BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Thirty One Number One February 2025 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 One Number One February 2025 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 31 Number 1 February 2025 Pages 1–841 ISI/BS

Volume 31 Number 1   February 2025   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 2025 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **31**(1), 2025, 1–29
https://doi.org/10.3150/23-BEJ1637

An efficient averaged stochastic Gauss-Newton algorithm for estimating parameters of nonlinear regressions models

PEGGY CÉNAC[1,a], ANTOINE GODICHON-BAGGIONI[2,b] and BRUNO PORTIER[3,c]

[1]*Institut de Mathématiques de Bourgogne, UMR 5584 CNRS, Université de Bourgogne, F-21000 Dijon, France,* [a]*peggy.cenac@u-bourgogne.fr*
[2]*Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, 75005 Paris, France,* [b]*antoine.godichon_baggioni@upmc.fr*
[3]*INSA Rouen Normandie, Laboratoire de Mathématiques de l'INSA Rouen, 76800 Saint Etienne du Rouvray, France,* [c]*bruno.portier@insa-rouen.fr*

Nonlinear regression models are a standard tool for modelling real phenomena, with several applications in machine learning, ecology, econometrics, and so forth. Estimating the parameters of these models has attracted a lot of attention during many years. We focus here on a recursive method for estimating parameters of nonlinear regressions. Indeed, these kinds of methods, whose most famous instances are probably the stochastic gradient algorithm and its averaged version, enable to deal efficiently with massive data arriving sequentially. Nevertheless, they can be, in practice, non-robust when the eigenvalues of the Hessian of the functional we would like to minimize are at different scales. To avoid this problem, we first focus on an online Stochastic Gauss-Newton algorithm. In order to improve the estimates behaviour in the case of bad initialization, we then introduce a new Averaged Stochastic Gauss-Newton algorithm and prove its asymptotic efficiency.

Keywords: Nonlinear regression model; online estimation; stochastic Gauss-Newton algorithm; stochastic Newton algorithm; stochastic optimization

References

Bach, F. (2014). Adaptivity of averaged stochastic gradient descent to local strong convexity for logistic regression. *J. Mach. Learn. Res.* **15** 595–627. MR3190851

Bach, F. and Moulines, E. (2013). Non-strongly-convex smooth stochastic approximation with convergence rate O (1/n). In *Adv. Neural. Inf. Process. Syst* 773–781.

Bates, D.M. and Watts, D.G. (1988). *Nonlinear Regression Analysis and Its Applications*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. MR1060528 https://doi.org/10.1002/9780470316757

Bercu, B., Godichon-Baggioni, A. and Portier, B. (2020). An efficient stochastic Newton algorithm for parameter estimation in logistic regressions. *SIAM J. Control Optim.* **58** 348–367. MR4054588 https://doi.org/10.1137/19M1261717

Byrd, R.H., Hansen, S.L., Nocedal, J. and Singer, Y. (2016). A stochastic quasi-Newton method for large-scale optimization. *SIAM J. Optim.* **26** 1008–1031. MR3485979 https://doi.org/10.1137/140954362

Cardot, H., Cénac, P. and Godichon-Baggioni, A. (2017). Online estimation of the geometric median in Hilbert spaces: Nonasymptotic confidence balls. *Ann. Statist.* **45** 591–614. MR3650394 https://doi.org/10.1214/16-AOS1460

Cardot, H. and Godichon-Baggioni, A. (2017). Fast estimation of the median covariation matrix with application to online robust principal components analysis. *TEST* **26** 461–480. MR3683984 https://doi.org/10.1007/s11749-016-0519-x

Cénac, P., Godichon-Baggioni, A. and Portier, B. (2025). Supplement to "An efficient averaged stochastic Gauss-Newton algorithm for estimating parameters of nonlinear regressions models." https://doi.org/10.3150/23-BEJ1637SUPP

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

Fabian, V. (1978). On asymptotically efficient recursive estimation. *Ann. Statist.* **6** 854–866. MR0478506

Gadat, S. and Panloup, F. (2023). Optimal non-asymptotic analysis of the Ruppert-Polyak averaging stochastic algorithm. *Stochastic Process. Appl.* **156** 312–348. MR4518751 https://doi.org/10.1016/j.spa.2022.11.012

Giurcăneanu, C.D., Tăbuş, I. and Astola, J. (2005). Clustering time series gene expression data based on sum-of-exponentials fitting. *EURASIP J. Appl. Signal Process.* **8** 1159–1173. MR2168617

Godichon-Baggioni, A. (2019a). Online estimation of the asymptotic variance for averaged stochastic gradient algorithms. *J. Statist. Plann. Inference* **203** 1–19. MR3950590 https://doi.org/10.1016/j.jspi.2019.01.001

Godichon-Baggioni, A. (2019b). L^p and almost sure rates of convergence of averaged stochastic gradient algorithms: Locally strongly convex objective. *ESAIM Probab. Stat.* **23** 841–873. MR4045544 https://doi.org/10.1051/ps/2019011

Huet, S., Bouvier, A., Poursat, M.-A. and Jolivet, E. (2004). *Statistical Tools for Nonlinear Regression: A Practical Guide with S-PLUS and R Examples*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2024808

Jakubowski, A. (1988). Tightness criteria for random measures with application to the principle of conditioning in Hilbert spaces. *Probab. Math. Statist.* **9** 95–114. MR0945679

Jennrich, R.I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Stat.* **40** 633–643. MR0238419 https://doi.org/10.1214/aoms/1177697731

Komori, O., Eguchi, S., Ikeda, S., Okamura, H., Ichinokawa, M. and Nakayama, S. (2016). An asymmetric logistic regression model for ecological data. *Methods Ecol. Evol.* **7** 249–260.

Kushner, H.J. and Yin, G.G. (2003). *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. *Applications of Mathematics (New York). Stochastic Modelling and Applied Probability* **35**. New York: Springer. MR1993642

Lai, T.L. (1994). Asymptotic properties of nonlinear least squares estimates in stochastic regression models. *Ann. Statist.* **22** 1917–1930. MR1329175 https://doi.org/10.1214/aos/1176325764

Lippmann, R. (1987). An introduction to computing with neural nets. *IEEE ASSP Mag.* **4** 4–22. https://doi.org/10.1109/MASSP.1987.1165576

Lucchi, A., McWilliams, B. and Hofmann, T. (2015). A variance reduced stochastic Newton method. arXiv preprint, arXiv:1503.08316.

Mokhtari, A. and Ribeiro, A. (2014). RES: Regularized stochastic BFGS algorithm. *IEEE Trans. Signal Process.* **62** 6089–6104. MR3281504 https://doi.org/10.1109/TSP.2014.2357775

Ngia, L.S. and Sjoberg, J. (2000). Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm. *IEEE Trans. Signal Process.* **48** 1915–1927.

Pelletier, M. (1998). On the almost sure asymptotic behaviour of stochastic algorithms. *Stochastic Process. Appl.* **78** 217–244. MR1654569 https://doi.org/10.1016/S0304-4149(98)00029-5

Pelletier, M. (2000). Asymptotic almost sure efficiency of averaged stochastic algorithms. *SIAM J. Control Optim.* **39** 49–72. MR1780908 https://doi.org/10.1137/S0363012998308169

Pollard, D. and Radchenko, P. (2006). Nonlinear least-squares estimation. *J. Multivariate Anal.* **97** 548–562. MR2234037 https://doi.org/10.1016/j.jmva.2005.04.002

Polyak, B.T. and Juditsky, A.B. (1992). Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.* **30** 838–855. MR1167814 https://doi.org/10.1137/0330046

Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400–407. MR0042668 https://doi.org/10.1214/aoms/1177729586

Rumelhart, D.E. and McClelland, J.L. (1987). A general framework for parallel distributed processing. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations* 45–76.

Ruppert, D. (1988). Efficient estimations from a slowly convergent Robbins-Monro process Technical Report, Cornell University Operations Research and Industrial Engineering.

Sherman, J. and Morrison, W.J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *Ann. Math. Stat.* **21** 124–127. MR0035118 https://doi.org/10.1214/aoms/1177729893

Skouras, K. (2000). Strong consistency in nonlinear stochastic regression models. *Ann. Statist.* **28** 871–879. MR1792791 https://doi.org/10.1214/aos/1015952002

Suárez, E., Pérez, C.M., Rivera, R. and Martínez, M.N. (2017). *Applications of Regression Models in Epidemiology*. New York: Wiley.

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

van de Geer, S. and Wegkamp, M. (1996). Consistency for the least squares estimator in nonparametric regression. *Ann. Statist.* **24** 2513–2523. MR1425964 https://doi.org/10.1214/aos/1032181165

Varian, H.R. (2014). Big data: New tricks for econometrics. *J. Econ. Perspect.* **28** 3–28.

Wu, C.-F. (1981). Asymptotic theory of nonlinear least squares estimation. *Ann. Statist.* **9** 501–513. MR0615427

Yang, Z., Wang, Z., Liu, H., Eldar, Y. and Zhang, T. (2016). Sparse nonlinear regression: Parameter estimation under nonconvexity. In *Proceedings of the 33rd International Conference on Machine Learning* (M.F. Balcan and K.Q. Weinberger, eds.). *Proc. Mach. Learn. Res.* **48** 2472–2481. New York, USA: PMLR.

Yao, J.-F. (2000). On least squares estimation for stable nonlinear AR processes. *Ann. Inst. Statist. Math.* **52** 316–331. MR1763565 https://doi.org/10.1023/A:1004117906532

Bernoulli **31**(1), 2025, 30–54
https://doi.org/10.3150/23-BEJ1688

Poisson approximation of Poisson-driven point processes and extreme values in stochastic geometry

MORITZ OTTO[a]

Department of Mathematics, Aarhus University, Aarhus, Denmark, [a]*otto@math.au.dk*

We study point processes that consist of certain centres of point tuples of an underlying Poisson process. Such processes arise in stochastic geometry in the study of exceedances of various functionals describing geometric properties of the Poisson process. We use a coupling of the point process with its Palm version to prove a general Poisson limit theorem. We then combine our general result with the theory of asymptotic shapes of large cells (Kendall's problem) in random mosaics and prove Poisson limit theorems for large cells (with respect to a general size functional) in the Poisson-Voronoi and Poisson-Delaunay mosaic. As a consequence, we establish scaling limits for maxima of concrete size functionals. This extends extreme value results from (*Extremes* **17** (2014) 359–385) and (*Stochastic Process. Appl.* **124** (2014) 2917–2953).

Keywords: Chen-Stein method; Delaunay mosaic; Kendall's problem; maximum cell; Palm distribution; point process approximation; Poisson process; stopping set; total variation distance; Voronoi mosaic

References

[1] Arratia, R., Goldstein, L. and Gordon, L. (1989). Two moments suffice for Poisson approximations: The Chen-Stein method. *Ann. Probab.* **17** 9–25. MR0972770

[2] Arratia, R., Goldstein, L. and Gordon, L. (1990). Poisson approximation and the Chen-Stein method. *Statist. Sci.* **5** 403–434. With comments and a rejoinder by the authors. MR1092983

[3] Bobrowski, O., Schulte, M. and Yogeshwaran, D. (2022). Poisson process approximation under stabilization and Palm coupling. *Ann. Henri Lebesgue* **5** 1489–1534. MR4526259 https://doi.org/10.5802/ahl.156

[4] Bonnet, G. (2018). Small cells in a Poisson hyperplane tessellation. *Adv. in Appl. Math.* **95** 31–52. MR3759210 https://doi.org/10.1016/j.aam.2017.11.002

[5] Bonnet, G., Calka, P. and Reitzner, M. (2018). Cells with many facets in a Poisson hyperplane tessellation. *Adv. Math.* **324** 203–240. MR3733885 https://doi.org/10.1016/j.aim.2017.11.016

[6] Bonnet, G. and Chenavier, N. (2020). The maximal degree in a Poisson-Delaunay graph. *Bernoulli* **26** 948–979. MR4058357 https://doi.org/10.3150/19-BEJ1123

[7] Calka, P. and Chenavier, N. (2014). Extreme values for characteristic radii of a Poisson-Voronoi tessellation. *Extremes* **17** 359–385. MR3252817 https://doi.org/10.1007/s10687-014-0184-y

[8] Chenavier, N. (2014). A general study of extremes of stationary tessellations with examples. *Stochastic Process. Appl.* **124** 2917–2953. MR3217429 https://doi.org/10.1016/j.spa.2014.04.009

[9] Chenavier, N. (2015). The extremal index for a random tessellation. In *Geometric Science of Information* (F. Nielsen and F. Barbaresco, eds.). *Lecture Notes in Computer Science* **9389** 171–178. Cham: Springer. MR3442197 https://doi.org/10.1007/978-3-319-25040-3_19

[10] Chenavier, N. and Hemsley, R. (2016). Extremes for the inradius in the Poisson line tessellation. *Adv. in Appl. Probab.* **48** 544–573. MR3511775

[11] Chenavier, N., Henze, N. and Otto, M. (2022). Limit laws for large kth-nearest neighbor balls. *J. Appl. Probab.* **59** 880–894. MR4480085 https://doi.org/10.1017/jpr.2021.92

[12] Chenavier, N. and Nagel, W. (2019). The largest order statistics for the inradius in an isotropic STIT tessellation. *Extremes* **22** 571–598. MR4031850 https://doi.org/10.1007/s10687-019-00356-0

[13] Chenavier, N. and Robert, C.Y. (2018). Cluster size distributions of extreme values for the Poisson-Voronoi tessellation. *Ann. Appl. Probab.* **28** 3291–3323. MR3861814 https://doi.org/10.1214/17-AAP1345
[14] Decreusefond, L., Schulte, M. and Thäle, C. (2016). Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Probab.* **44** 2147–2197. MR3502603 https://doi.org/10.1214/15-AOP1020
[15] Györfi, L., Henze, N. and Walk, H. (2019). The limit distribution of the maximum probability nearest-neighbour ball. *J. Appl. Probab.* **56** 574–589. MR3986953 https://doi.org/10.1017/jpr.2019.37
[16] Hug, D., Reitzner, M. and Schneider, R. (2004). Large Poisson-Voronoi cells and Crofton cells. *Adv. in Appl. Probab.* **36** 667–690. MR2079908 https://doi.org/10.1239/aap/1093962228
[17] Hug, D., Reitzner, M. and Schneider, R. (2004). The limit shape of the zero cell in a stationary Poisson hyperplane tessellation. *Ann. Probab.* **32** 1140–1167. MR2044676 https://doi.org/10.1214/aop/1079021474
[18] Hug, D. and Schneider, R. (2004). Large cells in Poisson-Delaunay tessellations. *Discrete Comput. Geom.* **31** 503–514. MR2053496 https://doi.org/10.1007/s00454-003-0818-3
[19] Hug, D. and Schneider, R. (2005). Large typical cells in Poisson-Delaunay mosaics. *Rev. Roumaine Math. Pures Appl.* **50** 657–670. MR2204143
[20] Hug, D. and Schneider, R. (2007). Asymptotic shapes of large cells in random tessellations. *Geom. Funct. Anal.* **17** 156–191. MR2306655 https://doi.org/10.1007/s00039-007-0592-0
[21] Hug, D. and Schneider, R. (2007). Typical cells in Poisson hyperplane tessellations. *Discrete Comput. Geom.* **38** 305–319. MR2343310 https://doi.org/10.1007/s00454-007-1340-9
[22] Kovalenko, I.N. (1997). Proof of David Kendall's conjecture concerning the shape of large random polygons. *Cybernet. Systems Anal.* **33** 461–467. MR1609157 https://doi.org/10.1007/BF02733102
[23] Last, G. and Otto, M. (2023). Disagreement coupling of Gibbs processes with an application to Poisson approximation. *Ann. Appl. Probab.* **33** 4091–4126. MR4663505 https://doi.org/10.1214/22-aap1916
[24] Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge: Cambridge Univ. Press. MR3791470
[25] Mitzenmacher, M. and Upfal, E. (2017). *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*, 2nd ed. Cambridge: Cambridge Univ. Press. MR3674428
[26] Otto, M. (2023). Extremal behavior of large cells in the Poisson hyperplane mosaic. *Electron. J. Probab.* **28** Paper No. 162, 29. MR4673664 https://doi.org/10.1214/23-ejp1049
[27] Otto, M. (2025). Supplement to "Poisson approximation of Poisson-driven point processes and extreme values in stochastic geometry." https://doi.org/10.3150/23-BEJ1688SUPP
[28] Otto, M. and Thäle, C. (2023). Large nearest neighbour balls in hyperbolic stochastic geometry. *Extremes* **26** 413–431. MR4627350 https://doi.org/10.1007/s10687-023-00470-0
[29] Penrose, M. and Yukich, J.E. (2005). Normal approximation in geometric probability. In *Stein's Method and Applications* (A.D. Barbour and L.H.Y. Chen, eds.). 37–58. Singapor: World Scientific.
[30] Penrose, M.D. (2007). Gaussian limits for random geometric measures. *Electron. J. Probab.* **12** 989–1035. MR2336596 https://doi.org/10.1214/EJP.v12-429
[31] Pianoforte, F. and Schulte, M. (2021). Poisson approximation with applications to stochastic geometry. *Electron. J. Probab.* **26** Paper No. 149, 36. MR4346678 https://doi.org/10.1214/21-ejp723
[32] Rathie, P.N. (1992). On the volume distribution of the typical Poisson-Delaunay cell. *J. Appl. Probab.* **29** 740–744. MR1174447 https://doi.org/10.1017/s0021900200043540
[33] Schneider, R. (2019). Small faces in stationary Poisson hyperplane tessellations. *Math. Nachr.* **292** 1811–1822. MR3994305 https://doi.org/10.1002/mana.201800366
[34] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. *Probability and Its Applications (New York)*. Berlin: Springer. MR2455326 https://doi.org/10.1007/978-3-540-78859-1
[35] Schuhmacher, D. (2005). Distance estimates for Poisson process approximations of dependent thinnings. *Electron. J. Probab.* **10** 165–201. MR2120242 https://doi.org/10.1214/EJP.v10-237
[36] Schuhmacher, D. (2009). Distance estimates for dependent thinnings of point processes with densities. *Electron. J. Probab.* **14** 1080–1116. MR2506126 https://doi.org/10.1214/EJP.v14-643
[37] Schulte, M. and Thäle, C. (2012). The scaling limit of Poisson-driven order statistics with applications in geometric probability. *Stochastic Process. Appl.* **122** 4096–4120. MR2971726 https://doi.org/10.1016/j.spa.2012.08.011

[38] Schulte, M. and Thäle, C. (2016). Poisson point process convergence and extreme values in stochastic geometry. In *Stochastic Analysis for Poisson Point Processes* (G. Peccati and M. Reitzner, eds.). *Bocconi Springer Ser.* **7** 255–294. Bocconi Univ. Press. MR3585403 https://doi.org/10.1007/978-3-319-05233-5_8

Bernoulli **31**(1), 2025, 55–80
https://doi.org/10.3150/23-BEJ1703

An edge CLT for the log determinant of Wigner ensembles

IAIN M. JOHNSTONE[1,a], YEGOR KLOCHKOV[2,b], ALEXEI ONATSKI[3,c] and DAMIAN PAVLYSHYN[4,d]

[1]*Department of Statistics, Stanford University, Stanford, USA,* [a]*imj@stanford.edu*
[2]*ByteDance AI Lab, London, UK,* [b]*yk376@cam.ac.uk*
[3]*Faculty of Economics, University of Cambridge, Cambridge, UK,* [c]*ao319@cam.ac.uk*
[4]*Modelling and Biostatistics Group, Burnet Institute, Melbourne, Australia,* [d]*damian.pavlyshyn@burnet.edu.au*

We derive a Central Limit Theorem (CLT) for $\log|\det(W_N - E_N)|$, where W_N is a Wigner matrix, and E_N is local to the edge of the semi-circle law. Precisely, $E_N = 2 + N^{-2/3}\sigma_N$ with σ_N being either a constant (possibly negative), or a sequence of positive real numbers, slowly diverging to infinity so that $\sigma_N \ll \log^2 N$. We also extend our CLT to cover spiked Wigner matrices. Our interest in the CLT is motivated by its applications to statistical testing in critically spiked models and to the fluctuations of the free energy in the spherical Sherrington-Kirkpatrick model of statistical physics.

Keywords: CLT; log determinant; edge of the semi-circle law; Wigner matrix

References

Albrecht, J.T., Chan, C.P. and Edelman, A. (2009). Sturm sequences and random eigenvalue distributions. *Found. Comput. Math.* **9** 461–483. MR2519867 https://doi.org/10.1007/s10208-008-9037-x

Anderson, G.W., Guionnet, A. and Zeitouni, O. (2010). *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge: Cambridge Univ. Press. MR2760897

Augeri, F., Butez, R. and Zeitouni, O. (2023). A CLT for the characteristic polynomial of random Jacobi matrices, and the GβE. *Probab. Theory Related Fields* **186** 1–89. MR4586217 https://doi.org/10.1007/s00440-023-01194-9

Bai, Z.D. and Yao, J. (2005). On the convergence of the spectral empirical process of Wigner matrices. *Bernoulli* **11** 1059–1092. MR2189081 https://doi.org/10.3150/bj/1137421640

Baik, J., Ben Arous, G. and Péché, S. (2005). Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.* **33** 1643–1697. MR2165575 https://doi.org/10.1214/009117905000000233

Baik, J. and Lee, J.O. (2016). Fluctuations of the free energy of the spherical Sherrington-Kirkpatrick model. *J. Stat. Phys.* **165** 185–224. MR3554380 https://doi.org/10.1007/s10955-016-1610-0

Baik, J. and Lee, J.O. (2017). Fluctuations of the free energy of the spherical Sherrington-Kirkpatrick model with ferromagnetic interaction. *Ann. Henri Poincaré* **18** 1867–1917. MR3649446 https://doi.org/10.1007/s00023-017-0562-5

Baik, J., Lee, J.O. and Wu, H. (2018). Ferromagnetic to paramagnetic transition in spherical spin glass. *J. Stat. Phys.* **173** 1484–1522. MR3878351 https://doi.org/10.1007/s10955-018-2150-6

Benaych-Georges, F. and Knowles, A. (2017). Local semicircle law for Wigner matrices. In *Advanced Topics in Random Matrices*. *Panor. Synthèses* **53** 1–90. Paris: Soc. Math. France. MR3792624

Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001

Bourgade, P. and Mody, K. (2019). Gaussian fluctuations of the determinant of Wigner matrices. *Electron. J. Probab.* **24** Paper No. 96. MR4017114 https://doi.org/10.1214/19-ejp356

1350-7265

Bourgade, P., Mody, K. and Pain, M. (2022). Optimal local law and central limit theorem for β-ensembles. *Comm. Math. Phys.* **390** 1017–1079. MR4389077 https://doi.org/10.1007/s00220-022-04311-2

Collins-Woodfin, E.W. and Le, H.G. (2024). An edge CLT for the log determinant of Laguerre beta ensembles. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at arXiv 2209.03271.

Duy, T.K. (2017). Distributions of the determinants of Gaussian beta ensembles. In *2023 Spectral and Scattering Theory and Related Topics* 77–85. RIMS Kokyuroku.

Forrester, P.J. (2010). *Log-Gases and Random Matrices*. *London Mathematical Society Monographs Series* **34**. Princeton, NJ: Princeton Univ. Press. MR2641363 https://doi.org/10.1515/9781400835416

Fyodorov, Y.V., Khoruzhenko, B.A. and Simm, N.J. (2016). Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian unitary ensemble. *Ann. Probab.* **44** 2980–3031. MR3531684 https://doi.org/10.1214/15-AOP1039

Götze, F. and Tikhomirov, A. (2005). The rate of convergence for spectra of GUE and LUE matrix ensembles. *Cent. Eur. J. Math.* **3** 666–704. MR2171668 https://doi.org/10.2478/BF02475626

Johansson, K. (1998). On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.* **91** 151–204. MR1487983 https://doi.org/10.1215/S0012-7094-98-09108-6

Johnstone, I.M. and Ma, Z. (2012). Fast approach to the Tracy-Widom law at the edge of GOE and GUE. *Ann. Appl. Probab.* **22** 1962–1988. MR3025686 https://doi.org/10.1214/11-AAP819

Johnstone, I.M. and Onatski, A. (2020). Testing in high-dimensional spiked models. *Ann. Statist.* **48** 1231–1254. MR4124321 https://doi.org/10.1214/18-AOS1697

Johnstone, I.M., Klochkov, Y., Onatski, A. and Pavlyshyn, D. (2024). Spin glass to paramagnetic transition and triple point in spherical SK model. *J. Stat. Phys.* **191** 99. https://doi.org/10.1007/s10955-024-03296-7

Johnstone, I.M., Klochkov, Y., Onatski, A. and Pavlyshyn, D. (2025). Supplement to "An edge CLT for the log determinant of Wigner ensembles." https://doi.org/10.3150/23-BEJ1703SUPP

Knowles, A. and Yin, J. (2013). The isotropic semicircle law and deformation of Wigner matrices. *Comm. Pure Appl. Math.* **66** 1663–1750. MR3103909 https://doi.org/10.1002/cpa.21450

Kosterlitz, J.M., Thouless, D.J. and Jones, R.C. (1976). Spherical model of a spin-glass. *Phys. Rev. Lett.* **36** 1217–1220.

Krasovsky, I.V. (2007). Correlations of the characteristic polynomials in the Gaussian unitary ensemble or a singular Hankel determinant. *Duke Math. J.* **139** 581–619. MR2350854 https://doi.org/10.1215/S0012-7094-07-13936-X

Lambert, G. and Paquette, E. (2020). Strong approximation of Gaussian β-ensemble characteristic polynomials: the edge regime and the stochastic Airy function. arXiv preprint. Available at arXiv:2009.05003.

