastic Epidemic Models with Inference. Part II* (T. Britton and E. Pardoux, eds.). *Lect. Notes Math.* **2255** 123–240. Springer.

Belomestny, D., Pilipauskaitė, V. and Podolskij, M. (2023). Semiparametric estimation of McKean-Vlasov SDEs. *Ann. Inst. Henri Poincaré B Probab. Stat.* **59** 79–96. MR4533721 https://doi.org/10.1214/22-aihp1261

Benachour, S., Roynette, B. and Vallois, P. (1998a). Nonlinear self-stabilizing processes – I Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl.* **75** 173–201.

Benachour, S., Roynette, B. and Vallois, P. (1998b). Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl.* **75** 203–224. MR1632197 https://doi.org/10.1016/S0304-4149(98)00019-2

Benedetto, D., Caglioti, E. and Pulvirenti, M. (1997). A kinetic equation for granular media. *RAIRO Modél. Math. Anal. Numér.* **31** 615–641. MR1471181 https://doi.org/10.1051/m2an/1997310506151

Carrillo, L.A., Choi, Y.-P. and Hauray, M. (2014). The derivation of swarming models: Mean-field limit and Wasserstein distances. Collective dynamics from bacteria to crowds. In *CISM Int. Cent. Mech. Sci.* (A. Muntean and F. Toschi, eds.) **553**. Springer.

Castellana, J.V. and Leadbetter, M.R. (1986). On smoothed probability density estimation for stationary processes. *Stochastic Process. Appl.* **21** 179–193. MR0833950 https://doi.org/10.1016/0304-4149(86)90095-5

Cattiaux, P., Guillin, A. and Malrieu, F. (2008). Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** 19–40. MR2357669 https://doi.org/10.1007/s00440-007-0056-3

Chen, X. (2021). Maximum likelihood estimation of potential energy in interacting particle systems from single-trajectory data. *Electron. Commun. Probab.* **26** 1–13. MR4284626 https://doi.org/10.1214/21-ecp416

Comte, F. (2017). *Estimation Nonparamétrique*, 2nd ed. Paris: Spartacus IDH.

Comte, F. and Genon-Catalot, V. (2023). Nonparametric adaptive estimation for interacting particle systems. *Scand. J. Stat.* On line. https://doi.org/10.1111/sjos.12661

Comte, F. and Merlevède, F. (2005). Super optimal rates for nonparametric density estimation via projection estimators. *Stochastic Process. Appl.* **115** 797–826. MR2132599 https://doi.org/10.1016/j.spa.2004.12.004

Dalalyan, A. and Reiß, M. (2006). Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** 248–282. MR2222384 https://doi.org/10.1007/s00440-004-0416-1

Dalalyan, A. and Reiß, M. (2007). Asymptotic statistical equivalence for ergodic diffusions: The multidimensional case. *Probab. Theory Related Fields* **137** 25–47. MR2278451 https://doi.org/10.1007/s00440-006-0502-7

Dawson, D.A. (1983). Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Stat. Phys.* **31** 29–85. MR0711469 https://doi.org/10.1007/BF01010922

Della Maestra, L. and Hoffmann, M. (2022). Nonparametric estimation for interacting particle systems: McKean-Vlasov models. *Probab. Theory Related Fields* **182** 551–613. MR4367954 https://doi.org/10.1007/s00440-021-01044-6

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. MR4503434 https://doi.org/10.1016/j.spa.2022.10.002

Forien, R. and Pardoux, É. (2022). Household epidemic models and McKean-Vlasov Poisson driven stochastic differential equations. *Ann. Appl. Probab.* **32** 1210–1233. MR4414704 https://doi.org/10.1214/21-aap1706

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000). Stochastic volatility models as hidden Markov models and statistical applications. *Bernoulli* **6** 1051–1079. MR1809735 https://doi.org/10.2307/3318471

Genon-Catalot, V. and Larédo, C. (2021a). Probabilistic properties and parametric inference of small variance nonlinear self-stabilizing stochastic differential equations. *Stochastic Process. Appl.* **142** 513–548. MR4324348 https://doi.org/10.1016/j.spa.2021.09.002

Genon-Catalot, V. and Larédo, C. (2021b). Parametric inference for small variance and long time horizon McKean-Vlasov diffusion models. *Electron. J. Stat.* **5** 5811–5854.

Genon-Catalot, V. and Larédo, C. (2023). Inference for ergodic McKean-Vlasov stochastic differential equations with polynomial interactions. *Ann. Inst. Henri Poincaré B, Probab. Stat.* Preprint HAL 03618498, v3.

Genon-Catalot, V. and Larédo, C. (2024). Supplement to "Parametric inference for ergodic McKean-Vlasov stochastic differential equations." https://doi.org/10.3150/23-BEJ1660SUPP

Giesecke, K., Schwenkler, G. and Sirignano, J.A. (2020). Inference for large financial systems. *Math. Finance* **30** 3–46. MR4067069 https://doi.org/10.1111/mafi.12222

Herrmann, S., Imkeller, P. and Peithmann, D. (2008). Large deviations and a Kramers' type law for self-stabilizing diffusions. *Ann. Appl. Probab.* **18** 1379–1423. MR2434175 https://doi.org/10.1214/07-AAP489

Herrmann, S. and Tugaut, J. (2010). Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Process. Appl.* **120** 1215–1246. MR2639745 https://doi.org/10.1016/j.spa.2010.03.009

Hoffmann, M. (1999). Adaptive estimation in diffusion processes. *Stochastic Process. Appl.* **79** 135–163. MR1670522 https://doi.org/10.1016/S0304-4149(98)00074-X

Höpfner, R. (2014). *Asymptotic Statistics with a View to Stochastic Processes*. *De Gruyter Graduate*. Berlin: de Gruyter. MR3185373 https://doi.org/10.1515/9783110250282

Iacus, S.M. (2010). *Simulation and Inference for Stochastic Differential Equations. With R Examples*. New York: Springer.

Kasonga, R.A. (1990). Maximum likelihood theory for large interacting systems. *SIAM J. Appl. Math.* **50** 865–875. MR1050917 https://doi.org/10.1137/0150050

Kessler, M., Lindner, A. and Sørensen, M. (2012). *Statistical Methods for Stochastic Differential Equations*. New York: CRC Press/CRC.

Kolokoltsov, V.N. (2010). *Nonlinear Markov Processes and Kinetic Equations*. *Cambridge Tracts in Mathematics* **182**. Cambridge: Cambridge Univ. Press. MR2680971 https://doi.org/10.1017/CBO9780511760303

Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. *Springer Series in Statistics*. London: Springer London, Ltd. MR2144185 https://doi.org/10.1007/978-1-4471-3866-2

Leblanc, F. (1997). Density estimation for a class of continuous time processes. *Math. Methods Statist.* **6** 171–199. MR1466626

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. MR4462406 https://doi.org/10.1007/s10208-021-09521-z

Malrieu, F. (2003). Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **3** 540–560.

Marie, N. and Rosier, A. (2023). Nadaraya-Watson estimator for I.I.D. paths of diffusion processes. *Scand. J. Stat.* **50** 589–637. MR4599926

Masuda, H. (2019). Non-Gaussian quasi-likelihood estimation of SDE driven by locally stable Lévy process. *Stochastic Process. Appl.* **129** 1013–1059. MR3913278 https://doi.org/10.1016/j.spa.2018.04.004

McKean, H.P. Jr. (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

Méléard, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995)*. *Lecture Notes in Math.* **1627** 42–95. Berlin: Springer. MR1431299 https://doi.org/10.1007/BFb0093177

Mogilner, A. and Edelstein-Keshet, L. (1999). A non-local model for a swarm. *J. Math. Biol.* **38** 534–570. MR1698215 https://doi.org/10.1007/s002850050158

Nickl, R. and Ray, K. (2020). Nonparametric statistical inference for drift vector fields of multi-dimensional diffusions. *Ann. Statist.* **48** 1383–1408. MR4124327 https://doi.org/10.1214/19-AOS1851

Pavliotis, G.A. and Zanoni, A. (2022a). Eigenfunction martingale estimators for interacting particle systems and their mean field limit. *SIAM J. Appl. Dyn. Syst.* **21** 2338–2370. MR4513316 https://doi.org/10.1137/21M1464348

Pavliotis, G.A. and Zanoni, A. (2022b). A method of moments for interacting particle systems and their mean-field limit. Preprint, arXiv:2212.00403v1.

Rogers, L.C.G. and Willians, D. (2000). *Diffusions, Markov Processes and Martingales, Volume 2, Itô Calculus*, 2nd ed. Cambridge: Cambridge Univ. Press.

Sharrock, L., Kantas, N., Parpas, P. and Pavliotis, G.A. (2023). Online parameter estimation for the McKean-Vlasov stochastic differential equation. *Stochastic Process. Appl.* **162** 481–546. MR4597535 https://doi.org/10.1016/j.spa.2023.05.002

Strauch, C. (2018). Adaptive invariant density estimation for ergodic diffusions over anisotropic classes. *Ann. Statist.* **46** 3451–3480. MR3852658 https://doi.org/10.1214/17-AOS1664

Sznitman, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX—1989*. *Lecture Notes in Math.* **1464** 165–251. Berlin: Springer. MR1108185 https://doi.org/10.1007/BFb0085169

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794

Bernoulli **30**(3), 2024, 1998–2025
https://doi.org/10.3150/23-BEJ1661

M-estimation for varying coefficient models with a functional response in a reproducing kernel Hilbert space

YAFEI WANG[1,a], BEI JIANG[1,b], LINGLONG KONG[1,c] and ZHONGZHAN ZHANG[2,d]

[1]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Canada,* [a]*yafei2@ualberta.ca,* [b]*bei1@ualberta.ca,* [c]*lkong@ualberta.ca*
[2]*Department of Statistics and Data Science, Beijing University of Technology, Beijing, China,* [d]*zzhang@bjut.edu.cn*

Modern neuroimaging research calls for statistical methods that can model dynamic relationships between a functional response and a set of covariates. Current methods, however, remain disparate and limited in their ability to robustly accommodate real-world data and integrate smoothness penalties. In this work, we propose an M-estimation framework for the varying-coefficient model with a functional response that encompasses both mean and quantile regression. To accommodate smoothness regularization and circumvent the stringent conditions on Fourier coefficients or the covariance operator's eigenvalues imposed by traditional fixed-basis representations, we assume that the functional coefficient resides in a reproducing kernel Hilbert space. We show that our proposed estimator is minimax rate optimal and establish convergence properties of our modified alternating direction method of multipliers algorithm. We further propose combining a weighted M-estimator and a copula model to quantify within-subject spatial dependence to improve estimation accuracy. Simulation studies and a real-world analysis demonstrate the robustness of our proposed methods to outliers.

Keywords: Alternating direction method of multipliers; copula model; functional response; M-estimator; minimax; reproducing kernel Hilbert space; varying coefficient model

References

Apostolova, L.G., Mosconi, L., Thompson, P.M., Green, A.E., Hwang, K.S., Ramirez, A., Mistur, R., Tsui, W.H. and de Leon, M.J. (2010). Subregional hippocampal atrophy predicts Alzheimer's dementia in the cognitively normal. *Neurobiol. Aging* **31** 1077–1088. https://doi.org/10.1016/j.neurobiolaging.2008.08.008

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. (2011). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3** 1–122.

Cai, T.T. and Hall, P. (2006). Prediction in functional linear regression. *Ann. Statist.* **34** 2159–2179. MR2291496 https://doi.org/10.1214/009053606000000830

Cai, X., Xue, L. and Cao, J. (2022). Robust estimation and variable selection for function-on-scalar regression. *Canad. J. Statist.* **50** 162–179. MR4389175 https://doi.org/10.1002/cjs.11661

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. MR2906870 https://doi.org/10.1214/11-AOS898

Cai, T.T. and Yuan, M. (2012). Minimax and adaptive prediction for functional linear regression. *J. Amer. Statist. Assoc.* **107** 1201–1216. MR3010906 https://doi.org/10.1080/01621459.2012.716337

Chen, X.R. and Wu, Y.H. (1988). Strong consistency of M-estimates in linear models. *J. Multivariate Anal.* **27** 116–130. MR0971177 https://doi.org/10.1016/0047-259X(88)90120-0

Crum, R.M., Anthony, J.C., Bassett, S.S. and Folstein, M.F. (1993). Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA* **269** 2386–2391.

Cuevas, A. (2014). A partial overview of the theory of statistics with functional data. *J. Statist. Plann. Inference* **147** 1–23. MR3151843 https://doi.org/10.1016/j.jspi.2013.04.002

DeVore, R.A. and Lorentz, G.G. (1993). *Constructive Approximation*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **303**. Berlin: Springer. MR1261635

Drineas, P. and Mahoney, M.W. (2005). On the Nyström method for approximating a Gram matrix for improved kernel-based learning. *J. Mach. Learn. Res.* **6** 2153–2175. MR2249884

Gu, C. (2013). *Smoothing Spline ANOVA Models*, 2nd ed. *Springer Series in Statistics* **297**. New York: Springer. MR3025869 https://doi.org/10.1007/978-1-4614-5369-7

Hobæk Haff, I., Aas, K. and Frigessi, A. (2010). On the simplified pair-copula construction—Simply useful or too simplistic? *J. Multivariate Anal.* **101** 1296–1310. MR2595309 https://doi.org/10.1016/j.jmva.2009.12.001

Hofert, M., Kojadinovic, I., Maechler, M. and Yan, J. (2022). copula: Multivariate Dependence with Copulas. R package version 1.1-0.

Hoover, D.R., Rice, J.A., Wu, C.O. and Yang, L.-P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika* **85** 809–822. MR1666699 https://doi.org/10.1093/biomet/85.4.809

Huang, J.Z., Wu, C.O. and Zhou, L. (2004). Polynomial spline estimation and inference for varying coefficient models with longitudinal data. *Statist. Sinica* **14** 763–788. MR2087972

Huber, P.J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* **1** 799–821. MR0356373

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In *Distributions with Fixed Marginals and Related Topics (Seattle, WA, 1993)*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **28** 120–141. Hayward, CA: IMS. MR1485527 https://doi.org/10.1214/lnms/1215452614

Koenker, R. and Bassett, G. Jr. (1978). Regression quantiles. *Econometrica* **46** 33–50. MR0474644 https://doi.org/10.2307/1913643

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces: Isoperimetry and Processes*. *Classics in Mathematics*. Berlin: Springer. MR2814399

Li, Y., Liu, Y. and Zhu, J. (2007). Quantile regression in reproducing kernel Hilbert spaces. *J. Amer. Statist. Assoc.* **102** 255–268. MR2293307 https://doi.org/10.1198/016214506000000979

Lian, H. (2015). Minimax prediction for functional linear regression with functional responses in reproducing kernel Hilbert spaces. *J. Multivariate Anal.* **140** 395–402. MR3372576 https://doi.org/10.1016/j.jmva.2015.06.005

Liang, H., Wu, H. and Carroll, R.J. (2003). The relationship between virologic and immunologic responses in AIDS clinical research using mixed-effects varying-coefficient models with measurement error. *Biostatistics* **4** 297–312.

Massart, P. (2000). About the constants in Talagrand's concentration inequalities for empirical processes. *Ann. Probab.* **28** 863–884. MR1782276 https://doi.org/10.1214/aop/1019160263

Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Avedissian, C., Madsen, S.K., Parikshak, N., Hua, X., Toga, A.W., Jack Jr, C.R., Schuff, N., Weiner, M.W., Thompson, P.M. and Alzheimer's Disease Neuroimaging Initiative (2009). Automated 3D mapping of hippocampal atrophy and its clinical correlates in 400 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls. *Hum. Brain Mapp.* **30** 2766–2788.

Morris, J.S. (2015). Functional regression. *Annu. Rev. Stat. Appl.* **2** 321–359.

Nosedal-Sanchez, A., Storlie, C.B., Lee, T.C.M. and Christensen, R. (2012). Reproducing kernel Hilbert spaces for penalized regression: A tutorial. *Amer. Statist.* **66** 50–60. MR2934741 https://doi.org/10.1080/00031305.2012.678196

Ramsay, J.O. and Silverman, B.W. (2005). *Functional Data Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2168993

Riesz, F. and Nagy, B.S. (2012). *Functional Analysis*. North Chelmsford: Courier Corporation.

Schumaker, L.L. (2007). *Spline Functions: Basic Theory*, 3rd ed. *Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. MR2348176 https://doi.org/10.1017/CBO9780511618994

Sklar, A. (1996). Random variables, distribution functions, and copulas—A personal look backward and forward. In *Distributions with Fixed Marginals and Related Topics (Seattle, WA, 1993)*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **28** 1–14. Hayward, CA: IMS. MR1485519 https://doi.org/10.1214/lnms/1215452606

Smith, M., Min, A., Almeida, C. and Czado, C. (2010). Modeling longitudinal data using a pair-copula decomposition of serial dependence. *J. Amer. Statist. Assoc.* **105** 1467–1479. MR2796564 https://doi.org/10.1198/jasa.2010.tm09572

Tang, Q. and Cheng, L. (2008). M-estimation and B-spline approximation for varying coefficient models with longitudinal data. *J. Nonparametr. Stat.* **20** 611–625.

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794

Wang, J.-L., Chiou, J.-M. and Müller, H.-G. (2016). Functional data analysis. *Annu. Rev. Stat. Appl.* **3** 257–295.

Wang, Y., Chan, T.F., Toga, A.W. and Thompson, P.M. (2009). Multivariate tensor-based brain anatomical surface morphometry via holomorphic one-forms. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 337–344. Berlin: Springer.

Wu, C.O., Chiang, C.-T. and Hoover, D.R. (1998). Asymptotic confidence regions for kernel smoothing of a varying-coefficient model with longitudinal data. *J. Amer. Statist. Assoc.* **93** 1388–1402. MR1666635 https://doi.org/10.2307/2670054

Yu, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam* 423–435. New York: Springer. MR1462963

Yuan, M. (2006). GACV for quantile smoothing splines. *Comput. Statist. Data Anal.* **50** 813–829. MR2207010 https://doi.org/10.1016/j.csda.2004.10.008

Yuan, M. and Cai, T.T. (2010). A reproducing kernel Hilbert space approach to functional linear regression. *Ann. Statist.* **38** 3412–3444. MR2766857 https://doi.org/10.1214/09-AOS772

Zhang, Z., Wang, X., Kong, L. and Zhu, H. (2022). High-dimensional spatial quantile function-on-scalar regression. *J. Amer. Statist. Assoc.* **117** 1563–1578. MR4480732 https://doi.org/10.1080/01621459.2020.1870984

Zhu, H., Li, R. and Kong, L. (2012). Multivariate varying coefficient model for functional responses. *Ann. Statist.* **40** 2634–2666. MR3097615 https://doi.org/10.1214/12-AOS1045

Bernoulli **30**(3), 2024, 2026–2051
https://doi.org/10.3150/23-BEJ1663

Towards standard imsets for maximal ancestral graphs

ZHONGYI HU[a] and ROBIN J. EVANS[b]

Department of Statistics, University of Oxford, Oxford, UK, [a]*zhongyi.hu@keble.ox.ac.uk,* [b]*evans@stats.ox.ac.uk*

The imsets of Studený (*Probabilistic Conditional Independence Structures* (2005) Springer) are an algebraic method for representing conditional independence models. They have many attractive properties when applied to such models, and they are particularly nice for working with directed acyclic graph (DAG) models. In particular, the 'standard' imset for a DAG is in one-to-one correspondence with the independences it induces, and hence is a label for its Markov equivalence class. We first present a proposed extension to standard imsets for maximal ancestral graph (MAG) models, using the parameterizing set representation of Hu and Evans (In *Proc. 36th Conf. Uncertainty in Artificial Intelligence* (2020) PMLR). We show that for many such graphs our proposed imset is *perfectly Markovian* with respect to the graph, including a class of graphs we refer to as *simple* MAGs, which includes DAGs as a special case. In these cases the imset provides a scoring criteria by measuring the discrepancy for a list of independences that define the model; this gives an alternative to the usual BIC score that is also consistent, and much easier to compute. We also show that, of independence models that do represent the MAG, the imset we give is minimal. Unfortunately, for some graphs the representation does not represent all the independences in the model, and in certain cases does not represent any at all. For these general MAGs, we refine the reduced ordered local Markov property (Richardson in (*Scand. J. Stat.* **30** (2003) 145–157)) by a novel graphical tool called *power DAGs*, and this results in an imset that induces the correct model and which, under a mild condition, can be constructed in polynomial time.

Keywords: Characteristic imset; Graphical models; maximal ancestral graphs; ordered local Markov property; standard imset

References

Andrews, B.J. (2022). Inducing Sets: A New Perspective for Ancestral Graph Markov Models. PhD thesis, Univ. Pittsburgh. MR4435328

Andrews, B.J., Cooper, G.F., Richardson, T.S. and Spirtes, P. (2022). The *m*-connecting imset and factorization for ADMG models. ArXiv preprint. Available at arXiv:2207.08963.

Chen, R., Dash, S. and Gao, T. (2021). Integer programming for causal structure learning in the presence of latent variables. In *Proc. 38th Int. Conf. Machine Learning* 1550–1560. PMLR.

Chickering, D.M. (2002). Optimal structure identification with greedy search. *J. Mach. Learn. Res.* **3** 507–554.

Claassen, T. and Bucur, I.G. (2022). Greedy equivalence search in the presence of latent confounders. In *Proc. 38th Conf. Uncertainty in Artificial Intelligence* PMLR.

Drton, M., Eichler, M. and Richardson, T.S. (2009). Computing maximum likelihood estimates in recursive linear models with correlated errors. *J. Mach. Learn. Res.* **10** 2329–2348. MR2563984

Evans, R.J. (2020). Model selection and local geometry. *Ann. Statist.* **48** 3513–3544. MR4185818 https://doi.org/10.1214/19-AOS1940

Evans, R.J. and Richardson, T.S. (2010). Maximum likelihood fitting of acyclic directed mixed graphs to binary data. In *Proc. 26th Conf. Uncertainty in Artificial Intelligence* PMLR.

Evans, R.J. and Richardson, T.S. (2013). Marginal log-linear parameters for graphical Markov models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 743–768. MR3091657 https://doi.org/10.1111/rssb.12020

Evans, R.J. and Richardson, T.S. (2014). Markovian acyclic directed mixed graphs for discrete data. *Ann. Statist.* **42** 1452–1482. MR3262457 https://doi.org/10.1214/14-AOS1206

Hemmecke, R., Morton, J., Shiu, A., Sturmfels, B. and Wienand, O. (2008). Three counter-examples on semi-graphoids. *Combin. Probab. Comput.* **17** 239–257. MR2396350 https://doi.org/10.1017/S0963548307008838

Hu, Z. and Evans, R. (2020). Faster algorithms for Markov equivalence. In *Proc. 36th Conf. Uncertainty in Artificial Intelligence* PMLR.

Hu, Z. and Evans, R.J. (2024). Supplement to "Towards standard imsets for maximal ancestral graphs." https://doi.org/10.3150/23-BEJ1663SUPP

Jaakkola, T., Sontag, D., Globerson, A. and Meila, M. (2010). Learning Bayesian network structure using LP relaxations. In *Proc. 13th Int. Conf. Artificial Intelligence and Statistics* 358–365. JMLR Workshop and Conference Proceedings.

Kashimura, T. and Takemura, A. (2015). Standard imsets for undirected and chain graphical models. *Bernoulli* **21** 1467–1493. MR3352051 https://doi.org/10.3150/14-BEJ611

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

Meek, C. (1997). Graphical Models: Selecting causal and statistical models Ph.D. thesis Carnegie Mellon Univ.

Ogarrio, J.M., Spirtes, P. and Ramsey, J. (2016). A hybrid causal search algorithm for latent variable models. In *Proc. 8th Int. Conf. Probabilistic Graph. Models* 368–379. PMLR.

Ramsey, J., Spirtes, P. and Zhang, J. (2006). Adjacency-faithfulness and conservative causal inference. In *Proc. 22nd Conf. Uncertainty in Artificial Intelligence* PMLR.

Rantanen, K., Hyttinen, A. and Järvisalo, M. (2021). Maximal ancestral graph structure learning via exact search. In *Proc. 37th Conf. Uncertainty in Artificial Intelligence* PMLR.