Lambert, G. and Paquette, E. (2023). Strong approximation of Gaussian β ensemble characteristic polynomials: The hyperbolic regime. *Ann. Appl. Probab.* **33** 549–612. MR4551558 https://doi.org/10.1214/22-aap1823

Landon, B. (2022). Free energy fluctuations of the two-spin spherical SK model at critical temperature. *J. Math. Phys.* **63** Paper No. 033301. MR4387140 https://doi.org/10.1063/5.0054298

Landon, B. and Sosoe, P. (2020). Applications of mesoscopic CLTs in random matrix theory. *Ann. Appl. Probab.* **30** 2769–2795. MR4187127 https://doi.org/10.1214/20-AAP1572

Li, Y., Schnelli, K. and Xu, Y. (2021). Central limit theorem for mesoscopic eigenvalue statistics of deformed Wigner matrices and sample covariance matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 506–546. MR4255183 https://doi.org/10.1214/20-aihp1086

Maïda, M. (2007). Large deviations for the largest eigenvalue of rank one deformations of Gaussian ensembles. *Electron. J. Probab.* **12** 1131–1150. MR2336602 https://doi.org/10.1214/EJP.v12-438

Onatski, A., Moreira, M.J. and Hallin, M. (2013). Asymptotic power of sphericity tests for high-dimensional data. *Ann. Statist.* **41** 1204–1231. MR3113808 https://doi.org/10.1214/13-AOS1100

Péché, S. (2006). The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probab. Theory Related Fields* **134** 127–173. MR2221787 https://doi.org/10.1007/s00440-005-0466-z

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

Tao, T. and Vu, V. (2012). A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** 74–101. MR2935384 https://doi.org/10.1016/j.aim.2012.05.006

Tracy, C.A. and Widom, H. (1998). Correlation functions, cluster functions, and spacing distributions for random matrices. *J. Stat. Phys.* **92** 809–835. MR1657844 https://doi.org/10.1023/A:1023084324803

Trotter, H.F. (1984). Eigenvalue distributions of large Hermitian matrices; Wigner's semicircle law and a theorem of Kac, Murdock, and Szegő. *Adv. Math.* **54** 67–82. MR0761763 https://doi.org/10.1016/0001-8708(84)90037-9

Wigner, E.P. (1965). Distribution laws for the roots of a random Hermitian matrix. *Stat. Theor. Spectra: Fluct.* 446–461.

Bernoulli **31**(1), 2025, 81–105
https://doi.org/10.3150/24-BEJ1719

On the expected number of critical points of locally isotropic Gaussian random fields

HAO XU[1,a], HAORAN YANG[2,c] and QIANG ZENG[1,b]

[1]*Department of Mathematics, University of Macau, Macau, China,* [a]*yc17446@connect.um.edu.mo,* [b]*qzeng.math@gmail.com*
[2]*School of Mathematical Sciences, Peking University, Beijing, China,* [c]*yanghr@pku.edu.cn*

We consider locally isotropic Gaussian random fields on the N-dimensional Euclidean space for fixed N. Using the so called Gaussian Orthogonally Invariant matrices first studied by Mallows in 1961 which include the celebrated Gaussian Orthogonal Ensemble (GOE), we establish the Kac–Rice representation of expected number of critical points of non-isotropic Gaussian fields, complementing the isotropic case obtained by Cheng and Schwartzman in 2018. In the limit $N = \infty$, we show that such a representation can be always given by GOE matrices, as conjectured by Auffinger and Zeng in 2020.

Keywords: Critical points; Gaussian random fields; GOE; GOI; isotropic increments; Kac–Rice formula

References

[1] Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. *Springer Monographs in Mathematics*. New York: Springer. MR2319516
[2] Auffinger, A. and Zeng, Q. (2023). Complexity of Gaussian random fields with isotropic increments. arXiv e-prints, arXiv:2206.13834.
[3] Auffinger, A. and Zeng, Q. (2023). Complexity of Gaussian random fields with isotropic increments. *Comm. Math. Phys.* **402** 951–993. MR4616691 https://doi.org/10.1007/s00220-023-04739-0
[4] Ben Arous, G., Bourgade, P. and McKenna, B. (2022). Exponential growth of random determinants beyond invariance. *Probab. Math. Phys.* **3** 731–789. MR4552227 https://doi.org/10.2140/pmp.2022.3.731
[5] Ben Arous, G., Bourgade, P. and McKenna, B. (2024). Landscape complexity beyond invariance and the elastic manifold. *Comm. Pure Appl. Math.* **77** 1302–1352. MR4673883 https://doi.org/10.1002/cpa.22146
[6] Cheng, D. and Schwartzman, A. (2018). Expected number and height distribution of critical points of smooth isotropic Gaussian random fields. *Bernoulli* **24** 3422–3446. MR3788177 https://doi.org/10.3150/17-BEJ964
[7] Engel, A. (1993). Replica symmetry breaking in zero dimension. *Nuclear Phys. B* **410** 617–646. https://doi.org/10.1016/0550-3213(93)90531-S
[8] Fyodorov, Y.V. (2004). Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Phys. Rev. Lett.* **92** 240601, 4. MR2115095 https://doi.org/10.1103/PhysRevLett.92.240601
[9] Fyodorov, Y.V. (2015). High-dimensional random fields and random matrix theory. *Markov Process. Related Fields* **21** 483–518. MR3469265
[10] Fyodorov, Y.V. and Bouchaud, J.-P. (2008). Statistical mechanics of a single particle in a multiscale random potential: Parisi landscapes in finite-dimensional Euclidean spaces. *J. Phys. A* **41** 324009, 25. MR2425780 https://doi.org/10.1088/1751-8113/41/32/324009
[11] Fyodorov, Y.V. and Le Doussal, P. (2018). Hessian spectrum at the global minimum of high-dimensional random landscapes. *J. Phys. A* **51** 474002, 27. MR3876574 https://doi.org/10.1088/1751-8121/aae74f
[12] Fyodorov, Y.V. and Le Doussal, P. (2020). Manifolds in a high-dimensional random landscape: Complexity of stationary points and depinning. *Phys. Rev. E* **101** 020101. https://doi.org/10.1103/PhysRevE.101.020101
[13] Fyodorov, Y.V. and Nadal, C. (2012). Critical behavior of the number of minima of a random landscape at the glass transition point and the Tracy-Widom distribution. *Phys. Rev. Lett.* **109** 167203. MR2115095 https://doi.org/10.1103/PhysRevLett.109.167203

1350-7265

[14] Fyodorov, Y.V. and Sommers, H.-J. (2007). Classical particle in a box with random potential: Exploiting rotational symmetry of replicated Hamiltonian. *Nuclear Phys. B* **764** 128–167. MR2293452 https://doi.org/10.1016/j.nuclphysb.2006.11.029

[15] Klimovsky, A. (2012). High-dimensional Gaussian fields with isotropic increments seen through spin glasses. *Electron. Commun. Probab.* **17** no. 17, 14. MR2915663 https://doi.org/10.1214/ECP.v17-1994

[16] Kolmogoroff, A. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynold's numbers. *C. R. (Dokl.) Acad. Sci. URSS* **30** 301–305. MR0004146

[17] Lazutkin, V.F. (1988). The signature of invertible symmetric matrices. *Mat. Zametki* **44** 202–207, 287. MR0969269 https://doi.org/10.1007/BF01159254

[18] Mallows, C.L. (1961). Latent vectors of random symmetric matrices. *Biometrika* **48** 133–149. MR0131312 https://doi.org/10.1093/biomet/48.1-2.133

[19] Mézard, M. and Parisi, G. (1991). Replica field theory for random manifolds. *J. Phys., I* **1** 809–836. https://doi.org/10.1051/jp1:1991171

[20] Schilling, R.L., Song, R. and Vondraček, Z. (2012). *Bernstein Functions: Theory and Applications*, 2nd ed. *De Gruyter Studies in Mathematics* **37**. Berlin: de Gruyter. MR2978140 https://doi.org/10.1515/9783110269338

[21] Schoenberg, I.J. (1938). Metric spaces and completely monotone functions. *Ann. of Math. (2)* **39** 811–841. MR1503439 https://doi.org/10.2307/1968466

[22] Xu, H. and Zeng, Q. (2022). Hessian spectrum at the global minimum and topology trivialization of locally isotropic Gaussian random fields. arXiv e-prints, arXiv:2210.15254. https://doi.org/10.48550/arXiv.2210.15254

[23] Yaglom, A.M. (1957). Certain types of random fields in n-dimensional space similar to stationary stochastic processes. *Teor. Veroyatn. Primen.* **2** 292–338. MR0094844

[24] Yaglom, A.M. (1987). *Correlation Theory of Stationary and Related Random Functions. Vol. I: Basic Results*. *Springer Series in Statistics*. New York: Springer. MR0893393

Bernoulli **31**(1), 2025, 106–129
https://doi.org/10.3150/24-BEJ1721

Distribution-free tests of multivariate independence based on center-outward quadrant, Spearman, Kendall, and van der Waerden statistics

HONGJIAN SHI[1,a], MATHIAS DRTON[1,b], MARC HALLIN[2,c] and FANG HAN[3,d]

[1]*Department of Mathematics, TUM School of Computation, Information and Technology, Technical University of Munich, 85748 Garching bei München, Germany,* [a]*hongjian.shi@tum.de,* [b]*mathias.drton@tum.de*
[2]*ECARES and Department of Mathematics, Université Libre de Bruxelles, Brussels, Belgium,* [c]*mhallin@ulb.ac.be*
[3]*Department of Statistics, University of Washington, Seattle, WA 98195, USA,* [d]*fanghan@uw.edu*

Due to the lack of a canonical ordering in $\mathbb{R}^d$ for $d > 1$, defining multivariate generalizations of the classical univariate ranks has been a long-standing open problem in statistics. Optimal transport has been shown to offer a solution in which multivariate ranks are obtained by transporting data points to a grid that approximates a uniform reference measure (*Ann. Statist.* **45** (2017) 223–256; Hallin (2017); *Ann. Statist.* **49** (2021) 1139–1165), thereby inducing ranks, signs, and a data-driven ordering of $\mathbb{R}^d$. We take up this new perspective to define and study multivariate analogues of the sign covariance/quadrant statistic, Spearman's rho, Kendall's tau, and van der Waerden covariances. The resulting tests of multivariate independence are fully distribution-free, hence uniformly valid irrespective of the actual (absolutely continuous) distribution of the observations. Our results provide the asymptotic distribution theory for these new test statistics, with asymptotic approximations to critical values to be used for testing independence between random vectors, as well as a power analysis of the resulting tests in an extension of the so-called (bivariate) Konijn model. This power analysis includes a multivariate Chernoff–Savage property guaranteeing that, under elliptical generalized Konijn models, the asymptotic relative efficiency of our van der Waerden tests with respect to Wilks' classical (pseudo-)Gaussian procedure is strictly larger than or equal to one, where equality is achieved under Gaussian distributions only. We similarly provide a lower bound for the asymptotic relative efficiency of our Spearman procedure with respect to Wilks' test, thus extending the classical result by Hodges and Lehmann on the asymptotic relative efficiency, in univariate location models, of Wilcoxon tests with respect to the Student ones.

Keywords: Center-outward ranks and signs; elliptical Chernoff–Savage property; multivariate independence test; Pitman asymptotic relative efficiency

References

Azadkia, M. and Chatterjee, S. (2021). A simple measure of conditional dependence. *Ann. Statist.* **49** 3070–3102. MR4352523 https://doi.org/10.1214/21-aos2073

Barbour, A.D. and Eagleson, G.K. (1986). Random association of symmetric arrays. *Stoch. Anal. Appl.* **4** 239–281. MR0857081 https://doi.org/10.1080/07362998608809090

Bhattacharya, R.N. and Ranga Rao, R. (1986). *Normal Approximation and Asymptotic Expansions*. Melbourne, FL: Robert E. Krieger Publishing Co., Inc. MR0855460

Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Stat.* **21** 593–600. MR0039190 https://doi.org/10.1214/aoms/1177729754

1350-7265

Carlier, G., Chernozhukov, V. and Galichon, A. (2016). Vector quantile regression: An optimal transport approach. *Ann. Statist.* **44** 1165–1192. MR3485957 https://doi.org/10.1214/15-AOS1401

Carlier, G., Chernozhukov, V. and Galichon, A. (2017). Vector quantile regression beyond the specified case. *J. Multivariate Anal.* **161** 96–102. MR3698117 https://doi.org/10.1016/j.jmva.2017.07.003

Chernoff, H. and Savage, I.R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Stat.* **29** 972–994. MR0100322 https://doi.org/10.1214/aoms/1177706436

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

Deb, N., Bhattacharya, B.B. and Sen, B. (2021). Efficiency lower bounds for distribution-free Hotelling-type two-sample tests based on optimal transport. Available at arXiv:2104.01986v2.

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

del Barrio, E., González-Sanz, A. and Hallin, M. (2020). A note on the regularity of optimal-transport-based center-outward distribution and quantile functions. *J. Multivariate Anal.* **180** 104671, 13. MR4147635 https://doi.org/10.1016/j.jmva.2020.104671

del Barrio, E., González-Sanz, A. and Hallin, M. (2024). Nonparametric multiple-output center-outward quantile regression. *J. Amer. Statist. Assoc.* To appear. https://doi.org/10.1080/01621459.2024.2366029

del Barrio, E., Cuesta-Albertos, J.A., Hallin, M. and Matrán, C. (2018). Smooth cyclically monotone interpolation and empirical center-outward distribution functions. Available at arXiv:1806.01238v1.

Fang, K.T., Kotz, S. and Ng, K.W. (1990). *Symmetric Multivariate and Related Distributions*. *Monographs on Statistics and Applied Probability* **36**. London: CRC Press. MR1071174 https://doi.org/10.1007/978-1-4899-2937-2

Figalli, A. (2018). On the continuity of center-outward distribution and quantile functions. *Nonlinear Anal.* **177** 413–421. MR3886582 https://doi.org/10.1016/j.na.2018.05.008

Friedman, J.H. and Rafsky, L.C. (1983). Graph-theoretic measures of multivariate association and prediction. *Ann. Statist.* **11** 377–391. MR0696054

Garel, B. and Hallin, M. (1995). Local asymptotic normality of multivariate ARMA processes with a linear trend. *Ann. Inst. Statist. Math.* **47** 551–579. MR1364260

Ghosal, P. and Sen, B. (2022). Multivariate ranks and quantiles using optimal transport: Consistency, rates and nonparametric testing. *Ann. Statist.* **50** 1012–1037. MR4404927 https://doi.org/10.1214/21-aos2136

Gieser, P.W. (1993). *A New Nonparametric Test for Independence Between Two Sets of Variates*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–University of Florida. MR2691505

Gieser, P.W. and Randles, R.H. (1997). A nonparametric test of independence between two vectors. *J. Amer. Statist. Assoc.* **92** 561–567. MR1467849 https://doi.org/10.2307/2965704

Gretton, A., Bousquet, O., Smola, A. and Schölkopf, B. (2005). Measuring statistical dependence with Hilbert-Schmidt norms. In *Algorithmic Learning Theory*. *Lecture Notes in Computer Science* **3734** 63–77. Berlin: Springer. MR2255909 https://doi.org/10.1007/11564089_7

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests*. New York-London: Academic Press. MR0229351

Hallin, M. (1994). On the Pitman non-admissibility of correlogram-based methods. *J. Time Series Anal.* **15** 607–611. MR1312324 https://doi.org/10.1111/j.1467-9892.1994.tb00215.x

Hallin, M. (2017). On distribution and quantile functions, ranks and signs in $\mathbb{R}^d$: A measure transportation approach.

Hallin, M. (2022). Measure transportation and statistical decision theory. *Annu. Rev. Stat. Appl.* **9** 401–424. MR4394914 https://doi.org/10.1146/annurev-statistics-040220-105948

Hallin, M., Hlubinka, D. and Hudecová, Š. (2023). Efficient fully distribution-free center-outward rank tests for multiple-output regression and MANOVA. *J. Amer. Statist. Assoc.* **118** 1923–1939. MR4646617 https://doi.org/10.1080/01621459.2021.2021921

Hallin, M., La Vecchia, D. and Liu, H. (2022). Center-outward R-estimation for semiparametric VARMA models. *J. Amer. Statist. Assoc.* **117** 925–938. MR4436323 https://doi.org/10.1080/01621459.2020.1832501

Hallin, M., La Vecchia, D. and Liu, H. (2023). Rank-based testing for semiparametric VAR models: A measure transportation approach. *Bernoulli* **29** 229–273. MR4497246 https://doi.org/10.3150/21-bej1456

Hallin, M. and Mordant, G. (2023). On the finite-sample performance of measure-transportation-based multivariate rank tests. In *Robust and Multivariate Statistical Methods—Festschrift in Honor of David E. Tyler* (M. Yi and K. Nordhausen, eds.) 87–119. Cham: Springer. MR4633219 https://doi.org/10.1007/978-3-031-22687-8_5

Hallin, M. and Paindaveine, D. (2002a). Optimal procedures based on interdirections and pseudo-Mahalanobis ranks for testing multivariate elliptic white noise against ARMA dependence. *Bernoulli* **8** 787–815. MR1963662

Hallin, M. and Paindaveine, D. (2002b). Optimal tests for multivariate location based on interdirections and pseudo-Mahalanobis ranks. *Ann. Statist.* **30** 1103–1133. MR1926170 https://doi.org/10.1214/aos/1031689019

Hallin, M. and Paindaveine, D. (2008). Chernoff-Savage and Hodges-Lehmann results for Wilks' test of multivariate independence. In *Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen*. *Inst. Math. Stat. (IMS) Collect.* **1** 184–196. Beachwood, OH: IMS. MR2462206 https://doi.org/10.1214/193940307000000130

Hallin, M. and Tribel, O. (2000). The efficiency of some nonparametric rank-based competitors to correlogram methods. In *Game Theory, Optimal Stopping, Probability and Statistics*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **35** 249–262. Beachwood, OH: IMS. MR1833863 https://doi.org/10.1214/lnms/1215089756

Hallin, M., del Barrio, E., Cuesta-Albertos, J. and Matrán, C. (2021). Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *Ann. Statist.* **49** 1139–1165. MR4255122 https://doi.org/10.1214/20-aos1996

Hannan, E.J. (1956). The asymptotic powers of certain tests based on multiple correlations. *J. Roy. Statist. Soc. Ser. B* **18** 227–233. MR0083228

Heller, R., Gorfine, M. and Heller, Y. (2012). A class of multivariate distribution-free tests of independence based on graphs. *J. Statist. Plann. Inference* **142** 3097–3106. MR2956796 https://doi.org/10.1016/j.jspi.2012.06.003

Heller, R., Heller, Y. and Gorfine, M. (2013). A consistent multivariate test of association based on ranks of distances. *Biometrika* **100** 503–510. MR3068450 https://doi.org/10.1093/biomet/ass070

Hodges, J.L. Jr. and Lehmann, E.L. (1956). The efficiency of some nonparametric competitors of the t-test. *Ann. Math. Stat.* **27** 324–335. MR0079383 https://doi.org/10.1214/aoms/1177728261

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

Hoeffding, W. (1951). A combinatorial central limit theorem. *Ann. Math. Stat.* **22** 558–566. MR0044058 https://doi.org/10.1214/aoms/1177729545

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* **30** 81–93.

Kim, I., Balakrishnan, S. and Wasserman, L. (2020). Robust multivariate nonparametric tests via projection averaging. *Ann. Statist.* **48** 3417–3441. MR4185814 https://doi.org/10.1214/19-AOS1936

Konijn, H.S. (1956). On the power of certain tests for independence in bivariate populations. *Ann. Math. Stat.* **27** 300–323. MR0079384 https://doi.org/10.1214/aoms/1177728260

Lehmann, E.L. and Romano, J.P. (2005). *Testing Statistical Hypotheses*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR2135927

Lin, J. (2017). *Copula Versions of Rkhs-Based and Distance-Based Criteria*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–The Pennsylvania State University. MR4172371

Lin, Z. and Han, F. (2023). On boosting the power of Chatterjee's rank correlation. *Biometrika* **110** 283–299. MR4589063 https://doi.org/10.1093/biomet/asac048

Lind, B. and Roussas, G. (1972). A remark on quadratic mean differentiability. *Ann. Math. Stat.* **43** 1030–1034. MR0307329 https://doi.org/10.1214/aoms/1177692570

Liu, R.Y. and Singh, K. (1993). A quality index based on data depth and multivariate rank tests. *J. Amer. Statist. Assoc.* **88** 252–260. MR1212489

McCann, R.J. (1995). Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80** 309–323. MR1369395 https://doi.org/10.1215/S0012-7094-95-08013-2

Moon, H. and Chen, K. (2022). Interpoint-ranking sign covariance for the test of independence. *Biometrika* **109** 165–179. MR4374647 https://doi.org/10.1093/biomet/asab011

Mordant, G. (2021). Transporting probability measures: Some contributions to statistical inference Ph.D. thesis UCL-Université Catholique de Louvain.

Oja, H. (2010). *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks*. *Lecture Notes in Statistics* **199**. New York: Springer. MR2598854 https://doi.org/10.1007/978-1-4419-0468-3

Oja, H., Paindaveine, D. and Taskinen, S. (2016). Affine-invariant rank tests for multivariate independence in independent component models. *Electron. J. Stat.* **10** 2372–2419. MR3544291 https://doi.org/10.1214/16-EJS1174

Paindaveine, D. (2004). A unified and elementary proof of serial and nonserial, univariate and multivariate, Chernoff-Savage results. *Stat. Methodol.* **1** 81–91. MR2160622 https://doi.org/10.1016/j.stamet.2004.08.001

Parzen, E. (2004). Quantile probability and statistical data modeling. *Statist. Sci.* **19** 652–662. MR2185587 https://doi.org/10.1214/088342304000000387

Pillai, K.C.S. (1955). Some new test criteria in multivariate analysis. *Ann. Math. Stat.* **26** 117–121. MR0067429 https://doi.org/10.1214/aoms/1177728599

Puri, M.L. and Sen, P.K. (1971). *Nonparametric Methods in Multivariate Analysis*. New York: Wiley. MR0298844

Randles, R.H. (1989). A distribution-free multivariate sign test based on interdirections. *J. Amer. Statist. Assoc.* **84** 1045–1050. MR1134492

Shi, H., Drton, M. and Han, F. (2022). Distribution-free consistent independence tests via center-outward ranks and signs. *J. Amer. Statist. Assoc.* **117** 395–410. MR4399094 https://doi.org/10.1080/01621459.2020.1782223

Shi, H., Hallin, M., Drton, M. and Han, F. (2022). On universally consistent and fully distribution-free rank tests of vector independence. *Ann. Statist.* **50** 1933–1959. MR4474478 https://doi.org/10.1214/21-aos2151

Shi, H., Drton, M., Hallin, M. and Han, F. (2025). Supplement to "Distribution-free tests of multivariate independence based on center-outward quadrant, Spearman, Kendall, and van der Waerden statistics." https://doi.org/10.3150/24-BEJ1721SUPP

Shorack, G.R. (2017). *Probability for Statisticians*, 2nd ed. *Springer Texts in Statistics*. Cham: Springer. MR3701383 https://doi.org/10.1007/978-3-319-52207-4

Spearman, C. (1904). The proof and measurement of association between two things. *Amer. J. Psychol.* **15** 72–101.

Székely, G.J. and Rizzo, M.L. (2014). Partial distance correlation with methods for dissimilarities. *Ann. Statist.* **42** 2382–2412. MR3269983 https://doi.org/10.1214/14-AOS1255

Székely, G.J., Rizzo, M.L. and Bakirov, N.K. (2007). Measuring and testing dependence by correlation of distances. *Ann. Statist.* **35** 2769–2794. MR2382665 https://doi.org/10.1214/009053607000000505

Taskinen, S., Kankainen, A. and Oja, H. (2003). Sign test of independence between two random vectors. *Statist. Probab. Lett.* **62** 9–21. MR1965367 https://doi.org/10.1016/S0167-7152(02)00399-1

Taskinen, S., Kankainen, A. and Oja, H. (2004). Rank scores tests of multivariate independence. In *Theory and Applications of Recent Robust Methods*. *Stat. Ind. Technol.* 329–341. Basel: Birkhäuser. MR2088309

Taskinen, S., Oja, H. and Randles, R.H. (2005). Multivariate nonparametric tests of independence. *J. Amer. Statist. Assoc.* **100** 916–925. MR2201019 https://doi.org/10.1198/016214505000000097

Weihs, L., Drton, M. and Meinshausen, N. (2018). Symmetric rank covariances: A generalized framework for nonparametric measures of dependence. *Biometrika* **105** 547–562. MR3842884 https://doi.org/10.1093/biomet/asy021

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biom. Bull.* **1** 80–83.

Wilks, S.S. (1935). On the independence of k sets of normally distributed statistical variables. *Econometrica* **3** 309–326.

Zhu, L., Xu, K., Li, R. and Zhong, W. (2017). Projection correlation between two random vectors. *Biometrika* **104** 829–843. MR3737307 https://doi.org/10.1093/biomet/asx043

Zuo, Y. and He, X. (2006). On the limiting distributions of multivariate depth-based rank sum statistics and related tests. *Ann. Statist.* **34** 2879–2896. MR2329471 https://doi.org/10.1214/009053606000000876

Bernoulli **31**(1), 2025, 130–161
https://doi.org/10.3150/24-BEJ1722

Self-normalized Cramér type moderate deviations for martingales and applications

XIEQUAN FAN[1,a] and QI-MAN SHAO[2,b]

[1]*School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao, China,* [a]*fanxiequan@126.com*
[2]*Department of Statistics and Data Science, SICM, National Center for Applied Mathematics Shenzhen, Southern University of Science and Technology, Shenzhen 518000, China,* [b]*shaoqm@sustech.edu.cn*

Cramér's moderate deviations give a quantitative estimate for the relative error of the normal approximation and provide theoretical justifications for many estimators used in statistics. In this paper, we establish self-normalized Cramér type moderate deviations for martingales under some mild conditions. The result extends an earlier work of Fan et al. (*Bernoulli* **25** (2019) 2793–2823). Moreover, applications of our result to Student's statistic, stationary martingale difference sequences and branching processes in a random environment are also discussed.

Keywords: Branching process in a random environment; Cramér's moderate deviations; martingales; self-normalized sequences; Student's statistic

References

Afanasyev, V.I., Böinghoff, Ch., Kersting, G. and Vatutin, V.A. (2014). Conditional limit theorems for intermediately subcritical branching processes in random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** 602–627. MR3189086 https://doi.org/10.1214/12-AIHP526

Athreya, K.B. and Karlin, S. (1971). Branching processes with random environments. II. Limit theorems. *Ann. Math. Stat.* **42** 1843–1858. MR0298781 https://doi.org/10.1214/aoms/1177693051

Bansaye, V. and Böinghoff, C. (2011). Upper large deviations for branching processes in random environment with heavy tails. *Electron. J. Probab.* **16** 1900–1933. MR2851050 https://doi.org/10.1214/EJP.v16-933

Bansaye, V. and Vatutin, V. (2017). On the survival probability for a class of subcritical branching processes in random environment. *Bernoulli* **23** 58–88. MR3556766 https://doi.org/10.3150/15-BEJ723

Bentkus, V. and Götze, F. (1996). The Berry-Esseen bound for Student's statistic. *Ann. Probab.* **24** 491–503. MR1387647 https://doi.org/10.1214/aop/1042644728

Bercu, B., Delyon, B. and Rio, E. (2015). *Concentration Inequalities for Sums and Martingales*. *SpringerBriefs in Mathematics*. Cham: Springer. MR3363542 https://doi.org/10.1007/978-3-319-22099-4

Böinghoff, C. (2014). Limit theorems for strongly and intermediately supercritical branching processes in random environment with linear fractional offspring distributions. *Stochastic Process. Appl.* **124** 3553–3577. MR3249347 https://doi.org/10.1016/j.spa.2014.05.009

Böinghoff, C. and Kersting, G. (2010). Upper large deviations of branching processes in a random environment—offspring distributions with geometrically bounded tails. *Stochastic Process. Appl.* **120** 2064–2077. MR2673988 https://doi.org/10.1016/j.spa.2010.05.017

Chang, J., Shao, Q.-M. and Zhou, W.-X. (2016). Cramér-type moderate deviations for Studentized two-sample *U*-statistics with applications. *Ann. Statist.* **44** 1931–1956. MR3546439 https://doi.org/10.1214/15-AOS1375

Chen, X., Shao, Q.-M., Wu, W.B. and Xu, L. (2016). Self-normalized Cramér-type moderate deviations under dependence. *Ann. Statist.* **44** 1593–1617. MR3519934 https://doi.org/10.1214/15-AOS1429

Chung, K.-L. (1946). The approximate distribution of Student's statistic. *Ann. Math. Stat.* **17** 447–465. MR0018390 https://doi.org/10.1214/aoms/1177730884

Cramér, H. (1938). Sur un nouveau théorème-limite de la théorie des probabilités. *Actualite's Sci. Indust.* **736** 5–23.

1350-7265

de la Peña, V.H., Lai, T.L. and Shao, Q.-M. (2009). *Self-Normalized Processes*. *Probability and Its Applications (New York)*. Berlin: Springer. Limit theory and statistical applications. MR2488094 https://doi.org/10.1007/978-3-540-85636-8

Fan, X., Grama, I. and Liu, Q. (2013). Cramér large deviation expansions for martingales under Bernstein's condition. *Stochastic Process. Appl.* **123** 3919–3942. MR3091094 https://doi.org/10.1016/j.spa.2013.06.010

Fan, X. and Shao, Q.M. (2023). Cramér's moderate deviations for martingales with applications. *Ann. Inst. Henri Poincaré Probab. Stat.* to appear.

Fan, X. and Shao, Q.-M. (2025). Supplement to "Self-normalized Cramér type moderate deviations for martingales and applications." https://doi.org/10.3150/24-BEJ1722SUPP

Fan, X., Grama, I., Liu, Q. and Shao, Q.-M. (2019). Self-normalized Cramér type moderate deviations for martingales. *Bernoulli* **25** 2793–2823. MR4003565 https://doi.org/10.3150/18-BEJ1071

Fan, X., Grama, I., Liu, Q. and Shao, Q.-M. (2020). Self-normalized Cramér type moderate deviations for stationary sequences and applications. *Stochastic Process. Appl.* **130** 5124–5148. MR4108484 https://doi.org/10.1016/j.spa.2020.03.001

Gao, L., Shao, Q.-M. and Shi, J. (2022). Refined Cramér-type moderate deviation theorems for general self-normalized sums with applications to dependent random variables and winsorized mean. *Ann. Statist.* **50** 673–697. MR4404916 https://doi.org/10.1214/21-aos2122

Giné, E., Götze, F. and Mason, D.M. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25** 1514–1531. MR1457629 https://doi.org/10.1214/aop/1024404523

Grama, I. and Haeusler, E. (2000). Large deviations for martingales via Cramér's method. *Stochastic Process. Appl.* **85** 279–293. MR1731027 https://doi.org/10.1016/S0304-4149(99)00079-4

Grama, I., Liu, Q. and Miqueu, E. (2017). Berry-Esseen's bound and Cramér's large deviation expansion for a supercritical branching process in a random environment. *Stochastic Process. Appl.* **127** 1255–1281. MR3619270 https://doi.org/10.1016/j.spa.2016.07.014

Grama, I., Liu, Q. and Miqueu, E. (2023). Asymptotics of the distribution and harmonic moments for a supercritical branching process in a random environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **59** 1934–1950. MR4663512 https://doi.org/10.1214/22-aihp1318

Huang, C. and Liu, Q. (2012). Moments, moderate and large deviations for a branching process in a random environment. *Stochastic Process. Appl.* **122** 522–545. MR2868929 https://doi.org/10.1016/j.spa.2011.09.001

Jing, B.-Y., Shao, Q.-M. and Wang, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann. Probab.* **31** 2167–2215. MR2016616 https://doi.org/10.1214/aop/1068646382

Liu, W., Shao, Q.-M. and Wang, Q. (2013). Self-normalized Cramér type moderate deviations for the maximum of sums. *Bernoulli* **19** 1006–1027. MR3079304 https://doi.org/10.3150/12-BEJ415

Liu, Q. and Watbled, F. (2009). Exponential inequalities for martingales and asymptotic properties of the free energy of directed polymers in a random environment. *Stochastic Process. Appl.* **119** 3101–3132. MR2568267 https://doi.org/10.1016/j.spa.2009.05.001

Lotka, A. (1939). Theorie analytique des assiciation biologiques. *Actualités Sci. Ind.* **780** 123–136.

Fan, X., Hu, H. and Xu, L.H. (2023). Normalized and self-normalized Cramér-type moderate deviations for Euler-Maruyama scheme for SDE. *Sci. China Math.* To appear.

Nagaev, A.V. (1967). On estimating the expected number of direct descendants of a particle in a branching process. *Theory Probab. Appl.* **12** 314–320.

Nakashima, M. (2013). Lower deviations of branching processes in random environment with geometrical offspring distributions. *Stochastic Process. Appl.* **123** 3560–3587. MR3071390 https://doi.org/10.1016/j.spa.2013.04.013

Novak, S.Y. (2012). *Extreme Value Methods with Applications to Finance*. *Monographs on Statistics and Applied Probability* **122**. Boca Raton, FL: CRC Press. MR2933280

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

Petrov, V.V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete [Results in Mathematics and Related Areas]*, **82**. New York: Springer. Translated from the Russian by A. A. Brown. MR0388499

Račkauskas, A. (1990). On probabilities of large deviations for martingales. *Liet. Mat. Rink.* **30** 784–795. MR1091658 https://doi.org/10.1007/BF00970833

Račkauskas, A. (1995). Large deviations for martingales with some applications. *Acta Appl. Math.* **38** 109–129. MR1326865 https://doi.org/10.1007/BF00992617

Saulis, L. and Statulevičius, V.A. (1978). *Limit Theorems for Large Deviations*. Dordrecht: Kluwer Academic.

Shao, Q.-M. (1997). Self-normalized large deviations. *Ann. Probab.* **25** 285–328. MR1428510 https://doi.org/10.1214/aop/1024404289

Shao, Q.-M. (1999). A Cramér type large deviation result for Student's t-statistic. *J. Theoret. Probab.* **12** 385–398. MR1684750 https://doi.org/10.1023/A:1021626127372

Shao, Q.-M. and Wang, Q. (2013). Self-normalized limit theorems: A survey. *Probab. Surv.* **10** 69–93. MR3161676 https://doi.org/10.1214/13-PS216

Shao, Q.-M. and Zhou, W.-X. (2016). Cramér type moderate deviation theorems for self-normalized processes. *Bernoulli* **22** 2029–2079. MR3498022 https://doi.org/10.3150/15-BEJ719

Slavova, V.V. (1985). On the Berry-Esseen bound for Student's statistic. In *Stability Problems for Stochastic Models (Uzhgorod, 1984)*. *Lecture Notes in Math.* **1155** 355–390. Berlin: Springer. MR0825335 https://doi.org/10.1007/BFb0074828

Smith, W.L. and Wilkinson, W.E. (1969). On branching processes in random environments. *Ann. Math. Stat.* **40** 814–827. MR0246380 https://doi.org/10.1214/aoms/1177697589

Tanny, D. (1988). A necessary and sufficient condition for a branching process in a random environment to grow like the product of its means. *Stochastic Process. Appl.* **28** 123–139. MR0936379 https://doi.org/10.1016/0304-4149(88)90070-1

Vatutin, V. and Zheng, X. (2012). Subcritical branching processes in a random environment without the Cramer condition. *Stochastic Process. Appl.* **122** 2594–2609. MR2926168 https://doi.org/10.1016/j.spa.2012.04.008

Wu, W.B. and Zhao, Z. (2008). Moderate deviations for stationary processes. *Statist. Sinica* **18** 769–782. MR2411619

Bernoulli **31**(1), 2025, 162–186
https://doi.org/10.3150/24-BEJ1723

Non-ergodic statistics and spectral density estimation for stationary real harmonizable symmetric α-stable processes

LY VIET HOANG[a] and EVGENY SPODAREV[b]

Institute of Stochastics, Ulm University, Ulm, Germany, [a]*ly.hoang@uni-ulm.de,* [b]*evgeny.spodarev@uni-ulm.de*

We consider a non-ergodic class of stationary real harmonizable symmetric α-stable processes $X = \{X(t) : t \in \mathbb{R}\}$ with a finite symmetric and absolutely continuous control measure. We refer to its density function as the spectral density of X. These processes admit a LePage series representation and are conditionally Gaussian, which allows us to derive the non-ergodic limit of sample functions on X. In particular, we give an explicit expression for the non-ergodic limits of the empirical characteristic function of X and the lag process $\{X(t + h) - X(t) : t \in \mathbb{R}\}$ with $h > 0$, respectively. The process admits an equivalent representation as a series of sinusoidal waves with random frequencies which are i.i.d. with the (normalized) spectral density of X as their probability density function. Based on strongly consistent frequency estimation using the periodogram we present a strongly consistent estimator of the spectral density. The computation of the periodogram is fast and efficient, and our method is not affected by the non-ergodicity of X.

Keywords: Fourier analysis; frequency estimation; harmonizable process; non-ergodic process; non-ergodic statistics; periodogram; spectral density estimation; stable process; stationary process

References

[1] Ayache, A. and Xiao, Y. (2016). Harmonizable fractional stable fields: Local nondeterminism and joint continuity of the local times. *Stochastic Process. Appl.* **126** 171–185. MR3426515 https://doi.org/10.1016/j.spa.2015.08.001
[2] Basse-O'Connor, A., Grønbæk, T. and Podolskij, M. (2021). Local asymptotic self-similarity for heavy tailed harmonizable fractional Lévy motions. *ESAIM Probab. Stat.* **25** 286–297. MR4281739 https://doi.org/10.1051/ps/2021011
[3] Biermé, H., Meerschaert, M.M. and Scheffler, H.-P. (2007). Operator scaling stable random fields. *Stochastic Process. Appl.* **117** 312–332. MR2290879 https://doi.org/10.1016/j.spa.2006.07.004
[4] Bowman, A.W. (1984). An alternative method of cross-validation for the smoothing of density estimates. *Biometrika* **71** 353–360. MR0767163 https://doi.org/10.1093/biomet/71.2.353
[5] 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
[6] Cambanis, S., Hardin, C.D. Jr. and Weron, A. (1987). Ergodic properties of stationary stable processes. *Stochastic Process. Appl.* **24** 1–18. MR0883599 https://doi.org/10.1016/0304-4149(87)90024-X
[7] Cambanis, S., Maejima, M. and Samorodnitsky, G. (1992). Characterization of linear and harmonizable fractional stable motions. *Stochastic Process. Appl.* **42** 91–110. MR1172509 https://doi.org/10.1016/0304-4149(92)90028-O
[8] Doob, J.L. (1990). *Stochastic Processes*. *Wiley Classics Library*. New York: Wiley. Reprint of the 1953 original, A Wiley-Interscience Publication. MR1038526
[9] Dym, H. and McKean, H.P. (1976). *Gaussian Processes, Function Theory, and the Inverse Spectral Problem*. *Probability and Mathematical Statistics, Vol. 31*. New York-London: Academic Press. MR0448523
[10] Einmahl, U. and Mason, D.M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. *Ann. Statist.* **33** 1380–1403. MR2195639 https://doi.org/10.1214/009053605000000129

1350-7265

[11] Fuller, W.A. (1996). *Introduction to Statistical Time Series*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1365746

[12] Giné, E. and Guillou, A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 907–921. MR1955344 https://doi.org/10.1016/S0246-0203(02)01128-7

[13] Hannan, E.J. (1973). The estimation of frequency. *J. Appl. Probab.* **10** 510–519. MR0370977 https://doi.org/10.2307/3212772

[14] Hoang, L.V. (2023). Matlab and R implementation for the spectral density estimation for stationary real harmonizable symmetric α-stable processes. https://github.com/lyviho/harmonizablestable.

[15] Hoang, L.V. and Spodarev, E. (2025). Supplement to "Non-ergodic statistics and spectral density estimation for stationary real harmonizable symmetric α-stable processes." https://doi.org/10.3150/24-BEJ1723SUPP

[16] Hoang, L.V. and Spodarev, E. (2021). Inversion of α-sine and α-cosine transforms on $\mathbb{R}$. *Inverse Probl.* **37** Paper No. 085008, 40. MR4295590 https://doi.org/10.1088/1361-6420/ac1327

[17] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications (New York)*. New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[18] Koutrouvelis, I.A. (1980). Regression-type estimation of the parameters of stable laws. *J. Amer. Statist. Assoc.* **75** 918–928. MR0600977

[19] Lamperti, J. (1962). Semi-stable stochastic processes. *Trans. Amer. Math. Soc.* **104** 62–78. MR0138128 https://doi.org/10.2307/1993933

[20] Lindgren, G. (2013). *Stationary Stochastic Processes: Theory and Applications*. *Chapman & Hall/CRC Texts in Statistical Science Series*. Boca Raton, FL: CRC Press. MR3024784

[21] McCulloch, J.H. (1986). Simple consistent estimators of stable distribution parameters. *Comm. Statist. Simulation Comput.* **15** 1109–1136. MR0876783 https://doi.org/10.1080/03610918608812563

[22] Oppenheim, A.V., Buck, J.R. and Schafer, R.W. (1999). *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Prentice Hall.

[23] Quinn, B.G. and Hannan, E.J. (2001). *The Estimation and Tracking of Frequency*. *Cambridge Series in Statistical and Probabilistic Mathematics* **9**. Cambridge: Cambridge Univ. Press. MR1813156 https://doi.org/10.1017/CBO9780511609602

[24] Rosiński, J. (1995). On the structure of stationary stable processes. *Ann. Probab.* **23** 1163–1187. MR1349166

[25] Rudemo, M. (1982). Empirical choice of histograms and kernel density estimators. *Scand. J. Stat.* **9** 65–78. MR0668683

[26] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. *Stochastic Modeling*. New York: CRC Press. MR1280932

[27] Sheather, S.J. (2004). Density estimation. *Statist. Sci.* **19** 588–597. MR2185580 https://doi.org/10.1214/088342304000000297

[28] Sheather, S.J. and Jones, M.C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *J. Roy. Statist. Soc. Ser. B* **53** 683–690. MR1125725

[29] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR0848134 https://doi.org/10.1007/978-1-4899-3324-9

[30] Ślezak, J. (2017). Asymptotic behaviour of time averages for non-ergodic Gaussian processes. *Ann. Physics* **383** 285–311. MR3682028 https://doi.org/10.1016/j.aop.2017.05.015

[31] Walker, A.M. (1971). On the estimation of a harmonic component in a time series with stationary independent residuals. *Biometrika* **58** 21–36. MR0275619 https://doi.org/10.1093/biomet/58.1.21

[32] Wied, D. and Weißbach, R. (2012). Consistency of the kernel density estimator: A survey. *Statist. Papers* **53** 1–21. MR2878587 https://doi.org/10.1007/s00362-010-0338-1

Bernoulli **31**(1), 2025, 187–212
https://doi.org/10.3150/24-BEJ1724

Simultaneous off-the-grid learning of mixtures issued from a continuous dictionary

CRISTINA BUTUCEA[1,a], JEAN-FRANÇOIS DELMAS[2,b], ANNE DUTFOY[3,d] and CLÉMENT HARDY[2,3,c]

[1]*CREST, ENSAE, IP Paris, Palaiseau, France,* [a]*cristina.butucea@ensae.fr*
[2]*CERMICS, École des Ponts, Champs-sur-Marne, France,* [b]*jean-francois.delmas@enpc.fr,*
[c]*clement.hardy@enpc.fr*
[3]*EDF R&D, Palaiseau, France,* [d]*anne.dutfoy@edf.fr*

In this paper we observe a set, possibly a continuum, of signals corrupted by noise. Each signal is a finite mixture of an unknown number of features belonging to a continuous dictionary. The continuous dictionary is parametrized by a real non-linear parameter. We shall assume that the signals share an underlying structure by assuming that each signal has its active features included in a finite and sparse set. We formulate regularized optimization problem to estimate simultaneously the linear coefficients in the mixtures and the non-linear parameters of the features. The optimization problem is composed of a data fidelity term and a (ℓ_1, L^p)-penalty. We call its solution the Group-Nonlinear-Lasso and provide high probability bounds on the prediction error using certificate functions. Following recent works on the geometry of off-the-grid methods, we show that such functions can be constructed provided the parameters of the active features are pairwise separated by a constant with respect to a Riemannian metric. When the number of signals is finite and the noise is assumed Gaussian, we give refinements of our results for $p = 1$ and $p = 2$ using tail bounds on suprema of Gaussian and χ^2 random processes. When $p = 2$, our prediction error reaches the rates obtained by the Group-Lasso estimator in the multi-task linear regression model. Furthermore, for $p = 2$ these prediction rates are faster than for $p = 1$ when all signals share most of the non-linear parameters.

Keywords: Continuous dictionary; group-nonlinear-lasso; interpolating certificates; mixture model; multi-task learning; non-linear regression model; off-the-grid methods; simultaneous recovery; sparse spike deconvolution

References

Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: Wiley. MR2478201 https://doi.org/10.1002/9780470434642

Bach, F.R. (2008). Consistency of the group lasso and multiple kernel learning. *J. Mach. Learn. Res.* **9** 1179–1225. MR2417268

Barber, R.F., Reimherr, M. and Schill, T. (2017). The function-on-scalar LASSO with applications to longitudinal GWAS. *Electron. J. Stat.* **11** 1351–1389. MR3635916 https://doi.org/10.1214/17-EJS1260

Beck, A. and Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2** 183–202. MR2486527 https://doi.org/10.1137/080716542

Bhaskar, B.N., Tang, G. and Recht, B. (2013). Atomic norm denoising with applications to line spectral estimation. *IEEE Trans. Signal Process.* **61** 5987–5999. MR3138795 https://doi.org/10.1109/TSP.2013.2273443

Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. MR2533469 https://doi.org/10.1214/08-AOS620

Boyd, N., Schiebinger, G. and Recht, B. (2017). The alternating descent conditional gradient method for sparse inverse problems. *SIAM J. Optim.* **27** 616–639. MR3634995 https://doi.org/10.1137/15M1035793

Boyer, C., De Castro, Y. and Salmon, J. (2017). Adapting to unknown noise level in sparse deconvolution. *Inf. Inference* **6** 310–348. MR3764527 https://doi.org/10.1093/imaiai/iaw024

Boyer, C., Chambolle, A., De Castro, Y., Duval, V., de Gournay, F. and Weiss, P. (2019). On representer theorems and convex regularization. *SIAM J. Optim.* **29** 1260–1281. MR3948246 https://doi.org/10.1137/18M1200750

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

Bunea, F., Tsybakov, A. and Wegkamp, M. (2007). Sparsity oracle inequalities for the Lasso. *Electron. J. Stat.* **1** 169–194. MR2312149 https://doi.org/10.1214/07-EJS008

Butucea, C., Delmas, J.-F., Dutfoy, A. and Hardy, C. (2021). Modeling infra-red spectra: An algorithm for an automatic and simultaneous analysis. In *Proceedings of the 31st European Safety and Reliability Conference* 3359–3366.

Butucea, C., Delmas, J.-F., Dutfoy, A. and Hardy, C. (2022). Off-the-grid learning of sparse mixtures from a continuous dictionary. Preprint. Available at arXiv:2207.00171.

Butucea, C., Delmas, J.-F., Dutfoy, A. and Hardy, C. (2025). Supplement to "Simultaneous off-the-grid learning of mixtures issued from a continuous dictionary." https://doi.org/10.3150/24-BEJ1724SUPP

Candès, E.J. and Fernandez-Granda, C. (2013). Super-resolution from noisy data. *J. Fourier Anal. Appl.* **19** 1229–1254. MR3132912 https://doi.org/10.1007/s00041-013-9292-3

Candès, E.J. and Fernandez-Granda, C. (2014). Towards a mathematical theory of super-resolution. *Comm. Pure Appl. Math.* **67** 906–956. MR3193963 https://doi.org/10.1002/cpa.21455

Candès, E.J. and Plan, Y. (2011). A probabilistic and RIPless theory of compressed sensing. *IEEE Trans. Inf. Theory* **57** 7235–7254. MR2883653 https://doi.org/10.1109/TIT.2011.2161794

Candes, E. and Tao, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n. *Ann. Statist.* **35** 2313–2351. MR2382644 https://doi.org/10.1214/009053606000001523

Chesneau, C. and Hebiri, M. (2008). Some theoretical results on the grouped variables Lasso. *Math. Methods Statist.* **17** 317–326. MR2483460 https://doi.org/10.3103/S1066530708040030

Chizat, L. (2022). Sparse optimization on measures with over-parameterized gradient descent. *Math. Program.* **194** 487–532. MR4445462 https://doi.org/10.1007/s10107-021-01636-z

de Castro, Y. and Gamboa, F. (2012). Exact reconstruction using Beurling minimal extrapolation. *J. Math. Anal. Appl.* **395** 336–354. MR2943626 https://doi.org/10.1016/j.jmaa.2012.05.011

Denoyelle, Q., Duval, V., Peyré, G. and Soubies, E. (2020). The sliding Frank–Wolfe algorithm and its application to super-resolution microscopy. *Inverse Probl.* **36** 014001, 42 pp. MR4040984 https://doi.org/10.1088/1361-6420/ab2a29

Diestel, J. and Uhl, J.J. Jr. (1977). *Vector Measures*. *Mathematical Surveys* **15**. Providence, RI: Amer. Math. Soc. With a foreword by B.J. Pettis. MR0453964

Duval, V. (2021). An epigraphical approach to the representer theorem. *J. Convex Anal.* **28** 819–836. MR4374321

Duval, V. and Peyré, G. (2015). Exact support recovery for sparse spikes deconvolution. *Found. Comput. Math.* **15** 1315–1355. MR3394712 https://doi.org/10.1007/s10208-014-9228-6

Duval, V. and Peyré, G. (2017). Sparse regularization on thin grids I: The Lasso. *Inverse Probl.* **33** 055008, 29 pp. MR3628904 https://doi.org/10.1088/1361-6420/aa5e12

Golbabaee, M. and Poon, C. (2022). An off-the-grid approach to multi-compartment magnetic resonance fingerprinting. *Inverse Probl.* **38** Paper No. 085002, 31 pp. MR4443913 https://doi.org/10.1088/1361-6420/ac70da

Huang, J. and Zhang, T. (2010). The benefit of group sparsity. *Ann. Statist.* **38** 1978–2004. MR2676881 https://doi.org/10.1214/09-AOS778

Liu, H. and Zhang, J. (2008). On the ℓ_1-ℓ_q regularized regression. Preprint. Available at arXiv:0802.1517.

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

Nardi, Y. and Rinaldo, A. (2008). On the asymptotic properties of the group lasso estimator for linear models. *Electron. J. Stat.* **2** 605–633. MR2426104 https://doi.org/10.1214/08-EJS200

Poon, C., Keriven, N. and Peyré, G. (2023). The geometry of off-the-grid compressed sensing. *Found. Comput. Math.* **23** 241–327. MR4546148 https://doi.org/10.1007/s10208-021-09545-5

Raskutti, G., Wainwright, M.J. and Yu, B. (2011). Minimax rates of estimation for high-dimensional linear regression over ℓ_q-balls. *IEEE Trans. Inf. Theory* **57** 6976–6994. MR2882274 https://doi.org/10.1109/TIT.2011.2165799

Tang, G., Bhaskar, B.N. and Recht, B. (2013). Sparse recovery over continuous dictionaries-just discretize. In *2013 Asilomar Conference on Signals, Systems and Computers* 1043–1047. IEEE.

Tang, G., Bhaskar, B.N. and Recht, B. (2015). Near minimax line spectral estimation. *IEEE Trans. Inf. Theory* **61** 499–512. MR3299978 https://doi.org/10.1109/TIT.2014.2368122

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. MR1379242

van de Geer, S.A. and Bühlmann, P. (2009). On the conditions used to prove oracle results for the Lasso. *Electron. J. Stat.* **3** 1360–1392. MR2576316 https://doi.org/10.1214/09-EJS506

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 49–67. MR2212574 https://doi.org/10.1111/j.1467-9868.2005.00532.x

Bernoulli **31**(1), 2025, 213–239
https://doi.org/10.3150/24-BEJ1725

Stability and sample complexity of divergence regularized optimal transport

ERHAN BAYRAKTAR[1,a], STEPHAN ECKSTEIN[2,b] and XIN ZHANG[3,c]

[1]*Department of Mathematics, University of Michigan, Michigan, USA,* [a]*erhan@umich.edu*
[2]*Department of Mathematics, University of Tübingen, Tübingen, Germany,* [b]*stephan.eckstein@uni-tuebingen.de*
[3]*Department of Mathematics, University of Vienna, Vienna, Austria,* [c]*xin.zhang@univie.ac.at*

We study stability and sample complexity properties of divergence regularized optimal transport (DOT). First, we obtain quantitative stability results for optimizers of DOT measured in Wasserstein distance, which are applicable to a wide class of divergences and simultaneously improve known results for entropic optimal transport. Second, we study the case of sample complexity, where the DOT problem is approximated using empirical measures of the marginals. We show that divergence regularization can improve the corresponding convergence rate compared to unregularized optimal transport. To this end, we prove upper bounds which exploit both the regularity of cost function and divergence functional, as well as the intrinsic dimension of the marginals. Along the way, we establish regularity properties of dual optimizers of DOT, as well as general limit theorems for empirical measures with suitable classes of test functions.

Keywords: Entropic optimal transport; f-divergence; intrinsic dimension; regularization

References

[1] Agrawal, R. and Horel, T. (2021). Optimal bounds between f-divergences and integral probability metrics. *J. Mach. Learn. Res.* **22** Paper No. 128, 59 pp. MR4279779

[2] Blanchet, J., Jambulapati, A., Kent, C. and Sidford, A. (2024). Towards optimal running timesfor optimal transport. *Oper. Res. Lett.* **52** 107054. MR4678768 https://doi.org/10.1016/j.orl.2023.11.007

[3] Blondel, M., Seguy, V. and Rolet, A. (2018). Smooth and sparse optimal transport. In *International Conference on Artificial Intelligence and Statistics* 880–889. PMLR.