Richardson, T. (2003). Markov properties for acyclic directed mixed graphs. *Scand. J. Stat.* **30** 145–157. MR1963898 https://doi.org/10.1111/1467-9469.00323

Richardson, T.S. (2009). A factorization criterion for acyclic directed mixed graphs. In *Proc. 25th Conf. Uncertainty in Artificial Intelligence*.

Richardson, T. and Spirtes, P. (2002). Ancestral graph Markov models. *Ann. Statist.* **30** 962–1030. MR1926166 https://doi.org/10.1214/aos/1031689015

Richardson, T.S., Evans, R.J., Robins, J.M. and Shpitser, I. (2023). Nested Markov properties for acyclic directed mixed graphs. *Ann. Statist.* **51** 334–361. MR4564859 https://doi.org/10.1214/22-aos2253

Sadeghi, K. and Lauritzen, S. (2014). Markov properties for mixed graphs. *Bernoulli* **20** 676–696. MR3178514 https://doi.org/10.3150/12-BEJ502

Spirtes, P., Glymour, C. and Scheines, R. (2000). *Causation, Prediction, and Search*, 2nd ed. *Adaptive Computation and Machine Learning*. Cambridge, MA: MIT Press. MR1815675

Studený, M. (2005). *Probabilistic Conditional Independence Structures*. *Information Science and Statistics*. London: Springer. MR3183760

Studený, M., Hemmecke, R. and Lindner, S. (2010). Characteristic imset: A simple algebraic representative of a Bayesian network structure. In *Proc. 5th Eur. WS. Probabilistic Graphical Models* 257–264. HIIT Publications.

Triantafillou, S. and Tsamardinos, I. (2016). Score-based vs constraint-based causal learning in the presence of confounders. In *Proc. 32rd Conf. Uncertainty in Artificial Intelligence* PMLR.

Bernoulli **30**(3), 2024, 2052–2073
https://doi.org/10.3150/23-BEJ1664

Large deviations of reflected weakly interacting particle systems

PING CHEN[1,a], RONG WEI[2,c] and TUSHENG ZHANG[1,b]

[1] *School of Mathematical Sciences, University of Science and Technology of China, Hefei, 230026, China,* [a]*chenping@mail.ustc.edu.cn,* [b]*tushengz@ustc.edu.cn*
[2] *School of Mathematics and Statistics, Anhui Normal University, Wuhu 241002, China,* [c]*rongw@ahnu.edu.cn*

In this paper, we prove a large deviation principle for the empirical measures of a system of weakly interacting diffusions with reflection. We adopt the weak convergence approach. To make this approach work, we show that the sequence of empirical measures of the controlled reflected system will converge to the weak solution of an associated reflected McKean–Vlasov equation.

Keywords: Interacting particle systems; large deviation; McKean–Vlasov equation; stochastic differential equation with reflection; sub-martingale problem; weak convergence

References

Adams, D., dos Reis, G., Ravaille, R., Salkeld, W. and Tugaut, J. (2022). Large deviations and exit-times for reflected McKean-Vlasov equations with self-stabilising terms and superlinear drifts. *Stochastic Process. Appl.* **146** 264–310. MR4374937 https://doi.org/10.1016/j.spa.2021.12.017

Arous, G.B. and Guionnet, A. (1995). Large deviations for Langevin spin glass dynamics. *Probab. Theory Related Fields* **102** 455–509. MR1346262 https://doi.org/10.1007/BF01198846

Budhiraja, A. and Conroy, M. (2022). Empirical measure and small noise asymptotics under large deviation scaling for interacting diffusions. *J. Theoret. Probab.* **35** 295–349. MR4379467 https://doi.org/10.1007/s10959-020-01071-4

Budhiraja, A. and Dupuis, P. (2000). A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Statist.* **20** 39–61. MR1785237

Budhiraja, A., Dupuis, P. and Fischer, M. (2012). Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* **40** 74–102. MR2917767 https://doi.org/10.1214/10-AOP616

Coghi, M., Dreyer, W., Friz, P.K., Gajewski, P., Guhlke, C. and Maurelli, M. (2022). A McKean-Vlasov SDE and particle system with interaction from reflecting boundaries. *SIAM J. Math. Anal.* **54** 2251–2294. MR4409227 https://doi.org/10.1137/21M1409421

Dawson, D.A. and Del Moral, P. (2005). Large deviations for interacting processes in the strong topology. In *Statistical Modeling and Analysis for Complex Data Problems*. *GERAD 25th Anniv. Ser.* **1** 179–208. New York: Springer. MR2189537 https://doi.org/10.1007/0-387-24555-3_10

Dawson, D.A. and Gärtner, J. (1987). Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. MR0885876 https://doi.org/10.1080/17442508708833446

Dawson, D.A. and Gärtner, J. (1994). Multilevel large deviations and interacting diffusions. *Probab. Theory Related Fields* **98** 423–487. MR1271106 https://doi.org/10.1007/BF01192835

Del Moral, P. and Guionnet, A. (1998). Large deviations for interacting particle systems: Applications to non-linear filtering. *Stochastic Process. Appl.* **78** 69–95. MR1653296 https://doi.org/10.1016/S0304-4149(98)00057-X

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

Djehiche, B. and Schied, A. (1998). Large deviations for hierarchical systems of interacting jump processes. *J. Theoret. Probab.* **11** 1–24. MR1607396 https://doi.org/10.1023/A:1021690707556

Dupuis, P. and Ellis, R.S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1431744 https://doi.org/10.1002/9781118165904

Hoeksema, J., Holding, T., Maurell, M. and Tse, O. (2020). Large deviations for singular interacting diffusions. Available at arXiv:2002.01295.

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

Léonard, C. (1995). Large deviations for long range interacting particle systems with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 289–323. MR1324810

Lions, P.-L. and Sznitman, A.-S. (1984). Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37** 511–537. MR0745330 https://doi.org/10.1002/cpa.3160370408

Stroock, D.W. and Varadhan, S.R.S. (1971). Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.* **24** 147–225. MR0277037 https://doi.org/10.1002/cpa.3160240206

Sznitman, A.-S. (1984). Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.* **56** 311–336. MR0743844 https://doi.org/10.1016/0022-1236(84)90080-6

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

Wei, R., Yang, S.S. and Zhai, J.L. (2022). McKean–Vlasov stochastic differential equations with oblique reflection on non-smooth time dependent domains. Available at arXiv:2208.10702.

Bernoulli **30**(3), 2024, 2074–2101
https://doi.org/10.3150/23-BEJ1665

Learning to reflect: A unifying approach for data-driven stochastic control strategies

SÖREN CHRISTENSEN[1,a], CLAUDIA STRAUCH[2,b] and LUKAS TROTTNER[3,c]

[1]*Department of Mathematics, Kiel University, Kiel, Germany,* [a]*christensen@math.uni-kiel.de*
[2]*Department of Mathematics, Aarhus University, Aarhus, Denmark,* [b]*strauch@math.au.dk*
[3]*Institute of Mathematics, University of Mannheim, Mannheim, Germany,* [c]*trottner@math.au.dk*

Stochastic optimal control problems have a long tradition in applied probability, with the questions addressed being of high relevance in a multitude of fields. Even though theoretical solutions are well understood in many scenarios, their practicability suffers from the assumption of known dynamics of the underlying stochastic process, raising the statistical challenge of developing purely data-driven strategies. For the mathematically separated classes of continuous diffusion processes and Lévy processes, we show that developing efficient strategies for related singular stochastic control problems can essentially be reduced to finding rate-optimal estimators with respect to the sup-norm risk of objects associated to the invariant distribution of ergodic processes which determine the theoretical solution of the control problem. From a statistical perspective, we exploit the exponential β-mixing property as the common factor of both scenarios to drive the convergence analysis, indicating that relying on general stability properties of Markov processes is a sufficiently powerful and flexible approach to treat complex applications requiring statistical methods. We show moreover that in the Lévy case—even though per se jump processes are more difficult to handle both in statistics and control theory—a fully data-driven strategy with regret of significantly better order than in the diffusion case can be constructed utilizing spatial ergodicity of a path-time transformation of the Lévy process in form of its overshoots.

Keywords: Data-driven singular control; exploration vs. exploitation; Lévy processes; overshoots; diffusion processes; reinforcement learning; nonparametric statistics; sup-norm risk

References

[1] Aeckerle-Willems, C. and Strauch, C. (2021). Concentration of scalar ergodic diffusions and some statistical implications. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 1857–1887. MR4328556 https://doi.org/10.1214/20-aihp1144

[2] Alvarez, L.H. (2018). A Class of Solvable Stationary Singular Stochastic Control Problems. Available at arXiv:1803.03464.

[3] Alvarez, L.H.R. and Shepp, L.A. (1998). Optimal harvesting of stochastically fluctuating populations. *J. Math. Biol.* **37** 155–177. MR1649508 https://doi.org/10.1007/s002850050124

[4] Alvarez, L.H.R. and Virtanen, J. (2006). A class of solvable stochastic dividend optimization problems: On the general impact of flexibility on valuation. *Econom. Theory* **28** 373–398. MR2217337 https://doi.org/10.1007/s00199-005-0627-4

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

[6] Asmussen, S. and Taksar, M. (1997). Controlled diffusion models for optimal dividend pay-out. *Insurance Math. Econom.* **20** 1–15. MR1466852 https://doi.org/10.1016/S0167-6687(96)00017-0

[7] Bertoin, J. (1996). *Lévy Processes*. *Cambridge Tracts in Mathematics* **121**. Cambridge: Cambridge Univ. Press. MR1406564

[8] Cadenillas, A., Sarkar, S. and Zapatero, F. (2007). Optimal dividend policy with mean-reverting cash reservoir. *Math. Finance* **17** 81–109. MR2281793 https://doi.org/10.1111/j.1467-9965.2007.00295.x

[9] Chow, Y.S. and Teicher, H. (1997). *Probability Theory*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. Independence, interchangeability, martingales. MR1476912 https://doi.org/10.1007/978-1-4612-1950-7

[10] Christensen, S. and Sohr, T. (2020). A solution technique for Lévy driven long term average impulse control problems. *Stochastic Process. Appl.* **130** 7303–7337. MR4167207 https://doi.org/10.1016/j.spa.2020.07.016

[11] Christensen, S. and Strauch, C. (2023). Nonparametric learning for impulse control problems—exploration vs. exploitation. *Ann. Appl. Probab.* **33** 1369–1387. MR4564434 https://doi.org/10.1214/22-aap1849

[12] Christensen, S., Strauch, C. and Trottner, L. (2024). Supplement to "Learning to reflect: A unifying approach for data-driven stochastic control strategies." https://doi.org/10.3150/23-BEJ1665SUPP

[13] Dalalyan, A. and Reiß, M. (2006). Asymptotic statistical equivalence for scalar ergodic diffusions. *Probab. Theory Related Fields* **134** 248–282. MR2222384 https://doi.org/10.1007/s00440-004-0416-1

[14] Davydov, J.A. (1973). Mixing conditions for Markov chains. *Teor. Veroyatn. Primen.* **18** 321–338. MR0321183

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

[16] Dexheimer, N., Strauch, C. and Trottner, L. (2022). Adaptive invariant density estimation for continuous-time mixing Markov processes under sup-norm risk. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** 2029–2064. MR4492970 https://doi.org/10.1214/21-aihp1235

[17] Döring, L. and Trottner, L. (to appear). Stability of overshoots of Markov additive processes. *Ann. Appl. Probab.*

[18] Gīhman, Ĭ.Ī. and Skorohod, A.V. (1972). *Stochastic Differential Equations*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete [Results in Mathematics and Related Areas], Band 72*. New York: Springer. Translated from the Russian by Kenneth Wickwire. MR0346904

[19] Grimmett, G.R. and Stirzaker, D.R. (2001). *Probability and Random Processes*, 3rd ed. New York: Oxford Univ. Press. MR2059709

[20] Hening, A., Nguyen, D.H., Ungureanu, S.C. and Wong, T.K. (2019). Asymptotic harvesting of populations in random environments. *J. Math. Biol.* **78** 293–329. MR3932647 https://doi.org/10.1007/s00285-018-1275-1

[21] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. *Springer Series in Statistics*. London: Springer. MR2144185 https://doi.org/10.1007/978-1-4471-3866-2

[22] Kuznetsov, A., Kyprianou, A.E. and Pardo, J.C. (2012). Meromorphic Lévy processes and their fluctuation identities. *Ann. Appl. Probab.* **22** 1101–1135. MR2977987 https://doi.org/10.1214/11-AAP787

[23] Kuznetsov, A. and Pardo, J.C. (2013). Fluctuations of stable processes and exponential functionals of hypergeometric Lévy processes. *Acta Appl. Math.* **123** 113–139. MR3010227 https://doi.org/10.1007/s10440-012-9718-y

[24] Kyprianou, A.E. (2014). *Fluctuations of Lévy Processes with Applications*, 2nd ed. *Universitext*. Heidelberg: Springer. Introductory lectures. MR3155252 https://doi.org/10.1007/978-3-642-37632-0

[25] Kyprianou, A.E., Pardo, J.C. and Rivero, V. (2010). Exact and asymptotic n-tuple laws at first and last passage. *Ann. Appl. Probab.* **20** 522–564. MR2650041 https://doi.org/10.1214/09-AAP626

[26] Kyprianou, A.E. and Surya, B.A. (2007). A note on a change of variable formula with local time-space for Lévy processes of bounded variation. In *Séminaire de Probabilités XL*. *Lecture Notes in Math.* **1899** 97–104. Berlin: Springer. MR2409000 https://doi.org/10.1007/978-3-540-71189-6_3

[27] Lande, R., Engen, S. and Saether, B.-E. (1994). Optimal harvesting, economic discounting and extinction risk in fluctuating populations. *Nature* **372** 88–90.

[28] Liang, G. and Zervos, M. (2020). Ergodic singular stochastic control motivated by the optimal sustainable exploitation of an ecosystem. Available at arXiv:2008.05576.

[29] Øksendal, B. and Sulem, A. (2019). *Applied Stochastic Control of Jump Diffusions*. *Universitext*. Cham: Springer. MR3931325 https://doi.org/10.1007/978-3-030-02781-0

[30] Pilipenko, A. (2014). *An Introduction to Stochastic Differential Equations with Reflection* **1**. Universitätsverlag Potsdam.

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

[32] Shimizu, Y. (2006). Density estimation of Lévy measures for discretely observed diffusion processes with jumps. *J. Japan Statist. Soc.* **36** 37–62. MR2266416 https://doi.org/10.14490/jjss.36.37

[33] Shimizu, Y. (2009). Functional estimation for Lévy measures of semimartingales with Poissonian jumps. *J. Multivariate Anal.* **100** 1073–1092. MR2508373 https://doi.org/10.1016/j.jmva.2008.10.006

[34] Sohr, T. (2020). Contributions to Optimal Stopping and Long-Term Average Impulse Control Ph.D. thesis CAU Kiel.

[35] Vershynin, R. (2018). *High-Dimensional Probability*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. An introduction with applications in data science. MR3837109 https://doi.org/10.1017/9781108231596

Bernoulli **30**(3), 2024, 2102–2126
https://doi.org/10.3150/23-BEJ1666

On large deviations and intersection of random interlacements

XINYI LI[1,a] and ZIJIE ZHUANG[2,b]

[1]*BICMR, Peking University, Beijing, China,* [a]*xinyili@bicmr.pku.edu.cn*
[2]*University of Pennsylvania, Philadelphia PA, USA,* [b]*zijie123@wharton.upenn.edu*

We investigate random interlacements on $\mathbb{Z}^d$ with $d \geq 3$, and derive the large deviation rate for the probability that the capacity of the interlacement set in a macroscopic box is much smaller than that of the box. As an application, we obtain the large deviation rate for the probability that two independent interlacements have empty intersections in a macroscopic box. Additionally, we prove that conditioning on this event, one of the interlacements will be sparse in terms of capacity within the box. This result is an example of the entropic repulsion phenomenon for random interlacements.

Keywords: Entropic repulsion; large deviations; random interlacements

References

Asselah, A. and Schapira, B. (2017). Moderate deviations for the range of a transient random walk: Path concentration. *Ann. Sci. Éc. Norm. Supér. (4)* **50** 755–786. MR3665554 https://doi.org/10.24033/asens.2331

Asselah, A. and Schapira, B. (2020). On the nature of the Swiss cheese in dimension 3. *Ann. Probab.* **48** 1002–1013. MR4089500 https://doi.org/10.1214/19-AOP1380

Bolthausen, E. and Deuschel, J.-D. (1993). Critical large deviations for Gaussian fields in the phase transition regime. I. *Ann. Probab.* **21** 1876–1920. MR1245293

Bolthausen, E., Deuschel, J.-D. and Zeitouni, O. (1995). Entropic repulsion of the lattice free field. *Comm. Math. Phys.* **170** 417–443. MR1334403

Chiarini, A. and Nitzschner, M. (2020a). Entropic repulsion for the occupation-time field of random interlacements conditioned on disconnection. *Ann. Probab.* **48** 1317–1351. MR4112716 https://doi.org/10.1214/19-AOP1393

Chiarini, A. and Nitzschner, M. (2020b). Entropic repulsion for the Gaussian free field conditioned on disconnection by level-sets. *Probab. Theory Related Fields* **177** 525–575. MR4095021 https://doi.org/10.1007/s00440-019-00957-7

Comets, F., Gallesco, C., Popov, S. and Vachkovskaia, M. (2013). On large deviations for the cover time of two-dimensional torus. *Electron. J. Probab.* **18** no. 96. MR3126579 https://doi.org/10.1214/EJP.v18-2856

Deuschel, J.-D. and Giacomin, G. (1999). Entropic repulsion for the free field: Pathwise characterization in $d \geq 3$. *Comm. Math. Phys.* **206** 447–462. MR1722113 https://doi.org/10.1007/s002200050713

Drewitz, A., Ráth, B. and Sapozhnikov, A. (2014). *An Introduction to Random Interlacements*. *SpringerBriefs in Mathematics*. Cham: Springer. MR3308116 https://doi.org/10.1007/978-3-319-05852-8

Duminil-Copin, H., Goswami, S., Rodriguez, P.-F. and Severo, F. (2023). Equality of critical parameters for percolation of Gaussian free field level sets. *Duke Math. J.* **172** 839–913. MR4568695 https://doi.org/10.1215/00127094-2022-0017

Li, X. (2017). A lower bound for disconnection by simple random walk. *Ann. Probab.* **45** 879–931. MR3630289 https://doi.org/10.1214/15-AOP1077

Li, X. and Sznitman, A.-S. (2014). A lower bound for disconnection by random interlacements. *Electron. J. Probab.* **19** no. 17. MR3164770 https://doi.org/10.1214/EJP.v19-3067

Nitzschner, M. and Sznitman, A.-S. (2020). Solidification of porous interfaces and disconnection. *J. Eur. Math. Soc. (JEMS)* **22** 2629–2672. MR4118617 https://doi.org/10.4171/JEMS/973

Phetpradap, P. (2011). Intersections of random walks Ph.D. thesis Univ. Bath.

1350-7265

Popov, S. and Teixeira, A. (2015). Soft local times and decoupling of random interlacements. *J. Eur. Math. Soc. (JEMS)* **17** 2545–2593. MR3420516 https://doi.org/10.4171/JEMS/565

Port, S.C. and Stone, C.J. (1978). *Brownian Motion and Classical Potential Theory*. *Probability and Mathematical Statistics*. New York-London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0492329

Sidoravicius, V. and Sznitman, A.-S. (2009). Percolation for the vacant set of random interlacements. *Comm. Pure Appl. Math.* **62** 831–858. MR2512613 https://doi.org/10.1002/cpa.20267

Sznitman, A.-S. (2010). Vacant set of random interlacements and percolation. *Ann. of Math. (2)* **171** 2039–2087. MR2680403 https://doi.org/10.4007/annals.2010.171.2039

Sznitman, A.-S. (2015). Disconnection and level-set percolation for the Gaussian free field. *J. Math. Soc. Japan* **67** 1801–1843. MR3417515 https://doi.org/10.2969/jmsj/06741801

Sznitman, A.-S. (2017). Disconnection, random walks, and random interlacements. *Probab. Theory Related Fields* **167** 1–44. MR3602841 https://doi.org/10.1007/s00440-015-0676-y

Sznitman, A.-S. (2019a). On macroscopic holes in some supercritical strongly dependent percolation models. *Ann. Probab.* **47** 2459–2493. MR3980925 https://doi.org/10.1214/18-AOP1312

Sznitman, A.-S. (2019b). On bulk deviations for the local behavior of random interlacements. *Ann. Sci. Éc. Norm. Supér*. To appear. Available at arXiv:1906.05809.

Sznitman, A.-S. (2021a). Excess deviations for points disconnected by random interlacements. *Probab. Math. Phys.* **2** 563–611. MR4408020 https://doi.org/10.2140/pmp.2021.2.563

Sznitman, A.-S. (2021b). On the C^1-property of the percolation function of random interlacements and a related variational problem. In *In and Out of Equilibrium 3. Celebrating Vladas Sidoravicius*. *Progress in Probability* **77** 775–796. Cham: Birkhäuser/Springer. MR4237292 https://doi.org/10.1007/978-3-030-60754-8_32

Sznitman, A.-S. (2023). On the cost of the bubble set for random interlacements. *Invent. Math.* **233** 903–950. MR4607724 https://doi.org/10.1007/s00222-023-01190-9

van den Berg, M., Bolthausen, E. and den Hollander, F. (2001). Moderate deviations for the volume of the Wiener sausage. *Ann. of Math. (2)* **153** 355–406. MR1829754 https://doi.org/10.2307/2661345

Zhuang, Z. (2021). On the percolative properties of the intersection of two independent interlacements. *ALEA Lat. Am. J. Probab. Math. Stat.* **18** 1061–1084. MR4282182 https://doi.org/10.30757/alea.v18-40

Bernoulli **30**(3), 2024, 2127–2153
https://doi.org/10.3150/23-BEJ1667

Sparse signal detection in heteroscedastic Gaussian sequence models: Sharp minimax rates

JULIEN CHHOR[1,2,a], RAJARSHI MUKHERJEE[1,b] and SUBHABRATA SEN[1,c]

[1]*Harvard University, Boston, MA, USA,* [a]*jchhor@hsph.harvard.edu,* [b]*ram521@mail.harvard.edu,* [c]*subhabratasen@fas.harvard.edu*
[2]*Toulouse School of Economics, University of Toulouse Capitole, Toulouse, France*

Given a heterogeneous Gaussian sequence model with unknown mean $\theta \in \mathbb{R}^d$ and known covariance matrix $\Sigma = \mathrm{diag}(\sigma_1^2, \dots, \sigma_d^2)$, we study the signal detection problem against sparse alternatives, for known sparsity s. Namely, we characterize how large $\epsilon^* > 0$ should be, in order to distinguish with high probability the null hypothesis $\theta = 0$ from the alternative composed of s-sparse vectors in $\mathbb{R}^d$, separated from 0 in L^t norm ($t \in [1, \infty]$) by at least ϵ^*. We find non-asymptotic minimax upper and lower bounds over the minimax separation radius ϵ^* and prove that they are always matching. We also derive the corresponding minimax tests achieving these bounds. Our results reveal new phase transitions regarding the behavior of ϵ^* with respect to the level of sparsity, to the L^t metric, and to the heteroscedasticity profile of Σ. In the case of the Euclidean (i.e. L^2) separation, we bridge the remaining gaps in the literature.

Keywords: Heteroscedasticity; non-Euclidean norms; signal detection; sparsity

References

Arias-Castro, E., Candès, E.J. and Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *Ann. Statist.* **39** 2533–2556. MR2906877 https://doi.org/10.1214/11-AOS910

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

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

Barnett, I., Mukherjee, R. and Lin, X. (2017). The generalized higher criticism for testing SNP-set effects in genetic association studies. *J. Amer. Statist. Assoc.* **112** 64–76. MR3646553 https://doi.org/10.1080/01621459.2016.1192039

Berrett, T. and Butucea, C. (2020). Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms. *Adv. Neural Inf. Process. Syst.* **33** 3164–3173.

Bhattacharya, B.B. and Mukherjee, R. (2021). Sparse Uniformity Testing. arXiv preprint arXiv:2109.10481.