[4] Bolley, F. and Villani, C. (2005). Weighted Csiszár-Kullback-Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse Math. (6)* **14** 331–352. MR2172583

[5] Carlier, G., Duval, V., Peyré, G. and Schmitzer, B. (2017). Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.* **49** 1385–1418. MR3635459 https://doi.org/10.1137/15M1050264

[6] Carlier, G. and Laborde, M. (2020). A differential approach to the multi-marginal Schrödinger system. *SIAM J. Math. Anal.* **52** 709–717. MR4062805 https://doi.org/10.1137/19M1253800

[7] Carlier, G., Pegon, P. and Tamanini, L. (2023). Convergence rate of general entropic optimal transport costs. *Calc. Var. Partial Differential Equations* **62** Paper No. 116, 28 pp. MR4565039 https://doi.org/10.1007/s00526-023-02455-0

[8] Chizat, L., Roussillon, P., Léger, F., Vialard, F.-X. and Peyré, G. (2020). Faster Wasserstein distance estimation with the Sinkhorn divergence. *Adv. Neural Inf. Process. Syst.* **33** 2257–2269.

[9] Cominetti, R. and San Martín, J. (1994). Asymptotic analysis of the exponential penalty trajectory in linear programming. *Math. Program.* **67** 169–187. MR1305565 https://doi.org/10.1007/BF01582220

[10] Conforti, G. and Tamanini, L. (2021). A formula for the time derivative of the entropic cost and applications. *J. Funct. Anal.* **280** Paper No. 108964, 48 pp. MR4232667 https://doi.org/10.1016/j.jfa.2021.108964

[11] 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

1350-7265

[12] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. *Adv. Neural Inf. Process. Syst.* **26**.
[13] del Barrio, E., González Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2023). An improved central limit theorem and fast convergence rates for entropic transportation costs. *SIAM J. Math. Data Sci.* **5** 639–669. MR4616887 https://doi.org/10.1137/22M149260X
[14] Deligiannidis, G., De Bortoli, V. and Doucet, A. (2021). Quantitative uniform stability of the iterative proportional fitting procedure. Preprint. Available at arXiv:2108.08129.
[15] Dessein, A., Papadakis, N. and Rouas, J.-L. (2018). Regularized optimal transport and the rot mover's distance. *J. Mach. Learn. Res.* **19** Paper No. 15, 53 pp. MR3862422
[16] Di Marino, S. and Gerolin, A. (2020). Optimal transport losses and Sinkhorn algorithm with general convex regularization. Preprint. Available at arXiv:2007.00976.
[17] Dudley, R.M. (1968). The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Stat.* **40** 40–50. MR0236977 https://doi.org/10.1214/aoms/1177697802
[18] Eckstein, S. and Kupper, M. (2021). Computation of optimal transport and related hedging problems via penalization and neural networks. *Appl. Math. Optim.* **83** 639–667. MR4239795 https://doi.org/10.1007/s00245-019-09558-1
[19] Eckstein, S., Kupper, M. and Pohl, M. (2020). Robust risk aggregation with neural networks. *Math. Finance* **30** 1229–1272. MR4154769 https://doi.org/10.1111/mafi.12280
[20] Eckstein, S. and Nutz, M. (2022). Quantitative stability of regularized optimal transport and convergence of Sinkhorn's algorithm. *SIAM J. Math. Anal.* **54** 5922–5948. MR4506579 https://doi.org/10.1137/21M145505X
[21] Eckstein, S. and Nutz, M. (2022). Convergence Rates for Regularized Optimal Transport via Quantization. Preprint. Available at arXiv:2208.14391.
[22] Essid, M. and Solomon, J. (2018). Quadratically regularized optimal transport on graphs. *SIAM J. Sci. Comput.* **40** A1961–A1986. MR3820384 https://doi.org/10.1137/17M1132665
[23] Folland, G.B. (1999). *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. *Pure and Applied Mathematics (New York)*. New York: Wiley. MR1681462
[24] Folland, G.B. (2005). Higher-order derivatives and Taylor's formula in several variables. Preprint, 1–4.
[25] 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
[26] Franklin, J. and Lorenz, J. (1989). On the scaling of multidimensional matrices. *Linear Algebra Appl.* **114/115** 717–735. MR0986904 https://doi.org/10.1016/0024-3795(89)90490-4
[27] Genevay, A., Chizat, L., Bach, F., Cuturi, M. and Peyré, G. (2019). Sample complexity of Sinkhorn divergences. In *The 22nd International Conference on Artificial Intelligence and Statistics* 1574–1583. PMLR.
[28] Ghosal, P., Nutz, M. and Bernton, E. (2022). Stability of entropic optimal transport and Schrödinger bridges. *J. Funct. Anal.* **283** Paper No. 109622, 22 pp. MR4460341 https://doi.org/10.1016/j.jfa.2022.109622
[29] Gigli, N. and Tamanini, L. (2021). Second order differentiation formula on $\mathrm{RCD}^*(K,N)$ spaces. *J. Eur. Math. Soc. (JEMS)* **23** 1727–1795. MR4244516 https://doi.org/10.4171/JEMS/1042
[30] Goldfeld, Z., Kato, K., Rioux, G. and Sadhu, R. (2022). Limit Theorems for Entropic Optimal Transport Maps and the Sinkhorn Divergence. Preprint. Available at arXiv:2207.08683.
[31] Goldfeld, Z., Kato, K., Rioux, G. and Sadhu, R. (2022). Statistical inference with regularized optimal transport. Preprint. Available at arXiv:2205.04283.
[32] Gonzalez-Sanz, A., Loubes, J.-M. and Niles-Weed, J. (2022). Weak limits of entropy regularized Optimal Transport; potentials, plans and divergences. Preprint. Available at arXiv:2207.07427.
[33] Graf, S. and Luschgy, H. (2000). *Foundations of Quantization for Probability Distributions*. *Lecture Notes in Math.* **1730**. Berlin: Springer. MR1764176 https://doi.org/10.1007/BFb0103945
[34] Keriven, N. (2023). Entropic Optimal Transport on Random Graphs. *SIAM J. Math. Data Sci.* **5** 1028–1050. MR4671658 https://doi.org/10.1137/22M1518281
[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] Kloeckner, B.R. (2020). Empirical measures: Regularity is a counter-curse to dimensionality. *ESAIM Probab. Stat.* **24** 408–434. MR4153634 https://doi.org/10.1051/ps/2019025

[37] Larsson, M., Park, J. and Wiesel, J. (2023). On concentration of the empirical measure for general transport costs. Preprint. Available at arXiv:2305.18636.

[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] Lin, T., Ho, N. and Jordan, M. (2019). On efficient optimal transport: An analysis of greedy and accelerated mirror descent algorithms. In *International Conference on Machine Learning* 3982–3991. PMLR.

[40] Lorenz, D. and Mahler, H. (2022). Orlicz space regularization of continuous optimal transport problems. *Appl. Math. Optim.* **85** Paper No. 14, 33 pp. MR4409806 https://doi.org/10.1007/s00245-022-09826-7

[41] Lorenz, D.A., Manns, P. and Meyer, C. (2021). Quadratically regularized optimal transport. *Appl. Math. Optim.* **83** 1919–1949. MR4261277 https://doi.org/10.1007/s00245-019-09614-w

[42] Manole, T. and Niles-Weed, J. (2021). Sharp convergence rates for empirical optimal transport with smooth costs. Preprint. Available at arXiv:2106.13181.

[43] McCann, R.J., Pass, B. and Warren, M. (2012). Rectifiability of optimal transportation plans. *Canad. J. Math.* **64** 924–934. MR2957236 https://doi.org/10.4153/CJM-2011-080-6

[44] Mena, G. and Niles-Weed, J. (2019). Statistical bounds for entropic optimal transport: Sample complexity and the central limit theorem. *Adv. Neural Inf. Process. Syst.* **32**.

[45] Müller, A. (1997). Integral probability metrics and their generating classes of functions. *Adv. in Appl. Probab.* **29** 429–443. MR1450938 https://doi.org/10.2307/1428011

[46] Nutz, M. and Wiesel, J. (2022). Entropic optimal transport: Convergence of potentials. *Probab. Theory Related Fields* **184** 401–424. MR4498514 https://doi.org/10.1007/s00440-021-01096-8

[47] Nutz, M. and Wiesel, J. (2023). Stability of Schrödinger potentials and convergence of Sinkhorn's algorithm. *Ann. Probab.* **51** 699–722. MR4546630 https://doi.org/10.1214/22-aop1611

[48] Pal, S. (2019). On the difference between entropic cost and the optimal transport cost. Preprint. Available at arXiv:1905.12206.

[49] Peyré, G., Cuturi, M. et al. (2019). Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** 355–607.

[50] Pooladian, A.-A. and Niles-Weed, J. (2021). Entropic estimation of optimal transport maps. Preprint. Available at arXiv:2109.12004.

[51] Rigollet, P. and Stromme, A.J. (2022). On the sample complexity of entropic optimal transport. Preprint. Available at arXiv:2206.13472.

[52] Rockafellar, R.T. and Wets, R.J.-B. (2009). *Variational Analysis* **317**. Berlin: Springer.

[53] Rudin, W. (1976). *Principles of Mathematical Analysis*, 3rd ed. *International Series in Pure and Applied Mathematics*. New York–Auckland–Düsseldorf: McGraw-Hill. MR0385023

[54] Schmitzer, B. (2019). Stabilized sparse scaling algorithms for entropy regularized transport problems. *SIAM J. Sci. Comput.* **41** A1443–A1481. MR3947294 https://doi.org/10.1137/16M1106018

[55] Seguy, V., Damodaran, B.B., Flamary, R., Courty, N., Rolet, A. and Blondel, M. (2017). Large-scale optimal transport and mapping estimation. Preprint. Available at arXiv:1711.02283.

[56] Sinkhorn, R. and Knopp, P. (1967). Concerning nonnegative matrices and doubly stochastic matrices. *Pacific J. Math.* **21** 343–348. MR0210731

[57] Sreekumar, S., Goldfeld, Z. and Kato, K. (2022). Limit distribution theory for f-Divergences. Preprint. Available at arXiv:2211.11184.

[58] Sriperumbudur, B.K., Fukumizu, K., Gretton, A., Schölkopf, B. and Lanckriet, G.R.G. (2012). On the empirical estimation of integral probability metrics. *Electron. J. Stat.* **6** 1550–1599. MR2988458 https://doi.org/10.1214/12-EJS722

[59] Taşkesen, B., Shafieezadeh-Abadeh, S. and Kuhn, D. (2023). Semi-discrete optimal transport: Hardness, regularization and numerical solution. *Math. Program.* **199** 1033–1106. MR4578388 https://doi.org/10.1007/s10107-022-01856-x

[60] Terjék, D. and González-Sánchez, D. (2022). Optimal transport with f-divergence regularization and generalized Sinkhorn algorithm. In *International Conference on Artificial Intelligence and Statistics* 5135–5165. PMLR.

[61] 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

[62] Weed, J. (2018). An explicit analysis of the entropic penalty in linear programming. In *Conference on Learning Theory* 1841–1855. PMLR.

[63] 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 **31**(1), 2025, 240–270
https://doi.org/10.3150/24-BEJ1726

Adaptive deep learning for nonlinear time series models

DAISUKE KURISU[1,a], RIKU FUKAMI[2,b] and YUTA KOIKE[2,c]

[1] *Center for Spatial Information Science, The University of Tokyo, 5-1-5, Kashiwanoha, Kashiwa-shi, Chiba 277-8568, Japan,* [a] *daisukekurisu@csis.u-tokyo.ac.jp*
[2] *Graduate School of Mathematical Science, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan,* [b] *rick.h.azuma@gmail.com,* [c] *kyuta@ms.u-tokyo.ac.jp*

In this paper, we develop a general theory for adaptive nonparametric estimation of the mean function of a non-stationary and nonlinear time series model using deep neural networks (DNNs). We first consider two types of DNN estimators, non-penalized and sparse-penalized DNN estimators, and establish their generalization error bounds for general non-stationary time series. We then derive minimax lower bounds for estimating mean functions belonging to a wide class of nonlinear autoregressive (AR) models that include nonlinear generalized additive AR, single index, and threshold AR models. Building upon the results, we show that the sparse-penalized DNN estimator is adaptive and attains the minimax optimal rates up to a poly-logarithmic factor for many nonlinear AR models. Through numerical simulations, we demonstrate the usefulness of the DNN methods for estimating nonlinear AR models with intrinsic low-dimensional structures and discontinuous or rough mean functions, which is consistent with our theory.

Keywords: Adaptive estimation; deep neural network; minimax optimality; nonlinear time series

References

Barlow, M.T., Jacka, S.D. and Yor, M. (1986). Inequalities for a pair of processes stopped at a random time. *Proc. Lond. Math. Soc. (3)* **52** 142–172. MR0812449 https://doi.org/10.1112/plms/s3-52.1.142

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

Bhattacharya, R. and Lee, C. (1995). On geometric ergodicity of nonlinear autoregressive models. *Statist. Probab. Lett.* **22** 311–315. MR1333189 https://doi.org/10.1016/0167-7152(94)00082-J

Chen, M. and Chen, G. (2000). Geometric ergodicity of nonlinear autoregressive models with changing conditional variances. *Canad. J. Statist.* **28** 605–613. MR1793114 https://doi.org/10.2307/3315968

Chen, R. and Tsay, R.S. (1993). Functional-coefficient autoregressive models. *J. Amer. Statist. Assoc.* **88** 298–308. MR1212492

Cline, D.B.H. and Pu, H.-M.H. (2004). Stability and the Lyapounov exponent of threshold AR-ARCH models. *Ann. Appl. Probab.* **14** 1920–1949. MR2099657 https://doi.org/10.1214/105051604000000431

de la Peña, V.H., Klass, M.J. and Lai, T.L. (2004). Self-normalized processes: Exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.* **32** 1902–1933. MR2073181 https://doi.org/10.1214/009117904000000397

Dicker, L., Huang, B. and Lin, X. (2013). Variable selection and estimation with the seamless-L_0 penalty. *Statist. Sinica* **23** 929–962. MR3086661

Elbrächter, D., Perekrestenko, D., Grohs, P. and Bölcskei, H. (2021). Deep neural network approximation theory. *IEEE Trans. Inf. Theory* **67** 2581–2623. MR4282380 https://doi.org/10.1109/TIT.2021.3062161

Exterkate, P. (2013). Model selection in kernel ridge regression. *Comput. Statist. Data Anal.* **68** 1–16. MR3103758 https://doi.org/10.1016/j.csda.2013.06.006

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

Fan, J. and Yao, Q. (2008). *Nonlinear Time Series: Nonparametric and Parametric Methods*. Berlin: Springer.

Granger, C.W. and Teräsvirta, T. (1993). *Modelling Nonlinear Economic Relationships*. London: Oxford Univ. Press.

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

Haggan, V. and Ozaki, T. (1981). Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika* **68** 189–196. MR0614955 https://doi.org/10.1093/biomet/68.1.189

Hairer, M. and Mattingly, J.C. (2011). Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*. *Progress in Probability* **63** 109–117. Basel: Birkhäuser/Springer Basel AG. MR2857021 https://doi.org/10.1007/978-3-0348-0021-1_7

Hansen, B.E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* **24** 726–748. MR2409261 https://doi.org/10.1017/S0266466608080304

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

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

Hoffmann, M. (1999). On nonparametric estimation in nonlinear AR(1)-models. *Statist. Probab. Lett.* **44** 29–45. MR1706307 https://doi.org/10.1016/S0167-7152(98)00289-2

Imaizumi, M. and Fukumizu, K. (2019). Deep neural networks learn non-smooth functions effectively. In *Proc. Mach. Learn. Res.* **89** 869–878. PMLR.

Kingma, D.P. and Ba, J.L. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings* (Y. Bengio and Y. LeCun, eds.).

Kohler, M. and Krzyżak, A. (2023). On the rate of convergence of a deep recurrent neural network estimate in a regression problem with dependent data. *Bernoulli* **29** 1663–1685. MR4550240 https://doi.org/10.3150/22-bej1516

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

Kulik, R. and Lorek, P. (2011). Some results on random design regression with long memory errors and predictors. *J. Statist. Plann. Inference* **141** 508–523. MR2719514 https://doi.org/10.1016/j.jspi.2010.06.030

Kurisu, D., Fukami, R. and Koike, Y. (2025). Supplement to "Adaptive deep learning for nonlinear time series models." https://doi.org/10.3150/24-BEJ1726SUPP

Kuznetsov, V. and Mohri, M. (2017). Generalization bounds for non-stationary mixing processes. *Mach. Learn.* **106** 93–117. MR3592151 https://doi.org/10.1007/s10994-016-5588-2

Liebscher, E. (2005). Towards a unified approach for proving geometric ergodicity and mixing properties of nonlinear autoregressive processes. *J. Time Series Anal.* **26** 669–689. MR2188304 https://doi.org/10.1111/j.1467-9892.2005.00412.x

Liu, W. and Wu, W.B. (2010). Simultaneous nonparametric inference of time series. *Ann. Statist.* **38** 2388–2421. MR2676893 https://doi.org/10.1214/09-AOS789

Lu, Z. and Jiang, Z. (2001). L_1 geometric ergodicity of a multivariate nonlinear AR model with an ARCH term. *Statist. Probab. Lett.* **51** 121–130. MR1808034 https://doi.org/10.1016/S0167-7152(00)00138-3

Masry, E. (1996a). Multivariate local polynomial regression for time series: Uniform strong consistency and rates. *J. Time Series Anal.* **17** 571–599. MR1424907 https://doi.org/10.1111/j.1467-9892.1996.tb00294.x

Masry, E. (1996b). Multivariate regression estimation—Local polynomial fitting for time series. *Stochastic Process. Appl.* **65** 81–101. MR1422881 https://doi.org/10.1016/S0304-4149(96)00095-6

Nakada, R. and Imaizumi, M. (2020). Adaptive approximation and generalization of deep neural network with intrinsic dimensionality. *J. Mach. Learn. Res.* **21** Paper No. 174, 38 pp. MR4209460

Oga, A. and Koike, Y. (2024). Drift estimation for a multi-dimensional diffusion process using deep neural networks. *Stochastic Process. Appl.* **170** Paper No. 104240, 25 pp. MR4682669 https://doi.org/10.1016/j.spa.2023.104240

Ogihara, T. (2021). Misspecified diffusion models with high-frequency observations and an application to neural networks. *Stochastic Process. Appl.* **142** 245–292. MR4318367 https://doi.org/10.1016/j.spa.2021.08.007

Ohn, I. and Kim, Y. (2022). Nonconvex sparse regularization for deep neural networks and its optimality. *Neural Comput.* **34** 476–517. MR4381798

Phandoidaen, N. and Richter, S. (2020). Forecasting time series with encoder-decoder neural networks. Preprint. Available at arXiv:2009.08848.

Probst, P. and Boulesteix, A.-L. (2017). To tune or not to tune the number of trees in random forest. *J. Mach. Learn. Res.* **18** Paper No. 181, 18 pp. MR3827069

Rio, E. (2013). Inequalities and limit theorems for weakly dependent sequences. Preprint. Available at https://cel.archives-ouvertes.fr/cel-00867106v2.

Robinson, P.M. (1983). Nonparametric estimators for time series. *J. Time Series Anal.* **4** 185–207. MR0732897 https://doi.org/10.1111/j.1467-9892.1983.tb00368.x

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

Schmidt-Hieber, J. (2019). Deep ReLU network approximation of functions on a manifold. Preprint. Available at arXiv:1908.00695.

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

Suzuki, T. (2019). Adaptivity of deep ReLU network for learning in Besov and mixed smooth Besov spaces: Optimal rate and curse of dimensionality. In *International Conference on Learning Representations*. Available at https://openreview.net/forum?id=H1ebTsActm.

Suzuki, T. and Nitanda, A. (2021). Deep learning is adaptive to intrinsic dimensionality of model smoothness in anisotropic Besov space. *Adv. Neural Inf. Process. Syst.* **34**.

Teräsvirta, T. (1994). Specification, estimation, and evaluation of smooth transition autoregressive models. *J. Amer. Statist. Assoc.* **89** 208–218.

Tjøstheim, D. (1990). Nonlinear time series and Markov chains. *Adv. in Appl. Probab.* **22** 587–611. MR1066965 https://doi.org/10.2307/1427459

Tran, L.T. (1993). Nonparametric function estimation for time series by local average estimators. *Ann. Statist.* **21** 1040–1057. MR1232531 https://doi.org/10.1214/aos/1176349163

Truong, Y.K. (1994). Nonparametric time series regression. *Ann. Inst. Statist. Math.* **46** 279–293. MR1293164 https://doi.org/10.1007/BF01720585

Tsuji, K. and Suzuki, T. (2021). Estimation error analysis of deep learning on the regression problem on the variable exponent Besov space. *Electron. J. Stat.* **15** 1869–1908. MR4255314 https://doi.org/10.1214/21-ejs1828

van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes—With Applications to Statistics*. *Springer Series in Statistics*. Cham: Springer. MR4628026 https://doi.org/10.1007/978-3-031-29040-4

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614 https://doi.org/10.1214/12-AOS1043

Xia, Y. and Li, W.K. (1999). On single-index coefficient regression models. *J. Amer. Statist. Assoc.* **94** 1275–1285. MR1731489 https://doi.org/10.2307/2669941

Xia, Y., Li, W.-K. and Tong, H. (2007). Threshold variable selection using nonparametric methods. *Statist. Sinica* **17** 265–287. MR2352510

Yang, Y. (2001). Nonparametric regression with dependent errors. *Bernoulli* **7** 633–655. MR1849372 https://doi.org/10.2307/3318730

Zhang, T. (2010a). Analysis of multi-stage convex relaxation for sparse regularization. *J. Mach. Learn. Res.* **11** 1081–1107. MR2629825

Zhang, C.-H. (2010b). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. MR2604701 https://doi.org/10.1214/09-AOS729

Zhang, T. and Wu, W.B. (2015). Time-varying nonlinear regression models: Nonparametric estimation and model selection. *Ann. Statist.* **43** 741–768. MR3319142 https://doi.org/10.1214/14-AOS1299

Zhao, Z. and Wu, W.B. (2008). Confidence bands in nonparametric time series regression. *Ann. Statist.* **36** 1854–1878. MR2435458 https://doi.org/10.1214/07-AOS533

Bernoulli **31**(1), 2025, 271–294
https://doi.org/10.3150/24-BEJ1727

Asymptotically optimal sequential multiple testing with asynchronous decisions

YIMING XING[1,a] and GEORGIOS FELLOURIS[2,b]

[1] *School of Mathematical Sciences, Tongji University, Shanghai, China,* [a]*yimingx4@tongji.edu.cn*
[2] *Department of Statistics, University of Illinois at Urbana-Champaign, Champaign, USA,* [b]*fellouri@illinois.edu*

The problem of simultaneously testing the marginal distributions of sequentially monitored, independent data streams is considered. The decisions for the various testing problems can be made at different times, using data from all streams, which can be monitored until all decisions have been made. Moreover, arbitrary a priori bounds are assumed on the number of signals, i.e., data streams in which the alternative hypothesis is correct. A novel sequential multiple testing procedure is proposed and it is shown to achieve the minimum expected decision time, simultaneously in every data stream and under every signal configuration, asymptotically as certain metrics of global error rates go to zero. This optimality property is established under general parametric composite hypotheses, various error metrics, and weak distributional assumptions that allow for temporal dependence. Furthermore, the limit of the factor by which the expected decision time in a data stream increases when one is limited to synchronous or decentralized procedures is evaluated. Finally, two existing sequential multiple testing procedures in the literature are compared with the proposed one in various simulation studies.

Keywords: Asymptotic optimality; asynchronous decisions; parametric composite hypotheses; prior information; sequential multiple testing

References

Almogi-Nadler, M., Oshman, Y. and Ben-Asher, J.Z. (2004). Boost-phase identification of theater ballistic missiles using radar measurements. *J. Guid. Control Dyn.* **27** 197–208.

Bartroff, J. and Lai, T.L. (2010). Multistage tests of multiple hypotheses. *Comm. Statist. Theory Methods* **39** 1597–1607.

Bartroff, J. and Song, J. (2014). Sequential tests of multiple hypotheses controlling type I and II familywise error rates. *J. Statist. Plann. Inference* **153** 100–114. MR3229025 https://doi.org/10.1016/j.jspi.2014.05.010

Bartroff, J. and Song, J. (2016). A rejection principle for sequential tests of multiple hypotheses controlling familywise error rates. *Scand. J. Stat.* **43** 3–19. MR3466990 https://doi.org/10.1111/sjos.12161

Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly detection: A survey. *ACM Comput. Surv.* **41** 1–58.

Chaudhuri, A. and Fellouris, G. (2022). Joint sequential detection and isolation for dependent data streams. arXiv preprint. Available at arXiv:2207.00120.

Cohen, K. and Zhao, Q. (2015). Asymptotically optimal anomaly detection via sequential testing. *IEEE Trans. Signal Process.* **63** 2929–2941. MR3346453 https://doi.org/10.1109/TSP.2015.2416674

De, S.K. and Baron, M. (2012a). Sequential Bonferroni methods for multiple hypothesis testing with strong control of family-wise error rates I and II. *Sequential Anal.* **31** 238–262. MR2911288 https://doi.org/10.1080/07474946.2012.665730

De, S.K. and Baron, M. (2012b). Step-up and step-down methods for testing multiple hypotheses in sequential experiments. *J. Statist. Plann. Inference* **142** 2059–2070. MR2903412 https://doi.org/10.1016/j.jspi.2012.02.005

De, S.K. and Baron, M. (2015). Sequential tests controlling generalized familywise error rates. *Stat. Methodol.* **23** 88–102. MR3278804 https://doi.org/10.1016/j.stamet.2014.10.001

Ding, Y., Markatou, M. and Ball, R. (2020). An evaluation of statistical approaches to postmarketing surveillance. *Stat. Med.* **39** 845–874. MR4075833 https://doi.org/10.1002/sim.8447

1350-7265

Gupta, M.S. and Kumar, K. (2019). Progression on spectrum sensing for cognitive radio networks: A survey, classification, challenges and future research issues. *J. Netw. Comput. Appl.* **143** 47–76.

He, X. and Bartroff, J. (2021). Asymptotically optimal sequential FDR and pFDR control with (or without) prior information on the number of signals. *J. Statist. Plann. Inference* **210** 87–99. MR4102184 https://doi.org/10.1016/j.jspi.2020.05.002

Hemo, B., Gafni, T., Cohen, K. and Zhao, Q. (2020). Searching for anomalies over composite hypotheses. *IEEE Trans. Signal Process.* **68** 1181–1196. MR4114616 https://doi.org/10.1109/TSP.2020.2971438

Huang, B., Cohen, K. and Zhao, Q. (2019). Active anomaly detection in heterogeneous processes. *IEEE Trans. Inf. Theory* **65** 2284–2301. MR3928201 https://doi.org/10.1109/TIT.2018.2866257

Jennison, C. and Turnbull, B.W. (2000). *Group Sequential Methods with Applications to Clinical Trials*. Boca Raton, FL: CRC Press/CRC. MR1710781

Lehmann, E.L. and Romano, J.P. (2005). Generalizations of the familywise error rate. *Ann. Statist.* **33** 1138–1154. MR2195631 https://doi.org/10.1214/009053605000000084

Malloy, M.L. and Nowak, R.D. (2014). Sequential testing for sparse recovery. *IEEE Trans. Inf. Theory* **60** 7862–7873. MR3285750 https://doi.org/10.1109/TIT.2014.2363846

Pavlov, I.V. (1991). Sequential procedure of testing composite hypotheses with applications to the Kiefer–Weiss problem. *Theory Probab. Appl.* **35** 280–292.

Prabhu, G.R., Bhashyam, S., Gopalan, A. and Sundaresan, R. (2022). Sequential multi-hypothesis testing in multi-armed bandit problems: An approach for asymptotic optimality. *IEEE Trans. Inf. Theory* **68** 4790–4817. MR4449074

Robbins, H. and Siegmund, D. (1974). The expected sample size of some tests of power one. *Ann. Statist.* **2** 415–436. MR0448750

Sarkar, S.K. (2007). Stepup procedures controlling generalized FWER and generalized FDR. *Ann. Statist.* **35** 2405–2420. MR2382652 https://doi.org/10.1214/009053607000000398

Sarkar, S.K., Chen, J. and Guo, W. (2013). Multiple testing in a two-stage adaptive design with combination tests controlling FDR. *J. Amer. Statist. Assoc.* **108** 1385–1401. MR3174715 https://doi.org/10.1080/01621459.2013.835662

Song, Y. and Fellouris, G. (2017a). Asymptotically optimal, sequential, multiple testing procedures with prior information on the number of signals. *Electron. J. Stat.* **11** 338–363. MR3606774 https://doi.org/10.1214/17-EJS1223

Song, Y. and Fellouris, G. (2019). Sequential multiple testing with generalized error control: An asymptotic optimality theory. *Ann. Statist.* **47** 1776–1803. MR3911130 https://doi.org/10.1214/18-AOS1737

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

Tsopelakos, A. and Fellouris, G. (2023). Sequential anomaly detection under sampling constraints. *IEEE Trans. Inf. Theory* **69** 8126–8146. MR4692659

Xing, Y. and Fellouris, G. (2023). Signal recovery with multistage tests and without sparsity constraints. *IEEE Trans. Inf. Theory* **69** 7220–7245. MR4660860

Xing, Y. and Fellouris, G. (2025). Supplement to "Asymptotically optimal sequential multiple testing with asynchronous decisions." https://doi.org/10.3150/24-BEJ1727SUPP

Xing, Y. and Fellouris, G. (to appear in 2025). Asymptotically optimal multistage tests for non-iid data. *Statist. Sinica*.

Zehetmayer, S., Bauer, P. and Posch, M. (2005). Two-stage designs for experiments with a large number of hypotheses. *Bioinformatics* **21** 3771–3777. https://doi.org/10.1093/bioinformatics/bti604

Bernoulli **31**(1), 2025, 295–311
https://doi.org/10.3150/24-BEJ1728

Independence preserving property of Kummer laws

ANGELO EFOÉVI KOUDOU[1,a] and JACEK WESOŁOWSKI[2,b]

[1]*Institut Elie Cartan de Lorraine, UMR CNRS 7502, Université de Lorraine, Nancy, France,*
[a]*efoevi.koudou@univ-lorraine.fr*
[2]*Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Warszawa, Poland,*
[b]*jacek.wesolowski@pw.edu.pl*

We prove that if X,Y are positive, independent, non-Dirac random variables and if for $\alpha,\beta \geq 0$, $\alpha \neq \beta$,

$$\psi_{\alpha,\beta}(x,y) = \left(y\,\frac{1+\beta(x+y)}{1+\alpha x+\beta y},\; x\,\frac{1+\alpha(x+y)}{1+\alpha x+\beta y}\right),$$

then the random variables U and V defined by $(U,V) = \psi_{\alpha,\beta}(X,Y)$ are independent if and only if X and Y follow Kummer distributions with suitably related parameters. In other words, any invariant measure for a lattice recursion model governed by $\psi_{\alpha,\beta}$ in the scheme introduced by Croydon and Sasada in (Croydon and Sasada (2020)) is necessarily a product measure with Kummer marginals. The result extends earlier characterizations of Kummer and gamma laws by independence of

$$U = \frac{Y}{1+X} \quad \text{and} \quad V = X\left(1 + \frac{Y}{1+X}\right),$$

which corresponds to the case of $\psi_{1,0}$. We also show, in the supplement, that this independence property of Kummer laws covers, as limiting cases, several independence models known in the literature: the Lukacs, the Kummer-Gamma, the Matsumoto-Yor and the discrete Korteweg de Vries models.

Keywords: Independence preserving property; Kummer distributions; Matsumoto-Yor property

References

[1] Abramowitz, M. and Stegund, I.A. (1984). *Pocketbook of Mathematical Functions*. Frankfurt am Main: Verlag Harri Deutsch-Thun-Frankfurt/Main.
[2] Bao, K.V. and Noack, C. (2021). Characterizations of the generalized inverse Gaussian, asymmetric Laplace, and shifted (truncated) exponential laws via independence properties. Available at arXiv:2107.01394.
[3] Barraquand, G. and Corwin, I. (2017). Random-walk in beta-distributed random environment. *Probab. Theory Related Fields* **167** 1057–1116. MR3627433 https://doi.org/10.1007/s00440-016-0699-z
[4] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396
[5] Chaumont, H. and Noack, C. (2018). Characterizing stationary 1 + 1 dimensional lattice polymer models. *Electron. J. Probab.* **23** 1–19. MR3806406 https://doi.org/10.1214/18-EJP163
[6] Crawford, G.B. (1966). Characterization of geometric and exponential distributions. *Ann. Math. Stat.* **37** 1790–1795. MR0207009 https://doi.org/10.1214/aoms/1177699167
[7] Croydon, D.A., Kato, T., Sasada, M. and Tsujimoto, S. (2023). Dynamics of the box-ball system with random initial conditions via Pitman's transformation. *Mem. Amer. Math. Soc.* **283** vii+99. MR4536407 https://doi.org/10.1090/memo/1398
[8] Croydon, D.A. and Sasada, M. (2020). Detailed balance and invariant mesures for discrete KdV- and Toda-type systems. Available at arXiv:2007.06203.
[9] Croydon, D.A., Sasada, M. and Tsujimoto, S. (2022). Bi-infinite solutions for KdV- and Toda-type discrete integrable systems based on path encodings. *Math. Phys. Anal. Geom.* **25** Paper No. 27, 71. MR4502797 https://doi.org/10.1007/s11040-022-09435-4

[10] Darroch, J.N. and Ratcliff, D. (1971). A characterization of the Dirichlet distribution. *J. Amer. Statist. Assoc.* **66** 641–643. MR0292213
[11] Ferguson, T.S. (1964). A characterization of the exponential distribution. *Ann. Math. Stat.* **35** 1199–1207. MR0168057 https://doi.org/10.1214/aoms/1177703277
[12] Hamza, M. and Vallois, P. (2016). On Kummer's distribution of type two and a generalized beta distribution. *Statist. Probab. Lett.* **118** 60–69. MR3531483 https://doi.org/10.1016/j.spl.2016.03.014
[13] Kołodziejek, B. and Piliszek, A. (2020). Independence characterization for Wishart and Kummer random matrices. *REVSTAT* **18** 357–373. MR4149493
[14] Koudou, A.E. and Vallois, P. (2012). Independence properties of the Matsumoto–Yor type. *Bernoulli* **18** 119–136. MR2888701 https://doi.org/10.3150/10-BEJ325
[15] Koudou, A.E. and Wesołowski, J. (2025). Supplement to "Independence preserving property of Kummer laws." https://doi.org/10.3150/24-BEJ1728SUPP
[16] Letac, G. and Wesołowski, J. (2000). An independence property for the product of GIG and gamma laws. *Ann. Probab.* **28** 1371–1383. MR1797878 https://doi.org/10.1214/aop/1019160339
[17] Letac, G. and Wesołowski, J. (2023). About an extension of the Matsumoto–Yor property. *Ann. Inst. Henri Poincaré* To appear.
[18] Lukacs, E. (1955). A characterization of the gamma distribution. *Ann. Math. Stat.* **26** 319–324. MR0069408 https://doi.org/10.1214/aoms/1177728549
[19] Matsumoto, H. and Yor, M. (2001). An analogue of Pitman's $2M - X$ theorem for exponential Wiener functionals. II. The role of the generalized inverse Gaussian laws. *Nagoya Math. J.* **162** 65–86. MR1836133 https://doi.org/10.1017/S0027763000007807
[20] Piliszek, A. (2022). Regression conditions that characterize free-Poisson and free-Kummer distributions. *Random Matrices Theory Appl.* **11** Paper No. 2250019, 28. MR4414197 https://doi.org/10.1142/S2010326322500198
[21] Piliszek, A. and Wesołowski, J. (2016). Kummer and gamma laws through independences on trees – another parallel with the Matsumoto–Yor property. *J. Multivariate Anal.* **152** 15–27. MR3554776 https://doi.org/10.1016/j.jmva.2016.07.004
[22] Piliszek, A. and Wesołowski, J. (2018). Change of measure technique in characterizations of the gamma and Kummer distributions. *J. Math. Anal. Appl.* **458** 967–979. MR3724710 https://doi.org/10.1016/j.jmaa.2017.10.011
[23] Sasada, M. and Uozumi, R. (2022). Yang-Baxter maps and independence preserving property. Available at arXiv:2212.00963 1.
[24] Seppäläinen, T. (2012). Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* **40** 19–73. MR2917766 https://doi.org/10.1214/10-AOP617
[25] Seshadri, V. and Wesołowski, J. (2003). Constancy of regressions for beta distributions. *Sankhyā* **65** 284–291. MR2028900
[26] Wesołowski, J. (2015). On the Matsumoto–Yor type regression characterization of the gamma and Kummer distributions. *Statist. Probab. Lett.* **107** 145–149. MR3412768 https://doi.org/10.1016/j.spl.2015.07.036

Bernoulli **31**(1), 2025, 312–332
https://doi.org/10.3150/24-BEJ1729

Nearest neighbor empirical processes

FRANÇOIS PORTIER[a]

Ensai, CREST, University of Rennes, France, [a]*francois.portier@gmail.com*

In the regression framework, the empirical measure based on the responses resulting from the nearest neighbors, among the covariates, to a given point x is introduced and studied as a central statistical quantity. First, the associated empirical process is shown to satisfy a uniform central limit theorem under a local bracketing entropy condition on the underlying class of functions reflecting the localizing nature of the nearest neighbor algorithm. Second a uniform non-asymptotic bound is established under a well-known condition, often referred to as Vapnik-Chervonenkis, on the uniform entropy numbers. The covariance of the Gaussian limit obtained in the uniform central limit theorem is simply equal to the conditional covariance operator given the covariate value. This suggests the possibility of using standard formulas to estimate the variance by using only the nearest neighbors instead of the full data. This is illustrated on two problems: the estimation of the conditional cumulative distribution function and local linear regression.

Keywords: Concentration inequality; empirical process theory; nearest neighbor algorithm; weak convergence

References

Ausset, G., Clémençon, S. and Portier, F. (2021). Nearest neighbour based estimates of gradients: Sharp nonasymptotic bounds and applications. In *AISTATS Proceedings* 532–540. PMLR.

Bartlett, P.L., Bousquet, O. and Mendelson, S. (2005). Local Rademacher complexities. *Ann. Statist.* **33** 1497–1537. MR2166554 https://doi.org/10.1214/009053605000000282

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical Report, Univ. California, Berkeley.

Biau, G., Cérou, F. and Guyader, A. (2010). On the rate of convergence of the bagged nearest neighbor estimate. *J. Mach. Learn. Res.* **11** 687–712. MR2600626

Biau, G. and Devroye, L. (2015). *Lectures on the Nearest Neighbor Method*. *Springer Series in the Data Sciences*. Cham: Springer. MR3445317 https://doi.org/10.1007/978-3-319-25388-6

Bickel, P.J. and Breiman, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. MR0682809

Boucheron, S., Bousquet, O. and Lugosi, G. (2005). Theory of classification: A survey of some recent advances. *ESAIM Probab. Stat.* **9** 323–375. MR2182250 https://doi.org/10.1051/ps:2005018

Bousquet, O. (2002). A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Math. Acad. Sci. Paris* **334** 495–500. MR1890640 https://doi.org/10.1016/S1631-073X(02)02292-6

Cannings, T.I., Berrett, T.B. and Samworth, R.J. (2020). Local nearest neighbour classification with applications to semi-supervised learning. *Ann. Statist.* **48** 1789–1814. MR4124344 https://doi.org/10.1214/19-AOS1868

Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **13** 21–27.

Datta, A., Banerjee, S., Finley, A.O. and Gelfand, A.E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *J. Amer. Statist. Assoc.* **111** 800–812. MR3538706 https://doi.org/10.1080/01621459.2015.1044091

Devroye, L.P. (1978). The uniform convergence of nearest neighbor regression function estimators and their application in optimization. *IEEE Trans. Inf. Theory* **24** 142–151. MR0518942 https://doi.org/10.1109/tit.1978.1055865

Devroye, L.P. and Wagner, T.J. (1977). The strong uniform consistency of nearest neighbor density estimates. *Ann. Statist.* **5** 536–540. MR0436442

Devroye, L. and Wise, G.L. (1980). Consistency of a recursive nearest neighbor regression function estimate. *J. Multivariate Anal.* **10** 539–550. MR0599688 https://doi.org/10.1016/0047-259X(80)90068-8

1350-7265

Devroye, L., Györfi, L., Krzyżak, A. and Lugosi, G. (1994). On the strong universal consistency of nearest neighbor regression function estimates. *Ann. Statist.* **22** 1371–1385. MR1311980 https://doi.org/10.1214/aos/1176325633

Devroye, L., Györfi, L., Lugosi, G. and Walk, H. (2017). On the measure of Voronoi cells. *J. Appl. Probab.* **54** 394–408. MR3668473 https://doi.org/10.1017/jpr.2017.7

Devroye, L., Györfi, L., Lugosi, G. and Walk, H. (2018). A nearest neighbor estimate of the residual variance. *Electron. J. Stat.* **12** 1752–1778. MR3811758 https://doi.org/10.1214/18-EJS1438

Döring, M., Györfi, L. and Walk, H. (2017). Rate of convergence of k-nearest-neighbor classification rule. *J. Mach. Learn. Res.* **18** Paper No. 227, 16. MR3827115

Dabrowska, D.M. (1989). Uniform consistency of the kernel conditional Kaplan-Meier estimate. *Ann. Statist.* **17** 1157–1167. MR1015143 https://doi.org/10.1214/aos/1176347261

Dudley, R.M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330. MR0220340 https://doi.org/10.1016/0022-1236(67)90017-1

Einmahl, U. and Mason, D.M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. *J. Theoret. Probab.* **13** 1–37. MR1744994 https://doi.org/10.1023/A:1007769924157

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. London: CRC Press. MR1383587

Fix, E. and Hodges, J.L. (1951). Discriminatory analysis. Nonparametric discrimination: Consistency properties. *Int. Stat. Rev.* **57** 238–247.

Gadat, S., Klein, T. and Marteau, C. (2016). Classification in general finite dimensional spaces with the k-nearest neighbor rule. *Ann. Statist.* **44** 982–1009. MR3485951 https://doi.org/10.1214/15-AOS1395

Giné, E. and Guillou, A. (2001). On consistency of kernel density estimators for randomly censored data: Rates holding uniformly over adaptive intervals. *Ann. Inst. Henri Poincaré Probab. Stat.* **37** 503–522. MR1876841 https://doi.org/10.1016/S0246-0203(01)01081-0

Giné, E. and Guillou, A. (2002). Rates of strong uniform consistency for multivariate kernel density estimators. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 907–921. En l'honneur de J. Bretagnolle, D. Dacunha-Castelle, I. Ibragimov. MR1955344 https://doi.org/10.1016/S0246-0203(02)01128-7

Giné, E. and Nickl, R. (2009). An exponential inequality for the distribution function of the kernel density estimator, with applications to adaptive estimation. *Probab. Theory Related Fields* **143** 569–596. MR2475673 https://doi.org/10.1007/s00440-008-0137-y

Györfi, L. (1981). The rate of convergence of k_n-NN regression estimates and classification rules. *IEEE Trans. Inf. Theory* **27** 362–364. MR0619124 https://doi.org/10.1109/TIT.1981.1056344

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

Hansen, B.E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory* **24** 726–748. MR2409261 https://doi.org/10.1017/S0266466608080304

Härdle, W., Janssen, P. and Serfling, R. (1988). Strong uniform consistency rates for estimators of conditional functionals. *Ann. Statist.* **16** 1428–1449. MR0964932 https://doi.org/10.1214/aos/1176351047

Härdle, W. and Tsybakov, A.B. (1988). Robust nonparametric regression with simultaneous scale curve estimation. *Ann. Statist.* **16** 120–135. MR0924860 https://doi.org/10.1214/aos/1176350694

Horváth, L. and Yandell, B.S. (1988). Asymptotics of conditional empirical processes. *J. Multivariate Anal.* **26** 184–206. MR0963831 https://doi.org/10.1016/0047-259X(88)90080-2

Jiang, H. (2019). Non-asymptotic uniform rates of consistency for k-NN regression. In *AAAI Proceedings* **33** 3999–4006.

Kpotufe, S. (2011). K-NN regression adapts to local intrinsic dimension. In *NeurIPS Proceedings* **24** 729–737.

Mack, Y.P. (1981). Local properties of k-NN regression estimates. *SIAM J. Algebr. Discrete Methods* **2** 311–323. MR0627598 https://doi.org/10.1137/0602035

Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.* **9** 141–142.

Nolan, D. and Pollard, D. (1987). U-processes: Rates of convergence. *Ann. Statist.* **15** 780–799. MR0888439 https://doi.org/10.1214/aos/1176350374

Plassier, V., Portier, F. and Segers, J. (2023). Risk bounds when learning infinitely many response functions by ordinary linear regression. *Ann. Inst. Henri Poincaré Probab. Stat.* **59** 53–78. MR4533720 https://doi.org/10.1214/22-aihp1259

Portier, F. (2025). Supplement to “Nearest neighbor empirical processes.” https://doi.org/10.3150/24-BEJ1729SUPP

Qiao, X., Duan, J. and Cheng, G. (2019). Rates of convergence for large-scale nearest neighbor classification. In *NeurIPS Proceedings* **32** 10769–10780.

Royall, R.M. (1966). A class of non-parametric estimates of a smooth regression function. ProQuest LLC, Ann Arbor, MI. Ph.D. thesis, Stanford University. MR2615964

Stone, C.J. (1977). Consistent nonparametric regression. *Ann. Statist.* **5** 595–645. With discussion and a reply by the author. MR0443204

Stone, C.J. (1982). Optimal global rates of convergence for nonparametric regression. *Ann. Statist.* **10** 1040–1053. MR0673642

Stute, W. (1986). Conditional empirical processes. *Ann. Statist.* **14** 638–647. MR0840519 https://doi.org/10.1214/aos/1176349943

Talagrand, M. (1996). New concentration inequalities in product spaces. *Invent. Math.* **126** 505–563. MR1419006 https://doi.org/10.1007/s002220050108

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

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

Vapnik, V.N. and Chervonenkis, A.Y. (2015). On the uniform convergence of relative frequencies of events to their probabilities. In *Measures of Complexity* 11–30. Cham: Springer. Reprint of Theor. Probability Appl. **16** (1971), 264–280. MR3408730

Veraverbeke, N., Omelka, M. and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scand. J. Stat.* **38** 766–780. MR2859749 https://doi.org/10.1111/j.1467-9469.2011.00744.x

Watson, G.S. (1964). Smooth regression analysis. *Sankhyā Ser. A* **26** 359–372. MR0185765

Weinberger, K.Q., Blitzer, J. and Saul, L.K. (2006). Distance metric learning for large margin nearest neighbor classification. In *NeurIPS Proceedings* **18** 1473–1480.

Wenocur, R.S. and Dudley, R.M. (1981). Some special Vapnik-Chervonenkis classes. *Discrete Math.* **33** 313–318. MR0602047 https://doi.org/10.1016/0012-365X(81)90274-0

Bernoulli **31**(1), 2025, 333–358
https://doi.org/10.3150/24-BEJ1730

Parameter estimation with increased precision for elliptic and hypo-elliptic diffusions

YUGA IGUCHI[1,a], ALEXANDROS BESKOS[1,b] and MATTHEW GRAHAM[2,c]

[1]*Department of Statistical Science, University College London, London, UK,* [a]*yuga.iguchi.21@ucl.ac.uk,* [b]*a.beskos@ucl.ac.uk*
[2]*Advanced Research Computing Centre, University College London, London, UK,* [c]*m.graham@ucl.ac.uk*

This work aims at making a comprehensive contribution in the general area of parametric inference for discretely observed diffusion processes. Established approaches for likelihood-based estimation invoke a time-discretisation scheme for the approximation of the intractable transition dynamics of the Stochastic Differential Equation (SDE) model over finite time periods. The scheme is applied for a step-size $\delta > 0$, that is either user-selected or determined by the data. Recent research has highlighted the critical effect of the choice of numerical scheme on the behaviour of derived parameter estimates in the setting of *hypo-elliptic* SDEs. In brief, in our work, first, we develop two weak second order *sampling schemes* (to cover both hypo-elliptic and elliptic SDEs) and produce a *small time expansion* for the density of the schemes to form a proxy for the true intractable SDE transition density. Then, we establish a collection of analytic results for likelihood-based parameter estimates obtained via the formed proxies, thus providing a theoretical framework that showcases advantages from the use of the developed methodology for SDE calibration. We present numerical results from carrying out classical or Bayesian inference, for both elliptic and hypo-elliptic SDEs.

Keywords: CLT; data augmentation; hypo-elliptic diffusion; small time density expansion; stochastic differential equation

References

Ableidinger, M., Buckwar, E. and Hinterleitner, H. (2017). A stochastic version of the Jansen and Rit neural mass model: Analysis and numerics. *J. Math. Neurosci.* **7** Paper No. 8, 35. MR3683994 https://doi.org/10.1186/s13408-017-0046-4

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica* **70** 223–262. MR1926260 https://doi.org/10.1111/1468-0262.00274

Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.* **36** 906–937. MR2396819 https://doi.org/10.1214/009053607000000622

Andrieu, C., Doucet, A. and Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 269–342. MR2758115 https://doi.org/10.1111/j.1467-9868.2009.00736.x

Anonymous (1978). News and notes: Influenza in a boarding school. *Br. Med. J.* **1** 586–590.

Beskos, A., Dureau, J. and Kalogeropoulos, K. (2015). Bayesian inference for partially observed stochastic differential equations driven by fractional Brownian motion. *Biometrika* **102** 809–827. MR3431555 https://doi.org/10.1093/biomet/asv051

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint. Available at arXiv:1701.02434.

Buckwar, E., Tamborrino, M. and Tubikanec, I. (2020). Spectral density-based and measure-preserving ABC for partially observed diffusion processes. An illustration on Hamiltonian SDEs. *Stat. Comput.* **30** 627–648. MR4065223 https://doi.org/10.1007/s11222-019-09909-6

Ditlevsen, S. and Samson, A. (2019). Hypoelliptic diffusions: Filtering and inference from complete and partial observations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 361–384. MR3928146

1350-7265

Gloter, A. and Yoshida, N. (2020). Adaptive and non-adaptive estimation for degenerate diffusion processes. arXiv preprint. Available at arXiv:2002.10164.

Gloter, A. and Yoshida, N. (2021). Adaptive estimation for degenerate diffusion processes. *Electron. J. Stat.* **15** 1424–1472. MR4255288 https://doi.org/10.1214/20-ejs1777

Gobet, E. and Labart, C. (2008). Sharp estimates for the convergence of the density of the Euler scheme in small time. *Electron. Commun. Probab.* **13** 352–363. MR2415143 https://doi.org/10.1214/ECP.v13-1393

Graham, M.M., Thiery, A.H. and Beskos, A. (2022). Manifold Markov chain Monte Carlo methods for Bayesian inference in diffusion models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **84** 1229–1256. MR4494159 https://doi.org/10.1111/rssb.12497

Hoffman, M.D. and Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15** 1593–1623. MR3214779

Iguchi, Y., Beskos, A. and Graham, M. (2025). Supplement to "Parameter estimation with increased precision for elliptic and hypo-elliptic diffusions." https://doi.org/10.3150/24-BEJ1730SUPP

Iguchi, Y. and Yamada, T. (2021). Operator splitting around Euler–Maruyama scheme and high order discretization of heat kernels. *ESAIM Math. Model. Numer. Anal.* **55** S323–S367. MR4221306 https://doi.org/10.1051/m2an/2020043

Iguchi, Y. and Yamada, T. (2022). Weak approximation of SDEs for tempered distributions and applications. *Adv. Comput. Math.* **48** Paper No. 52, 43. MR4462780 https://doi.org/10.1007/s10444-022-09960-4

Ikeda, N. and Watanabe, S. (2014). *Stochastic Differential Equations and Diffusion Processes*. Amsterdam: Elsevier.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Stat.* **24** 211–229. MR1455868 https://doi.org/10.1111/1467-9469.00059

Kingma, D.P. and Ba, J. (2015). Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR* (Y. Bengio and Y. LeCun, eds.).