Blais, E., Canonne, C.L. and Gur, T. (2019). Distribution testing lower bounds via reductions from communication complexity. *ACM Trans. Comput. Theory* **11** Art. 6, 37. MR3940784 https://doi.org/10.1145/3305270

Burnashev, M. (1979). On the minimax detection of an inaccurately known signal in a white Gaussian noise background. *Theory Probab. Appl.* **24** 107–119. MR0522240

Butucea, C. and Issartel, Y. (2021). Locally differentially private estimation of functionals of discrete distributions. *Adv. Neural Inf. Process. Syst.* **34** 24753–24764.

Cai, T.T., Guo, Z. and Ma, R. (2023). Statistical inference for high-dimensional generalized linear models with binary outcomes. *J. Amer. Statist. Assoc.* **118** 1319–1332. MR4595497 https://doi.org/10.1080/01621459.2021.1990769

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

1350-7265

Cai, T., Liu, W. and Xia, Y. (2013). Two-sample covariance matrix testing and support recovery in high-dimensional and sparse settings. *J. Amer. Statist. Assoc.* **108** 265–277. MR3174618 https://doi.org/10.1080/01621459.2012.758041

Cai, T.T., Liu, W. and Xia, Y. (2014). Two-sample test of high dimensional means under dependence. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 349–372. MR3164870 https://doi.org/10.1111/rssb.12034

Cai, T.T. and Low, M.G. (2011). Testing composite hypotheses, Hermite polynomials and optimal estimation of a nonsmooth functional. *Ann. Statist.* **39** 1012–1041. MR2816346 https://doi.org/10.1214/10-AOS849

Cai, T.T. and Xia, Y. (2014). High-dimensional sparse MANOVA. *J. Multivariate Anal.* **131** 174–196. MR3252643 https://doi.org/10.1016/j.jmva.2014.07.002

Canonne, C.L. (2020). A survey on distribution testing: Your data is big. But is it blue? *Theory Comput.* 1–100.

Canonne, C.L. (2022). Topics and Techniques in Distribution Testing: A Biased but Representative Sample.

Cao, Y., Lin, W. and Li, H. (2018). Two-sample tests of high-dimensional means for compositional data. *Biometrika* **105** 115–132. MR3768869 https://doi.org/10.1093/biomet/asx060

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

Carpentier, A., Collier, O., Comminges, L., Tsybakov, A.B. and Wang, Y. (2019). Minimax rate of testing in sparse linear regression. *Autom. Remote Control* **80** 1817–1834. MR4032408 https://doi.org/10.1134/s0005231019100040

Carpentier, A., Collier, O., Comminges, L., Tsybakov, A.B. and Wang, Y. (2022). Estimation of the ℓ_2-norm and testing in sparse linear regression with unknown variance. *Bernoulli* **28** 2744–2787. MR4474561 https://doi.org/10.3150/21-bej1436

Chang, J., Zheng, C., Zhou, W.-X. and Zhou, W. (2017). Simulation-based hypothesis testing of high dimensional means under covariance heterogeneity. *Biometrics* **73** 1300–1310. MR3744543 https://doi.org/10.1111/biom.12695

Chhor, J. and Carpentier, A. (2021). Goodness-of-Fit Testing for Hölder-Continuous Densities: Sharp Local Minimax Rates. arXiv preprint arXiv:2109.04346.

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

Chhor, J., Mukherjee, R. and Sen, S. (2024). Supplement to "Sparse signal detection in heteroscedastic Gaussian sequence models: sharp minimax rates." https://doi.org/10.3150/23-BEJ1667SUPP

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., Comminges, L. and Tsybakov, A.B. (2020). On estimation of nonsmooth functionals of sparse normal means. *Bernoulli* **26** 1989–2020. MR4091099 https://doi.org/10.3150/19-BEJ1180

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.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

Fukuchi, K. and Sakuma, J. (2017). Minimax optimal estimators for additive scalar functionals of discrete distributions. In *2017 IEEE Int. Symp. Inf. Theory - Proc. (ISIT)* 2103–2107. IEEE.

Gerber, P.R. and Polyanskiy, Y. (2022). Likelihood-free hypothesis testing. arXiv preprint arXiv:2211.01126.

Gutzeit, M. (2019). Topics in statistical minimax hypothesis testing.

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

Han, Y., Jiao, J. and Mukherjee, R. (2020). On estimation of L_r-norms in Gaussian white noise models. *Probab. Theory Related Fields* **177** 1243–1294. MR4126939 https://doi.org/10.1007/s00440-020-00982-x

Hartl, D.L., Clark, A.G. and Clark, A.G. (1997). *Principles of Population Genetics* **116**. Sunderland, MA: Sinauer Associates.

Ingster, Y.I. (1982). Minimax nonparametric detection of signals in white Gaussian noise. *Problemy Peredachi Informatsii* **18** 61–73. MR0689340

Ingster, Y.I. (1987). Minimax testing of nonparametric hypotheses on a distribution density in the L_p metrics. *Theory Probab. Appl.* **31** 333–337. MR0851000

Ingster, Y.I. (1997). Some problems of hypothesis testing leading to infinitely divisible distributions. *Math. Methods Statist.* **6** 47–69. MR1456646

Ingster, Y.I. and Suslina, I.A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. New York: Springer. 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

Jiao, J., Venkat, K., Han, Y. and Weissman, T. (2015). Minimax estimation of functionals of discrete distributions. *IEEE Trans. Inf. Theory* **61** 2835–2885. MR3342309 https://doi.org/10.1109/TIT.2015.2412945

Johnstone, I.M. (2002). Function estimation and gaussian sequence models. Unpublished Manuscript 2 2.

Kipnis, A. (2021). Unification of rare/weak detection models using moderate deviations analysis and log-chisquared p-values. arXiv preprint arXiv:2103.03999.

Kipnis, A. (2022). Higher criticism for discriminating word-frequency tables and authorship attribution. *Ann. Appl. Stat.* **16** 1236–1252. MR4438832 https://doi.org/10.1214/21-aoas1544

Kipnis, A. and Donoho, D.L. (2021). Two-sample testing of discrete distributions under rare/weak perturbations. In *2021 IEEE Int. Symp. Inf. Theory - Proc. (ISIT)* 3314–3319. IEEE.

Kotekal, S. and Gao, C. (2021). Minimax rates for sparse signal detection under correlation. arXiv preprint arXiv:2110.12966.

Kraft, P. and Hunter, D.J. (2009). Genetic risk prediction–are we there yet? *N. Engl. J. Med.* **360** 1701–1703. https://doi.org/10.1056/NEJMp0810107

Laird, N.M. and Lange, C. (2011). *The Fundamentals of Modern Statistical Genetics*. *Statistics for Biology and Health*. New York: Springer. MR2762582 https://doi.org/10.1007/978-1-4419-7338-2

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

Lee, S., Abecasis, G.R., Boehnke, M. and Lin, X. (2014). Rare-variant association analysis: Study designs and statistical tests. *Am. J. Hum. Genet.* **95** 5–23.

Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the L_r norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. MR1670867 https://doi.org/10.1007/s004409970006

Li, B. and Leal, S.M. (2008). Methods for detecting associations with rare variants for common diseases: Application to analysis of sequence data. *Am. J. Hum. Genet.* **83** 311–321. https://doi.org/10.1016/j.ajhg.2008.06.024

Li, Z., Qin, S. and Li, Q. (2021). A novel test by combining the maximum and minimum values among a large number of dependent Z-scores with application to genome wide association study. *Stat. Med.* **40** 2422–2434. MR4242805 https://doi.org/10.1002/sim.8912

Liu, H., Gao, C. and Samworth, R.J. (2021). Minimax rates in sparse, high-dimensional change point detection. *Ann. Statist.* **49** 1081–1112. MR4255120 https://doi.org/10.1214/20-aos1994

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A. et al. (2009). Finding the missing heritability of complex diseases. *Nature* **461** 747–753.

McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P.A. and Hirschhorn, J.N. (2008). Genome-wide association studies for complex traits: Consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9** 356–369. https://doi.org/10.1038/nrg2344

Mukherjee, R., Mukherjee, S. and Yuan, M. (2018). Global testing against sparse alternatives under Ising models. *Ann. Statist.* **46** 2062–2093. MR3845011 https://doi.org/10.1214/17-AOS1612

Stephens, M. (2017). False discovery rates: A new deal. *Biostatistics* **18** 275–294. MR3824755 https://doi.org/10.1093/biostatistics/kxw041

Sun, R. and Lin, X. (2020). Genetic variant set-based tests using the generalized Berk–Jones statistic with application to a genome-wide association study of breast cancer. *J. Amer. Statist. Assoc.* **115** 1079–1091. MR4143451 https://doi.org/10.1080/01621459.2019.1660170

Talagrand, M. (2014). *Upper and Lower Bounds for Stochastic Processes: Modern Methods and Classical Problems*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*

[Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics] **60**. Heidelberg: Springer. MR3184689 https://doi.org/10.1007/978-3-642-54075-2

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

Visscher, P.M., Brown, M.A., McCarthy, M.I. and Yang, J. (2012). Five years of GWAS discovery. *Am. J. Hum. Genet.* **90** 7–24.

Waggoner, B. (2015). ℓ_p testing and learning of discrete distributions. In *ITCS'15—Proceedings of the 6th Innovations in Theoretical Computer Science* 347–356. New York: ACM. MR3419028 https://doi.org/10.1145/2688073.2688095

Wu, Y. and Yang, P. (2016). Minimax rates of entropy estimation on large alphabets via best polynomial approximation. *IEEE Trans. Inf. Theory* **62** 3702–3720. MR3506758 https://doi.org/10.1109/TIT.2016.2548468

Wu, Y. and Yang, P. (2019). Chebyshev polynomials, moment matching, and optimal estimation of the unseen. *Ann. Statist.* **47** 857–883. MR3909953 https://doi.org/10.1214/17-AOS1665

Xia, Y., Cai, T. and Cai, T.T. (2018). Two-sample tests for high-dimensional linear regression with an application to detecting interactions. *Statist. Sinica* **28** 63–92. MR3752252

Xue, K. and Yao, F. (2020). Distribution and correlation-free two-sample test of high-dimensional means. *Ann. Statist.* **48** 1304–1328. MR4124324 https://doi.org/10.1214/19-AOS1848

Yu, X., Li, D., Xue, L. and Li, R. (2022). Power-enhanced simultaneous test of high-dimensional mean vectors and covariance matrices with application to gene-set testing. *J. Amer. Statist. Assoc.* 1–14.

Zhang, X. and Cheng, G. (2017). Simultaneous inference for high-dimensional linear models. *J. Amer. Statist. Assoc.* **112** 757–768. MR3671768 https://doi.org/10.1080/01621459.2016.1166114

Zhang, H. and Wu, Z. (2022). The general goodness-of-fit tests for correlated data. *Comput. Statist. Data Anal.* **167** Paper No. 107379, 16. MR4333746 https://doi.org/10.1016/j.csda.2021.107379

Bernoulli **30**(3), 2024, 2154–2184
https://doi.org/10.3150/23-BEJ1668

Empirical Bayes inference for the block maxima method

SIMONE A. PADOAN[1,a] and STEFANO RIZZELLI[2,b]

[1]*Department of Decision Sciences, Bocconi University, Milano, Italy,* [a]*simone.padoan@unibocconi.it*
[2]*Department of Statistical Sciences, Catholic University, Milano, Italy,* [b]*stefano.rizzelli@unicatt.it*

The block maxima method is one of the most popular approaches for extreme value analysis with independent and identically distributed observations in the domain of attraction of an extreme value distribution. The lack of a rigorous study on the Bayesian inference in this context has limited its use for statistical analysis of extremes. In this paper we propose an empirical Bayes procedure for inference on the block maxima law and its related quantities. We show that the posterior distributions of the tail index of the data distribution and of the return levels (representative of future extreme episodes) are consistent and asymptotically normal. These properties guarantee the reliability of posterior-based inference. We also establish contraction rates of the posterior predictive distribution, the key tool in Bayesian probabilistic forecasting. Posterior computations are readily obtained via an efficient adaptive Metropolis-Hasting type of algorithm. Simulations show its excellent inferential performances already with modest sample sizes. The utility of our proposal is showcased analysing extreme winds generated by hurricanes in Southeastern US.

Keywords: Contraction rate; extreme quantiles; posterior consistency; return levels; tail index; wind speed

References

[1] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR2108013 https://doi.org/10.1002/0470012382

[2] Bickel, P.J. and Kleijn, B.J.K. (2012). The semiparametric Bernstein–von Mises theorem. *Ann. Statist.* **40** 206–237. MR3013185 https://doi.org/10.1214/11-AOS921

[3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR0898871 https://doi.org/10.1017/CBO9780511721434

[4] Bücher, A. and Segers, J. (2017). On the maximum likelihood estimator for the generalized extreme-value distribution. *Extremes* **20** 839–872. MR3737387 https://doi.org/10.1007/s10687-017-0292-6

[5] Choudhuri, N., Ghosal, S. and Roy, A. (2004). Bayesian estimation of the spectral density of a time series. *J. Amer. Statist. Assoc.* **99** 1050–1059. MR2109494 https://doi.org/10.1198/016214504000000557

[6] Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. *Springer Series in Statistics*. London: Springer London, Ltd. MR1932132 https://doi.org/10.1007/978-1-4471-3675-0

[7] Coles, S. and Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **52** 405–416. MR2012566 https://doi.org/10.1111/1467-9876.00413

[8] Coles, S.G. and Powell, E.A. (1996). Bayesian methods in extreme value modelling: A review and new developments. *Int. Stat. Rev.* **64** 119–136.

[9] Coles, S.G. and Tawn, J.A. (1996). A Bayesian analysis of extreme rainfall data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **45** 463–478.

[10] de Carvalho, M., Pereira, S., Pereira, P. and de Zea Bermudez, P. (2022). An extreme value Bayesian Lasso for the conditional left and right tails. *J. Agric. Biol. Environ. Stat.* **27** 222–239. MR4416781 https://doi.org/10.1007/s13253-021-00469-9

[11] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3
[12] de Haan, L. and Resnick, S. (1996). Second-order regular variation and rates of convergence in extreme-value theory. *Ann. Probab.* **24** 97–124. MR1387628 https://doi.org/10.1214/aop/1042644709
[13] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Aust. Math. Soc. A* **61** 381–395. MR1420345 https://doi.org/10.1017/S144678870000046X
[14] Devroye, L., Mehrabian, A. and Reddad, T. (2020). The total variation distance between high-dimensional Gaussians. ArXive E-prints. Available at https://arxiv.org/pdf/1810.08693.pdf.
[15] Diebolt, J., Guillou, A., Naveau, P. and Ribereau, P. (2008). Improving probability-weighted moment methods for the generalized extreme value distribution. *REVSTAT* **6** 35–50. MR2386298 https://doi.org/10.57805/revstat.v6i1.56
[16] Dombry, C. (2015). Existence and consistency of the maximum likelihood estimators for the extreme value index within the block maxima framework. *Bernoulli* **21** 420–436. MR3322325 https://doi.org/10.3150/13-BEJ573
[17] Dombry, C. and Ferreira, A. (2019). Maximum likelihood estimators based on the block maxima method. *Bernoulli* **25** 1690–1723. MR3961227 https://doi.org/10.3150/18-BEJ1032
[18] Ferreira, A. and de Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *Ann. Statist.* **43** 276–298. MR3285607 https://doi.org/10.1214/14-AOS1280
[19] Garthwaite, P.H., Fan, Y. and Sisson, S.A. (2016). Adaptive optimal scaling of Metropolis–Hastings algorithms using the Robbins–Monro process. *Comm. Statist. Theory Methods* **45** 5098–5111. MR3520297 https://doi.org/10.1080/03610926.2014.936562
[20] Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR3587782 https://doi.org/10.1017/9781139029834
[21] Haario, H., Saksman, E. and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli* **7** 223–242. MR1828504 https://doi.org/10.2307/3318737
[22] Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27** 251–261. MR0797563 https://doi.org/10.2307/1269706
[23] Jenkinson, A. (1969). Statistics of extemes. In *Estimation of Maximum Floods, WMO Tech. Note 98* 183–228.
[24] Knight, K. (2000). *Mathematical Statistics*. *Chapman & Hall/CRC Texts in Statistical Science Series*. Boca Raton, FL: CRC Press/CRC. MR1739674 https://doi.org/10.1201/9780367805319
[25] Krüger, F., Lerch, S., Thorarinsdottir, T. and Gneiting, T. (2021). Predictive inference based on Markov chain Monte Carlo output. *Int. Stat. Rev.* **89** 274–301. MR4411906 https://doi.org/10.1111/insr.12405
[26] Lehmann, E.L. and Casella, G. (1998). *Theory of Point Estimation*, 2nd ed. *Springer Texts in Statistics*. New York: Springer. MR1639875 https://doi.org/10.1007/b98854
[27] Maribe, G., Verster, A. and Beirlant, J. (2016). Reducing MSE in estimation of heavy tails: A Bayesian approach. ArXiv preprint. Available at arXiv:1606.05687.
[28] Northrop, P.J. and Attalides, N. (2016). Posterior propriety in Bayesian extreme value analyses using reference priors. *Statist. Sinica* **26** 721–743. MR3497768 https://doi.org/10.5705/ss.2014.034
[29] Padoan, S.A. and Rizzelli, S. (2022). Consistency of Bayesian inference for multivariate max-stable distributions. *Ann. Statist.* **50** 1490–1518. MR4441129 https://doi.org/10.1214/21-AOS2160
[30] Padoan, S.A. and Rizzelli, S. (2024). Supplement to "Empirical Bayes inference for the block maxima method." https://doi.org/10.3150/23-BEJ1668SUPP
[31] Prescott, P. and Walden, A.T. (1980). Maximum likelihood estimation of the parameters of the generalized extreme-value distribution. *Biometrika* **67** 723–724. MR0601119 https://doi.org/10.1093/biomet/67.3.723
[32] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2364939 https://doi.org/10.1007/978-0-387-75953-1
[33] Robert, C.P. (2007). *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*, 2nd ed. *Springer Texts in Statistics*. New York: Springer. MR2723361 https://doi.org/10.1007/0-387-71599-1

[34] Roberts, G.O., Gelman, A. and Gilks, W.R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.* **7** 110–120. MR1428751 https://doi.org/10.1214/aoap/1034625254
[35] Smith, R.L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72** 67–90. MR0790201 https://doi.org/10.1093/biomet/72.1.67
[36] Stephenson, A. and Tawn, J. (2004). Bayesian inference for extremes: Accounting for the three extremal types. *Extremes* **7** 291–307. MR2212389 https://doi.org/10.1007/s10687-004-3479-6
[37] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

Bernoulli **30**(3), 2024, 2185–2206
https://doi.org/10.3150/23-BEJ1669

Limit theorems for random Motzkin paths near boundary

WŁODZIMIERZ BRYC[a] and YIZAO WANG[b]

Department of Mathematical Sciences, University of Cincinnati, 2815 Commons Way, Cincinnati, OH, 45221-0025, USA, [a]*wlodzimierz.bryc@uc.edu,* [b]*yizao.wang@uc.edu*

We consider Motzkin paths of length L, not fixed at zero at both end points, with constant weights on the edges and general weights on the end points. We investigate, as the length L tends to infinity, the limit behaviors of (a) boundary measures induced by the weights on both end points and (b) the segments of the sampled Motzkin path viewed as a process starting from each of the two end points, referred to as boundary processes. Our first result concerns the case when the induced boundary measures have finite first moments. Our second result concerns when the boundary measure on the right end point is a generalized geometric measure with parameter $\rho_1 \geq 1$, so that this is an infinite measure and yet it induces a probability measure for random Motzkin path when ρ_1 is not too large. The two cases under investigation reveal a phase transition. In particular, we show that the limit left boundary processes in the two cases have the same transition probabilities as random walks conditioned to stay non-negative.

Keywords: Discrete Bessel process; matrix ansatz; Motzkin paths; random walk conditioned to stay positive; Viennot's formula

References

Barraquand, G. and Le Doussal, P. (2022). Steady state of the KPZ equation on an interval and Liouville quantum mechanics. *Europhys. Lett.* **137** 61003. https://doi.org/10.1209/0295-5075/ac25a9. ArXiv preprint with Supplementary material available at https://arxiv.org/abs/2105.15178.

Barraquand, G. and Le Doussal, P. (2023). Stationary measures of the KPZ equation on an interval from Enaud-Derrida's matrix product ansatz representation. *J. Phys. A* **56** Paper No. 144003, 14. MR4562516

Bertoin, J. and Doney, R.A. (1994). On conditioning a random walk to stay nonnegative. *Ann. Probab.* **22** 2152–2167. MR1331218

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

Bryc, W. and Kuznetsov, A. (2022). Markov limits of steady states of the KPZ equation on an interval. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 1329–1351. MR4512149 https://doi.org/10.30757/alea.v19-53

Bryc, W. and Wang, Y. (2019). Fluctuations of random Motzkin paths. *Adv. in Appl. Math.* **106** 96–116. MR3915365 https://doi.org/10.1016/j.aam.2019.02.003

Bryc, W. and Wesołowski, J. (2017). Asymmetric simple exclusion process with open boundaries and quadratic harnesses. *J. Stat. Phys.* **167** 383–415. MR3626634 https://doi.org/10.1007/s10955-017-1747-5

Bryc, W. and Wesołowski, J. (2023). Pitman's discrete $2M - X$ theorem for arbitrary initial laws and continuous time limits. arXiv preprint. Available at http://arxiv.org/abs/2304.07144.

Charalambides, C.A. (2016). *Discrete q-Distributions*. Hoboken, NJ: Wiley. MR3559356 https://doi.org/10.1002/9781119119128

Derrida, B., Enaud, C. and Lebowitz, J.L. (2004). The asymmetric exclusion process and Brownian excursions. *J. Stat. Phys.* **115** 365–382. MR2070099 https://doi.org/10.1023/B:JOSS.0000019833.35328.b4

Flajolet, P. and Sedgewick, R. (2009). *Analytic Combinatorics*. Cambridge: Cambridge Univ. Press. MR2483235 https://doi.org/10.1017/CBO9780511801655

1350-7265

Ismail, M.E.H. (2009). *Classical and Quantum Orthogonal Polynomials in One Variable*. *Encyclopedia of Mathematics and Its Applications* **98**. Cambridge: Cambridge Univ. Press. MR2542683

Kaigh, W.D. (1976). An invariance principle for random walk conditioned by a late return to zero. *Ann. Probab.* **4** 115–121. MR0415706 https://doi.org/10.1214/aop/1176996189

Liggett, T.M. (1975). Ergodic theorems for the asymmetric simple exclusion process. *Trans. Amer. Math. Soc.* **213** 237–261. MR0410986 https://doi.org/10.2307/1998046

Miyazaki, H. and Tanaka, H. (1989). A theorem of Pitman type for simple random walks on $\mathbf{Z}^d$. *Tokyo J. Math.* **12** 235–240. MR1001744 https://doi.org/10.3836/tjm/1270133560

Pitman, J.W. (1975). One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. in Appl. Probab.* **7** 511–526. MR0375485 https://doi.org/10.2307/1426125

Viennot, G. (1985). A combinatorial theory for general orthogonal polynomials with extensions and applications. In *Orthogonal Polynomials and Applications (Bar-le-Duc, 1984)*. *Lecture Notes in Math.* **1171** 139–157. Berlin: Springer. MR0838979 https://doi.org/10.1007/BFb0076539

Bernoulli **30**(3), 2024, 2207–2227
https://doi.org/10.3150/23-BEJ1670

Poincaré inequalities and integrated curvature-dimension criterion for generalised Cauchy and convex measures

BAPTISTE HUGUET[a]

ENS Rennes, Univ. Rennes, Rennes, France, [a]*baptiste.huguet@math.cnrs.fr*

We obtain new sharp weighted Poincaré inequalities on Riemannian manifolds for a general class of measures. When specialised to generalised Cauchy measures, this gives a unified and simple proof of the weighted Poincaré inequality for the whole range of parameters, with the optimal spectral gap, the error term and the extremal functions.