Kloeden, P.E. and Platen, E. (1992). *Numerical Solution of Stochastic Differential Equations*. *Applications of Mathematics (New York)* **23**. Berlin: Springer. MR1214374 https://doi.org/10.1007/978-3-662-12616-5

Li, C. (2013). Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.* **41** 1350–1380. MR3113814 https://doi.org/10.1214/13-AOS1118

Milstein, G.N. and Tretyakov, M.V. (2021). *Stochastic Numerics for Mathematical Physics*. *Scientific Computation*. Cham: Springer. Second edition [of 2069903]. MR4369963 https://doi.org/10.1007/978-3-030-82040-4

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

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

Pavliotis, G.A. (2014). *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. *Texts in Applied Mathematics* **60**. New York: Springer. MR3288096 https://doi.org/10.1007/978-1-4939-1323-7

Pokern, Y., Stuart, A.M. and Wiberg, P. (2009). Parameter estimation for partially observed hypoelliptic diffusions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 49–73. MR2655523 https://doi.org/10.1111/j.1467-9868.2008.00689.x

Prakasa Rao, B.L.S. (1988). Statistical inference from sampled data for stochastic processes. In *Statistical Inference from Stochastic Processes (Ithaca, NY, 1987)*. *Contemp. Math.* **80** 249–284. Providence, RI: Amer. Math. Soc. MR0999016 https://doi.org/10.1090/conm/080/999016

Stramer, O., Bognar, M. and Schneider, P. (2010). Bayesian inference for discretely sampled Markov processes with closed-form likelihood expansions. *J. Financ. Econom.* **8** 450–480.

Uchida, M. and Yoshida, N. (2012). Adaptive estimation of an ergodic diffusion process based on sampled data. *Stochastic Process. Appl.* **122** 2885–2924. MR2931346 https://doi.org/10.1016/j.spa.2012.04.001

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B. and Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Anal.* **16** 667–718. MR4298989 https://doi.org/10.1214/20-ba1221

Watanabe, S. (1987). Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels. *Ann. Probab.* **15** 1–39. MR0877589

Bernoulli **31**(1), 2025, 359–387
https://doi.org/10.3150/24-BEJ1731

Poisson hulls

GÜNTER LAST[1,a] and ILYA MOLCHANOV[2,b]

[1]*Institute for Stochastics, Karlsruhe Institute of Technology, Englerstrasse 5, 76131 Karlsruhe, Germany,* [a]*guenter.last@kit.edu*
[2]*Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, CH-3012 Bern, Switzerland,* [b]*ilya.molchanov@stat.unibe.ch*

We introduce a hull operator on Poisson point processes, the easiest example being the convex hull of the support of a point process in Euclidean space. Assuming that the intensity measure of the process is known on the set generated by the hull operator, we discuss estimation of an expected linear statistic built on the Poisson process. In special cases, our general scheme yields an estimator of the volume of a convex body or an estimator of an integral of a Hölder function. We show that the estimation error is given by the Kabanov–Skorohod integral with respect to the underlying Poisson process. A crucial ingredient of our approach is a spatial strong Markov property of the underlying Poisson process with respect to the hull. We derive the rate of normal convergence for the estimation error, and illustrate it on an application to estimators of integrals of a Hölder function. We also discuss estimation of higher order symmetric statistics.

Keywords: Convex hull; hull operator; Kabanov–Skorohod integral; Poisson process; symmetric statistics

References

[1] Baldin, N. and Reiß, M. (2016). Unbiased estimation of the volume of a convex body. *Stochastic Process. Appl.* **126** 3716–3732. MR3565474 https://doi.org/10.1016/j.spa.2016.04.014
[2] Beermann, M. and Reitzner, M. (2015). Beyond the Efron-Buchta identities: Distributional results for Poisson polytopes. *Discrete Comput. Geom.* **53** 226–244. MR3293497 https://doi.org/10.1007/s00454-014-9649-7
[3] Brunel, V.-E. (2018). Methods for estimation of convex sets. *Statist. Sci.* **33** 615–632. MR3881211 https://doi.org/10.1214/18-STS669
[4] Cuevas, A. and Fraiman, R. (2010). Set estimation. In *New Perspectives in Stochastic Geometry* (W.S. Kendall and I. Molchanov, eds.) 374–397. Oxford: Oxford Univ. Press. MR2654684
[5] Gierz, G., Hofmann, K.H., Keimel, K., Lawson, J.D., Mislove, M. and Scott, D.S. (2003). *Continuous Lattices and Domains*. *Encyclopedia of Mathematics and Its Applications* **93**. Cambridge: Cambridge Univ. Press. MR1975381 https://doi.org/10.1017/CBO9780511542725
[6] Grygierek, J. (2021). Multivariate normal approximation for functionals of random polytopes. *J. Theoret. Probab.* **34** 897–922. MR4259453 https://doi.org/10.1007/s10959-020-00985-3
[7] Hug, D. and Schneider, R. (2020). Poisson hyperplane processes and approximation of convex bodies. *Mathematika* **66** 713–732. MR4134241 https://doi.org/10.1112/mtk.12040
[8] Kabluchko, Z., Marynych, A., Temesvari, D. and Thäle, C. (2019). Cones generated by random points on half-spheres and convex hulls of Poisson point processes. *Probab. Theory Related Fields* **175** 1021–1061. MR4026612 https://doi.org/10.1007/s00440-019-00907-3
[9] Kallenberg, O. (2017). *Random Measures, Theory and Applications*. *Probability Theory and Stochastic Modelling* **77**. Cham: Springer. MR3642325 https://doi.org/10.1007/978-3-319-41598-7
[10] Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann. Appl. Probab.* **29** 931–993. MR3910021 https://doi.org/10.1214/18-AAP1405
[11] Last, G. (2016). Stochastic analysis for Poisson processes. In *Stochastic Analysis for Poisson Point Processes* (G. Peccati and M. Reitzner, eds.). *Bocconi Springer Ser.* **7** 1–36. Bocconi Univ. Press, [place of publication not identified]. MR3585396 https://doi.org/10.1007/978-3-319-05233-5_1
[12] Last, G. and Molchanov, I. (2025). Supplement to “Poisson hulls.” https://doi.org/10.3150/24-BEJ1731SUPP

1350-7265

[13] Last, G., Molchanov, I. and Schulte, M. (2024). Normal approximation of stochastic integrals on Poisson spaces. *J. Theor. Probab.* **37**. To appear. https://doi.org/10.1007/s10959-023-01287-0

[14] Last, G., Peccati, G. and Yogeshwaran, D. (2023). Phase transitions and noise sensitivity on the Poisson space via stopping sets and decision trees. *Random Structures Algorithms* **63** 457–511. MR4622333 https://doi.org/10.1002/rsa.21136

[15] Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge: Cambridge Univ. Press. MR3791470

[16] Privault, N. (2012). Invariance of Poisson measures under random transformations. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** 947–972. MR3052400 https://doi.org/10.1214/11-AIHP422

[17] Reiß, M. and Selk, L. (2017). Efficient estimation of functionals in nonparametric boundary models. *Bernoulli* **23** 1022–1055. MR3606758 https://doi.org/10.3150/15-BEJ768

[18] Reiss, M. and Wahl, M. (2019). Functional estimation and hypothesis testing in nonparametric boundary models. *Bernoulli* **25** 2597–2619. MR4003559 https://doi.org/10.3150/18-BEJ1064

[19] Reitzner, M. and Schulte, M. (2013). Central limit theorems for U-statistics of Poisson point processes. *Ann. Probab.* **41** 3879–3909. MR3161465 https://doi.org/10.1214/12-AOP817

[20] Ripley, B.D. and Rasson, J.-P. (1977). Finding the edge of a Poisson forest. *J. Appl. Probab.* **14** 483–491. MR0451339 https://doi.org/10.2307/3213451

[21] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. *Probability and Its Applications (New York)*. Berlin: Springer. MR2455326 https://doi.org/10.1007/978-3-540-78859-1

[22] Zuyev, S. (1999). Stopping sets: Gamma-type results and hitting properties. *Adv. in Appl. Probab.* **31** 355–366. MR1724557 https://doi.org/10.1239/aap/1029955139

Bernoulli **31**(1), 2025, 388–411
https://doi.org/10.3150/24-BEJ1732

Cross-validation for change-point regression: Pitfalls and solutions

FLORIAN PEIN[1,a] and RAJEN D. SHAH[2,b]

[1]*Lancaster University, Lancaster, UK,* [a]*f.pein@lancaster.ac.uk*
[2]*University of Cambridge, Cambridge, UK,* [b]*r.shah@statslab.cam.ac.uk*

Cross-validation is the standard approach for tuning parameter selection in many non-parametric regression problems. However its use is less common in change-point regression, perhaps as its prediction error-based criterion may appear to permit small spurious changes and hence be less well-suited to estimation of the number and location of change-points. We show that in fact the problems of cross-validation with squared error loss are more severe and can lead to systematic under- or over-estimation of the number of change-points, and highly suboptimal estimation of the mean function in simple settings where changes are easily detectable. We propose two simple approaches to remedy these issues, the first involving the use of absolute error rather than squared error loss, and the second involving modifying the holdout sets used. For the latter, we provide conditions that permit consistent estimation of the number of change-points for a general change-point estimation procedure. We show these conditions are satisfied for least squares estimation using new results on its performance when supplied with the incorrect number of change-points. Numerical experiments show that our new approaches are competitive with common change-point methods using classical tuning parameter choices when error distributions are well-specified, but can substantially outperform these in misspecified models. An implementation of our methodology is available in the R package `crossvalidationCP` on CRAN.

Keywords: Change-point regression; cross-validation; segment neighbourhood; sample-splitting; selection consistency; tuning parameter selection

References

Antoch, J., Hušková, M. and Veraverbeke, N. (1995). Change-point problem and bootstrap. *J. Nonparametr. Stat.* **5** 123–144. MR1346891 https://doi.org/10.1080/10485259508832639

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Stat. Surv.* **4** 40–79. MR2602303 https://doi.org/10.1214/09-SS054

Arlot, S. and Celisse, A. (2011). Segmentation of the mean of heteroscedastic data via cross-validation. *Stat. Comput.* **21** 613–632. MR2826696 https://doi.org/10.1007/s11222-010-9196-x

Auger, I.E. and Lawrence, C.E. (1989). Algorithms for the optimal identification of segment neighborhoods. *Bull. Math. Biol.* **51** 39–54. MR0978902 https://doi.org/10.1016/S0092-8240(89)80047-3

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *J. Appl. Econometrics* **18** 1–22.

Chan, H.P. and Walther, G. (2013). Detection with the scan and the average likelihood ratio. *Statist. Sinica* **23** 409–428. MR3076173

Chetverikov, D., Liao, Z. and Chernozhukov, V. (2021). On cross-validated Lasso in high dimensions. *Ann. Statist.* **49** 1300–1317. MR4298865 https://doi.org/10.1214/20-aos2000

D'Angelo, M.F.S.V., Palhares, R.M., Takahashi, R.H.C. and Loschi, R.H. (2011). Fuzzy/Bayesian change point detection approach to incipient fault detection. *IET Control Theory Appl.* **5** 539–551. MR2858289 https://doi.org/10.1049/iet-cta.2009.0033

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

1350-7265

Du, C., Kao, C.-L.M. and Kou, S.C. (2016). Stepwise signal extraction via marginal likelihood. *J. Amer. Statist. Assoc.* **111** 314–330. MR3494662 https://doi.org/10.1080/01621459.2015.1006365

Dümbgen, L. and Spokoiny, V.G. (2001). Multiscale testing of qualitative hypotheses. *Ann. Statist.* **29** 124–152. MR1833961 https://doi.org/10.1214/aos/996986504

Fearnhead, P. (2006). Exact and efficient Bayesian inference for multiple changepoint problems. *Stat. Comput.* **16** 203–213. MR2227396 https://doi.org/10.1007/s11222-006-8450-8

Fearnhead, P. and Rigaill, G. (2019). Changepoint detection in the presence of outliers. *J. Amer. Statist. Assoc.* **114** 169–183. MR3941246 https://doi.org/10.1080/01621459.2017.1385466

Fearnhead, P. and Rigaill, G. (2020). Relating and comparing methods for detecting changes in mean. *Stat* **9** e291, 11. MR4156473 https://doi.org/10.1002/sta4.291

Frick, K., Munk, A. and Sieling, H. (2014). Multiscale change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 495–580. With 32 discussions by 47 authors and a rejoinder by the authors. MR3210728 https://doi.org/10.1111/rssb.12047

Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *Ann. Statist.* **42** 2243–2281. MR3269979 https://doi.org/10.1214/14-AOS1245

Fryzlewicz, P. (2020). Detecting possibly frequent change-points: Wild Binary Segmentation 2 and steepest-drop model selection. *J. Korean Statist. Soc.* **49** 1027–1070. MR4183039 https://doi.org/10.1007/s42952-020-00060-x

Garreau, D. and Arlot, S. (2018). Consistent change-point detection with kernels. *Electron. J. Stat.* **12** 4440–4486. MR3892345 https://doi.org/10.1214/18-ejs1513

Harchaoui, Z., Vallet, F., Lung-Yut-Fong, A. and Cappé, O. (2009). A regularized kernel-based approach to unsupervised audio segmentation. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing* 1665–1668. IEEE.

Hušková, M. and Kirch, C. (2008). Bootstrapping confidence intervals for the change-point of time series. *J. Time Series Anal.* **29** 947–972. MR2464948 https://doi.org/10.1111/j.1467-9892.2008.00589.x

Jackson, B., Scargle, J.D., Barnes, D., Arabhi, S., Alt, A., Gioumousis, P., Gwin, E., Sangtrakulcharoen, P., Tan, L. and Tsai, T.T. (2005). An algorithm for optimal partitioning of data on an interval. *IEEE Signal Process. Lett.* **12** 105–108.

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

Kim, C.-J., Morley, J.C. and Nelson, C.R. (2005). The structural break in the equity premium. *J. Bus. Econom. Statist.* **23** 181–191. MR2157269 https://doi.org/10.1198/073500104000000352

Kovács, S., Bühlmann, P., Li, H. and Munk, A. (2023). Seeded binary segmentation: A general methodology for fast and optimal changepoint detection. *Biometrika* **110** 249–256. MR4565454 https://doi.org/10.1093/biomet/asac052

Li, H., Guo, Q. and Munk, A. (2019). Multiscale change-point segmentation: Beyond step functions. *Electron. J. Stat.* **13** 3254–3296. MR4010980 https://doi.org/10.1214/19-ejs1608

Li, H., Munk, A. and Sieling, H. (2016). FDR-control in multiscale change-point segmentation. *Electron. J. Stat.* **10** 918–959. MR3486421 https://doi.org/10.1214/16-EJS1131

Liehrmann, A. and Rigaill, G. (2023). Ms. FPOP: An Exact and Fast Segmentation Algorithm With a Multiscale Penalty. arXiv preprint. Available at arXiv:2303.08723.

Lin, K., Sharpnack, J., Rinaldo, A. and Tibshirani, R.J. (2016). Approximate Recovery in Changepoint Problems, from l_2 Estimation Error Rates. arXiv preprint. Available at arXiv:1606.06746.

Maidstone, R., Hocking, T., Rigaill, G. and Fearnhead, P. (2017). On optimal multiple changepoint algorithms for large data. *Stat. Comput.* **27** 519–533. MR3599687 https://doi.org/10.1007/s11222-016-9636-3

Niu, Y.S., Hao, N. and Zhang, H. (2016). Multiple change-point detection: A selective overview. *Statist. Sci.* **31** 611–623. MR3598742 https://doi.org/10.1214/16-STS587

Olshen, A.B., Venkatraman, E.S., Lucito, R. and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5** 557–572. https://doi.org/10.1093/biostatistics/kxh008

Pein, F., Eltzner, B. and Munk, A. (2021). Analysis of patchclamp recordings: Model-free multiscale methods and software. *Eur. Biophys. J.* **50** 187–209. https://doi.org/10.1007/s00249-021-01506-8

Pein, F. and Shah, R.D. (2024). Supplement to "Cross-validation for change-point regression: pitfalls and solutions." https://doi.org/10.3150/24-BEJ1732SUPP

Pein, F., Sieling, H. and Munk, A. (2017). Heterogeneous change point inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1207–1227. MR3689315 https://doi.org/10.1111/rssb.12202

Rigaill, G. (2015). A pruned dynamic programming algorithm to recover the best segmentations with 1 to $K_{\max}$ change-points. *J. SFdS* **156** 180–205. MR3436653

Sharipov, O., Tewes, J. and Wendler, M. (2016). Sequential block bootstrap in a Hilbert space with application to change point analysis. *Canad. J. Statist.* **44** 300–322. MR3536199 https://doi.org/10.1002/cjs.11293

Truong, C., Oudre, L. and Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Process.* **167** 107299.

Verzelen, N., Fromont, M., Lerasle, M. and Reynaud-Bouret, P. (2023). Optimal change-point detection and localization. *Ann. Statist.* **51** 1586–1610. MR4658569 https://doi.org/10.1214/23-aos2297

Vostrikova, L.J. (1981). Discovery of "discord" in multidimensional random processes. *Dokl. Akad. Nauk SSSR* **259** 270–274. MR0625215

Wong, W.H. (1983). On the consistency of cross-validation in kernel nonparametric regression. *Ann. Statist.* **11** 1136–1141. MR0720259 https://doi.org/10.1214/aos/1176346327

Yang, Y. (2007). Consistency of cross validation for comparing regression procedures. *Ann. Statist.* **35** 2450–2473. MR2382654 https://doi.org/10.1214/009053607000000514

Yao, Y.-C. (1988). Estimating the number of change-points via Schwarz' criterion. *Statist. Probab. Lett.* **6** 181–189. MR0919373 https://doi.org/10.1016/0167-7152(88)90118-6

Yao, Y.-C. and Au, S.T. (1989). Least-squares estimation of a step function. *Sankhyā Ser. A* **51** 370–381. MR1175613

Yu, Y. and Feng, Y. (2014). Modified cross-validation for penalized high-dimensional linear regression models. *J. Comput. Graph. Statist.* **23** 1009–1027. MR3270708 https://doi.org/10.1080/10618600.2013.849200

Zhang, N.R. and Siegmund, D.O. (2007). A modified Bayes information criterion with applications to the analysis of comparative genomic hybridization data. *Biometrics* **63** 22–32. MR2345571 https://doi.org/10.1111/j.1541-0420.2006.00662.x

Zou, C., Wang, G. and Li, R. (2020). Consistent selection of the number of change-points via sample-splitting. *Ann. Statist.* **48** 413–439. MR4065168 https://doi.org/10.1214/19-AOS1814

Bernoulli **31**(1), 2025, 412–431
https://doi.org/10.3150/24-BEJ1733

Yule's "nonsense correlation": Moments and density

PHILIP A. ERNST[1,a], L.C.G. ROGERS[2,b] and QUAN ZHOU[3,c]

[1]*Department of Mathematics, Imperial College London, London, UK,* [a]*p.ernst@imperial.ac.uk*
[2]*Department of Mathematics, University of Cambridge, Cambridge, UK,* [b]*chris@statslab.cam.ac.uk*
[3]*Department of Statistics, Texas A&M University, College Station, TX, USA,* [c]*quan@stat.tamu.edu*

In 1926, G. Udny Yule (*J. R. Stat. Soc.* **89** (1926) 1–63) considered the following problem: given a sequence of pairs of random variables $\{X_k, Y_k\}$ $(k = 1, 2, \dots, n)$, and letting $X_i = S_i$ and $Y_i = S'_i$ where S_i and S'_i are the partial sums of two independent random walks, what is the distribution of the empirical correlation coefficient

$$\rho_n = \frac{\sum_{i=1}^n S_i S'_i - \frac{1}{n}(\sum_{i=1}^n S_i)(\sum_{i=1}^n S'_i)}{\sqrt{\sum_{i=1}^n S_i^2 - \frac{1}{n}(\sum_{i=1}^n S_i)^2}\sqrt{\sum_{i=1}^n (S'_i)^2 - \frac{1}{n}(\sum_{i=1}^n S'_i)^2}}?$$

Yule empirically observed the distribution of this statistic to be heavily dispersed and frequently large in absolute value, leading him to call it "nonsense correlation." This unexpected finding led to his formulation of two concrete questions, each of which would remain open for more than ninety years: (i) Find (analytically) the variance of ρ_n as $n \to \infty$ and (ii): Find (analytically) the higher order moments and the density of ρ_n as $n \to \infty$. In 2017, Ernst, Shepp and Wyner (*Ann. Statist.* **45** (2017) 1789–1809) considered the empirical correlation coefficient

$$\rho := \frac{\int_0^1 W_1(t)W_2(t)dt - \int_0^1 W_1(t)dt \int_0^1 W_2(t)dt}{\sqrt{\int_0^1 W_1^2(t)dt - \left(\int_0^1 W_1(t)dt\right)^2}\sqrt{\int_0^1 W_2^2(t)dt - \left(\int_0^1 W_2(t)dt\right)^2}}$$

of two *independent* Wiener processes W_1, W_2, the limit to which ρ_n converges weakly, as was first shown by P.C.B. Phillips (*J. Econometrics* **33** (1986) 311–340). Using tools from integral equation theory, Ernst, Shepp and Wyner (*Ann. Statist.* **45** (2017) 1789–1809) closed question (i) by explicitly calculating the second moment of ρ to be .240522. This paper adopts a completely different approach to the same question, rooted in an earlier literature on the laws of quadratic functionals of Gaussian diffusions (in particular, (*Adv. in Appl. Probab.* **25** (1993) 570–584; *Stoch. Stoch. Rep.* **41** (1992) 201–218)). This allows us to develop an Itô-formula approach from which we calculate expressions for the Laplace transform of ρ, leading to expressions for the moments which we evaluate up to order 16, thereby closing question (ii). This leads, for the first time, to an approximation to the density of Yule's nonsense correlation. The broad applicability of this approach is demonstrated by answering the corresponding questions when the pair of independent Brownian motions is replaced by a pair of correlated Brownian motions, or by two independent Ornstein-Uhlenbeck processes, or by two independent Brownian bridges. We conclude by extending the definition of ρ to the time interval $[0, T]$ for any $T > 0$ and prove a Central Limit Theorem for the case of two independent Ornstein-Uhlenbeck processes.

Keywords: Nonsense correlation; Ornstein-Uhlenbeck processes; volatile correlation; Wiener processes

References

[1] Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statist. Sci.* **10** 364–376.
[2] Billingsley, P. (2008). *Probability and Measure*. New York: Wiley.
[3] Burden, R.L. and Faires, J.D. (1997). Numerical Analysis. *Cole, Belmont*.

1350-7265

[4] Chan, T. (1991). Indefinite quadratic functionals of Gaussian processes and least-action paths. *Ann. Inst. Henri Poincaré Probab. Stat.* **27** 239–271. MR1118937

[5] Chan, T., Dean, D.S., Jansons, K.M. and Rogers, L.C.G. (1994). On polymer conformations in elongational flows. *Comm. Math. Phys.* **160** 239–257. MR1262196

[6] Cressie, N., Davis, A.S., Folks, J.L. and Policello, G.E. II (1981). The moment-generating function and negative integer moments. *Amer. Statist.* **35** 148–150. MR0632425 https://doi.org/10.2307/2683982

[7] Donati-Martin, C. and Yor, M. (1993). On some examples of quadratic functionals of Brownian motion. *Adv. in Appl. Probab.* **25** 570–584. MR1234297 https://doi.org/10.2307/1427524

[8] Donati-Martin, C. and Yor, M. (1997). Some Brownian functionals and their laws. *Ann. Probab.* **25** 1011–1058. MR1457611 https://doi.org/10.1214/aop/1024404505

[9] Douissi, S., Es-Sebaiy, K. and Viens, F. (2022). Asymptotics of Yule's nonsense correlation for Ornstein-Uhlenbeck paths: A Wiener chaos approach. *Electron. J. Stat.* **16** 3176–3211. MR4420439 https://doi.org/10.1214/22-ejs2021

[10] Dynkin, E.B. (1980). Markov processes and random fields. *Bull. Amer. Math. Soc. (N.S.)* **3** 975–999. MR0585179 https://doi.org/10.1090/S0273-0979-1980-14831-4

[11] Ernst, P.A., Rogers, L.C.G. and Zhou, Q. (2019). The distribution of Yule's "nonsense correlation". ArXiv preprint. Available at arXiv:1909.02546.

[12] Ernst, P.A., Shepp, L.A. and Wyner, A.J. (2017). Yule's "nonsense correlation" solved! *Ann. Statist.* **45** 1789–1809. MR3670196 https://doi.org/10.1214/16-AOS1509

[13] Fixman, M. (1962). Radius of gyration of polymer chains. *J. Chem. Phys.* **36** 306–310.

[14] Granger, C. and Newbold, D. (1974). Spurious regression in econometrics. *J. Econometrics* **2** 111–120.

[15] Hendry, D.F. (1986). Economic modelling with cointegrated variables: An overview. *Oxf. Bull. Econ. Stat.* **48** 201–212.

[16] Jones, M.C. (1987). Inverse factorial moments. *Statist. Probab. Lett.* **6** 37–42. MR0907257 https://doi.org/10.1016/0167-7152(87)90056-3

[17] Lévy, P. (1951). Wiener's random function, and other Laplacian random functions. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* **2** 171–188. The Regents of the Univ. California.

[18] Mac Aonghusa, P. and Pulé, J.V. (1989). An extension of Lévy's stochastic area formula. *Stoch. Stoch. Rep.* **26** 247–255. MR1056073 https://doi.org/10.1080/17442508908833560

[19] Phillips, P.C.B. (1986). Understanding spurious regressions in econometrics. *J. Econometrics* **33** 311–340. MR0867979 https://doi.org/10.1016/0304-4076(86)90001-1

[20] Phillips, P.C.B. (1998). New tools for understanding spurious regressions. *Econometrica* **66** 1299–1325. MR1654311 https://doi.org/10.2307/2999618

[21] Provost, S.B. (2005). Moment-based density approximants. *Math. J.* **9** 727–756.

[22] Rogers, L.C.G. and Shi, Z. (1992). Quadratic functionals of Brownian motion, optimal control, and the "Colditz" example. *Stoch. Stoch. Rep.* **41** 201–218. MR1275583 https://doi.org/10.1080/17442509208833803

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

[24] Sawa, T. (1972). Finite-sample properties of the k-class estimators. *Econometrica* **40** 653–680. MR0329179 https://doi.org/10.2307/1912960

[25] Saxena, R.B. (1967). Expansion of continuous differentiable functions in Fourier Legendre series. *Canad. J. Math.* **19** 823–827. MR0212489 https://doi.org/10.4153/CJM-1967-076-1

[26] Suetin, P.K. (1964). On the representation of continuous and differentiable functions by Fourier series in Legendre polynomials. *Dokl. Akad. Nauk SSSR* **158** 1275–1277. MR0167778

[27] Wang, H. (2021). How much faster does the best polynomial approximation converge than Legendre projection? *Numer. Math.* **147** 481–503. MR4215344 https://doi.org/10.1007/s00211-021-01173-z

[28] Wang, H. and Xiang, S. (2012). On the convergence rates of Legendre approximation. *Math. Comp.* **81** 861–877. MR2869040 https://doi.org/10.1090/S0025-5718-2011-02549-4

[29] Yule, G.U. (1926). Why do we sometimes get nonsense-correlations between time-series? – A study in sampling and the nature of time-series. *J. R. Stat. Soc.* **89** 1–63.