Keywords: Curvature-dimension criterion; generalised Cauchy measures; heavy tails; Poincaré inequality

References

Arnaudon, M., Bonnefont, M. and Joulin, A. (2018). Intertwinings and generalized Brascamp-Lieb inequalities. *Rev. Mat. Iberoam.* **34** 1021–1054. MR3850277 https://doi.org/10.4171/RMI/1014

Bakry, D. and Émery, M. (1985). Diffusions hypercontractives. In *Séminaire de Probabilités, XIX, 1983/84*. *Lecture Notes in Math.* **1123** 177–206. Berlin: Springer. MR0889476 https://doi.org/10.1007/BFb0075847

Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9

Bakry, D., Gentil, I. and Scheffer, G. (2020). Sharp Beckner-type inequalities for Cauchy and spherical distributions. *Studia Math.* **251** 219–245. MR4048609 https://doi.org/10.4064/sm180503-17-1

Blanchet, A., Bonforte, M., Dolbeault, J., Grillo, G. and Vázquez, J.-L. (2007). Hardy-Poincaré inequalities and applications to nonlinear diffusions. *C. R. Math. Acad. Sci. Paris* **344** 431–436. MR2320246 https://doi.org/10.1016/j.crma.2007.01.011

Bobkov, S.G. and Ledoux, M. (2009). Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.* **37** 403–427. MR2510011 https://doi.org/10.1214/08-AOP407

Bobkov, S.G., Gozlan, N., Roberto, C. and Samson, P.-M. (2014). Bounds on the deficit in the logarithmic Sobolev inequality. *J. Funct. Anal.* **267** 4110–4138. MR3269872 https://doi.org/10.1016/j.jfa.2014.09.016

Bonnefont, M., Joulin, A. and Ma, Y. (2016a). A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. *ESAIM Probab. Stat.* **20** 18–29. MR3519678 https://doi.org/10.1051/ps/2015019

Bonnefont, M., Joulin, A. and Ma, Y. (2016b). Spectral gap for spherically symmetric log-concave probability measures, and beyond. *J. Funct. Anal.* **270** 2456–2482. MR3464047 https://doi.org/10.1016/j.jfa.2016.02.007

Cattiaux, P. and Guillin, A. (2023). A journey with the integrated Γ_2 criterion and its weak forms. In *Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2020-2022*. *Lecture Notes in Math.* **2327**.

Chafaï, D. and Lehec, J. (2020). On Poincaré and logarithmic Sobolev inequalities for a class of singular Gibbs measures. In *Geometric Aspects of Functional Analysis. Vol. I*. *Lecture Notes in Math.* **2256** 219–246. Cham: Springer. MR4175749 https://doi.org/10.1007/978-3-030-36020-7_10

Denzler, J. and McCann, R.J. (2005). Fast diffusion to self-similarity: Complete spectrum, long-time asymptotics, and numerology. *Arch. Ration. Mech. Anal.* **175** 301–342. MR2126633 https://doi.org/10.1007/s00205-004-0336-3

Gentil, I. and Zugmeyer, S. (2021). A family of Beckner inequalities under various curvature-dimension conditions. *Bernoulli* **27** 751–771. MR4255214 https://doi.org/10.3150/20-bej1228

1350-7265

Huguet, B. (2022). Intertwining relations for diffusions in manifolds and applications to functional inequalities. *Stochastic Process. Appl.* **145** 1–28. MR4354402 https://doi.org/10.1016/j.spa.2021.11.004

Ledoux, M. (1992). On an integral criterion for hypercontractivity of diffusion semigroups and extremal functions. *J. Funct. Anal.* **105** 444–465. MR1160084 https://doi.org/10.1016/0022-1236(92)90084-V

Nguyen, V.H. (2014). Dimensional variance inequalities of Brascamp-Lieb type and a local approach to dimensional Prékopa's theorem. *J. Funct. Anal.* **266** 931–955. MR3132733 https://doi.org/10.1016/j.jfa.2013.11.003

Obata, M. (1962). Certain conditions for a Riemannian manifold to be isometric with a sphere. *J. Math. Soc. Japan* **14** 333–340. MR0142086 https://doi.org/10.2969/jmsj/01430333

Saumard, A. (2019). Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one. *Bernoulli* **25** 3978–4006. MR4010979 https://doi.org/10.3150/19-bej1117

Bernoulli **30**(3), 2024, 2228–2252
https://doi.org/10.3150/23-BEJ1671

Sequential change diagnosis revisited and the Adaptive Matrix CuSum

AUSTIN WARNER[a] and GEORGIOS FELLOURIS[b]

Department of Statistics, University of Illinois Urbana-Champaign, Urbana, U.S.A., [a]*awarner5@illinois.edu,* [b]*fellouri@illinois.edu*

The problem of sequential change diagnosis is considered, where observations are obtained on-line, an abrupt change occurs in their distribution, and the goal is to quickly detect the change and accurately identify the post-change distribution, while controlling the false alarm rate. A finite set of alternatives is postulated for the post-change regime, but no prior information is assumed for the unknown change point. A drawback of many algorithms that have been proposed for this problem is the implicit use of pre-change data for determining the post-change distribution. This can lead to very large conditional probabilities of misidentification, given that there was no false alarm, unless the change occurs soon after monitoring begins. A novel, recursive algorithm is proposed and shown to resolve this issue without the use of additional tuning parameters and without sacrificing control of the worst-case delay in Lorden's sense. A theoretical analysis is conducted for a general family of sequential change diagnosis procedures, which supports the proposed algorithm and revises certain state-of-the-art results. Additionally, a novel, comprehensive method is proposed for the design and evaluation of sequential change diagnosis algorithms. This method is illustrated with simulation studies, where existing procedures are compared to the proposed.

Keywords: CuSum; Lorden's criterion; sequential change detection; sequential change diagnosis

References

Bakhache, B. and Nikiforov, I. (1999). Reliable detection of faults in navigation systems. In *Proceedings of the 38th IEEE Conference on Decision and Control (Cat. No. 99CH36304)* **5** 4976–4981. IEEE.

Bissell, A. (1969). Cusum techniques for quality control. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **18** 1–25.

Chen, Y., Wang, T. and Samworth, R.J. (2022). High-dimensional, multiscale online changepoint detection. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 234–266. MR4400396 https://doi.org/10.1111/rssb.12447

Chen, Y.C., Banerjee, T., Dominguez-Garcia, A.D. and Veeravalli, V.V. (2015). Quickest line outage detection and identification. *IEEE Trans. Power Syst.* **31** 749–758.

Dayanik, S., Goulding, C. and Poor, H.V. (2008). Bayesian sequential change diagnosis. *Math. Oper. Res.* **33** 475–496. MR2416004 https://doi.org/10.1287/moor.1070.0307

Dayanik, S., Powell, W.B. and Yamazaki, K. (2013). Asymptotically optimal Bayesian sequential change detection and identification rules. *Ann. Oper. Res.* **208** 337–370. MR3100637 https://doi.org/10.1007/s10479-012-1121-6

Fellouris, G. and Sokolov, G. (2016). Second-order asymptotic optimality in multisensor sequential change detection. *IEEE Trans. Inf. Theory* **62** 3662–3675. MR3506755 https://doi.org/10.1109/TIT.2016.2549042

Fienberg, S.E. and Shmueli, G. (2005). Statistical issues and challenges associated with rapid detection of bio-terrorist attacks. *Stat. Med.* **24** 513–529. MR2134521 https://doi.org/10.1002/sim.2032

Gösmann, J., Stoehr, C., Heiny, J. and Dette, H. (2022). Sequential change point detection in high dimensional time series. *Electron. J. Stat.* **16** 3608–3671. MR4444665 https://doi.org/10.1214/22-ejs2027

Han, D. and Tsung, F. (2007). Detection and diagnosis of unknown abrupt changes using CUSUM multi-chart schemes. *Sequential Anal.* **26** 225–249. MR2327978 https://doi.org/10.1080/07474940701404765

Hawkins, D.M., Qiu, P. and Kang, C.W. (2003). The changepoint model for statistical process control. *J. Qual. Technol.* **35** 355–366.

Hinkley, D.V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika* **57** 1–17. MR0273727 https://doi.org/10.1093/biomet/57.1.1

1350-7265

Huang, Y.-C., Huang, Y.-J. and Lin, S.-C. (2021). Asymptotic optimality in Byzantine distributed quickest change detection. *IEEE Trans. Inf. Theory* **67** 5942–5962. MR4345046 https://doi.org/10.1109/TIT.2021.3100423

Joe Qin, S. (2003). Statistical process monitoring: Basics and beyond. *J. Chemom.* **17** 480–502.

Lai, T.L. (1998). Information bounds and quick detection of parameter changes in stochastic systems. *IEEE Trans. Inf. Theory* **44** 2917–2929. MR1672051 https://doi.org/10.1109/18.737522

Lai, T.L. (2000). Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems. *IEEE Trans. Inf. Theory* **46** 595–608. https://doi.org/10.1109/18.825826

Lai, T.L. and Shan, J.Z. (1999). Efficient recursive algorithms for detection of abrupt changes in signals and control systems. *IEEE Trans. Automat. Control* **44** 952–966. MR1690539 https://doi.org/10.1109/9.763211

Lai, T.L. and Xing, H. (2010). Sequential change-point detection when the pre- and post-change parameters are unknown. *Sequential Anal.* **29** 162–175. MR2747518 https://doi.org/10.1080/07474941003741078

Lau, T.S., Tay, W.P. and Veeravalli, V.V. (2019). A binning approach to quickest change detection with unknown post-change distribution. *IEEE Trans. Signal Process.* **67** 609–621. MR3912276 https://doi.org/10.1109/TSP.2018.2881666

Liang, Y. and Veeravalli, V.V. (2022). Non-parametric quickest mean-change detection. *IEEE Trans. Inf. Theory* **68** 8040–8052. MR4544930

Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Stat.* **42** 1897–1908. MR0309251 https://doi.org/10.1214/aoms/1177693055

Ma, X., Lai, L. and Cui, S. (2021). Two-stage Bayesian sequential change diagnosis. *IEEE Trans. Signal Process.* **69** 6131–6147. MR4352880 https://doi.org/10.1109/TSP.2021.3115426

Malladi, D.P. and Speyer, J.L. (1999). A generalized Shiryayev sequential probability ratio test for change detection and isolation. *IEEE Trans. Automat. Control* **44** 1522–1534. MR1707049 https://doi.org/10.1109/9.780416

Mei, Y. (2006). Sequential change-point detection when unknown parameters are present in the pre-change distribution. *Ann. Statist.* **34** 92–122. MR2275236 https://doi.org/10.1214/009053605000000859

Mei, Y. (2010). Efficient scalable schemes for monitoring a large number of data streams. *Biometrika* **97** 419–433. MR2650748 https://doi.org/10.1093/biomet/asq010

Moustakides, G.V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist.* **14** 1379–1387. MR0868306 https://doi.org/10.1214/aos/1176350164

Nikiforov, I.V. (1995). A generalized change detection problem. *IEEE Trans. Inf. Theory* **41** 171–187. https://doi.org/10.1109/18.370109

Nikiforov, I.V. (2000). A simple recursive algorithm for diagnosis of abrupt changes in random signals. *IEEE Trans. Inf. Theory* **46** 2740–2746. MR1807403 https://doi.org/10.1109/18.887891

Nikiforov, I.V. (2003). A lower bound for the detection/isolation delay in a class of sequential tests. *IEEE Trans. Inf. Theory* **49** 3037–3047. MR2027584 https://doi.org/10.1109/TIT.2003.818398

Nikiforov, I.V. (2016). Sequential detection/isolation of abrupt changes. *Sequential Anal.* **35** 268–301. MR3547963 https://doi.org/10.1080/07474946.2016.1206354

Nikiforov, I., Varavva, V. and Kireichikov, V. (1993). Application of statistical fault detection algorithms to navigation systems monitoring. *Automatica* **29** 1275–1290.

Oskiper, T. and Poor, H.V. (2002). Online activity detection in a multiuser environment using the matrix CUSUM algorithm. *IEEE Trans. Inf. Theory* **48** 477–493. MR1891258 https://doi.org/10.1109/18.979323

Page, E.S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115. MR0088850 https://doi.org/10.1093/biomet/41.1-2.100

Pergamenchtchikov, S.M., Tartakovsky, A.G. and Spivak, V.S. (2022). Minimax and pointwise sequential change-point detection and identification for general stochastic models. *J. Multivariate Anal.* **190** Paper No. 104977, 22. MR4396577 https://doi.org/10.1016/j.jmva.2022.104977

Pollak, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13** 206–227. MR0773162 https://doi.org/10.1214/aos/1176346587

Rolka, H., Burkom, H., Cooper, G.F., Kulldorff, M., Madigan, D. and Wong, W.-K. (2007). Issues in applied statistics for public health bioterrorism surveillance using multiple data streams: Research needs. *Stat. Med.* **26** 1834–1856. MR2359196 https://doi.org/10.1002/sim.2793

Ru, J., Jilkov, V.P., Li, X.R. and Bashi, A. (2009). Detection of target maneuver onset. *IEEE Trans. Aerosp. Electron. Syst.* **45** 536–554.

Shewhart, W.A. (1931). *Economic Control of Quality of Manufactured Product*. London: Macmillan And Co Ltd.

Shin, J., Ramdas, A. and Rinaldo, A. (2022). E-detectors: A nonparametric framework for online changepoint detection. arXiv preprint arXiv:2203.03532.

Shiryaev, A.N. (1963). On optimum methods in quickest detection problems. *Theory Probab. Appl.* **8** 22–46. MR0155708 https://doi.org/10.1137/1108002

Shiryaev, A.N. (2008). *Optimal Stopping Rules*. *Stochastic Modelling and Applied Probability* **8**. Berlin: Springer. Translated from the 1976 Russian second edition by A. B. Aries, Reprint of the 1978 translation. MR2374974

Siegmund, D. (1979). Corrected diffusion approximations in certain random walk problems. *Adv. in Appl. Probab.* **11** 701–719. MR0544191 https://doi.org/10.2307/1426855

Siegmund, D. (2013). Change-points: From sequential detection to biology and back. *Sequential Anal.* **32** 2–14. MR3023983 https://doi.org/10.1080/07474946.2013.751834

Siegmund, D. and Venkatraman, E.S. (1995). Using the generalized likelihood ratio statistic for sequential detection of a change-point. *Ann. Statist.* **23** 255–271. MR1331667 https://doi.org/10.1214/aos/1176324466

Tartakovsky, A.G. (2008). Multidecision quickest change-point detection: Previous achievements and open problems. *Sequential Anal.* **27** 201–231. MR2422959 https://doi.org/10.1080/07474940801989202

Tartakovsky, A.G. (2021). An asymptotic theory of joint sequential changepoint detection and identification for general stochastic models. *IEEE Trans. Inf. Theory* **67** 4768–4783. MR4306295 https://doi.org/10.1109/TIT.2021.3064344

Tartakovsky, A.G., Li, X.R. and Yaralov, G. (2003). Sequential detection of targets in multichannel systems. *IEEE Trans. Inf. Theory* **49** 425–445. MR1966790 https://doi.org/10.1109/TIT.2002.807288

Tartakovsky, A., Nikiforov, I. and Basseville, M. (2015). *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. *Monographs on Statistics and Applied Probability* **136**. Boca Raton, FL: CRC Press. MR3241619

Tartakovsky, A.G., Rozovskii, B.L., Blazek, R.B. and Kim, H. (2006b). A novel approach to detection of intrusions in computer networks via adaptive sequential and batch-sequential change-point detection methods. *IEEE Trans. Signal Process.* **54** 3372–3382.

Tartakovsky, A.G., Rozovskii, B.L., Blažek, R.B. and Kim, H. (2006a). Detection of intrusions in information systems by sequential change-point methods. *Stat. Methodol.* **3** 252–293. MR2240956 https://doi.org/10.1016/j.stamet.2005.05.003

Warner, A. and Fellouris, G. (2022). CuSum for sequential change diagnosis. In *2022 IEEE International Symposium on Info. Theory* 486–491. https://doi.org/10.1109/ISIT50566.2022.9834755

Warner, A. and Fellouris, G. (2024). Supplement to "Sequential change diagnosis revisited and the Adaptive Matrix CuSum." https://doi.org/10.3150/23-BEJ1671SUPP

Wei, S. and Xie, Y. (2022). Online kernel CUSUM for change-point detection. arXiv preprint arXiv:2211.15070.

Xie, L., Moustakides, G.V. and Xie, Y. (2023). Window-limited CUSUM for sequential change detection. *IEEE Trans. Inf. Theory* **69** 5990–6005. MR4635159

Xie, Y. and Siegmund, D. (2013). Sequential multi-sensor change-point detection. *Ann. Statist.* **41** 670–692. MR3099117 https://doi.org/10.1214/13-AOS1094

Xie, L., Zou, S., Xie, Y. and Veeravalli, V.V. (2021). Sequential (quickest) change detection: Classical results and new directions. *IEEE J. Sel. Areas Inf. Theory* **2** 494–514.

Yang, H., Hadjiliadis, O. and Ludkovski, M. (2017). Quickest detection in the Wiener disorder problem with post-change uncertainty. *Stochastics* **89** 654–685. MR3607746 https://doi.org/10.1080/17442508.2016.1276908

Bernoulli **30**(3), 2024, 2253–2274
https://doi.org/10.3150/23-BEJ1672

A trajectorial approach to entropy dissipation for degenerate parabolic equations

DONGHAN KIM[1,a] and LANE CHUN YEUNG[2,b]

[1]*Department of Mathematics, University of Michigan, Ann Arbor, USA,* [a]*donghank@umich.edu*
[2]*Department of Industrial Engineering & Operations Research, Columbia University, New York, USA,* [b]*l.yeung@columbia.edu*

We consider degenerate diffusion equations of the form $\partial_t p_t = \Delta f(p_t)$ on a bounded domain and subject to no-flux boundary conditions, for a class of nonlinearities f that includes the porous medium equation. We derive for them a trajectorial analogue of the entropy dissipation identity, which describes the rate of entropy dissipation along every path of the diffusion. In line with the recent work (*Theory Probab. Appl.* **66** (2022) 668–707), our approach is based on applying stochastic analysis to the underlying probabilistic representations, which in our context are stochastic differential equations with normal reflection on the boundary. This trajectorial approach also leads to a new derivation of the Wasserstein gradient flow property for nonlinear diffusions, as well as to a simple proof of the HWI inequality in the present context.

Keywords: Degenerate diffusion; entropy dissipation; gradient flow; HWI inequality; porous medium equation

References

Agueh, M., Ghoussoub, N. and Kang, X. (2004). Geometric inequalities via a general comparison principle for interacting gases. *Geom. Funct. Anal.* **14** 215–244. MR2053603 https://doi.org/10.1007/s00039-004-0455-x

Alt, H.W. and Luckhaus, S. (1983). Quasilinear elliptic-parabolic differential equations. *Math. Z.* **183** 311–341. MR0706391 https://doi.org/10.1007/BF01176474

Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2401600

Barbu, V., Röckner, M. and Russo, F. (2011). Probabilistic representation for solutions of an irregular porous media type equation: The degenerate case. *Probab. Theory Related Fields* **151** 1–43. MR2834711 https://doi.org/10.1007/s00440-010-0291-x

Benachour, S., Chassaing, P., Roynette, B. and Vallois, P. (1996). Processus associés à l'équation des milieux poreux. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (4)* **23** 793–832. MR1469575

Blanchard, P., Röckner, M. and Russo, F. (2010). Probabilistic representation for solutions of an irregular porous media type equation. *Ann. Probab.* **38** 1870–1900. MR2722788 https://doi.org/10.1214/10-AOP526

Carrillo, J.A. and Toscani, G. (2000). Asymptotic L^1-decay of solutions of the porous medium equation to self-similarity. *Indiana Univ. Math. J.* **49** 113–142. MR1777035 https://doi.org/10.1512/iumj.2000.49.1756

Carrillo, J.A., Jüngel, A., Markowich, P.A., Toscani, G. and Unterreiter, A. (2001). Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities. *Monatsh. Math.* **133** 1–82. MR1853037 https://doi.org/10.1007/s006050170032

Ciotir, I. and Russo, F. (2014). Probabilistic representation for solutions of a porous media type equation with Neumann boundary condition: The case of the half-line. *Differential Integral Equations* **27** 181–200. MR3161601

Erbar, M. and Maas, J. (2014). Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.* **34** 1355–1374. MR3117845 https://doi.org/10.3934/dcds.2014.34.1355

Hu, Y., Qian, Z. and Zhang, Z. (2017). Gradient estimates for porous medium and fast diffusion equations by martingale method. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1793–1820. MR3729635 https://doi.org/10.1214/16-AIHP771

Itô, S. (1957). A boundary value problem of partial differential equations of parabolic type. *Duke Math. J.* **24** 299–312. MR0090751

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. MR4273512 https://doi.org/10.4310/CIS.2021.v21.n4.a1

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. MR4466405

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

Lions, P.-L. and Sznitman, A.-S. (1984). Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* **37** 511–537. MR0745330 https://doi.org/10.1002/cpa.3160370408

Lisini, S., Mainini, E. and Segatti, A. (2018). A gradient flow approach to the porous medium equation with fractional pressure. *Arch. Ration. Mech. Anal.* **227** 567–606. MR3740382 https://doi.org/10.1007/s00205-017-1168-2

Otto, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations* **26** 101–174. MR1842429 https://doi.org/10.1081/PDE-100002243

Pilipenko, A. (2014). *An Introduction to Stochastic Differential Equations with Reflection* **1**. Universitätsverlag Potsdam.

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

Sato, K. and Tanaka, H. (1962). Local times on the boundary for multi-dimensional reflecting diffusion. *Proc. Jpn. Acad.* **38** 699–702. MR0148109

Soner, H.M. (2007). Stochastic representations for nonlinear parabolic PDEs. In *Handbook of Differential Equations: Evolutionary Equations. Vol. III*. *Handb. Differ. Equ.* 477–526. Amsterdam: Elsevier/North-Holland. MR2549373 https://doi.org/10.1016/S1874-5717(07)80009-0

Teschl, G. (2012). *Ordinary Differential Equations and Dynamical Systems*. *Graduate Studies in Mathematics* **140**. Providence, RI: Amer. Math. Soc. MR2961944 https://doi.org/10.1090/gsm/140

Tschiderer, B. and Chun Yeung, L. (2023). A trajectorial approach to relative entropy dissipation of McKean-Vlasov diffusions: Gradient flows and HWBI inequalities. *Bernoulli* **29** 725–756. MR4497265 https://doi.org/10.3150/22-bej1476

Vázquez, J.L. (2007). *The Porous Medium Equation: Mathematical Theory*. *Oxford Mathematical Monographs*. Oxford: The Clarendon Press. MR2286292

Villani, C. (2003). *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Providence, RI: Amer. Math. Soc. MR1964483 https://doi.org/10.1090/gsm/058

Bernoulli **30**(3), 2024, 2275–2300
https://doi.org/10.3150/23-BEJ1673

Adaptive inference over Besov spaces in the white noise model using *p*-exponential priors

SERGIOS AGAPIOU[a] and AIMILIA SAVVA[b]

Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus, [a]*agapiou.sergios@ucy.ac.cy,* [b]*savva.emilia@ucy.ac.cy*

In many scientific applications the aim is to infer a function which is smooth in some areas, but rough or even discontinuous in other areas of its domain. Such *spatially inhomogeneous* functions can be modelled in Besov spaces with suitable integrability parameters. In this work we study adaptive Bayesian inference over Besov spaces, in the white noise model from the point of view of rates of contraction, using *p-exponential* priors, which range between Laplace and Gaussian and possess regularity and scaling hyper-parameters. To achieve adaptation, we employ empirical and hierarchical Bayes approaches for tuning these hyper-parameters. Our results show that, while it is known that Gaussian priors can attain the minimax rate *only* in Besov spaces of spatially homogeneous functions, Laplace priors lead to adaptive or nearly adaptive procedures in *both* Besov spaces of spatially homogeneous functions *and* Besov spaces permitting spatial inhomogeneities.