Bernoulli **31**(1), 2025, 432–456
https://doi.org/10.3150/24-BEJ1734

Fréchet mean set estimation in the Hausdorff metric, via relaxation

MOÏSE BLANCHARD[1,a] and ADAM QUINN JAFFE[2,b]

[1]*Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, USA,* [a]*moiseb@mit.edu*
[2]*Department of Statistics, University of California, Berkeley, CA, USA,* [b]*aqjaffe@berkeley.edu*

This work resolves the following question in non-Euclidean statistics: Is it possible to consistently estimate the Fréchet mean set of an unknown population distribution, with respect to the Hausdorff metric, when given access to independent identically-distributed samples? Our affirmative answer is based on a careful analysis of the "relaxed empirical Fréchet mean set estimators" which identify the set of near-minimizers of the empirical Fréchet functional and where the amount of "relaxation" vanishes as the number of data tends to infinity. On the theoretical side, our results include exact descriptions of which relaxation rates give weak consistency and which give strong consistency, as well as a description of an estimator which (assuming only the finiteness of certain moments and a mild condition on the metric entropy of the underlying metric space) adaptively finds the fastest possible relaxation rate for strongly consistent estimation. On the applied side, we consider the problem of estimating the set of Fermat-Weber points of an unknown distribution in the space of equidistant trees endowed with the tropical projective metric; in this setting, we provide an algorithm that provably implements our adaptive estimator, and we apply this method to real phylogenetic data.

Keywords: Computational phylogenetics; Fermat-Weber point; Fréchet mean; Hausdorff metric; medoids; non-Euclidean statistics; random sets; stochastic optimization

References

[1] Afsari, B. (2011). Riemannian L^p center of mass: Existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.* **139** 655–673. MR2736346 https://doi.org/10.1090/S0002-9939-2010-10541-5
[2] Bačák, M. (2014). Computing medians and means in Hadamard spaces. *SIAM J. Optim.* **24** 1542–1566. MR3264572 https://doi.org/10.1137/140953393
[3] Bagaria, V., Kamath, G., Ntranos, V., Zhang, M. and Tse, D. (2018). Medoids in almost-linear time via multi-armed bandits (A. Storkey and F. Perez-Cruz, eds.). *Proceedings of Machine Learning Research* **84** 500–509. Playa Blanca, Lanzarote, Canary Islands: PMLR.
[4] Baharav, T. and Tse, D. (2019). Ultra fast medoid identification via correlated sequential halving. In *Advances in Neural Information Processing Systems* 3655–3664.
[5] Barnhill, D. and Yoshida, R. (2023). Clustering methods over the tropical projective torus. *Mathematics* **11** 3433.
[6] Bhattacharya, A. (2008). Statistical analysis on manifolds: A nonparametric approach for inference on shape spaces. *Sankhyā* **70** 223–266. MR2551814
[7] Bhattacharya, A. and Bhattacharya, R. (2012). *Nonparametric Inference on Manifolds: With Applications to Shape Spaces*. *Institute of Mathematical Statistics (IMS) Monographs* **2**. Cambridge: Cambridge Univ. Press. MR2934285 https://doi.org/10.1017/CBO9781139094764
[8] Bhattacharya, R. and Lin, L. (2017). Omnibus CLTs for Fréchet means and nonparametric inference on non-Euclidean spaces. *Proc. Amer. Math. Soc.* **145** 413–428. MR3565392 https://doi.org/10.1090/proc/13216
[9] Bhattacharya, R. and Patrangenaru, V. (2003). Large sample theory of intrinsic and extrinsic sample means on manifolds. I. *Ann. Statist.* **31** 1–29. MR1962498 https://doi.org/10.1214/aos/1046294456
[10] Bhattacharya, R. and Patrangenaru, V. (2005). Large sample theory of intrinsic and extrinsic sample means on manifolds. II. *Ann. Statist.* **33** 1225–1259. MR2195634 https://doi.org/10.1214/009053605000000093

1350-7265

[11] Billera, L.J., Holmes, S.P. and Vogtmann, K. (2001). Geometry of the space of phylogenetic trees. *Adv. in Appl. Math.* **27** 733–767. MR1867931 https://doi.org/10.1006/aama.2001.0759
[12] Blanchard, M. and Jaffe, A.Q. (2024). Supplement to "Fréchet mean set estimation in the Hausdorff metric, via relaxation." https://doi.org/10.3150/24-BEJ1734SUPP
[13] Cao, Y. and Monod, A. (2022). A geometric condition for uniqueness of Fréchet means of persistence diagrams.
[14] Cazals, F., Delmas, B. and O'Donnell, T. (2021). Fréchet mean and p-mean on the unit circle: Decidability, algorithm, and applications to clustering on the flat torus. In *19th International Symposium on Experimental Algorithms*. *LIPIcs. Leibniz Int. Proc. Inform.* **190** 15. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4288694
[15] Chakraborty, R. and Vemuri, B.C. (2015). Recursive Fréchet mean computation on the Grassmannian and its applications to computer vision. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
[16] Chakraborty, R. and Vemuri, B.C. (2019). Statistics on the Stiefel manifold: Theory and applications. *Ann. Statist.* **47** 415–438. MR3909937 https://doi.org/10.1214/18-AOS1692
[17] Dryden, I.L., Koloydenko, A. and Zhou, D. (2009). Non-Euclidean statistics for covariance matrices, with applications to diffusion tensor imaging. *Ann. Appl. Stat.* **3** 1102–1123. MR2750388 https://doi.org/10.1214/09-AOAS249
[18] Dudley, R.M. (2014). *Uniform Central Limit Theorems*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **142**. New York: Cambridge Univ. Press. MR3445285
[19] Eichfelder, G., Hotz, T. and Wieditz, J. (2019). An algorithm for computing Fréchet means on the sphere. *Optim. Lett.* **13** 1523–1533. MR4002314 https://doi.org/10.1007/s11590-019-01415-y
[20] Evans, S.N. and Jaffe, A.Q. (2024). Limit theorems for Fréchet mean sets. *Bernoulli* **30** 419–447. MR4665584 https://doi.org/10.3150/23-bej1603
[21] Feller, W. (1968). An extention of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* **18** 343–355.
[22] Feragen, A., Hauberg, S., Nielsen, M. and Lauze, F. (2011). Means in spaces of tree-like shapes. In *2011 International Conference on Computer Vision* 736–746. IEEE.
[23] Ferguson, D. and Meyer, F.G. (2022). Computation of the sample Fréchet mean for sets of large graphs with applications to regression. *SIAM Int. Conf. Data Min. (SDM 2022)*.
[24] Ferguson, D. and Meyer, F.G. (2023). Theoretical analysis and computation of the sample Fréchet mean of sets of large graphs for various metrics. *Inf. Inference* **12** iaad002. MR4581448 https://doi.org/10.1093/imaiai/iaad002
[25] Fiori, S. (2009). Learning the Fréchet mean over the manifold of symmetric positive-definite matrices. *Cogn. Comput.* **1** 279–291.
[26] Ginestet, C.E. (2012). Strong consistency of Fréchet sample mean sets for graph-valued random variables.
[27] Huckemann, S.F. (2011). Intrinsic inference on the mean geodesic of planar shapes and tree discrimination by leaf growth. *Ann. Statist.* **39** 1098–1124. MR2816349 https://doi.org/10.1214/10-AOS862
[28] Huckemann, S.F. (2015). (Semi-)Intrinsic Statistical Analysis on Non-Euclidean Spaces. In *Advances in Complex Data Modeling and Computational Methods in Statistics* 103–118. Cham, Switzerland: Springer.
[29] Jain, B.J. and Obermayer, K. (2012). Learning in Riemannian orbifolds.
[30] Kang, Q., Lin, B., Monod, A. and Yoshida, R. (2018). Tropical geometry of phylogenetic tree space: A statistical perspective.
[31] Kolaczyk, E.D., Lin, L., Rosenberg, S., Walters, J. and Xu, J. (2020). Averages of unlabeled networks: Geometric characterization and asymptotic behavior. *Ann. Statist.* **48** 514–538. MR4065172 https://doi.org/10.1214/19-AOS1820
[32] Krakowski, K., Hüper, K. and Manton, J. (2007). On the computation of the Karcher mean on spheres and special orthogonal groups. In *RoboMat 2007: Workshop on Robotics and Mathematics, Coimbra, Portugal, 17th – 19th September, 2007*. Coimbra, Portugal: Centro Internacional de Matematica.
[33] Le, H. and Kume, A. (2000). The Fréchet mean shape and the shape of the means. *Adv. in Appl. Probab.* **32** 101–113. MR1765168 https://doi.org/10.1239/aap/1013540025
[34] Lin, B., Monod, A. and Yoshida, R. (2022). Tropical geometric variation of tree shapes. *Discrete Comput. Geom.* **68** 817–849. MR4481323 https://doi.org/10.1007/s00454-022-00410-y

[35] Lin, B. and Yoshida, R. (2018). Tropical Fermat-Weber points. *SIAM J. Discrete Math.* **32** 1229–1245. MR3810501 https://doi.org/10.1137/16M1071122
[36] Maclagan, D. and Sturmfels, B. (2021). *Introduction to Tropical Geometry* **161**. Providence: Amer. Math. Soc.
[37] Miolane, N., Guigui, N., Le Brigant, A. et al. (2020). Geomstats: A Python package for Riemannian geometry in machine learning. *J. Mach. Learn. Res.* **21** 223. MR4209509 https://doi.org/10.1080/14697688.2020.1825781
[38] Rockafellar, R.T. and Wets, R.J.-B. (1998). *Variational Analysis*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **317**. Berlin: Springer. MR1491362 https://doi.org/10.1007/978-3-642-02431-3
[39] Schötz, C. (2022). Strong laws of large numbers for generalizations of Fréchet mean sets. *Statistics* **56** 34–52. MR4392146 https://doi.org/10.1080/02331888.2022.2032063
[40] Skwerer, S., Provan, S. and Marron, J.S. (2018). Relative optimality conditions and algorithms for treespace Fréchet means. *SIAM J. Optim.* **28** 959–988. MR3782405 https://doi.org/10.1137/15M1050914
[41] Sverdrup-Thygeson, H. (1981). Strong law of large numbers for measures of central tendency and dispersion of random variables in compact metric spaces. *Ann. Statist.* **9** 141–145. MR0600540
[42] Turner, K., Mileyko, Y., Mukherjee, S. and Harer, J. (2014). Fréchet means for distributions of persistence diagrams. *Discrete Comput. Geom.* **52** 44–70. MR3231030 https://doi.org/10.1007/s00454-014-9604-7
[43] Yoshida, R. (2021). Tropical balls and its applications to K nearest neighbor over the space of phylogenetic trees. *Mathematics* **9** 779.
[44] Ziezold, H. (1977). On expected figures and a strong law of large numbers for random elements in quasi-metric spaces. In *Transactions of the Seventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes and of the Eighth European Meeting of Statisticians (Tech. Univ. Prague, Prague, 1974), Vol. A* 591–602. Dordrecht-Boston, MA: Reidel. MR0501230

Bernoulli **31**(1), 2025, 457–483
https://doi.org/10.3150/24-BEJ1735

Inference in balanced community modulated recursive trees

ANNA BEN-HAMOU[1,a] and VASILIKI VELONA[2,b]

[1]*LPSM, Sorbonne University, Paris, France,* [a]*anna.ben_hamou@sorbonne-universite.fr*
[2]*Einstein Institute of Mathematics, Hebrew University of Jerusalem, Israel,* [b]*vasiliki.velona@mail.huji.ac.il*

We introduce a random recursive tree model with two communities, called *balanced community modulated random recursive tree*, or BCMRT in short. In this setting, pairs of nodes of different type appear sequentially. Each node of the pair decides independently to attach to their own type with probability $1-q$, or to the other type with probability q, and then chooses its parent uniformly within the set of existing nodes with the selected type. We find that the limiting degree distributions coincide for different q. Therefore, as far as inference is concerned, other statistics have to be studied. We first consider the setting where the time-labels of the nodes, i.e., their time of arrival, are observed but their type is not. In this setting, we design a consistent estimator for q and provide bounds for the feasibility of testing between two different values of q. Moreover, we show that if q is small enough, then it is possible to cluster the nodes in a way correlated with the true partition, even though the algorithm is exponential in time (in passing, we show that our clustering procedure is intimately connected to the NP-hard problem of minimum fair bisection). In the unlabelled setting, i.e., when only the tree structure is observed, we show that it is possible to test between different values of q in a strictly better way than by random guessing. This follows from a delicate analysis of the sum-of-distances statistic.

Keywords: Clustering; combinatorial statistics; community detection; community modulated recursive trees; minimum fair bisection; parameter testing; random recursive trees; Wiener index

References

[1] Abbe, E. and Sandon, C. (2015). Community detection in general stochastic block models: Fundamental limits and efficient algorithms for recovery. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science—FOCS 2015* 670–688. Los Alamitos, CA: IEEE Computer Soc. MR3473334 https://doi.org/10.1109/FOCS.2015.47
[2] Addario-Berry, L., Devroye, L., Lugosi, G. and Velona, V. (2022). Broadcasting on random recursive trees. *Ann. Appl. Probab.* **32** 497–528. MR4386534 https://doi.org/10.1214/21-aap1686
[3] Banerjee, S. and Bhamidi, S. (2022). Root finding algorithms and persistence of Jordan centrality in growing random trees. *Ann. Appl. Probab.* **32** 2180–2210. MR4430011 https://doi.org/10.1214/21-aap1731
[4] Banerjee, S. and Huang, X. (2023). Degree centrality and root finding in growing random networks. *Electron. J. Probab.* **28** Paper No. 42. MR4559329 https://doi.org/10.1214/23-ejp930
[5] Banks, J., Moore, C., Neeman, J. and Netrapalli, P. (2016). Information-theoretic thresholds for community detection in sparse networks. In *Conf. Learning Theory* 383–416.
[6] Ben-Hamou, A. and Velona, V. (2024). Supplement to "Inference in balanced community modulated recursive trees." https://doi.org/10.3150/24-BEJ1735SUPP
[7] Bhamidi, S., Fan, R., Fraiman, N. and Nobel, A. (2022). Community modulated recursive trees and population dependent branching processes. *Random Structures Algorithms* **60** 201–232. MR4359934 https://doi.org/10.1002/rsa.21027
[8] Bonsma, P., Epping, T. and Hochstättler, W. (2006). Complexity results on restricted instances of a paint shop problem for words. *Discrete Appl. Math.* **154** 1335–1343. MR2221553 https://doi.org/10.1016/j.dam.2005.05.033

1350-7265

[9] Briend, S., Calvillo, F. and Lugosi, G. (2023). Archaeology of random recursive dags and Cooper–Frieze random networks. *Combin. Probab. Comput.* **32** 859–873. MR4653727 https://doi.org/10.1017/s0963548323000184
[10] Bubeck, S., Devroye, L. and Lugosi, G. (2017). Finding Adam in random growing trees. *Random Structures Algorithms* **50** 158–172. MR3607120 https://doi.org/10.1002/rsa.20649
[11] Bubeck, S., Eldan, R., Mossel, E. and Rácz, M.Z. (2017). From trees to seeds: On the inference of the seed from large trees in the uniform attachment model. *Bernoulli* **23** 2887–2916. MR3648049 https://doi.org/10.3150/16-BEJ831
[12] Bubeck, S., Mossel, E. and Rácz, M.Z. (2015). On the influence of the seed graph in the preferential attachment model. *IEEE Trans. Netw. Sci. Eng.* **2** 30–39. MR3361606 https://doi.org/10.1109/TNSE.2015.2397592
[13] Casel, K., Friedrich, T., Schirneck, M. and Wietheger, S. (2023). Fair correlation clustering in forests. In *4th Symposium on Foundations of Responsible Computing* (K. Talwar, ed.). *LIPIcs. Leibniz Int. Proc. Inform.* **256** 9:1–9:12. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4603967 https://doi.org/10.4230/lipics.forc.2023.9
[14] Casse, J. (2023). Siblings in d-dimensional nearest neighbour trees. ArXiv preprint. Available at arXiv:2302.10795.
[15] Crane, H. and Xu, M. (2021). Inference on the history of a randomly growing tree. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 639–668. MR4319997 https://doi.org/10.1111/rssb.12428
[16] Crane, H. and Xu, M. (2023). Root and community inference on latent network growth processes using noisy attachment models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*
[17] Curien, N., Duquesne, T., Kortchemski, I. and Manolescu, I. (2015). Scaling limits and influence of the seed graph in preferential attachment trees. *J. Éc. Polytech. Math.* **2** 1–34. MR3326003 https://doi.org/10.5802/jep.15
[18] Devroye, L. and Reddad, T. (2019). On the discovery of the seed in uniform attachment trees. *Internet Math.* 1–29. MR3934290
[19] Drmota, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Vienna: Springer. MR2484382 https://doi.org/10.1007/978-3-211-75357-6
[20] Efron, B. and Stein, C. (1981). The jackknife estimate of variance. *Ann. Statist.* **9** 586–596. MR0615434
[21] Gao, F., van der Vaart, A., Castro, R. and van der Hofstad, R. (2017). Consistent estimation in general sublinear preferential attachment trees. *Electron. J. Stat.* **11** 3979–3999. MR3714305 https://doi.org/10.1214/17-EJS1356
[22] Garey, M.R., Johnson, D.S. and Stockmeyer, L. (1974). Some simplified NP-complete problems. In *Sixth Annual ACM Symposium on Theory of Computing (Seattle, Wash., 1974)* 47–63. New York: Association for Computing Machinery. MR0418510
[23] Hajek, B. and Sankagiri, S. (2019). Community recovery in a preferential attachment graph. *IEEE Trans. Inf. Theory* **65** 6853–6874. MR4030863 https://doi.org/10.1109/TIT.2019.2927624
[24] Janson, S. (2005). Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms* **26** 69–83. MR2116576 https://doi.org/10.1002/rsa.20046
[25] Kesten, H. and Stigum, B.P. (1966). Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Stat.* **37** 1463–1481. MR0200979 https://doi.org/10.1214/aoms/1177699139
[26] Khim, J. and Loh, P.-L. (2017). Confidence sets for the source of a diffusion in regular trees. *IEEE Trans. Netw. Sci. Eng.* **4** 27–40. MR3625953 https://doi.org/10.1109/TNSE.2016.2627502
[27] Levin, D.A. and Peres, Y. (2017). *Markov Chains and Mixing Times*. Providence, RI: Amer. Math. Soc. MR3726904 https://doi.org/10.1090/mbk/107
[28] Lugosi, G. (2017). Lectures on combinatorial statistics. *47th Probability Summer School, Saint-Flour* 1–91.
[29] Lugosi, G. and Pereira, A.S. (2019). Finding the seed of uniform attachment trees. *Electron. J. Probab.* **24** Paper No. 18. MR3925458 https://doi.org/10.1214/19-EJP268
[30] Massoulié, L. (2014). Community detection thresholds and the weak Ramanujan property. In *STOC'14—Proceedings of the 2014 ACM Symposium on Theory of Computing* 694–703. New York: ACM. MR3238997
[31] McDiarmid, C. (1989). On the method of bounded differences. In *Surveys in Combinatorics, 1989 (Norwich, 1989). London Mathematical Society Lecture Note Series* **141** 148–188. Cambridge: Cambridge Univ. Press. MR1036755

[32] Moore, C. (2017). The computer science and physics of community detection: Landscapes, phase transitions, and hardness. *Bull. Eur. Assoc. Theor. Comput. Sci. EATCS* **121** 26–61. MR3699594
[33] Mossel, E., Neeman, J. and Sly, A. (2015). Reconstruction and estimation in the planted partition model. *Probab. Theory Related Fields* **162** 431–461. MR3383334 https://doi.org/10.1007/s00440-014-0576-6
[34] Mossel, E., Neeman, J. and Sly, A. (2018). A proof of the block model threshold conjecture. *Combinatorica* **38** 665–708. MR3876880 https://doi.org/10.1007/s00493-016-3238-8
[35] Mossel, E. and Peres, Y. (2003). Information flow on trees. *Ann. Appl. Probab.* **13** 817–844. MR1994038 https://doi.org/10.1214/aoap/1060202828
[36] Na, H.S. and Rapoport, A. (1970). Distribution of nodes of a tree by degree. *Math. Biosci.* **6** 313–329. MR0278985 https://doi.org/10.1016/0025-5564(70)90071-4
[37] Neininger, R. (2002). The Wiener index of random trees. *Combin. Probab. Comput.* **11** 587–597. MR1940122 https://doi.org/10.1017/S0963548302005321
[38] Rácz, M.Z. and Bubeck, S. (2017). Basic models and questions in statistical network analysis. *Stat. Surv.* **11** 1–47. MR3696007 https://doi.org/10.1214/17-SS117
[39] Rácz, M.Z. and Sridhar, A. (2022). Correlated randomly growing graphs. *Ann. Appl. Probab.* **32** 1058–1111. MR4414701 https://doi.org/10.1214/21-aap1703
[40] Steele, J.M. (1986). An Efron-Stein inequality for nonsymmetric statistics. *Ann. Statist.* **14** 753–758. MR0840528 https://doi.org/10.1214/aos/1176349952
[41] Tang, W., Guo, X. and Tang, F. (2020). The Buckley-Osthus model and the block preferential attachment model: Statistical analysis and application. In *Int. Conf. Machine Learning* 9377–9386.
[42] Van Der Hofstad, R. (2009). Random graphs and complex networks. **11** 60. Available on http://www.win.tue.nl/rhofstad/NotesRGCN.pdf.

Bernoulli **31**(1), 2025, 484–508
https://doi.org/10.3150/24-BEJ1736

Estimation of a pure-jump stable Cox-Ingersoll-Ross process

ELISE BAYRAKTAR[a] and EMMANUELLE CLÉMENT[b]

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

We consider a pure-jump stable Cox-Ingersoll-Ross (α-stable CIR) process driven by a non-symmetric stable Lévy process with jump activity $\alpha \in (1,2)$ and we address the joint estimation of drift, scaling and jump activity parameters from high-frequency observations of the process on a fixed time period. We first prove the existence of a consistent, rate optimal and asymptotically conditionally Gaussian estimator based on an approximation of the likelihood function. Moreover, uniqueness of the drift estimators is established assuming that the scaling coefficient and the jump activity are known or consistently estimated. Next we propose easy-to-implement preliminary estimators of all parameters and we improve them by a one-step procedure.

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

References

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

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

[3] Bayraktar, E. and Clément, E. (2024). Supplement to "Estimation of a pure-jump stable Cox-Ingersoll-Ross process." https://doi.org/10.3150/24-BEJ1736SUPP

[4] Brouste, A. and Masuda, H. (2018). Efficient estimation of stable Lévy process with symmetric jumps. *Stat. Inference Stoch. Process.* **21** 289–307. MR3824969 https://doi.org/10.1007/s11203-018-9181-0

[5] Clément, E. and Gloter, A. (2020). Joint estimation for SDE driven by locally stable Lévy processes. *Electron. J. Stat.* **14** 2922–2956. MR4134348 https://doi.org/10.1214/20-EJS1737

[6] Clément, E., Gloter, A. and Nguyen, H. (2019). LAMN property for the drift and volatility parameters of a SDE driven by a stable Lévy process. *ESAIM Probab. Stat.* **23** 136–175. MR3945580 https://doi.org/10.1051/ps/2018007

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

[8] Foucart, C. and Uribe Bravo, G. (2014). Local extinction in continuous-state branching processes with immigration. *Bernoulli* **20** 1819–1844. MR3263091 https://doi.org/10.3150/13-BEJ543

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

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

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

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

[13] Jiao, Y., Ma, C., Scotti, S. and Zhou, C. (2021). The alpha-Heston stochastic volatility model. *Math. Finance* **31** 943–978. MR4273541 https://doi.org/10.1111/mafi.12306
[14] Kawazu, K. and Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Teor. Veroyatn. Primen.* **16** 34–51. MR0290475
[15] Li, Z. (2020). Continuous-state branching processes with immigration. In *From Probability to Finance—Lecture Notes of BICMR Summer School on Financial Mathematics*. *Math. Lect. Peking Univ.* 1–69. Singapore: Springer. MR4339413 https://doi.org/10.1007/978-981-15-1576-7_1
[16] Li, Z. and Ma, C. (2015). Asymptotic properties of estimators in a stable Cox-Ingersoll-Ross model. *Stochastic Process. Appl.* **125** 3196–3233. MR3343292 https://doi.org/10.1016/j.spa.2015.03.002
[17] Li, Z. and Mytnik, L. (2011). Strong solutions for stochastic differential equations with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 1055–1067. MR2884224 https://doi.org/10.1214/10-AIHP389
[18] Long, H. (2010). Parameter estimation for a class of stochastic differential equations driven by small stable noises from discrete observations. *Acta Math. Sci. Ser. B Engl. Ed.* **30** 645–663. MR2675775 https://doi.org/10.1016/S0252-9602(10)60067-7
[19] Masuda, H. (2009). Joint estimation of discretely observed stable Lévy processes with symmetric Lévy density. *J. Japan Statist. Soc.* **39** 49–75. MR2571802 https://doi.org/10.14490/jjss.39.49
[20] 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
[21] Masuda, H. (2023). Optimal stable Ornstein-Uhlenbeck regression. *Jpn. J. Stat. Data Sci.* **6** 573–605. MR4594508 https://doi.org/10.1007/s42081-023-00197-z
[22] Nolan, J.P. (1997). Numerical calculation of stable densities and distribution functions. *Commun. Stat., Stoch. Models* **13** 759–774. Heavy tails and highly volatile phenomena. MR1482292 https://doi.org/10.1080/15326349708807450
[23] 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
[24] Sørensen, M. (1999). On asymptotics of estimating functions. *Braz. J. Probab. Stat.* **13** 111–136. MR1803041
[25] Todorov, V. (2013). Power variation from second order differences for pure jump semimartingales. *Stochastic Process. Appl.* **123** 2829–2850. MR3054547 https://doi.org/10.1016/j.spa.2013.04.005
[26] Todorov, V. and Tauchen, G. (2011). Limit theorems for power variations of pure-jump processes with application to activity estimation. *Ann. Appl. Probab.* **21** 546–588. MR2807966 https://doi.org/10.1214/10-AAP700
[27] Woerner, J.H.C. (2003). Purely discontinuous Lévy processes and power variation: Inference for integrated volatility and the scale parameter. OFRC Working Papers.
[28] Yang, X. (2017). Maximum likelihood type estimation for discretely observed CIR model with small α-stable noises. *Statist. Probab. Lett.* **120** 18–27. MR3567917 https://doi.org/10.1016/j.spl.2016.09.014
[29] Zolotarev, V.M. (1986). *One-Dimensional Stable Distributions*. *Translations of Mathematical Monographs* **65**. Providence, RI: Amer. Math. Soc. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR0854867 https://doi.org/10.1090/mmono/065

Bernoulli **31**(1), 2025, 509–536
https://doi.org/10.3150/24-BEJ1737

Sampling using adaptive regenerative processes

HECTOR MCKIMM[1,a], ANDI WANG[2,b], MURRAY POLLOCK[3,e], CHRISTIAN ROBERT[2,4,c] and GARETH ROBERTS[2,d]

[1]*Department of Mathematics, Imperial College London, London, United Kingdom,* [a]*h.mckimm@imperial.ac.uk*
[2]*Department of Statistics, University of Warwick, Coventry, United Kingdom,* [b]*andi.wang@warwick.ac.uk,* [c]*C.A.M.Robert@warwick.ac.uk,* [d]*Gareth.O.Roberts@warwick.ac.uk*
[3]*School of Mathematics, Newcastle University, Newcastle-upon-Tyne, United Kingdom,* [e]*murray.pollock@newcastle.ac.uk*
[4]*CEREMADE, Paris Dauphine University, Paris, France*

Enriching Brownian motion with regenerations from a fixed regeneration distribution μ at a particular regeneration rate κ results in a Markov process that has a target distribution π as its invariant distribution. For the purpose of Monte Carlo inference, implementing such a scheme requires firstly selection of regeneration distribution μ, and secondly computation of a specific constant C. Both of these tasks can be very difficult in practice for good performance. We introduce a method for adapting the regeneration distribution, by adding point masses to it. This allows the process to be simulated with as few regenerations as possible and obviates the need to find said constant C. Moreover, the choice of fixed μ is replaced with the choice of the initial regeneration distribution, which is considerably less difficult. We establish convergence of this resulting self-reinforcing process and explore its effectiveness at sampling from a number of target distributions. The examples show that adapting the regeneration distribution guards against poor choices of fixed regeneration distribution and can reduce the error of Monte Carlo estimates of expectations of interest, especially when π is skewed.