Keywords: Adaptation; Besov spaces; empirical Bayes; Gaussian prior; hierarchical Bayes; Laplace prior; posterior contraction rates; spatially inhomogeneous functions; white noise model

References

[1] Abramovich, F., Sapatinas, T. and Silverman, B.W. (1998). Wavelet thresholding via a Bayesian approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 725–749. MR1649547 https://doi.org/10.1111/1467-9868.00151

[2] Agapiou, S., Bardsley, J.M., Papaspiliopoulos, O. and Stuart, A.M. (2014). Analysis of the Gibbs sampler for hierarchical inverse problems. *SIAM/ASA J. Uncertain. Quantificat.* **2** 511–544. MR3283919 https://doi.org/10.1137/130944229

[3] Agapiou, S., Burger, M., Dashti, M. and Helin, T. (2018). Sparsity-promoting and edge-preserving maximum *a posteriori* estimators in non-parametric Bayesian inverse problems. *Inverse Probl.* **34** 045002. MR3774703 https://doi.org/10.1088/1361-6420/aaacac

[4] Agapiou, S., Dashti, M. and Helin, T. (2021). Rates of contraction of posterior distributions based on *p*-exponential priors. *Bernoulli* **27** 1616–1642. MR4278794 https://doi.org/10.3150/20-bej1285

[5] Agapiou, S. and Savva, A. (2024). Supplement to "Adaptive inference over Besov spaces in the white noise model using *p*-exponential priors." https://doi.org/10.3150/23-BEJ1673SUPP

[6] Agapiou, S. and Wang, S. (2021). Laplace priors and spatial inhomogeneity in Bayesian inverse problems. *Bernoulli*, To appear. Available at arXiv:2112.05679 (preprint).

[7] Andrieu, C. and Roberts, G.O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. *Ann. Statist.* **37** 697–725. MR2502648 https://doi.org/10.1214/07-AOS574

[8] Aurzada, F. (2007). On the lower tail probabilities of some random sequences in l_p. *J. Theoret. Probab.* **20** 843–858. MR2359058 https://doi.org/10.1007/s10959-007-0095-9

[9] Belitser, E. and Enikeeva, F. (2008). Empirical Bayesian test of the smoothness. *Math. Methods Statist.* **17** 1–18. MR2400361 https://doi.org/10.3103/S1066530708010018

[10] Beskos, A., Girolami, M., Lan, S., Farrell, P.E. and Stuart, A.M. (2017). Geometric MCMC for infinite-dimensional inverse problems. *J. Comput. Phys.* **335** 327–351. MR3612501 https://doi.org/10.1016/j.jcp.2016.12.041

[11] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958 https://doi.org/10.1214/aos/1032181159

1350-7265

[12] Chen, V., Dunlop, M.M., Papaspiliopoulos, O. and Stuart, A.M. (2018). Robust MCMC sampling with non-Gaussian and hierarchical priors in high dimensions. Available at arXiv:1803.03344 (preprint).
[13] Cotter, S.L., Roberts, G.O., Stuart, A.M. and White, D. (2013). MCMC methods for functions: Modifying old algorithms to make them faster. *Statist. Sci.* **28** 424–446. MR3135540 https://doi.org/10.1214/13-STS421
[14] Cui, T., Law, K.J.H. and Marzouk, Y.M. (2016). Dimension-independent likelihood-informed MCMC. *J. Comput. Phys.* **304** 109–137. MR3422405 https://doi.org/10.1016/j.jcp.2015.10.008
[15] Dashti, M., Harris, S. and Stuart, A. (2012). Besov priors for Bayesian inverse problems. *Inverse Probl. Imaging* **6** 183–200. MR2942737 https://doi.org/10.3934/ipi.2012.6.183
[16] Dashti, M. and Stuart, A.M. (2017). The Bayesian approach to inverse problems. In *Handbook of Uncertainty Quantification. Vol. 1, 2, 3* 311–428. Cham: Springer. MR3839555
[17] Donnet, S., Rivoirard, V., Rousseau, J. and Scricciolo, C. (2018). Posterior concentration rates for empirical Bayes procedures with applications to Dirichlet process mixtures. *Bernoulli* **24** 231–256. MR3706755 https://doi.org/10.3150/16-BEJ872
[18] Donoho, D.L. and Johnstone, I.M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika* **81** 425–455. MR1311089 https://doi.org/10.1093/biomet/81.3.425
[19] Donoho, D.L. and Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.* **90** 1200–1224. MR1379464
[20] Donoho, D.L. and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. MR1635414 https://doi.org/10.1214/aos/1024691081
[21] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228
[22] Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist.* **35** 192–223. MR2332274 https://doi.org/10.1214/009053606000001172
[23] Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR3587782 https://doi.org/10.1017/9781139029834
[24] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics*. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862
[25] Giordano, M. Besov priors in density estimation: Optimal posterior contraction rates and adaptation. Available at arXiv:2208.14350 (preprint).
[26] Giordano, M. and Nickl, R. (2020). Consistency of Bayesian inference with Gaussian process priors in an elliptic inverse problem. *Inverse Probl.* **36** 085001. MR4151406 https://doi.org/10.1088/1361-6420/ab7d2a
[27] Giordano, M. and Ray, K. (2022). Nonparametric Bayesian inference for reversible multidimensional diffusions. *Ann. Statist.* **50** 2872–2898. MR4500628 https://doi.org/10.1214/22-aos2213
[28] Johnstone, I.M. (2019). *Gaussian Estimation: Sequence and Wavelet Models*. Book draft.
[29] Knapik, B.T., Szabó, B.T., van der Vaart, A.W. and van Zanten, J.H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probab. Theory Related Fields* **164** 771–813. MR3477780 https://doi.org/10.1007/s00440-015-0619-7
[30] Kolehmainen, V., Lassas, M., Niinimäki, K. and Siltanen, S. (2012). Sparsity-promoting Bayesian inversion. *Inverse Probl.* **28** 025005. MR2876856 https://doi.org/10.1088/0266-5611/28/2/025005
[31] Lassas, M., Saksman, E. and Siltanen, S. (2009). Discretization-invariant Bayesian inversion and Besov space priors. *Inverse Probl. Imaging* **3** 87–122. MR2558305 https://doi.org/10.3934/ipi.2009.3.87
[32] Lassas, M. and Siltanen, S. (2004). Can one use total variation prior for edge-preserving Bayesian inversion? *Inverse Probl.* **20** 1537–1563. MR2109134 https://doi.org/10.1088/0266-5611/20/5/013
[33] Lepski, O.V., Mammen, E. and Spokoiny, V.G. (1997). Optimal spatial adaptation to inhomogeneous smoothness: An approach based on kernel estimates with variable bandwidth selectors. *Ann. Statist.* **25** 929–947. MR1447734 https://doi.org/10.1214/aos/1069362731
[34] Papaspiliopoulos, O., Roberts, G.O. and Sköld, M. (2007). A general framework for the parametrization of hierarchical models. *Statist. Sci.* **22** 59–73. MR2408661 https://doi.org/10.1214/088342307000000014
[35] Rockova, V. and Rousseau, J. (2021). Ideal Bayesian spatial adaptation. Available at arXiv:2105.12793 (preprint).

[36] Rousseau, J. and Szabo, B. (2017). Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator. *Ann. Statist.* **45** 833–865. MR3650402 https://doi.org/10.1214/16-AOS1469

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

[38] Szabó, B.T., van der Vaart, A.W. and van Zanten, J.H. (2013). Empirical Bayes scaling of Gaussian priors in the white noise model. *Electron. J. Stat.* **7** 991–1018. MR3044507 https://doi.org/10.1214/13-EJS798

[39] Talagrand, M. (1994). The supremum of some canonical processes. *Amer. J. Math.* **116** 283–325. MR1269606 https://doi.org/10.2307/2374931

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

Bernoulli **30**(3), 2024, 2301–2325
https://doi.org/10.3150/23-BEJ1674

An asymptotic Peskun ordering and its application to lifted samplers

PHILIPPE GAGNON[a] and FLORIAN MAIRE[b]

Department of Mathematics and Statistics, Université de Montréal, Montréal, Canada,
[a]*philippe.gagnon.3@umontreal.ca,* [b]*florian.maire@umontreal.ca*

A Peskun ordering between two samplers, implying a dominance of one over the other, is known among the Markov chain Monte Carlo community for being a remarkably strong result. It is however also known for being a result that is notably difficult to establish. Indeed, one has to prove that the probability to reach a state $\mathbf{y}$ from a state $\mathbf{x}$, using a sampler, is greater than or equal to the probability using the other sampler, and this must hold for all pairs $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{x} \neq \mathbf{y}$. We provide in this paper a weaker version that does not require an inequality between the probabilities for all these states: essentially, the dominance holds asymptotically, as a varying parameter grows without bound, as long as the states for which the probabilities are greater than or equal to belong to a mass-concentrating set. The weak ordering turns out to be useful to compare *lifted* samplers for *partially-ordered* discrete state-spaces with their Metropolis–Hastings counterparts. An analysis in great generality yields a qualitative conclusion: they asymptotically perform better in certain situations (and we are able to identify them), but not necessarily in others (and the reasons why are made clear). A quantitative study in a specific context of graphical-model simulation is also conducted.

Keywords: Bayesian statistics; binary random variables; Ising model; Markov chain Monte Carlo methods; variable selection

References

Andrieu, C., Lee, A. and Vihola, M. (2018). Uniform ergodicity of the iterated conditional SMC and geometric ergodicity of particle Gibbs samplers. *Bernoulli* **24** 842–872. MR3706778 https://doi.org/10.3150/15-BEJ785

Andrieu, C. and Livingstone, S. (2021). Peskun–Tierney ordering for Markovian Monte Carlo: Beyond the reversible scenario. *Ann. Statist.* **49** 1958–1981. MR4319237 https://doi.org/10.1214/20-aos2008

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

Barker, A.A. (1965). Monte Carlo calculations of the radial distribution functions for a proton–electron plasma. *Aust. J. Phys.* **18** 119–134.

Bierkens, J. (2016). Non-reversible Metropolis–Hastings. *Stat. Comput.* **26** 1213–1228. MR3538633 https://doi.org/10.1007/s11222-015-9598-x

Chen, F., Lovász, L. and Pak, I. (1999). Lifting Markov chains to speed up mixing. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)* 275–281. New York: ACM. MR1798046 https://doi.org/10.1145/301250.301315

Diaconis, P., Holmes, S. and Neal, R.M. (2000). Analysis of a nonreversible Markov chain sampler. *Ann. Appl. Probab.* **10** 726–752. MR1789978 https://doi.org/10.1214/aoap/1019487508

Faizi, F., Deligiannidis, G. and Rosta, E. (2020). Efficient irreversible Monte Carlo samplers. *J. Chem. Theory Comput.* **16** 2124–2138. https://doi.org/10.1021/acs.jctc.9b01135

Gagnon, P. and Doucet, A. (2021). Nonreversible jump algorithms for Bayesian nested model selection. *J. Comput. Graph. Statist.* **30** 312–323. MR4270506 https://doi.org/10.1080/10618600.2020.1826955

Gagnon, P. and Maire, F. (2024). Supplement to "An asymptotic Peskun ordering and its application to lifted samplers." https://doi.org/10.3150/23-BEJ1674SUPP

Gustafson, P. (1998). A guided walk Metropolis algorithm. *Stat. Comput.* **8** 357–364.

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

Herschlag, G., Mattingly, J.C., Sachs, M. and Wyse, E. (2020). Non-reversible Markov chain Monte Carlo for sampling of districting maps. arXiv:2008.07843.

Horowitz, A.M. (1991). A generalized guided Monte Carlo algorithm. *Phys. Lett. B* **268** 247–252.

Kamatani, K. and Song, X. (2023). Non-reversible guided Metropolis kernel. *J. Appl. Probab.* **60** 955–981. MR4624051 https://doi.org/10.1017/jpr.2022.109

Kleijn, B.J.K. and Van der Vaart, A.W. (2012). The Bernstein–Von-Mises theorem under misspecification. *Electron. J. Stat.* **6** 354–381.

Livingstone, S. and Zanella, G. (2022). The Barker proposal: Combining robustness and efficiency in gradient-based MCMC. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 496–523. MR4412995 https://doi.org/10.1111/rssb.12482

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21** 1087.

Peskun, P.H. (1973). Optimum Monte-Carlo sampling using Markov chains. *Biometrika* **60** 607–612. MR0362823 https://doi.org/10.1093/biomet/60.3.607

Power, S. and Goldman, J.V. (2019). Accelerated sampling on discrete spaces with non-reversible Markov Processes. arXiv:1912.04681.

Sakai, Y. and Hukushima, K. (2016a). Irreversible simulated tempering. *J. Phys. Soc. Jpn.* **85** 104002.

Sakai, Y. and Hukushima, K. (2016b). Eigenvalue analysis of an irreversible random walk with skew detailed balance conditions. *Phys. Rev. E* **93** 043318. https://doi.org/10.1103/PhysRevE.93.043318

Syed, S., Bouchard-Côté, A., Deligiannidis, G. and Doucet, A. (2022). Non-reversible parallel tempering: A scalable highly parallel MCMC scheme. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 321–350. MR4412989

Tierney, L. (1998). A note on Metropolis–Hastings kernels for general state spaces. *Ann. Appl. Probab.* **8** 1–9. MR1620401 https://doi.org/10.1214/aoap/1027961031

van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

Yang, J. and Rosenthal, J.S. (2023). Complexity results for MCMC derived from quantitative bounds. *Ann. Appl. Probab.* **33** 1259–1300. MR4564431 https://doi.org/10.1214/22-aap1846

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

Bernoulli **30**(3), 2024, 2326–2345
https://doi.org/10.3150/23-BEJ1675

Low-rank matrix recovery under heavy-tailed errors

MYEONGHUN YU[1,a], QIANG SUN[2,b] and WEN-XIN ZHOU[3,c]

[1]*Department of Mathematics, University of California, San Diego, La Jolla, CA, 92093, USA,* [a]*myyu@ucsd.edu*
[2]*Department of Statistical Sciences, University of Toronto, Toronto, ON M5G 1Z5, Canada,*
[b]*qiang.sun@utoronto.ca*
[3]*Department of Information and Decision Sciences, University of Illinois at Chicago, Chicago, IL, 60607, USA,*
[c]*wenxinz@uic.edu*

This paper proposes convex relaxation based robust methods to recover approximately low-rank matrices in the presence of heavy-tailed and asymmetric errors, allowing for heteroscedasticity. We focus on three archetypal applications in matrix recovery: matrix compressed sensing, matrix completion and multitask regression. Statistically, we provide sub-Gaussian-type deviation bounds when the noise variables only have bounded variances in each aforementioned setting. Improving upon the earlier results in Fan, Wang and Zhu (*Ann. Statist.* **49** (2021) 1239–1266), the convergence rates of our estimators are proportional to the noise scale under matrix sensing and multitask regression settings, and thus diminish to 0 in the noiseless case. Computationally, we propose a matrix version of the local adaptive majorize-minimization algorithm, which is much faster than the alternating direction method of multiplier used in previous work and is scalable to large datasets. Numerical experiments demonstrate the advantage of our methods over their non-robust counterparts and corroborate the theoretical findings that the convergence rates are proportional to the noise scale.

Keywords: Heavy-tailed data; Huber loss; low-rank matrix recovery; nuclear norm; trace regression

References

[1] Argyriou, A., Evgeniou, T. and Pontil, M. (2008). Convex multi-task feature learning. *Mach. Learn.* **73** 243–272.
[2] Burer, S. and Monteiro, R.D.C. (2003). A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Math. Program.* **95** 329–357. MR1976484 https://doi.org/10.1007/s10107-002-0352-8
[3] Cai, J.-F., Candès, E.J. and Shen, Z. (2010). A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* **20** 1956–1982. MR2600248 https://doi.org/10.1137/080738970
[4] Candès, E.J. and Plan, Y. (2009). Matrix completion with noise. *Proc. IEEE* **98** 925–936.
[5] Candès, E.J. and Plan, Y. (2011). Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements. *IEEE Trans. Inf. Theory* **57** 2342–2359. MR2809094 https://doi.org/10.1109/TIT.2011.2111771
[6] Candès, E.J. and Recht, B. (2009). Exact matrix completion via convex optimization. *Found. Comput. Math.* **9** 717–772. MR2565240 https://doi.org/10.1007/s10208-009-9045-5
[7] Chen, J., Liu, D. and Li, X. (2020). Nonconvex rectangular matrix completion via gradient descent without $\ell_{2,\infty}$ regularization. *IEEE Trans. Inf. Theory* **66** 5806–5841. MR4158648 https://doi.org/10.1109/TIT.2020.2992234
[8] Chen, Y., Jalali, A., Sanghavi, S. and Caramanis, C. (2013). Low-rank matrix recovery from errors and erasures. *IEEE Trans. Inf. Theory* **59** 4324–4337.
[9] Dalalyan, A. and Thompson, P. (2019). Outlier-robust estimation of a sparse linear model using ℓ_1-penalized Huber's M-estimator. In *Advances in Neural Information Processing Systems* **32** 13188–13198.

[10] Elsener, A. and van de Geer, S. (2018). Robust low-rank matrix estimation. *Ann. Statist.* **46** 3481–3509. MR3852659 https://doi.org/10.1214/17-AOS1666

[11] Fan, J., Li, Q. and Wang, Y. (2017). Estimation of high dimensional mean regression in the absence of symmetry and light tail assumptions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 247–265. MR3597972 https://doi.org/10.1111/rssb.12166

[12] Fan, J., Liu, H., Sun, Q. and Zhang, T. (2018). I-LAMM for sparse learning: Simultaneous control of algorithmic complexity and statistical error. *Ann. Statist.* **46** 814–841. MR3782385 https://doi.org/10.1214/17-AOS1568

[13] Fan, J., Wang, W. and Zhu, Z. (2021). A shrinkage principle for heavy-tailed data: High-dimensional robust low-rank matrix recovery. *Ann. Statist.* **49** 1239–1266. MR4298863 https://doi.org/10.1214/20-aos1980

[14] Goldberg, D., Nichols, D., Oki, B.M. and Terry, D. (1992). Using collaborative filtering to weave an information tapestry. *Commun. ACM* **35** 61–70.

[15] Gross, D., Liu, Y.-K., Flammia, S.T., Becker, S. and Eisert, J. (2010). Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105** 150401. https://doi.org/10.1103/PhysRevLett.105.150401

[16] Huber, P.J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Ann. Statist.* **1** 799–821. MR0356373

[17] Izenman, A.J. (1975). Reduced-rank regression for the multivariate linear model. *J. Multivariate Anal.* **5** 248–264. MR0373179 https://doi.org/10.1016/0047-259X(75)90042-1

[18] Klopp, O. (2014). Noisy low-rank matrix completion with general sampling distribution. *Bernoulli* **20** 282–303. MR3160583 https://doi.org/10.3150/12-BEJ486

[19] Klopp, O., Lounici, K. and Tsybakov, A.B. (2017). Robust matrix completion. *Probab. Theory Related Fields* **169** 523–564. MR3704775 https://doi.org/10.1007/s00440-016-0736-y

[20] Koltchinskii, V., Lounici, K. and Tsybakov, A.B. (2011). Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Ann. Statist.* **39** 2302–2329. MR2906869 https://doi.org/10.1214/11-AOS894

[21] Li, X. (2013). Compressed sensing and matrix completion with constant proportion of corruptions. *Constr. Approx.* **37** 73–99. MR3010211 https://doi.org/10.1007/s00365-012-9176-9

[22] Li, Y., Ma, T. and Zhang, H. (2018). Algorithmic regularization in over-parameterized matrix sensing and neural networks with quadratic activations. In *Conference on Learning Theory* 2–47.

[23] Lounici, K., Pontil, M., van de Geer, S. and Tsybakov, A.B. (2011). Oracle inequalities and optimal inference under group sparsity. *Ann. Statist.* **39** 2164–2204. MR2893865 https://doi.org/10.1214/11-AOS896

[24] Luan, X., Fang, B., Liu, L., Yang, W. and Qian, J. (2014). Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion. *Pattern Recognit.* **47** 495–508.

[25] Ma, C., Wang, K., Chi, Y. and Chen, Y. (2020). Implicit regularization in nonconvex statistical estimation: Gradient descent converges linearly for phase retrieval, matrix completion, and blind deconvolution. *Found. Comput. Math.* **20** 451–632. MR4099988 https://doi.org/10.1007/s10208-019-09429-9

[26] Ma, J. and Fattahi, S. (2023). Global convergence of sub-gradient method for robust matrix recovery: Small initialization, noisy measurements, and over-parameterization. *J. Mach. Learn. Res.* **24** Paper No. [96], 84. MR4582518

[27] Minsker, S. (2018). Sub-Gaussian estimators of the mean of a random matrix with heavy-tailed entries. *Ann. Statist.* **46** 2871–2903. MR3851758 https://doi.org/10.1214/17-AOS1642

[28] Negahban, S. and Wainwright, M.J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *Ann. Statist.* **39** 1069–1097. MR2816348 https://doi.org/10.1214/10-AOS850

[29] Negahban, S. and Wainwright, M.J. (2012). Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *J. Mach. Learn. Res.* **13** 1665–1697. MR2930649

[30] Recht, B., Fazel, M. and Parrilo, P.A. (2010). Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **52** 471–501. MR2680543 https://doi.org/10.1137/070697835

[31] Rohde, A. and Tsybakov, A.B. (2011). Estimation of high-dimensional low-rank matrices. *Ann. Statist.* **39** 887–930. MR2816342 https://doi.org/10.1214/10-AOS860

[32] She, Y. and Chen, K. (2017). Robust reduced-rank regression. *Biometrika* **104** 633–647. MR3694587 https://doi.org/10.1093/biomet/asx032

[33] Shen, Y., Li, J., Cai, J. and Xia, D. (2022). Computationally efficient and statistically optimal robust low-rank matrix estimation. arXiv:2203.00953.

[34] Sun, Q., Zhou, W.-X. and Fan, J. (2020). Adaptive Huber regression. *J. Amer. Statist. Assoc.* **115** 254–265. MR4078461 https://doi.org/10.1080/01621459.2018.1543124
[35] Tan, K.M., Sun, Q. and Witten, D. (2022). Sparse reduced rank Huber regression in high dimensions. *J. Amer. Statist. Assoc.* To appear.
[36] Thompson, P. (2020). Outlier-robust sparse/low-rank least-squares regression and robust matrix completion. arXiv:2012.06750.
[37] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. MR1379242
[38] Tong, T., Ma, C. and Chi, Y. (2021). Accelerating ill-conditioned low-rank matrix estimation via scaled gradient descent. *J. Mach. Learn. Res.* **22** Paper No. 150, 63. MR4318506 https://doi.org/10.1080/15502287.2020.1856971
[39] Trefethen, L.N. and Bau, D. III (1997). *Numerical Linear Algebra*. Philadelphia, PA: SIAM. MR1444820 https://doi.org/10.1137/1.9780898719574
[40] Wang, B. and Fan, J. (2022). Robust matrix completion with heavy-tailed noise. arXiv:2206.04276.
[41] Wei, K., Cai, J.-F., Chan, T.F. and Leung, S. (2016). Guarantees of Riemannian optimization for low rank matrix recovery. *SIAM J. Matrix Anal. Appl.* **37** 1198–1222. MR3543156 https://doi.org/10.1137/15M1050525
[42] Yu, M., Sun, Q. and Zhou, W.-X. (2024). Supplement to “Low-rank matrix recovery under heavy-tailed errors.” https://doi.org/10.3150/23-BEJ1675SUPP
[43] Zhang, J., Fattahi, S. and Zhang, R. (2021). Preconditioned gradient descent for over-parameterized nonconvex matrix factorization. In *Advances in Neural Information Processing Systems* **34** 5985–5996.

Bernoulli **30**(3), 2024, 2346–2371
https://doi.org/10.3150/23-BEJ1676

When scattering transform meets non-Gaussian random processes, a double scaling limit result

GI-REN LIU[1,a], YUAN-CHUNG SHEU[2,b] and HAU-TIENG WU[3,c]

[1]*Department of Mathematics, National Cheng-Kung University, Tainan, Taiwan,* [a]*girenliu@ncku.edu.tw*
[2]*Department of Applied Mathematics, National Yang Ming Chiao Tung University, Hsinchu, Taiwan,* [b]*sheu@math.nctu.edu.tw*
[3]*Department of Mathematics and Department of Statistical Science, Duke University, Durham, NC, USA,* [c]*hauwu@math.duke.edu*

Let T be a function of Hermite rank one and let $\{G(t)\}_{t\in\mathbb{R}}$ be a mean-square continuous stationary Gaussian process with long-range dependence. We explore the finite-dimensional distributions of the second-order scattering transform of the process $X = \{T(G(t))\}_{t\in\mathbb{R}}$ when all the scale parameters go to infinity simultaneously. For frequently used wavelets, we find a constraint on the ratio of the scale parameters of the wavelet transform within the first and second layers such that the limit exists. The constraint is explicitly expressed in terms of the Hurst index of the long-range dependent inputs and the gap between the indices of the first and second non-zero coefficients in the Hermite expansion of the function T. Under the constraint on the ratio of the scale parameters, we prove that the rescaled second-order scattering transform converges in the finite-dimensional distribution sense to a chi process of degree one. The limiting process is expressed in terms of the Fourier transform of mother wavelet and the Hurst index of long-range dependence.

Keywords: Double scaling limits; Feynman diagram; long-range dependent processes; non-Gaussian processes; scattering transform; wavelet transform; Wiener-Itô decomposition

References

[1] Andén, J. and Mallat, S. (2011). Multiscale scattering for audio classification. In *Proc. ISMIR* 657–662. Miami, FL, USA.

[2] Andén, J. and Mallat, S. (2014). Deep scattering spectrum. *IEEE Trans. Signal Process.* **62** 4114–4128. MR3260414 https://doi.org/10.1109/TSP.2014.2326991

[3] Anh, V., Leonenko, N. and Olenko, A. (2015). On the rate of convergence to Rosenblatt-type distribution. *J. Math. Anal. Appl.* **425** 111–132. MR3299653 https://doi.org/10.1016/j.jmaa.2014.12.016

[4] Anh, V.V., Leonenko, N.N. and Ruiz-Medina, M.D. (2013). Macroscaling limit theorems for filtered spatiotemporal random fields. *Stoch. Anal. Appl.* **31** 460–508. MR3049079 https://doi.org/10.1080/07362994.2013.777280

[5] Bai, S. and Taqqu, M.S. (2018). How the instability of ranks under long memory affects large-sample inference. *Statist. Sci.* **33** 96–116. MR3757507 https://doi.org/10.1214/17-STS633

[6] Bai, S. and Taqqu, M.S. (2019). Sensitivity of the Hermite rank. *Stochastic Process. Appl.* **129** 822–840. MR3913269 https://doi.org/10.1016/j.spa.2018.03.020

[7] Balestriero, R. and Glotin, H. (2017). Linear time complexity deep Fourier scattering network and extension to nonlinear invariants. Available at: https://arxiv.org/abs/1707.05841.

[8] Beran, J., Möhrle, S. and Ghosh, S. (2016). Testing for Hermite rank in Gaussian subordination processes. *J. Comput. Graph. Statist.* **25** 917–934. MR3533645 https://doi.org/10.1080/10618600.2015.1056345

[9] Berry, R.B., Brooks, R., Gamaldo, C.E., Harding, S.M., Marcus, C. and Vaughn, B.V. (2012). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules*. *Terminol. Tech. Specificat.* Darien, IL: American Academy of Sleep Medicine.

1350-7265

[10] Bleher, P. and Its, A. (2003). Double scaling limit in the random matrix model: The Riemann-Hilbert approach. *Comm. Pure Appl. Math.* **56** 433–516. MR1949138 https://doi.org/10.1002/cpa.10065
[11] Bleher, P.M. and Kuijlaars, A.B.J. (2007). Large *n* limit of Gaussian random matrices with external source. III. Double scaling limit. *Comm. Math. Phys.* **270** 481–517. MR2276453 https://doi.org/10.1007/s00220-006-0159-1
[12] Breuer, P. and Major, P. (1983). Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13** 425–441. MR0716933 https://doi.org/10.1016/0047-259X(83)90019-2
[13] Bruna, J., Mallat, S., Bacry, E. and Muzy, J.-F. (2015). Intermittent process analysis with scattering moments. *Ann. Statist.* **43** 323–351. MR3311862 https://doi.org/10.1214/14-AOS1276
[14] Chudáček, V., Andén, J., Mallat, S., Abry, P. and Doret, M. (2013). Scattering transform for intrapartum fetal heart rate variability fractal analysis: A case-control study. *IEEE Trans. Biomed. Eng.* **61** 1100–1108.
[15] Couillet, R. and McKay, M. (2014). Large dimensional analysis and optimization of robust shrinkage covariance matrix estimators. *J. Multivariate Anal.* **131** 99–120. MR3252638 https://doi.org/10.1016/j.jmva.2014.06.018
[16] Daubechies, I. (1992). *Ten Lectures on Wavelets*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **61**. Philadelphia, PA: SIAM. MR1162107 https://doi.org/10.1137/1.9781611970104
[17] Giveon, A. and Kutasov, D. (1999). Little string theory in a double scaling limit. *J. High Energy Phys.* **10** Paper 34, 22. MR1726444 https://doi.org/10.1088/1126-6708/1999/10/034
[18] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M. et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **101** 215–220.
[19] Houdré, C., Pérez-Abreu, V. and Üstünel, A.S. (1994). Multiple Wiener-Itô integrals: An introductory survey. In *Chaos Expansions, Multiple Wiener-Itô Integrals and Their Applications (Guanajuato, 1992)*. *Probab. Stochastics Ser.* 1–33. Boca Raton, FL: CRC. MR1278036
[20] Itô, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–169. MR0044064
[21] Ivanov, A.V., Leonenko, N., Ruiz-Medina, M.D. and Savich, I.N. (2013). Limit theorems for weighted nonlinear transformations of Gaussian stationary processes with singular spectra. *Ann. Probab.* **41** 1088–1114. MR3077537 https://doi.org/10.1214/12-AOP775
[22] Ivanov, A.V. and Leonenko, N.N. (2012). *Statistical Analysis of Random Fields 28*. Berlin: Springer. MR1009786
[23] Kawabata, N. (1973). A nonstationary analysis of the electroencephalogram. *IEEE Trans. Biomed. Eng.* **20** 444–452. https://doi.org/10.1109/TBME.1973.324218
[24] Krylov, N.V. (2002). *Introduction to the Theory of Random Processes*. *Graduate Studies in Mathematics* **43**. Providence, RI: Amer. Math. Soc. MR1885884 https://doi.org/10.1090/gsm/043
[25] Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Statist.* **40** 1024–1060. MR2985942 https://doi.org/10.1214/12-AOS989
[26] Leonenko, N. (1999). *Limit Theorems for Random Fields with Singular Spectrum*. *Mathematics and Its Applications* **465**. Dordrecht: Kluwer Academic. MR1687092 https://doi.org/10.1007/978-94-011-4607-4
[27] Li, J., Ke, L., Du, Q., Ding, X., Chen, X. and Wang, D. (2019). Heart sound signal classification algorithm: A combination of wavelet scattering transform and twin support vector machine. *IEEE Access* **7** 179339–179348.
[28] Liu, G.-R., Lin, T.-Y., Wu, H.-T., Sheu, Y.-C., Liu, C.-L., Liu, W.-T. et al. (2021). Large-scale assessment of consistency in sleep stage scoring rules among multiple sleep centers using an interpretable machine learning algorithm. *J. Clin. Sleep Med.* **17** 159–166.
[29] Liu, G.-R., Lo, Y.-L., Malik, J., Sheu, Y.-C. and Wu, H.-T. (2020). Diffuse to fuse EEG spectra-intrinsic geometry of sleep dynamics for classification. *Biomed. Signal Process. Control* **55** 101576.
[30] Liu, G.-R., Sheu, Y.-C. and Wu, H.-T. (2022). Asymptotic analysis of higher-order scattering transform of Gaussian processes. *Electron. J. Probab.* **27** Paper No. 48, 27. MR4408124 https://doi.org/10.1214/22-ejp766
[31] Liu, G.-R., Sheu, Y.-C. and Wu, H.-T. (2023). Central and noncentral limit theorems arising from the scattering transform and its neural activation generalization. *SIAM J. Math. Anal.* **55** 1170–1213. MR4579725 https://doi.org/10.1137/21M1454511
[32] Liu, G.-R. and Shieh, N.-R. (2015). Multi-scaling limits for relativistic diffusion equations with random initial data. *Trans. Amer. Math. Soc.* **367** 3423–3446. MR3314812 https://doi.org/10.1090/S0002-9947-2014-06498-2

[33] Major, P. (2014). *Multiple Wiener-Itô Integrals*, 2nd ed. *Lecture Notes in Math.* **849**. Cham: Springer. MR3155040 https://doi.org/10.1007/978-3-319-02642-8

[34] Mallat, S. (2012). Group invariant scattering. *Comm. Pure Appl. Math.* **65** 1331–1398. MR2957703 https://doi.org/10.1002/cpa.21413

[35] Menéndez, P., Ghosh, S., Künsch, H.R. and Tinner, W. (2013). On trend estimation under monotone Gaussian subordination with long-memory: Application to fossil pollen series. *J. Nonparametr. Stat.* **25** 765–785. MR3174296 https://doi.org/10.1080/10485252.2013.826357

[36] Meyer, Y. (1992). *Wavelets and Operators*. *Cambridge Studies in Advanced Mathematics* **37**. Cambridge: Cambridge Univ. Press. MR1228209

[37] Nualart, D. and Peccati, G. (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** 177–193. MR2118863 https://doi.org/10.1214/009117904000000621

[38] Peccati, G. and Tudor, C.A. (2005). Gaussian limits for vector-valued multiple stochastic integrals. In *Séminaire de Probabilités XXXVIII*. *Lecture Notes in Math.* **1857** 247–262. Berlin: Springer. MR2126978 https://doi.org/10.1007/978-3-540-31449-3_17

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

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

[41] Taqqu, M.S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrsch. Verw. Gebiete* **31** 287–302. MR0400329 https://doi.org/10.1007/BF00532868

[42] Wiatowski, T. and Bölcskei, H. (2018). A mathematical theory of deep convolutional neural networks for feature extraction. *IEEE Trans. Inf. Theory* **64** 1845–1866. MR3766318 https://doi.org/10.1109/TIT.2017.2776228

[43] Wu, H.-T., Talmon, R. and Lo, Y.-L. (2015). Assess sleep stage by modern signal processing techniques. *IEEE Trans. Biomed. Eng.* **62** 1159–1168. https://doi.org/10.1109/TBME.2014.2375292

Bernoulli **30**(3), 2024, 2372–2398
https://doi.org/10.3150/23-BEJ1677

Multiscale jump testing and estimation under complex temporal dynamics

WEICHI WU[1,a] and ZHOU ZHOU[2,b]

[1]*Center for Statistical Science, Department of Industrial Engineering, Tsinghua University, China,* [a]*wuweichi@mail.tsinghua.edu*
[2]*Department of Statistical Science, University of Toronto, Canada,* [b]*zhou@utstat.toronto.edu*

We consider the problem of detecting jumps in an otherwise smoothly evolving trend whilst the covariance and higher-order structures of the system can experience both smooth and abrupt changes over time. The number of jump points is allowed to diverge to infinity with the jump sizes possibly shrinking to zero. The method is based on a multiscale application of an optimal jump-pass filter to the time series, where the scales are dense between admissible lower and upper bounds. For a wide class of non-stationary time series models and trend functions, the proposed method is shown to be able to detect all jump points within a nearly optimal range with a prescribed probability asymptotically under mild conditions. For a time series of length n, the computational complexity of the proposed method is $O(n)$ for each scale and $O(n\log^{1+\epsilon} n)$ overall, where ϵ is an arbitrarily small positive constant. Numerical studies show that the proposed jump testing and estimation method performs robustly and accurately under complex temporal dynamics.

Keywords: Diverging number of jumps; local CUSUM procedure; nonstationary time series; optimal estimation accuracy

References

[1] Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory* **13** 315–352. MR1455175 https://doi.org/10.1017/S0266466600005831
[2] Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica* **66** 47–78. MR1616121 https://doi.org/10.2307/2998540
[3] Beibel, M. (1996). A note on Ritov's Bayes approach to the minimax property of the cusum procedure. *Ann. Statist.* **24** 1804–1812. MR1416661 https://doi.org/10.1214/aos/1032298296
[4] Chen, B. and Hong, Y. (2012). Testing for smooth structural changes in time series models via nonparametric regression. *Econometrica* **80** 1157–1183. MR2963885 https://doi.org/10.3982/ECTA7990
[5] Chen, L., Wang, W. and Wu, W.B. (2022). Inference of breakpoints in high-dimensional time series. *J. Amer. Statist. Assoc.* **117** 1951–1963. MR4528482 https://doi.org/10.1080/01621459.2021.1893178
[6] Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *Ann. Statist.* **25** 1–37. MR1429916 https://doi.org/10.1214/aos/1034276620
[7] Dahlhaus, R. and Subba Rao, S. (2006). Statistical inference for time-varying ARCH processes. *Ann. Statist.* **34** 1075–1114. MR2278352 https://doi.org/10.1214/009053606000000227
[8] Daubechies, I., Lu, J. and Wu, H.-T. (2011). Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* **30** 243–261. MR2754779 https://doi.org/10.1016/j.acha.2010.08.002
[9] Dette, H., Eckle, T. and Vetter, M. (2020). Multiscale change point detection for dependent data. *Scand. J. Stat.* **47** 1243–1274. MR4178193 https://doi.org/10.1111/sjos.12465
[10] Dette, H. and Wied, D. (2016). Detecting relevant changes in time series models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 371–394. MR3454201 https://doi.org/10.1111/rssb.12121
[11] Dette, H., Wu, W. and Zhou, Z. (2019). Change point analysis of correlation in non-stationary time series. *Statist. Sinica* **29** 611–643. MR3931381

1350-7265

[12] Dümbgen, L. (1991). The asymptotic behavior of some nonparametric change-point estimators. *Ann. Statist.* **19** 1471–1495. MR1126333 https://doi.org/10.1214/aos/1176348257

[13] Eubank, R.L. and Speckman, P.L. (1994). Nonparametric estimation of functions with jump discontinuities. In *Change-Point Problems (South Hadley, MA, 1992). Institute of Mathematical Statistics Lecture Notes—Monograph Series* **23** 130–144. Hayward, CA: IMS. MR1477919 https://doi.org/10.1214/lnms/1215463119

[14] Fan, J. and Marron, J.S. (1994). Fast implementations of nonparametric curve estimators. *J. Comput. Graph. Statist.* **3** 35–56.

[15] Frick, K., Munk, A. and Sieling, H. (2014). Multiscale change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 495–580. MR3210728 https://doi.org/10.1111/rssb.12047

[16] Gao, J., Gijbels, I. and Van Bellegem, S. (2008). Nonparametric simultaneous testing for structural breaks. *J. Econometrics* **143** 123–142. MR2384436 https://doi.org/10.1016/j.jeconom.2007.08.009

[17] Gijbels, I., Hall, P. and Kneip, A. (1999). On the estimation of jump points in smooth curves. *Ann. Inst. Statist. Math.* **51** 231–251. MR1707773 https://doi.org/10.1023/A:1003802007064

[18] Hájek, P. and Johanis, M. (2010). Smooth approximations. *J. Funct. Anal.* **259** 561–582. MR2644097 https://doi.org/10.1016/j.jfa.2010.04.020

[19] Horowitz, J.L. and Spokoiny, V.G. (2001). An adaptive, rate-optimal test of a parametric mean-regression model against a nonparametric alternative. *Econometrica* **69** 599–631. MR1828537 https://doi.org/10.1111/1468-0262.00207

[20] Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C. and Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **454** 903–995. MR1631591 https://doi.org/10.1098/rspa.1998.0193

[21] Khismatullina, M. and Vogt, M. (2020). Multiscale inference and long-run variance estimation in nonparametric regression with time series errors. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 5–37. MR4060975

[22] Killick, R., Fearnhead, P. and Eckley, I.A. (2012). Optimal detection of changepoints with a linear computational cost. *J. Amer. Statist. Assoc.* **107** 1590–1598. MR3036418 https://doi.org/10.1080/01621459.2012.737745

[23] Langrené, N. and Warin, X. (2019). Fast and stable multivariate kernel density estimation by fast sum updating. *J. Comput. Graph. Statist.* **28** 596–608. MR4007743 https://doi.org/10.1080/10618600.2018.1549052

[24] Liu, W., Xiao, H. and Wu, W.B. (2013). Probability and moment inequalities under dependence. *Statist. Sinica* **23** 1257–1272. MR3114713

[25] Loader, C.R. (1996). Change point estimation using nonparametric regression. *Ann. Statist.* **24** 1667–1678. MR1416655 https://doi.org/10.1214/aos/1032298290

[26] Mikosch, T. and Nagaev, A.V. (1998). Large deviations of heavy-tailed sums with applications in insurance. *Extremes* **1** 81–110. MR1652936 https://doi.org/10.1023/A:1009913901219

[27] Müller, H.-G. (1992). Change-points in nonparametric regression analysis. *Ann. Statist.* **20** 737–761. MR1165590 https://doi.org/10.1214/aos/1176348654

[28] Müller, H.-G. and Song, K.-S. (1997). Two-stage change-point estimators in smooth regression models. *Statist. Probab. Lett.* **34** 323–335. MR1467437 https://doi.org/10.1016/S0167-7152(96)00197-6

[29] Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*. *Springer Series in Statistics*. New York: Springer. MR1707286 https://doi.org/10.1007/978-1-4612-1554-7

[30] Qiu, P. (2003). A jump-preserving curve fitting procedure based on local piecewise-linear kernel estimation. *J. Nonparametr. Stat.* **15** 437–453. MR2017479 https://doi.org/10.1080/10485250310001595083

[31] Qu, Z. (2008). Testing for structural change in regression quantiles. *J. Econometrics* **146** 170–184. MR2459652 https://doi.org/10.1016/j.jeconom.2008.08.006

[32] Rho, Y. and Shao, X. (2019). Bootstrap-assisted unit root testing with piecewise locally stationary errors. *Econometric Theory* **35** 142–166. MR3904174 https://doi.org/10.1017/S0266466618000038

[33] Ritov, Y. (1990). Decision theoretic optimality of the CUSUM procedure. *Ann. Statist.* **18** 1464–1469. MR1062720 https://doi.org/10.1214/aos/1176347761

[34] Schmidt-Hieber, J., Munk, A. and Dümbgen, L. (2013). Multiscale methods for shape constraints in deconvolution: Confidence statements for qualitative features. *Ann. Statist.* **41** 1299–1328. MR3113812 https://doi.org/10.1214/13-AOS1089

[35] Seifert, B., Brockmann, M., Engel, J. and Gasser, T. (1994). Fast algorithms for nonparametric curve estimation. *J. Comput. Graph. Statist.* **3** 192–213. MR1278843 https://doi.org/10.2307/1390668
[36] Shao, X. (2010). A self-normalized approach to confidence interval construction in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 343–366. MR2758116 https://doi.org/10.1111/j.1467-9868.2009.00737.x
[37] Shao, X. and Zhang, X. (2010). Testing for change points in time series. *J. Amer. Statist. Assoc.* **105** 1228–1240. MR2752617 https://doi.org/10.1198/jasa.2010.tm10103
[38] Siegmund, D. (1988). Confidence sets in change-point problems. *Int. Stat. Rev.* **56** 31–48. MR0963139 https://doi.org/10.2307/1403360
[39] Stărică, C. and Granger, C. (2005). Nonstationarities in stock returns. *Rev. Econ. Stat.* **87** 503–522.
[40] Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields. *Ann. Probab.* **21** 34–71. MR1207215
[41] Sun, J. and Loader, C.R. (1994). Simultaneous confidence bands for linear regression and smoothing. *Ann. Statist.* **22** 1328–1345. MR1311978 https://doi.org/10.1214/aos/1176325631
[42] Weyl, H. (1939). On the Volume of Tubes. *Amer. J. Math.* **61** 461–472. MR1507388 https://doi.org/10.2307/2371513
[43] Wu, W. and Zhou, Z. (2018). Gradient-based structural change detection for nonstationary time series M-estimation. *Ann. Statist.* **46** 1197–1224. MR3798001 https://doi.org/10.1214/17-AOS1582
[44] Wu, W. and Zhou, Z. (2024). Supplement to "Multiscale jump testing and estimation under complex temporal dynamics." https://doi.org/10.3150/23-BEJ1677SUPP
[45] Zhang, C.M. (2003). Adaptive tests of regression functions via multiscale generalized likelihood ratios. *Canad. J. Statist.* **31** 151–171. MR2016225 https://doi.org/10.2307/3316065
[46] Zhang, J. and Fan, J. (2000). Minimax kernels for nonparametric curve estimation. *Int. J. Comput. Math.* **12** 417–445. MR1760716 https://doi.org/10.1080/10485250008832816
[47] Zhang, T. (2016). Testing for jumps in the presence of smooth changes in trends of nonstationary time series. *Electron. J. Stat.* **10** 706–735. MR3477739 https://doi.org/10.1214/16-EJS1127
[48] Zhao, Z., Jiang, F. and Shao, X. (2022). Segmenting time series via self-normalisation. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1699–1725. MR4515555
[49] Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *J. Amer. Statist. Assoc.* **108** 726–740. MR3174655 https://doi.org/10.1080/01621459.2013.787184
[50] Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. MR2541444 https://doi.org/10.1214/08-AOS636
[51] Zhu, Y.-K. and Hayes, W.B. (2010). Algorithm 908: Online exact summation of floating-point streams. *ACM Trans. Math. Software* **37** 1–13.

Bernoulli **30**(3), 2024, 2399–2422
https://doi.org/10.3150/23-BEJ1678

The motion of the tagged particle in the asymmetric exclusion process with long jumps

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 prove a law of large numbers and invariance principles for the tagged particle in the asymmetric exclusion process with long jumps when the process starts from its equilibrium measure.

Keywords: Asymmetric exclusion process; central limit theorems; long jumps; tagged particle

References

Arratia, R. (1983). The motion of a tagged particle in the simple symmetric exclusion system on **Z**. *Ann. Probab.* **11** 362–373. MR0690134

Bäumler, J. (2023). Recurrence and transience of symmetric random walks with long-range jumps. *Electron. J. Probab.* **28** Paper No. 106, 24. MR4632146 https://doi.org/10.1214/23-ejp998

Bernardin, C., Gonçalves, P. and Sethuraman, S. (2016). Occupation times of long-range exclusion and connections to KPZ class exponents. *Probab. Theory Related Fields* **166** 365–428. MR3547742 https://doi.org/10.1007/s00440-015-0661-5

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

Gonçalves, P. and Jara, M. (2018). Density fluctuations for exclusion processes with long jumps. *Probab. Theory Related Fields* **170** 311–362. MR3748326 https://doi.org/10.1007/s00440-017-0758-0

Jara, M. (2009a). Hydrodynamic limit of particle systems with long jumps. Preprint. Available at arXiv:0805.1326.

Jara, M. (2009b). Nonequilibrium scaling limit for a tagged particle in the simple exclusion process with long jumps. *Comm. Pure Appl. Math.* **62** 198–214. MR2468608 https://doi.org/10.1002/cpa.20253

Jara, M.D. and Landim, C. (2006). Nonequilibrium central limit theorem for a tagged particle in symmetric simple exclusion. *Ann. Inst. Henri Poincaré Probab. Stat.* **42** 567–577. MR2259975 https://doi.org/10.1016/j.anihpb.2005.04.007

Jara, M.D. and Landim, C. (2008). Quenched non-equilibrium central limit theorem for a tagged particle in the exclusion process with bond disorder. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 341–361. MR2446327 https://doi.org/10.1214/07-AIHP112

Kipnis, C. (1986). Central limit theorems for infinite series of queues and applications to simple exclusion. *Ann. Probab.* **14** 397–408. MR0832016

Kipnis, C. and Landim, C. (1999). *Scaling Limits of Interacting Particle Systems*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **320**. Berlin: Springer. MR1707314 https://doi.org/10.1007/978-3-662-03752-2

Kipnis, C. and Varadhan, S.R.S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** 1–19. MR0834478

Komorowski, T., Landim, C. and Olla, S. (2012). *Fluctuations in Markov Processes*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **345**. Heidelberg: Springer. Time symmetry and martingale approximation. MR2952852 https://doi.org/10.1007/978-3-642-29880-6

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

1350-7265

Peligrad, M. and Sethuraman, S. (2008). On fractional Brownian motion limits in one dimensional nearest-neighbor symmetric simple exclusion. *ALEA Lat. Am. J. Probab. Math. Stat.* **4** 245–255. MR2448774

Rezakhanlou, F. (1994). Evolution of tagged particles in non-reversible particle systems. *Comm. Math. Phys.* **165** 1–32. MR1298939

Saada, E. (1987). A limit theorem for the position of a tagged particle in a simple exclusion process. *Ann. Probab.* **15** 375–381. MR0877609

Sethuraman, S. (2000). Central limit theorems for additive functionals of the simple exclusion process. *Ann. Probab.* **28** 277–302. MR1756006 https://doi.org/10.1214/aop/1019160120

Sethuraman, S. (2006). Diffusive variance for a tagged particle in $d \leq 2$ asymmetric simple exclusion. *ALEA Lat. Am. J. Probab. Math. Stat.* **1** 305–332. MR2249659

Sethuraman, S. and Shahar, D. (2018). Hydrodynamic limits for long-range asymmetric interacting particle systems. *Electron. J. Probab.* **23** Paper No. 130, 54. MR3896867 https://doi.org/10.1214/18-EJP237

Sethuraman, S. and Varadhan, S.R.S. (2013). Large deviations for the current and tagged particle in 1D nearest-neighbor symmetric simple exclusion. *Ann. Probab.* **41** 1461–1512. MR3098682 https://doi.org/10.1214/11-AOP703

Sethuraman, S., Varadhan, S.R.S. and Yau, H.-T. (2000). Diffusive limit of a tagged particle in asymmetric simple exclusion processes. *Comm. Pure Appl. Math.* **53** 972–1006. MR1755948 https://doi.org/10.1002/1097-0312(200008)53:8<972::AID-CPA2>3.0.CO;2-#

Spitzer, F. (1970). Interaction of Markov processes. *Adv. Math.* **5** 246–290 (1970). MR0268959 https://doi.org/10.1016/0001-8708(70)90034-4

Varadhan, S.R.S. (1995). Self-diffusion of a tagged particle in equilibrium for asymmetric mean zero random walk with simple exclusion. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 273–285. MR1340041

Xue, X. F. and Zhao, L. J. (2022). Moderate deviations for the current and tagged particle in symmetric simple exclusion processes. Preprint. Available at arXiv:2203.05260.

Zhao, L. (2024). Supplement to "The motion of the tagged particle in the asymmetric exclusion process with long jumps." https://doi.org/10.3150/23-BEJ1678SUPP

Bernoulli **30**(3), 2024, 2423–2449
https://doi.org/10.3150/23-BEJ1679

Conditional hazard rate estimation for right censored data

SAM EFROMOVICH[a]

Department of Mathematical Sciences, UTDallas, Richardson, USA, [a]*efrom@utdallas.edu*

Theory and methodology of nonparametric sharp minimax estimation of the conditional hazard rate function of a right censored lifetime given a continuous covariate are developed. The theory, using an oracle's approach, shows how the conditional hazard and nuisance functions affect rate and constant of the mean integrated squared error (MISE) convergence. The methodology suggests a data-driven estimator matching performance of the oracle. Further, if the lifetime is independent of the covariate, the estimator recognizes that and the MISE converges with the univariate rate. Then the setting is extended to a vector of continuous and ordinal/nominal categorical predictors, and an estimator performing adaptation to smoothness and dimensionality of conditional hazard is suggested. Practical examples devoted to reducing potent greenhouse gas emissions are presented.

Keywords: Adaptation; dimension reduction; nonparametric; nuisance function; sharp minimax

References

Albu, N., Barani, N. and Constantin, M. (2021). Choosing an conomical solution for water aeration. *Hydraulica* **145** 32–37.

Balan, T.A. and Putter, H. (2020). A tutorial on frailty models. *Stat. Methods Med. Res.* **29** 3424–3454. MR4156864 https://doi.org/10.1177/0962280220921889

Cai, J., Fan, J., Jiang, J. and Zhou, H. (2008). Partially linear hazard regression with varying coefficients for multivariate survival data. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 141–158. MR2412635 https://doi.org/10.1111/j.1467-9868.2007.00630.x

Comte, F., Gaïffas, S. and Guilloux, A. (2011). Adaptive estimation of the conditional intensity of marker-dependent counting processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 1171–1196. MR2884230 https://doi.org/10.1214/10-AIHP386

Cox, D.R. and Oakes, D. (1984). *Analysis of Survival Data*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR0751780

Cui, Y. and Hannig, J. (2019). Nonparametric generalized fiducial inference for survival functions under censoring. *Biometrika* **106** 501–518. MR3992384 https://doi.org/10.1093/biomet/asz016

Drewnowski, J., Remiszewska-Skwarek, A., Duda, S. and Lagod, G. (2019). Aeration process in bioreactors as the main energy consumer in a wastewater treatment plant. Review of solutions and methods of process optimization. *Processes* **7** 1–21.

Efromovich, S.Y. (1985). Nonparametric estimation of the density with unknown smoothness. *Theory Probab. Appl.* **30** 524–534. MR1005732 https://doi.org/10.1137/1130067

Efromovich, S.Y. (1989). On sequential nonparametric estimation of density. *Theory Probab. Appl.* **34** 228–239. MR1005732 https://doi.org/10.1137/1134019

Efromovich, S. (1999). *Nonparametric Curve Estimation: Methods, Theory, and Applications*. *Springer Series in Statistics*. New York: Springer. MR1705298

Efromovich, S. (2001). Density estimation under random censorship and order restrictions: From asymptotic to small samples. *J. Amer. Statist. Assoc.* **96** 667–684. MR1946433 https://doi.org/10.1198/016214501753168334

Efromovich, S. (2013). Nonparametric regression with the scale depending on auxiliary variable. *Ann. Statist.* **41** 1542–1568. MR3113821 https://doi.org/10.1214/13-AOS1126

1350-7265

Efromovich, S. (2016). Minimax theory of nonparametric hazard rate estimation: Efficiency and adaptation. *Ann. Inst. Statist. Math.* **68** 25–75. MR3440214 https://doi.org/10.1007/s10463-014-0487-4

Efromovich, S. (2018). *Missing and Modified Data in Nonparametric Estimation: With R Examples*. *Monographs on Statistics and Applied Probability* **156**. Boca Raton, FL: CRC Press. MR3752670

Efromovich, S. (2021). Sharp minimax distribution estimation for current status censoring with or without missing. *Ann. Statist.* **49** 568–589. MR4206691 https://doi.org/10.1214/20-AOS1970

Efromovich, S. (2024). Supplement to "Conditional hazard rate estimation for right censored data." https://doi.org/10.3150/23-BEJ1679SUPP

Fleming, T.R. and Harrington, D.P. (2011). *Counting Processes and Survival Analysis*. New York: Wiley. MR1100924

Gill, R. (2006). *Lectures on Survival Analysis*. New York: Springer.

Gneyou, K.E. (2014). A strong linear representation for the maximum conditional hazard rate estimator in survival analysis. *J. Multivariate Anal.* **128** 10–18. MR3199824 https://doi.org/10.1016/j.jmva.2014.02.013

Golubev, G.K. (1991). LAN in problems of non-parametric estimation of functions and lower bounds for quadratic risks. *Probl. Inf. Transm.* **36** 152–157. MR1109023 https://doi.org/10.1137/1136014

Hoffmann, M. and Lepski, O. (2002). Random rates in anisotropic regression. *Ann. Statist.* **30** 325–396. With discussions and a rejoinder by the authors. MR1902892 https://doi.org/10.1214/aos/1021379858

Huang, J.Z. and Su, Y. (2021). Asymptotic properties of penalized spline estimators in concave extended linear models: Rates of convergence. *Ann. Statist.* **49** 3383–3407. MR4352534 https://doi.org/10.1214/21-aos2088

Ibragimov, I.A. and Khasminskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. *Applications of Mathematics* **16**. New York–Berlin: Springer. Translated from the Russian by Samuel Kotz. MR0620321

Kahane, J.-P. (1985). *Some Random Series of Functions*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **5**. Cambridge: Cambridge Univ. Press. MR0833073

Kalbfleisch, J.D. and Prentice, R.L. (2011). *The Statistical Analysis of Failure Time Data*. New York: Wiley. MR0570114

Kang, S., Lu, W. and Zhang, J. (2018). On estimation of the optimal treatment regime with the additive hazards model. *Statist. Sinica* **28** 1539–1560. MR3821017

Klein, J. and Moeschberger, M. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer.

Kooperberg, C., Stone, C.J. and Truong, Y.K. (1995). Hazard regression. *J. Amer. Statist. Assoc.* **90** 78–94. MR1325116

Lee, E.T. and Wang, J.W. (2003). *Statistical Methods for Survival Data Analysis*, 3rd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR1968483 https://doi.org/10.1002/0471458546

Lee, E.T. and Wang, J.W. (2013). *Statistical Methods for Survival Data Analysis*, 4th ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR3235949

Legrand, C. (2021). *Advanced Survival Models*. *Chapman & Hall/CRC Biostatistics Series*. Boca Raton, FL: CRC Press. MR4596197

Li, G. and Doss, H. (1995). An approach to nonparametric regression for life history data using local linear fitting. *Ann. Statist.* **23** 787–823. MR1345201 https://doi.org/10.1214/aos/1176324623

Li, D., Lu, W., Shu, D., Toh, S. and Wang, W. (2023). Distributed Cox proportional hazards regression using summary-level information. *Biostatistics* **24** 776–794. https://doi.org/10.1093/biostatistics/kxac006

Lu, M., Lu, T. and Li, C.-S. (2018). Efficient estimation of partially linear additive Cox model under monotonicity constraint. *J. Statist. Plann. Inference* **192** 18–34. MR3697122 https://doi.org/10.1016/j.jspi.2017.07.003

McKeague, I.W. and Utikal, K.J. (1990). Inference for a nonlinear counting process regression model. *Ann. Statist.* **18** 1172–1187. MR1062704 https://doi.org/10.1214/aos/1176347745

Miller, R.G. Jr. (2011). *Survival Analysis*. New York: Wiley. MR0634228

Moore, D. (2016). *Applied Survival Analysis Using R*. New York: Springer.

Pinsker, M.S. (1980). Optimal filtration of square-integrable signals in Gaussian noise. *Problemy Peredachi Informatsii* **16** 52–68. MR0624591

Prentice, R.L. and Zhao, S. (2019). *The Statistical Analysis of Multivariate Failure Time Data: A Marginal Modeling Approach*. *Monographs on Statistics and Applied Probability* **163**. Boca Raton, FL: CRC Press. MR3966434 https://doi.org/10.1201/9780429162367

Samuelson, O., Björka, A. and Carlsson, B. (2021). Model-based monitoring of diffuser fouling using standard sensors. *Water Res.* **13** 100–118.

Slavov, A.K. (2017). General characteristics and treatment possibilities of dairy wastewater - a review. *Food Technol. Biotechnol.* **55** 14–28. https://doi.org/10.17113/ftb.55.01.17.4520

Spierdijk, L. (2008). Nonparametric conditional hazard rate estimation: A local linear approach. *Comput. Statist. Data Anal.* **52** 2419–2434. MR2411948 https://doi.org/10.1016/j.csda.2007.08.007

Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proc. Natl. Acad. Sci. USA* **72** 20–22. MR0356425 https://doi.org/10.1073/pnas.72.1.20

Van Keilegom, I. and Veraverbeke, N. (2001). Hazard rate estimation in nonparametric regression with censored data. *Ann. Inst. Statist. Math.* **53** 730–745. MR1880808 https://doi.org/10.1023/A:1014696717644

Vinardella, S., Astalsa, S., Pecesb, M., Cardetea, M., Fernandez, I., Mata-Alvareza, J. and Dosta, J. (2020). Advances in anaerobic membrane bioreactor technology for municipal wastewater treatment. *Renew. Sustain. Energy Rev.* **130** 109–136.

Wasserman, L. (2006). *All of Nonparametric Statistics*. *Springer Texts in Statistics*. New York: Springer. MR2172729

Zhao, L. and Feng, D. (2020). Deep neural networks for survival analysis using pseudo values. *IEEE J. Biomed. Health Inform.* **24** 3308–3314.

Zhong, Q., Mueller, J. and Wang, J.-L. (2022). Deep learning for the partially linear Cox model. *Ann. Statist.* **50** 1348–1375. MR4441123 https://doi.org/10.1214/21-aos2153

Zhou, M. (2016). *Empirical Likelihood Method in Survival Analysis*. *Chapman & Hall/CRC Biostatistics Series*. Boca Raton, FL: CRC Press. MR3616660

Bernoulli **30**(3), 2024, 2450–2474
https://doi.org/10.3150/23-BEJ1680

Detecting long-range dependence for time-varying linear models

LUJIA BAI[a] and WEICHI WU[b]

Center for Statistical Science, Department of Industrial Engineering, Tsinghua University, Beijing, China, [a]*blj20@mails.tsinghua.edu.cn,* [b]*wuweichi@mail.tsinghua.edu.cn*

We consider the problem of testing for long-range dependence in time-varying coefficient regression models, where the covariates and errors are locally stationary, allowing complex temporal dynamics and heteroscedasticity. We develop KPSS, R/S, V/S, and K/S-type statistics based on the nonparametric residuals. Under the null hypothesis, the local alternatives as well as the fixed alternatives, we derive the limiting distributions of the test statistics. As the four types of test statistics could degenerate when the time-varying mean, variance, long-run variance of errors, covariates, and the intercept lie in certain hyperplanes, we show the bootstrap-assisted tests are consistent under both degenerate and non-degenerate scenarios. In particular, in the presence of covariates the exact local asymptotic power of the bootstrap-assisted tests can enjoy the same order as that of the classical KPSS test of long memory for strictly stationary series. The asymptotic theory is built on a new Gaussian approximation technique for locally stationary long-memory processes with short-memory covariates, which is of independent interest. The effectiveness of our tests is demonstrated by extensive simulation studies and real data analysis.

Keywords: Long-range dependence; locally stationary process; spurious long memory; time-varying models

References

Bai, L. and Wu, W. (2023). Difference-based covariance matrix estimate in time series nonparametric regression with applications to specification tests. ArXiv preprint. Available at arXiv:2303.16599.

Bai, L. and Wu, W. (2024). Supplement to "Detecting long-range dependence for time-varying linear models." https://doi.org/10.3150/23-BEJ1680SUPP

Beran, J. (2009). On parameter estimation for locally stationary long-memory processes. *J. Statist. Plann. Inference* **139** 900–915. MR2479836 https://doi.org/10.1016/j.jspi.2008.05.047

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

Berkes, I., Horváth, L., Kokoszka, P. and Shao, Q.-M. (2006). On discriminating between long-range dependence and changes in mean. *Ann. Statist.* **34** 1140–1165. MR2278354 https://doi.org/10.1214/009053606000000254

Caporale, G.M. and Gil-Alana, L.A. (2013). Long memory and fractional integration in high frequency data on the US dollar/British pound spot exchange rate. *Int. Rev. Financ. Anal.* **29** 1–9.

Caporale, G.M., Gil-Alana, L.A. and Lovcha, Y. (2016). Testing unemployment theories: A multivariate long memory approach. *J. Appl. Econometrics* **19** 95–112.

Cavaliere, G., Nielsen, M.Ø. and Taylor, A.M.R. (2022). Adaptive inference in heteroscedastic fractional time series models. *J. Bus. Econom. Statist.* **40** 50–65. MR4356557 https://doi.org/10.1080/07350015.2020.1773275

Chen, M. and Song, Q. (2015). Simultaneous inference of the mean of functional time series. *Electron. J. Stat.* **9** 1779–1798. MR3391119 https://doi.org/10.1214/15-EJS1052

Chen, X.B., Gao, J., Li, D. and Silvapulle, P. (2018). Nonparametric estimation and forecasting for time-varying coefficient realized volatility models. *J. Bus. Econom. Statist.* **36** 88–100. MR3750911 https://doi.org/10.1080/07350015.2016.1138118

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

Dahlhaus, R., Richter, S. and Wu, W.B. (2019). Towards a general theory for nonlinear locally stationary processes. *Bernoulli* **25** 1013–1044. MR3920364 https://doi.org/10.3150/17-bej1011

Davis, R.A. and Yau, C.Y. (2013). Consistency of minimum description length model selection for piecewise stationary time series models. *Electron. J. Stat.* **7** 381–411. MR3020426 https://doi.org/10.1214/13-EJS769

Dehling, H. and Taqqu, M.S. (1989). The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17** 1767–1783. MR1026312 https://doi.org/10.1214/aos/1176347394

Dette, H., Preuss, P. and Sen, K. (2017). Detecting long-range dependence in non-stationary time series. *Electron. J. Stat.* **11** 1600–1659. MR3638972 https://doi.org/10.1214/17-EJS1262

Dette, H., Preuss, P. and Vetter, M. (2011). A measure of stationarity in locally stationary processes with applications to testing. *J. Amer. Statist. Assoc.* **106** 1113–1124. MR2894768 https://doi.org/10.1198/jasa.2011.tm10811

Dette, H. and Wu, W. (2019). Detecting relevant changes in the mean of nonstationary processes—a mass excess approach. *Ann. Statist.* **47** 3578–3608. MR4025752 https://doi.org/10.1214/19-AOS1811

Duffy, J.A. and Kasparis, I. (2021). Estimation and inference in the presence of fractional $d = 1/2$ and weakly nonstationary processes. *Ann. Statist.* **49** 1195–1217. MR4255124 https://doi.org/10.1214/20-aos1998

Fan, J. (1993). Local linear regression smoothers and their minimax efficiencies. *Ann. Statist.* **21** 196–216. MR1212173 https://doi.org/10.1214/aos/1176349022

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. London: CRC Press. MR1383587

Fan, J. and Zhang, W. (2000). Simultaneous confidence bands and hypothesis testing in varying-coefficient models. *Scand. J. Stat.* **27** 715–731. MR1804172 https://doi.org/10.1111/1467-9469.00218

Ferreira, G., Piña, N. and Porcu, E. (2018). Estimation of slowly time-varying trend function in long memory regression models. *J. Stat. Comput. Simul.* **88** 1903–1920. MR3799169 https://doi.org/10.1080/00949655.2018.1466141

Giraitis, L., Kokoszka, P. and Leipus, R. (2001). Testing for long memory in the presence of a general trend. *J. Appl. Probab.* **38** 1033–1054. MR1876557 https://doi.org/10.1017/s0021900200019215

Giraitis, L., Kokoszka, P., Leipus, R. and Teyssière, G. (2003). Rescaled variance and related tests for long memory in volatility and levels. *J. Econometrics* **112** 265–294. MR1951145 https://doi.org/10.1016/S0304-4076(02)00197-5

Harris, D. and Kew, H. (2017). Adaptive long memory testing under heteroskedasticity. *Econometric Theory* **33** 755–778. MR3637975 https://doi.org/10.1017/S0266466615000481

Harris, D., McCabe, B. and Leybourne, S. (2008). Testing for long memory. *Econometric Theory* **24** 143–175. MR2408862 https://doi.org/10.1017/S0266466608080080

Hu, L., Huang, T. and You, J. (2019). Estimation and identification of a varying-coefficient additive model for locally stationary processes. *J. Amer. Statist. Assoc.* **114** 1191–1204. MR4011772 https://doi.org/10.1080/01621459.2018.1482753

Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civ. Eng.* **116** 770–799.

Kokoszka, P.S. and Taqqu, M.S. (1995). Fractional ARIMA with stable innovations. *Stochastic Process. Appl.* **60** 19–47. MR1362317 https://doi.org/10.1016/0304-4149(95)00034-8

Koutsoyiannis, D. (2013). Hydrology and change. *Hydrol. Sci. J.* **58** 1177–1197.

Kulik, R. and Wichelhaus, C. (2012). Conditional variance estimation in regression models with long memory. *J. Time Series Anal.* **33** 468–483. MR2915097 https://doi.org/10.1111/j.1467-9892.2012.00782.x

Lee, D. and Schmidt, P. (1996). On the power of the KPSS test of stationarity against fractionally-integrated alternatives. *J. Econometrics* **73** 285–302. MR1410008 https://doi.org/10.1016/0304-4076(95)01741-0

Lima, L. and Xiao, Z. (2004). Robustness of stationary tests under long-memory alternatives Technical Report EPGE Brazilian School of Economics and Finance-FGV EPGE (Brazil).

Marinucci, D. and Robinson, P.M. (1999). Alternative forms of fractional Brownian motion. *J. Statist. Plann. Inference* **80** 111–122. MR1713794 https://doi.org/10.1016/S0378-3758(98)00245-6

McCloskey, A. and Perron, P. (2013). Memory parameter estimation in the presence of level shifts and deterministic trends. *Econometric Theory* **29** 1196–1237. MR3148830 https://doi.org/10.1017/S0266466613000042

Nason, G.P., von Sachs, R. and Kroisandt, G. (2000). Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 271–292. MR1749539 https://doi.org/10.1111/1467-9868.00231

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

Preuß, P. and Vetter, M. (2013). Discriminating between long-range dependence and non-stationarity. *Electron. J. Stat.* **7** 2241–2297. MR3108814 https://doi.org/10.1214/13-EJS836

Qu, Z. (2011). A test against spurious long memory. *J. Bus. Econom. Statist.* **29** 423–438. MR2848513 https://doi.org/10.1198/jbes.2010.09153

Roueff, F. and von Sachs, R. (2011). Locally stationary long memory estimation. *Stochastic Process. Appl.* **121** 813–844. MR2770908 https://doi.org/10.1016/j.spa.2010.12.004

Shao, J. (2003). *Mathematical Statistics*. Berlin: Springer.

Shao, X. and Wu, W.B. (2007). Local asymptotic powers of nonparametric and semiparametric tests for fractional integration. *Stochastic Process. Appl.* **117** 251–261. MR2290195 https://doi.org/10.1016/j.spa.2006.08.002

Sibbertsen, P., Leschinski, C. and Busch, M. (2018). A multivariate test against spurious long memory. *J. Econometrics* **203** 33–49. MR3758326 https://doi.org/10.1016/j.jeconom.2017.07.005

Veitch, D., Gorst-Rasmussen, A. and Gefferth, A. (2013). Why FARIMA models are brittle. *Fractals* **21** 1350012. MR3092054 https://doi.org/10.1142/S0218348X13500126

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614 https://doi.org/10.1214/12-AOS1043

Vogt, M. and Dette, H. (2015). Detecting gradual changes in locally stationary processes. *Ann. Statist.* **43** 713–740. MR3319141 https://doi.org/10.1214/14-AOS1297

Wang, Q., Lin, Y.-X. and Gulati, C.M. (2003). Strong approximation for long memory processes with applications. *J. Theoret. Probab.* **16** 377–389. MR1982033 https://doi.org/10.1023/A:1023570510824

Wu, W. and Zhou, Z. (2018). Simultaneous quantile inference for non-stationary long-memory time series. *Bernoulli* **24** 2991–3012. MR3779708 https://doi.org/10.3150/17-BEJ951

Wu, W. and Zhou, Z. (2018). Gradient-based structural change detection for nonstationary time series M-estimation. *Ann. Statist.* **46** 1197–1224. MR3798001 https://doi.org/10.1214/17-AOS1582

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. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320. MR2353389 https://doi.org/10.1214/009117907000000060

Wu, W.B. and Shao, X. (2006). Invariance principles for fractionally integrated nonlinear processes. In *Recent Developments in Nonparametric Inference and Probability*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **50** 20–30. Beachwood, OH: IMS. MR2409061 https://doi.org/10.1214/074921706000000572

Wu, W.B. and Zhao, Z. (2007). Inference of trends in time series. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 391–410. MR2323759 https://doi.org/10.1111/j.1467-9868.2007.00594.x

Wu, W.B. and Zhou, Z. (2011). Gaussian approximations for non-stationary multiple time series. *Statist. Sinica* **21** 1397–1413. MR2827528 https://doi.org/10.5705/ss.2008.223

Zhang, Q., Zhou, Y., Singh, V.P. and Chen, Y.D. (2011). Comparison of detrending methods for fluctuation analysis in hydrology. *J. Hydrol.* **400** 121–132.

Zhou, Z. and Wu, W.B. (2009). Local linear quantile estimation for nonstationary time series. *Ann. Statist.* **37** 2696–2729. MR2541444 https://doi.org/10.1214/08-AOS636

Zhou, Z. and Wu, W.B. (2010). Simultaneous inference of linear models with time varying coefficients. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 513–531. MR2758526 https://doi.org/10.1111/j.1467-9868.2010.00743.x

Bernoulli **30**(3), 2024, 2475–2500
https://doi.org/10.3150/23-BEJ1681

Sparse M-estimators in semi-parametric copula models

JEAN-DAVID FERMANIAN[1,a] and BENJAMIN POIGNARD[2,b]

[1] *CREST-Ensae, 5, avenue Henry Le Chatelier, 91120 Palaiseau, France,* [a] *jean-david.fermanian@ensae.fr*
[2] *Osaka University, Graduate School of Economics, 1-7, Machikaneyama, Toyonaka-Shi, Osaka-Fu, 560-0043, Japan. Jointly affiliated at RIKEN Center for Advanced Intelligence Project (AIP) and CREST-LFA,* [b] *bpoignard@econ.osaka-u.ac.jp*

We study the large-sample properties of sparse M-estimators in the presence of pseudo-observations. Our framework covers a broad class of semi-parametric copula models, for which the marginal distributions are unknown and replaced by their empirical counterparts. It is well known that the latter modification significantly alters the limiting laws compared to usual M-estimation. We establish the consistency and the asymptotic normality of our sparse penalized M-estimator and we prove the asymptotic oracle property with pseudo-observations, possibly in the case when the number of parameters is diverging. Our framework allows to manage copula-based loss functions that are potentially unbounded. Additionally, we state the weak limit of multivariate rank statistics for an arbitrary dimension and the weak convergence of empirical copula processes indexed by maps. We apply our inference method to Canonical Maximum Likelihood losses with Gaussian copulas, mixtures of copulas or conditional copulas. The theoretical results are illustrated by two numerical experiments.

Keywords: Copulas; M-estimation; pseudo-observations; sparsity

References

[1] Abegaz, F., Gijbels, I. and Veraverbeke, N. (2012). Semiparametric estimation of conditional copulas. *J. Multivariate Anal.* **110** 43–73. MR2927509 https://doi.org/10.1016/j.jmva.2012.04.001
[2] Aistleitner, C. and Dick, J. (2015). Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arith.* **167** 143–171. MR3312093 https://doi.org/10.4064/aa167-2-4
[3] Berghaus, B., Bücher, A. and Volgushev, S. (2017). Weak convergence of the empirical copula process with respect to weighted metrics. *Bernoulli* **23** 743–772. MR3556791 https://doi.org/10.3150/15-BEJ751
[4] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Methods, Theory and Applications*. *Springer Series in Statistics*. Heidelberg: Springer. MR2807761 https://doi.org/10.1007/978-3-642-20192-9
[5] Cai, Z. and Wang, X. (2014). Selection of mixed copula model via penalized likelihood. *J. Amer. Statist. Assoc.* **109** 788–801. MR3223750 https://doi.org/10.1080/01621459.2013.873366
[6] Chen, X. and Fan, Y. (2005). Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. *Canad. J. Statist.* **33** 389–414. MR2193982 https://doi.org/10.1002/cjs.5540330306
[7] Chen, X. and Fan, Y. (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *J. Econometrics* **135** 125–154. MR2328398 https://doi.org/10.1016/j.jeconom.2005.07.027
[8] Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas: A Practical Guide with R*. *Lecture Notes in Statistics* **222**. Cham: Springer. MR3931334 https://doi.org/10.1007/978-3-030-13785-4
[9] Dehling, H., Durieu, O. and Tusche, M. (2014). Approximating class approach for empirical processes of dependent sequences indexed by functions. *Bernoulli* **20** 1372–1403. MR3217447 https://doi.org/10.3150/13-BEJ525
[10] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. MR1946581 https://doi.org/10.1198/016214501753382273

1350-7265

[11] Fan, J. and Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Ann. Statist.* **32** 928–961. MR2065194 https://doi.org/10.1214/009053604000000256

[12] Fermanian, J.-D. and Lopez, O. (2018). Single-index copulas. *J. Multivariate Anal.* **165** 27–55. MR3768751 https://doi.org/10.1016/j.jmva.2017.11.004

[13] Fermanian, J.-D. and Poignard, B. (2024). Supplement to "Sparse M-estimators in semi-parametric copula models." https://doi.org/10.3150/23-BEJ1681SUPP

[14] Fermanian, J.-D., Radulović, D. and Wegkamp, M. (2004). Weak convergence of empirical copula processes. *Bernoulli* **10** 847–860. MR2093613 https://doi.org/10.3150/bj/1099579158

[15] Fermanian, J.-D. and Wegkamp, M.H. (2012). Time-dependent copulas. *J. Multivariate Anal.* **110** 19–29. MR2927507 https://doi.org/10.1016/j.jmva.2012.02.018

[16] Fermanian, J.D., Radulović, D. and Wegkamp, M. (2002). Weak convergence of empirical copula processes. Working Paper No. 2002-06, Center for Research in Economics and Statistics.

[17] Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82** 543–552. MR1366280 https://doi.org/10.1093/biomet/82.3.543

[18] Ghoudi, K. and Rémillard, B. (2004). Empirical processes based on pseudo-observations. II. The multivariate case. In *Asymptotic Methods in Stochastics*. *Fields Inst. Commun.* **44** 381–406. Providence, RI: Amer. Math. Soc. MR2106867

[19] Hamori, S., Motegi, K. and Zhang, Z. (2019). Calibration estimation of semiparametric copula models with data missing at random. *J. Multivariate Anal.* **173** 85–109. MR3920997 https://doi.org/10.1016/j.jmva.2019.02.003

[20] Hastie, T., Tibshirani, R. and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2722294 https://doi.org/10.1007/978-0-387-84858-7

[21] Loh, P.-L. and Wainwright, M.J. (2015). Regularized M-estimators with nonconvexity: Statistical and algorithmic theory for local optima. *J. Mach. Learn. Res.* **16** 559–616. MR3335800

[22] Loh, P.-L. and Wainwright, M.J. (2017). Support recovery without incoherence: A case for nonconvex regularization. *Ann. Statist.* **45** 2455–2482. MR3737898 https://doi.org/10.1214/16-AOS1530

[23] Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2197664 https://doi.org/10.1007/0-387-28678-0

[24] Radulović, D., Wegkamp, M. and Zhao, Y. (2017). Weak convergence of empirical copula processes indexed by functions. *Bernoulli* **23** 3346–3384. MR3654809 https://doi.org/10.3150/16-BEJ849

[25] Ruymgaart, F.H. (1974). Asymptotic normality of nonparametric tests for independence. *Ann. Statist.* **2** 892–910. MR0386140 https://doi.org/10.1214/aos/1176342812

[26] Ruymgaart, F.H., Shorack, G.R. and van Zwet, W.R. (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Stat.* **43** 1122–1135. MR0339397 https://doi.org/10.1214/aoms/1177692465

[27] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli* **18** 764–782. MR2948900 https://doi.org/10.3150/11-BEJ387

[28] Shih, J.H. and Louis, T.A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics* **51** 1384–1399. MR1381050 https://doi.org/10.2307/2533269

[29] Tsukahara, H. (2005). Semiparametric estimation in copula models. *Canad. J. Statist.* **33** 357–375. MR2193980 https://doi.org/10.1002/cjs.5540330304

[30] van der Vaart, A.W. and Wellner, J.A. (2007). Empirical processes indexed by estimated functions. In *Asymptotics: Particles, Processes and Inverse Problems*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **55** 234–252. Beachwood, OH: IMS. MR2459942 https://doi.org/10.1214/074921707000000382

[31] Yang, B., Hafner, C.M., Liu, G. and Long, W. (2021). Semiparametric estimation and variable selection for single-index copula models. *J. Appl. Econometrics* **36** 962–988. MR4362603 https://doi.org/10.1002/jae.2812

[32] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. MR2604701 https://doi.org/10.1214/09-AOS729

[33] Zhang, S., Okhrin, O., Zhou, Q.M. and Song, P.X.-K. (2016). Goodness-of-fit test for specification of semi-parametric copula dependence models. *J. Econometrics* **193** 215–233. MR3500185 https://doi.org/10.1016/j.jeconom.2016.02.017

[34] Zou, H. (2006). The adaptive lasso and its oracle properties. *J. Amer. Statist. Assoc.* **101** 1418–1429. MR2279469 https://doi.org/10.1198/016214506000000735

Bernoulli **30**(3), 2024, 2501–2520
https://doi.org/10.3150/23-BEJ1682

Moment inequalities for sums of weakly dependent random fields

GILLES BLANCHARD[1,a], ALEXANDRA CARPENTIER[2,b] and OLEKSANDR ZADOROZHNYI[3,c]

[1]*Institut de Mathématiques d'Orsay, Université Paris-Saclay, Paris, France,* [a]*gilles.blanchard@universite-paris-saclay.fr*
[2]*Institut für Mathematik, Universität Potsdam, Germany,* [b]*carpentier@uni-potsdam.de*
[3]*Lerhstuhl für Mathematische Statistik, TUM School of Computation, Information and Technology Technical University of Munich, Germany,* [c]*oleksandr.zadorozhnyi@tum.de*

We derive both Azuma-Hoeffding and Burkholder-type inequalities for partial sums over a rectangular grid of dimension d of a random field satisfying a weak dependency assumption of projective type: the difference between the expectation of an element of the random field and its conditional expectation given the rest of the field at a distance more than δ is bounded, in L_p distance, by a known decreasing function of δ. The analysis is based on the combination of a multi-scale approximation of random sums by martingale difference sequences, and of a careful decomposition of the domain. The obtained results extend previously known bounds under comparable hypotheses, and do not use the assumption of commuting filtrations.

Keywords: Burkholder-type inequalities; concentration inequalities; weakly dependent random fields

References

Azuma, K. (1967). Weighted sums of certain dependent random variables. *Tohoku Math. J. (2)* **19** 357–367. MR0221571 https://doi.org/10.2748/tmj/1178243286

Basu, A.K. and Dorea, C.C.Y. (1979). On functional central limit theorem for stationary martingale random fields. *Acta Math. Acad. Sci. Hung.* **33** 307–316. MR0542479 https://doi.org/10.1007/BF01902565

Bickel, P.J. and Bühlmann, P. (1999). A new mixing notion and functional central limit theorems for a sieve bootstrap in time series. *Bernoulli* **5** 413–446. MR1693612 https://doi.org/10.2307/3318711

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

Cairoli, R. and Walsh, J.B. (1975). Stochastic integrals in the plane. *Acta Math.* **134** 111–183. MR0420845 https://doi.org/10.1007/BF02392100

Dedecker, J. (1998). A central limit theorem for stationary random fields. *Probab. Theory Related Fields* **110** 397–426. MR1616496 https://doi.org/10.1007/s004400050153

Dedecker, J. (2001). Exponential inequalities and functional central limit theorems for a random fields. *ESAIM Probab. Stat.* **5** 77–104. MR1875665 https://doi.org/10.1051/ps:2001103

Dedecker, J. (2020). Personal communication.

Dedecker, J. and Merlevède, F. (2015). Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.* **125** 3401–3429. MR3357614 https://doi.org/10.1016/j.spa.2015.05.002

Dedecker, J. and Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Related Fields* **132** 203–236. MR2199291 https://doi.org/10.1007/s00440-004-0394-3

Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence: With Examples and Applications*. *Lecture Notes in Statistics* **190**. New York: Springer. MR2338725

Dehling, H. and Philipp, W. (1982). Almost sure invariance principles for weakly dependent vector-valued random variables. *Ann. Probab.* **10** 689–701. MR0659538

1350-7265

Doukhan, P. (1994). *Mixing: Properties and Examples*. *Lecture Notes in Statistics* **85**. New York: Springer. MR1312160 https://doi.org/10.1007/978-1-4612-2642-0

Doukhan, P., León, J. and Portal, F. (1984). Vitesse de convergence dans le théorème central limite pour des variables aléatoires mélangeantes à valeurs dans un espace de Hilbert. *C. R. Acad. Sci. Paris Sér. I Math.* **298** 305–308. MR0765429

El Machkouri, M. and Giraudo, D. (2016). Orthomartingale-coboundary decomposition for stationary random fields. *Stoch. Dyn.* **16** 1650017. MR3522451 https://doi.org/10.1142/S0219493716500179

El Machkouri, M., Volný, D. and Wu, W.B. (2013). A central limit theorem for stationary random fields. *Stochastic Process. Appl.* **123** 1–14. MR2988107 https://doi.org/10.1016/j.spa.2012.08.014

Fazekas, I. (2005). Burkholder's inequality for multiindex martingales. *Ann. Math. Inform.* **32** 45–51. MR2264866

Giraudo, D. (2018). Invariance principle via orthomartingale approximation. *Stoch. Dyn.* **18** 1850043. MR3869881 https://doi.org/10.1142/S0219493718500430

Giraudo, D. (2019). Deviation inequalities for Banach space valued martingales differences sequences and random fields. *ESAIM Probab. Stat.* **23** 922–946. MR4046858 https://doi.org/10.1051/ps/2019016

Giraudo, D. (2021a). Maximal function associated to the bounded law of the iterated logarithms via orthomartingale approximation. *J. Math. Anal. Appl.* **496** 124792. MR4186670 https://doi.org/10.1016/j.jmaa.2020.124792

Giraudo, D. (2021b). An exponential inequality for orthomartingale differences random fields and some applications. Available at arXiv:2106.13128.

Giraudo, D. (2022). Bound on the maximal function associated to the law of the iterated logarithms for Bernoulli random fields. *Stochastics* **94** 248–276. MR4372657 https://doi.org/10.1080/17442508.2021.1920942

Gundy, R.F. and Varopoulos, N.T. (1976). A martingale that occurs in harmonic analysis. *Ark. Mat.* **14** 179–187. MR0448539 https://doi.org/10.1007/BF02385833

Khoshnevisan, D. (2002). *Multiparameter Processes. an Introduction to Random Fields*. *Springer Monographs in Mathematics*. New York: Springer. MR1914748 https://doi.org/10.1007/b97363

Klicnarová, J., Volný, D. and Wang, Y. (2016). Limit theorems for weighted Bernoulli random fields under Hannan's condition. *Stochastic Process. Appl.* **126** 1819–1838. MR3483738 https://doi.org/10.1016/j.spa.2015.12.006

Maume-Deschamps, V. (2006). Exponential inequalities and functional estimations for weak dependent data; applications to dynamical systems. *Stoch. Dyn.* **6** 535–560. MR2285515 https://doi.org/10.1142/S0219493706001876

McLeish, D.L. (1975). Invariance principles for dependent variables. *Z. Wahrsch. Verw. Gebiete* **32** 165–178. MR0388483 https://doi.org/10.1007/BF00532611

Métraux, C. (1978). Quelques inégalités pour martingales à paramètre bidimensionnel. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*. *Lecture Notes in Math.* **649** 170–179. Berlin: Springer. MR0520006

Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33** 798–815. MR2123210 https://doi.org/10.1214/009117904000001035

Peligrad, M., Utev, S. and Wu, W.B. (2007). A maximal $\mathbb{L}_p$-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* **135** 541–550. MR2255301 https://doi.org/10.1090/S0002-9939-06-08488-7

Pinelis, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** 1679–1706. MR1331198

Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **29** 587–597. MR1251142

Rio, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. *Mathématiques & Applications (Berlin) [Mathematics & Applications]* **31**. Berlin: Springer. MR2117923

Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions. *J. Theoret. Probab.* **22** 146–163. MR2472010 https://doi.org/10.1007/s10959-008-0155-9

Sang, H. and Xiao, Y. (2018). Exact moderate and large deviations for linear random fields. *J. Appl. Probab.* **55** 431–449. MR3832897 https://doi.org/10.1017/jpr.2018.28

Volný, D. (2015). A central limit theorem for fields of martingale differences. *C. R. Math. Acad. Sci. Paris* **353** 1159–1163. MR3427925 https://doi.org/10.1016/j.crma.2015.09.017

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

Bernoulli **30**(3), 2024, 2521–2545
https://doi.org/10.3150/23-BEJ1683

Optimal choice of bootstrap block length for periodically correlated time series

PATRICE BERTAIL[1,a] and ANNA E. DUDEK[2,b]

[1]*MODAL'X, UPL, Univ. Paris Nanterre, CNRS, F92000 Nanterre, France,* [a]*patrice.bertail@gmail.com*
[2]*Department of Applied Mathematics, AGH University of Krakow, al. Mickiewicza 30, 30-059 Krakow, Poland,* [b]*aedudek@agh.edu.pl*

This paper discusses the problem of choosing the optimal block length for two block bootstrap methods designed for periodically correlated processes. These are the Generalized Seasonal Block Bootstrap and the Extension of Moving Block Bootstrap. Two estimation problems are considered: the overall mean and the seasonal means. In both cases, the optimal block length is obtained by minimizing the mean squared error of the corresponding bootstrap variance estimator and in all cases it is proportional to the cube root of the sample size and should be a multiple of the period length plus one observation to avoid some bias. Finally, the results of the performed simulation are presented, in which optimal blocks lengths are estimated for several periodically correlated time series.

Keywords: Asymptotic expansion; bifrequency square; periodic time series; spectral estimation

References

Antoni, J. (2009). Cyclostationarity by examples. *Mech. Syst. Signal Process.* **23** 987–1036.

Bertail, P. (2011). Comments on: Subsampling weakly dependent time series and application to extremes [MR2864705]. *TEST* **20** 487–490. MR2864708 https://doi.org/10.1007/s11749-011-0272-0

Bertail, P. and Dudek, A.E. (2024). Supplement to "Optimal choice of bootstrap block length for periodically correlated time series." https://doi.org/10.3150/23-BEJ1683SUPP

Bickel, P.J. and Sakov, A. (2008). On the choice of *m* in the *m* out of *n* bootstrap and confidence bounds for extrema. *Statist. Sinica* **18** 967–985. MR2440400

Brillinger, D.R. (2001). *Time Series: Data Analysis and Theory*. *Classics in Applied Mathematics* **36**. Philadelphia, PA: SIAM. Reprint of the 1981 edition. MR1853554 https://doi.org/10.1137/1.9780898719246

Bühlmann, P. and Künsch, H. R. (1999). Block length selection in the bootstrap for time series. *Comput. Statist. Data Anal.* **31** 295–310.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14** 1171–1179. MR0856813 https://doi.org/10.1214/aos/1176350057

Chan, V., Lahiri, S.N. and Meeker, W.Q. (2004). Block bootstrap estimation of the distribution of cumulative outdoor degradation. *Technometrics* **46** 215–224. MR2060017 https://doi.org/10.1198/004017004000000266

de Sousa, B. and Michailidis, G. (2012). A diagnostic plot for estimating the tail index of a distribution. *J. Comput. Graph. Statist.* **13** 974–995. MR2109061 https://doi.org/10.1198/106186004X12335

Dudek, A.E. (2015). Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series. *Metrika* **78** 313–335. MR3320900 https://doi.org/10.1007/s00184-014-0505-9

Dudek, A.E. (2018). Block bootstrap for periodic characteristics of periodically correlated time series. *J. Nonparametr. Stat.* **30** 87–124. MR3756234 https://doi.org/10.1080/10485252.2017.1404060

Dudek, A.E., Hurd, H. and Wójtowicz, W. (2015a). PARMA models with applications in R. In *Cyclostationarity: Theory and Methods - II* (F. Chaari et al., eds.). *Applied Condition Monitoring* **3** 131–153. Switzerland: Springer. Chapter 7. https://doi.org/10.1007/978-3-319-16330-7_7

Dudek, A.E., Hurd, H. and Wójtowicz, W. (2015b). perARMA: Package for periodic time series analysis. R package version 1.5. Available at http://cran.r-project.org/web/packages/perARMA.

Dudek, A.E., Hurd, H. and Wójtowicz, W. (2016). Periodic autoregressive moving average methods based on Fourier representation of periodic coefficients. *Wiley Interdiscip. Rev.: Comput. Stat.* **8** 130–149. MR3511662 https://doi.org/10.1002/wics.1380

Dudek, A.E., Paparoditis, E. and Politis, D.N. (2016). Generalized seasonal tapered block bootstrap. *Statist. Probab. Lett.* **115** 27–35. MR3498365 https://doi.org/10.1016/j.spl.2016.03.022

Dudek, A.E. and Potorski, P. (2020). Bootstrapping the autocovariance of PC time series - a simulation study. In *Cyclostationarity: Theory and Methods - IV; Contributions to the 10th Workshop on Cyclostationary Systems and Their Applications, February 2017, Grodek, Poland* (F. Chaari, J. Leskow, R. Zimroz, A. Wylomanska and A. Dudek, eds.). *Applied Condition Monitoring* **16** 41–55. Cham: Springer.

Dudek, A.E., Leśkow, J., Paparoditis, E. and Politis, D.N. (2014). A generalized block bootstrap for seasonal time series. *J. Time Series Anal.* **35** 89–114. MR3166348 https://doi.org/10.1002/jtsa.12053

Gardner, W.A. (1994). *Cyclostationarity in Communications and Signal Processing*. IEEE Press.

Gardner, W.A., Napolitano, A. and Paura, L. (2006). Cyclostationarity: Half a century of research. *Signal Process.* **86** 639–697.

Gladyshev, E.G. (1961). Periodically correlated random sequences. *Sov. Math.* **2** 385–388.

Hall, P., Horowitz, J.L. and Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82** 561–574. MR1366282 https://doi.org/10.1093/biomet/82.3.561

Hannan, E.J. (1955). A test for singularities in Sydney rainfall. *Aust. J. Phys.* **8** 289–297.

Hurd, H.L. and Miamee, A. (2007). *Periodically Correlated Random Sequences: Spectral Theory and Practice*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR2348769 https://doi.org/10.1002/9780470182833

Jones, R.H. and Brelsford, W.M. (1967). Time series with periodic structure. *Biometrika* **54** 403–408. MR0223041 https://doi.org/10.1093/biomet/54.3-4.403

Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241. MR1015147 https://doi.org/10.1214/aos/1176347265

Lahiri, S.N. (1999). Theoretical comparisons of block bootstrap methods. *Ann. Statist.* **27** 386–404. MR1701117 https://doi.org/10.1214/aos/1018031117

Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*. *Springer Series in Statistics*. New York: Springer. MR2001447 https://doi.org/10.1007/978-1-4757-3803-2

Lahiri, S.N., Furukawa, K. and Lee, Y.-D. (2007). A nonparametric plug-in rule for selecting optimal block lengths for block bootstrap methods. *Stat. Methodol.* **4** 292–321. MR2380557 https://doi.org/10.1016/j.stamet.2006.08.002

Lenart, Ł. (2011). Asymptotic distributions and subsampling in spectral analysis for almost periodically correlated time series. *Bernoulli* **17** 290–319. MR2797993 https://doi.org/10.3150/10-BEJ269

Liu, R.Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap (East Lansing, MI, 1990). Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.* 225–248. New York: Wiley. MR1197787

Loève, M. (1963). *Probability Theory*, 3rd ed. Princeton, N.J.-Toronto, Ont.-London: D. Van Nostrand Co., Inc. MR0203748

Napolitano, A. (2012). *Generalizations of Cyclostationary Signal Processing: Spectral Analysis and Applications*. Wiley-IEEE Press.

Napolitano, A. (2016). Cyclostationarity: New trends and applications. *Signal Process.* **120** 385–408.

Napolitano, A. (2019). *Cyclostationary Processes and Time Series: Theory, Applications, and Generalizations*. Academic Press.

Nematollahi, A.R. and Subba Rao, T. (2005). On the spectral density estimation of periodically correlated (cyclostationary) time series. *Sankhyā* **67** 568–589. MR2235579

Nordman, D.J. (2009). A note on the stationary bootstrap's variance. *Ann. Statist.* **37** 359–370. MR2488355 https://doi.org/10.1214/07-AOS567

Nordman, D.J. and Lahiri, S.N. (2014). Convergence rates of empirical block length selectors for block bootstrap. *Bernoulli* **20** 958–978. MR3178523 https://doi.org/10.3150/13-BEJ511

Paparoditis, E. and Politis, D.N. (2001). Tapered block bootstrap. *Biometrika* **88** 1105–1119. MR1872222 https://doi.org/10.1093/biomet/88.4.1105

Politis, D.N. (2001). Resampling time series with seasonal components. In *Frontiers in Data Mining and Bioinformatics: Proceedings of the 33rd Symposium on the Interface of Computing Science and Statistics* 619–621.

Politis, D.N. and Romano, J.P. (1992). A circular block-resampling procedure for stationary data. In *Exploring the Limits of Bootstrap (East Lansing, MI, 1990). Wiley Ser. Probab. Math. Statist. Probab. Math. Statist.* 263–270. New York: Wiley. MR1197789

Synowiecki, R. (2007). Consistency and application of moving block bootstrap for non-stationary time series with periodic and almost periodic structure. *Bernoulli* **13** 1151–1178. MR2364230 https://doi.org/10.3150/07-BEJ102