Keywords: Adaptive algorithm; Markov process; MCMC; normalizing constant; regeneration distribution; Restore sampler; sampling; simulation

References

Ahn, S., Chen, Y. and Welling, M. (2013). Distributed and adaptive darting Monte Carlo through regenerations. In *Proc. Sixteenth Int. Conf. Artif. Intel. Stat.* (C.M. Carvalho and P. Ravikumar, eds.). *Proceedings of Machine Learning Research* **31** 108–116. Scottsdale, Arizona, USA: PMLR.

Aldous, D., Flannery, B. and Palacios, J.L. (1988). Two applications of urn processes: The fringe analysis of search trees and the simulation of quasi-stationary distributions of Markov chains. *Probab. Engrg. Inform. Sci.* **2** 293–307.

Andrieu, C. and Thoms, J. (2008). A tutorial on adaptive MCMC. *Stat. Comput.* **18** 343–373. MR2461882 https://doi.org/10.1007/s11222-008-9110-y

Benaïm, M. (1999). Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*. *Lecture Notes in Math.* **1709** 1–68. Berlin: Springer. MR1767993 https://doi.org/10.1007/BFb0096509

Benaïm, M. and Cloez, B. (2015). A stochastic approximation approach to quasi-stationary distributions on finite spaces. *Electron. Commun. Probab.* **20** 37. MR3352332 https://doi.org/10.1214/ECP.v20-3956

Benaïm, M., Cloez, B. and Panloup, F. (2018). Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Ann. Appl. Probab.* **28** 2370–2416. MR3843832 https://doi.org/10.1214/17-AAP1360

Bierkens, J., Fearnhead, P. and Roberts, G. (2019). The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *Ann. Statist.* **47** 1288–1320. MR3911113 https://doi.org/10.1214/18-AOS1715

Blanchet, J., Glynn, P. and Zheng, S. (2016). Analysis of a stochastic approximation algorithm for computing quasi-stationary distributions. *Adv. in Appl. Probab.* **48** 792–811. MR3568892 https://doi.org/10.1017/apr.2016.28

1350-7265

Bouchard-Côté, A., Vollmer, S.J. and Doucet, A. (2018). The bouncy particle sampler: A nonreversible rejection-free Markov chain Monte Carlo method. *J. Amer. Statist. Assoc.* **113** 855–867. MR3832232 https://doi.org/10.1080/01621459.2017.1294075

Carlin, B.P. and Gelfand, A.E. (1991). An iterative Monte Carlo method for nonconjugate Bayesian analysis. *Stat. Comput.* **1** 119–128. https://doi.org/10.1007/BF01889986

Del Moral, P., Doucet, A. and Jasra, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 411–436. MR2278333 https://doi.org/10.1111/j.1467-9868.2006.00553.x

Demuth, M. and van Casteren, J.A. (2000). *Stochastic Spectral Theory for Selfadjoint Feller Operators: A Functional Integration Approach*. *Probability and Its Applications*. Basel: Birkhäuser. MR1772266 https://doi.org/10.1007/978-3-0348-8460-0

Gabrié, M., Rotskoff, G.M. and Vanden-Eijnden, E. (2021). Efficient Bayesian sampling using normalizing flows to assist Markov chain Monte Carlo methods. In *Third Workshop on Invertible Neural Networks, Normalizing Flows, and Explicit Likelihood Models (ICML 2021)*.

Gelman, A. and Meng, X.-L. (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statist. Sci.* **13** 163–185. MR1647507 https://doi.org/10.1214/ss/1028905934

Geyer, C.J. (1991). Markov chain Monte Carlo maximum likelihood. *Proceedings of the 23rd Symposium on the Interface* 156–163.

Gilks, W.R., Roberts, G.O. and Sahu, S.K. (1998). Adaptive Markov chain Monte Carlo through regeneration. *J. Amer. Statist. Assoc.* **93** 1045–1054. MR1649199 https://doi.org/10.2307/2669848

Haario, H., Saksman, E. and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli* **7** 223–242. MR1828504 https://doi.org/10.2307/3318737

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

Ihler, A.T., Fisher, J.W., Moses, R.L. and Willsky, A.S. (2005). Nonparametric belief propagation for self-localization of sensor networks. *IEEE J. Sel. Areas Commun.* **23** 809–819. https://doi.org/10.1109/JSAC.2005.843548

Kass, R.E. and Raftery, A.E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795. MR3363402 https://doi.org/10.1080/01621459.1995.10476572

Kolb, M. and Steinsaltz, D. (2012). Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.* **40** 162–212. MR2917771 https://doi.org/10.1214/10-AOP623

Kumar, D. (2019). On a quasi-stationary approach to Bayesian computation, with application to tall data Ph.D. thesis, Univ. Warwick.

Lan, S., Streets, J. and Shahbaba, B. (2014). Wormhole Hamiltonian Monte Carlo. In *Proc. Twenty-Eighth AAAI Conf. Artif. Intel. AAAI'14* 1953–1959. AAAI Press.

Lewis, P.A.W. and Shedler, G.S. (1979). Simulation of nonhomogeneous Poisson processes by thinning. *Nav. Res. Logist. Q.* **26** 403–413. MR0546120 https://doi.org/10.1002/nav.3800260304

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

Mailler, C. and Villemonais, D. (2020). Stochastic approximation on noncompact measure spaces and application to measure-valued Pólya processes. *Ann. Appl. Probab.* **30** 2393–2438. MR4149532 https://doi.org/10.1214/20-AAP1561

Marinari, E. and Parisi, G. (1992). Simulated tempering: A new Monte Carlo scheme. *Europhys. Let. (EPL)* **19** 451–458. https://doi.org/10.1209/0295-5075/19/6/002

McKimm, H., Wang, A., Pollock, M., Robert, C. and Roberts, G. (2024). Supplement to "Sampling using adaptive regenerative processes." https://doi.org/10.3150/24-BEJ1737SUPP

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** 1087–1092.

Mykland, P., Tierney, L. and Yu, B. (1995). Regeneration in Markov chain samplers. *J. Amer. Statist. Assoc.* **90** 233–241. MR1325131

Neal, R.M. (1993). Probabilistic inference using Markov Chain Monte Carlo methods. Technical Report, University of Toronto.

Nummelin, E. (1978). A splitting technique for Harris recurrent Markov chains. *Z. Wahrsch. Verw. Gebiete* **43** 309–318. MR0501353 https://doi.org/10.1007/BF00534764

Ogata, Y. (1989). A Monte Carlo method for high-dimensional integration. *Numer. Math.* **55** 137–157. MR0987382 https://doi.org/10.1007/BF01406511

Papamakarios, G., Nalisnick, E., Rezende, D.J., Mohamed, S. and Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.* **22** 57. MR4253750

Pollock, M., Fearnhead, P., Johansen, A.M. and Roberts, G.O. (2020). Quasi-stationary Monte Carlo and the ScaLE algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 1167–1221. MR4176340

Pompe, E., Holmes, C. and Łatuszyński, K. (2020). A framework for adaptive MCMC targeting multimodal distributions. *Ann. Statist.* **48** 2930–2952. MR4152629 https://doi.org/10.1214/19-AOS1916

Robert, C.P. and Casella, G. (2004). *Monte Carlo Statistical Methods*, 2nd ed. *Springer Texts in Statistics*. New York: Springer. MR2080278 https://doi.org/10.1007/978-1-4757-4145-2

Roberts, G.O. and Rosenthal, J.S. (2009). Examples of adaptive MCMC. *J. Comput. Graph. Statist.* **18** 349–367. MR2749836 https://doi.org/10.1198/jcgs.2009.06134

Roberts, G.O. and Tweedie, R.L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* **2** 341–363. MR1440273 https://doi.org/10.2307/3318418

Rudolf, D. and Wang, A.Q. (2021). Perturbation theory for killed Markov processes and quasi-stationary distributions. http://arxiv.org/abs/2109.13819.

Skilling, J. (2006). Nested sampling for general Bayesian computation. *Bayesian Anal.* **1** 833–859. MR2282208 https://doi.org/10.1214/06-BA127

Tjelmeland, H. and Hegstad, B.K. (2001). Mode jumping proposals in MCMC. *Scand. J. Stat.* **28** 205–223. MR1844357 https://doi.org/10.1111/1467-9469.00232

Wang, A.Q. (2020). Theory of killing and regeneration in continuous-time Monte Carlo sampling. Ph.D. thesis, Univ. Oxford.

Wang, A.Q., Roberts, G.O. and Steinsaltz, D. (2020). An approximation scheme for quasi-stationary distributions of killed diffusions. *Stochastic Process. Appl.* **130** 3193–3219. MR4080743 https://doi.org/10.1016/j.spa.2019.09.010

Wang, A.Q., Kolb, M., Roberts, G.O. and Steinsaltz, D. (2019). Theoretical properties of quasi-stationary Monte Carlo methods. *Ann. Appl. Probab.* **29** 434–457. MR3910008 https://doi.org/10.1214/18-AAP1422

Wang, A.Q., Pollock, M., Roberts, G.O. and Steinsaltz, D. (2021). Regeneration-enriched Markov processes with application to Monte Carlo. *Ann. Appl. Probab.* **31** 703–735. MR4254493 https://doi.org/10.1214/20-aap1602

Bernoulli **31**(1), 2025, 537–560
https://doi.org/10.3150/24-BEJ1738

Maximum interpoint distance of high-dimensional random vectors

JOHANNES HEINY[1,a] and CAROLIN KLEEMANN[2,b]

[1]*Department of Mathematics, Stockholm University, Stockholm, Sweden,* [a]*johannes.heiny@math.su.se*
[2]*Department of Mathematics, Ruhr University Bochum, Bochum, Germany,* [b]*carolin.kleemann@rub.de*

A limit theorem for the largest interpoint distance of p independent and identically distributed points in $\mathbb{R}^n$ to the Gumbel distribution is proved, where the number of points $p = p_n$ tends to infinity as the dimension of the points $n \to \infty$. The theorem holds under moment assumptions and corresponding conditions on the growth rate of p. We obtain a plethora of ancillary results such as the joint convergence of maximum and minimum interpoint distances. Using the inherent sum structure of interpoint distances, our result is generalized to maxima of dependent random walks with non-decaying correlations and we also derive point process convergence. An application of the maximum interpoint distance to testing the equality of means for high-dimensional random vectors is presented. Moreover, we study the largest off-diagonal entry of a sample covariance matrix. The proofs are based on the Chen-Stein Poisson approximation method and Gaussian approximation to large deviation probabilities.

Keywords: Extreme value theory; Gumbel distribution; high dimension; independence test; maximum under dependence; p-norms

References

[1] Appel, M.J.B., Najim, C.A. and Russo, R.P. (2002). Limit laws for the diameter of a random point set. *Adv. in Appl. Probab.* **34** 1–10. MR1895327 https://doi.org/10.1239/aap/1019160946
[2] Arratia, R., Goldstein, L. and Gordon, L. (1989). Two moments suffice for Poisson approximations: The Chen–Stein method. *Ann. Probab.* **17** 9–25. MR0972770
[3] Biau, G. and Mason, D.M. (2015). High-dimensional p-norms. In *Mathematical Statistics and Limit Theorems* 21–40. Cham: Springer. MR3380729
[4] Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **106** 672–684. MR2847949 https://doi.org/10.1198/jasa.2011.tm10560
[5] Cai, T.T. (2017). Global testing and large-scale multiple testing for high-dimensional covariance structures. *Annu. Rev. Stat. Appl.* **4** 423–446.
[6] Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors. *Ann. Statist.* **41** 2786–2819. MR3161448 https://doi.org/10.1214/13-AOS1161
[7] Demichel, Y., Fermin, A.-K. and Soulier, P. (2015). The diameter of an elliptical cloud. *Electron. J. Probab.* **20** 1–32. MR3325097 https://doi.org/10.1214/EJP.v20-3777
[8] Donoho, D. (2000). High-dimensional data analysis: The curses and blessings of dimensionality Technical Report, Stanford Univ.
[9] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. *Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613 https://doi.org/10.1007/978-3-642-33483-2
[10] Fan, J. and Jiang, T. (2019). Largest entries of sample correlation matrices from equi-correlated normal populations. *Ann. Probab.* **47** 3321–3374. MR4021253 https://doi.org/10.1214/19-AOP1341
[11] 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

[12] Han, F., Chen, S. and Liu, H. (2017). Distribution-free tests of independence in high dimensions. *Biometrika* **104** 813–828. MR3737306 https://doi.org/10.1093/biomet/asx050

[13] Heiny, J. and Kleemann, C. (2024). Supplement to "Maximum interpoint distance of high-dimensional random vectors." https://doi.org/10.3150/24-BEJ1738SUPP

[14] Heiny, J., Mikosch, T. and Yslas, J. (2021). Point process convergence for the off-diagonal entries of sample covariance matrices. *Ann. Appl. Probab.* **31** 538–560. MR4254488 https://doi.org/10.1214/20-aap1597

[15] Henze, N. and Klein, T. (1996). The limit distribution of the largest interpoint distance from a symmetric Kotz sample. *J. Multivariate Anal.* **57** 228–239. MR1391170 https://doi.org/10.1006/jmva.1996.0031

[16] Jammalamadaka, S.R. and Janson, S. (1986). Limit theorems for a triangular scheme of U-statistics with applications to inter-point distances. *Ann. Probab.* **14** 1347–1358. MR0866355

[17] Jammalamadaka, S.R. and Janson, S. (2015). Asymptotic distribution of the maximum interpoint distance in a sample of random vectors with a spherically symmetric distribution. *Ann. Appl. Probab.* **25** 3571–3591. MR3404644 https://doi.org/10.1214/14-AAP1082

[18] Jiang, T. (2004). The asymptotic distributions of the largest entries of sample correlation matrices. *Ann. Appl. Probab.* **14** 865–880. MR2052906 https://doi.org/10.1214/105051604000000143

[19] Johnstone, I.M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327. MR1863961 https://doi.org/10.1214/aos/1009210544

[20] Johnstone, I.M. and Titterington, D.M. (2009). Statistical challenges of high-dimensional data. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **367** 4237–4253. MR2546386 https://doi.org/10.1098/rsta.2009.0159

[21] Kallenberg, O. (1983). *Random Measures*, 3rd ed. Berlin: Akademie-Verlag, Academic Press [Harcourt Brace Jovanovich, Publishers], London. MR0818219

[22] Kuba, M. and Panholzer, A. (2016). On moment sequences and mixed Poisson distributions. *Probab. Surv.* **13** 89–155. MR3550936 https://doi.org/10.1214/14-PS244

[23] Lao, W. and Mayer, M. (2008). U-max-statistics. *J. Multivariate Anal.* **99** 2039–2052. MR2466550 https://doi.org/10.1016/j.jmva.2008.02.001

[24] Li, D., Liu, W.-D. and Rosalsky, A. (2010). Necessary and sufficient conditions for the asymptotic distribution of the largest entry of a sample correlation matrix. *Probab. Theory Related Fields* **148** 5–35. MR2653220 https://doi.org/10.1007/s00440-009-0220-z

[25] Li, D., Qi, Y. and Rosalsky, A. (2012). On Jiang's asymptotic distribution of the largest entry of a sample correlation matrix. *J. Multivariate Anal.* **111** 256–270. MR2944420 https://doi.org/10.1016/j.jmva.2012.04.002

[26] Li, J. (2018). Asymptotic normality of interpoint distances for high-dimensional data with applications to the two-sample problem. *Biometrika* **105** 529–546. MR3842883 https://doi.org/10.1093/biomet/asy020

[27] Linnik, J.V. (1960). On the probability of large deviations for the sums of independent variables. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob., Vol. II* 289–306. Berkeley-Los Angeles, Calif.: Univ. California Press. MR0137142

[28] Liu, S.-H. and Zhang, Z.-S. (2023). Cramér-type moderate deviations under local dependence. *Ann. Appl. Probab.* **33** 4747–4797. MR4674063 https://doi.org/10.1214/23-aap1931

[29] Matthews, P.C. and Rukhin, A.L. (1993). Asymptotic distribution of the normal sample range. *Ann. Appl. Probab.* **3** 454–466. MR1221161

[30] Mayer, M. and Molchanov, I. (2007). Limit theorems for the diameter of a random sample in the unit ball. *Extremes* **10** 129–150. MR2394205 https://doi.org/10.1007/s10687-007-0038-y

[31] Petrov, V.V. (1975). *Sums of Independent Random Variables*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete [Results in Mathematics and Related Areas], Band 82*. New York: Springer. Translated from the Russian by A. A. Brown. MR0388499

[32] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. Reprint of the 1987 original. MR2364939

[33] Schrempp, M. (2016). The limit distribution of the largest interpoint distance for distributions supported by a d-dimensional ellipsoid and generalizations. *Adv. in Appl. Probab.* **48** 1256–1270. MR3595774 https://doi.org/10.1017/apr.2016.73

[34] Shao, Q.-M. and Zhou, W.-X. (2014). Necessary and sufficient conditions for the asymptotic distributions of coherence of ultra-high dimensional random matrices. *Ann. Probab.* **42** 623–648. MR3178469 https://doi.org/10.1214/13-AOP837

[35] Tang, P., Lu, R. and Xie, J. (2022). Asymptotic distribution of the maximum interpoint distance for high-dimensional data. *Statist. Probab. Lett.* **190** Paper No. 109567, 7. MR4461681 https://doi.org/10.1016/j.spl.2022.109567

[36] Zaĭtsev, A.Y. (1987). On the Gaussian approximation of convolutions under multidimensional analogues of S.N. Bernstein's inequality conditions. *Probab. Theory Related Fields* **74** 535–566. MR0876255 https://doi.org/10.1007/BF00363515

[37] Zhu, C. and Shao, X. (2021). Interpoint distance based two sample tests in high dimension. *Bernoulli* **27** 1189–1211. MR4255231 https://doi.org/10.3150/20-bej1270

Bernoulli **31**(1), 2025, 561–583
https://doi.org/10.3150/24-BEJ1739

Synchronisation for scalar conservation laws via Dirichlet boundary

ANA DJURDJEVAC[1,a] and TOMMASO ROSATI[2,b]

[1]*Freie Universität Berlin, Berlin, Germany,* [a]*adjurdjevac@zedat.fu-berlin.de*
[2]*University of Warwick, Coventry, UK,* [b]*t.rosati@warwick.ac.uk*

We provide an elementary proof of geometric synchronisation for scalar conservation laws on a domain with Dirichlet boundary conditions. Unlike previous results, our proof does not rely on a strict maximum principle, and builds instead on a quantitative estimate of the dissipation at the boundary. We identify a coercivity condition under which the estimates are uniform over all initial conditions, via the construction of suitable super- and sub-solutions. In lack of such coercivity our results build on L^p energy estimates and a Lyapunov structure.

Keywords: Burgers; mixing; scalar conservation laws; synchronisation

References

Boritchev, A.A. (2014). Turbulence in the generalised Burgers equation. *Uspekhi Mat. Nauk* **69** 3–44. MR3400554 https://doi.org/10.4213/rm9629

Boritchev, A. (2016). Multidimensional potential Burgers turbulence. *Comm. Math. Phys.* **342** 441–489. MR3459157 https://doi.org/10.1007/s00220-015-2521-7

Boritchev, A. (2018). Exponential convergence to the stationary measure for a class of 1D Lagrangian systems with random forcing. *Stoch. Partial Differ. Equ. Anal. Comput.* **6** 109–123. MR3768996 https://doi.org/10.1007/s40072-017-0104-7

Corwin, I. and Knizel, A. (2024). Stationary measure for the open KPZ equation. *Comm. Pure Appl. Math.* **77** 2183–2267. MR4705293

Debussche, A. and Vovelle, J. (2010). Scalar conservation laws with stochastic forcing. *J. Funct. Anal.* **259** 1014–1042. MR2652180 https://doi.org/10.1016/j.jfa.2010.02.016

Debussche, A. and Vovelle, J. (2015). Invariant measure of scalar first-order conservation laws with stochastic forcing. *Probab. Theory Related Fields* **163** 575–611. MR3418750 https://doi.org/10.1007/s00440-014-0599-z

Djurdjevac, A. and Shirikyan, A. (2022). Stabilisation of a viscous conservation law by a one-dimensional external force. arXiv preprint. Available at arXiv:2204.03427.

Drivas, T.D., Dunlap, A., Graham, C., La, J. and Ryzhik, L. (2023). Invariant measures for stochastic conservation laws on the line. *Nonlinearity* **36** 4553–4584. MR4619242 https://doi.org/10.1088/1361-6544/acdb3a

Dunlap, A., Graham, C. and Ryzhik, L. (2021). Stationary solutions to the stochastic Burgers equation on the line. *Comm. Math. Phys.* **382** 875–949. MR4227165 https://doi.org/10.1007/s00220-021-04025-x

E, W., Khanin, K., Mazel, A. and Sinai, Y. (2000). Invariant measures for Burgers equation with stochastic forcing. *Ann. of Math. (2)* **151** 877–960. MR1779561 https://doi.org/10.2307/121126

Gubinelli, M. and Perkowski, N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269. MR3592748 https://doi.org/10.1007/s00220-016-2788-3

Gyöngy, I. and Rovira, C. (2000). On L^p-solutions of semilinear stochastic partial differential equations. *Stochastic Process. Appl.* **90** 83–108. MR1787126 https://doi.org/10.1016/S0304-4149(00)00033-8

Hairer, M. (2013). Solving the KPZ equation. *Ann. of Math. (2)* **178** 559–664. MR3071506 https://doi.org/10.4007/annals.2013.178.2.4

Lunardi, A. (1995). *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. *Modern Birkhäuser Classics*. Basel: Birkhäuser/Springer Basel AG. [2013 reprint of the 1995 original] [MR1329547]. MR3012216

Mourrat, J.-C. and Weber, H. (2017). The dynamic Φ^4_3 model comes down from infinity. *Comm. Math. Phys.* **356** 673–753. MR3719541 https://doi.org/10.1007/s00220-017-2997-4

Perkowski, N. and Rosati, T.C. (2019). The KPZ equation on the real line. *Electron. J. Probab.* **24** 1–56. MR4029420 https://doi.org/10.1214/19-ejp362

Rosati, T.C. (2022). Synchronization for KPZ. *Stoch. Dyn.* **22** Paper No. 2250010, 46. MR4439607 https://doi.org/10.1142/S0219493722500101

Rosati, T. (2024). Lyapunov exponents in a slow environment. *Stochastic Process. Appl.* **170** Paper No. 104296, 32. MR4688670 https://doi.org/10.1016/j.spa.2024.104296

Shirikyan, A. (2018a). Mixing for the Burgers equation driven by a localized two-dimensional stochastic forcing. In *Evolution Equations: Long Time Behavior and Control*. *London Mathematical Society Lecture Note Series* **439** 179–194. Cambridge: Cambridge Univ. Press. MR3586898

Shirikyan, A. (2018b). Control theory for the Burgers equation: Agrachev–Sarychev approach. *Pure Appl. Funct. Anal.* **3** 219–240. MR3777005

Sinaĭ, Y.G. (1991). Two results concerning asymptotic behavior of solutions of the Burgers equation with force. *J. Stat. Phys.* **64** 1–12. MR1117645 https://doi.org/10.1007/BF01057866

Zhang, X., Zhu, R. and Zhu, X. (2022). Singular HJB equations with applications to KPZ on the real line. *Probab. Theory Related Fields* **183** 789–869. MR4453317 https://doi.org/10.1007/s00440-022-01137-w

Bernoulli **31**(1), 2025, 584–608
https://doi.org/10.3150/24-BEJ1740

Tractably modelling dependence in networks beyond exchangeability

WEICHI WU[1,a], SOFIA OLHEDE[2,b] and PATRICK WOLFE[3,c]

[1] *Center for Statistical Sciences,Department of Industrial Engineering, Tsinghua University, Beijing, China,* [a]*wuweichi@mail.tsinghua.edu.cn*
[2] *Institute of Mathematics, Chair of Statistical Data Science, EPFL, Lausanne, Switzerland,* [b]*sofia.olhede@epfl.ch*
[3] *Departments of Statistics and Computer Science, Purdue University, West Lafayette, USA,* [c]*patrick@purdue.edu*

We propose a general framework for modelling network data that is designed to describe aspects of non-exchangeability, with an explicit parameter describing the degree of non-exchangeability. Conditional on latent (unobserved) variables, the edges of the network are generated by their finite growth history (via a latent order) while the marginal probabilities of the adjacency matrix are modeled by a generalization of a graph limit function (or a graphon). In particular, we study the estimation, clustering and degree behavior of the network in this setting. We determine (i) the least squares estimator of a composite graphon attaining the minimax rate under weak dependence with respect to squared error loss; (ii) that spectral clustering is able to consistently detect the latent membership when the block-wise constant composite graphon is considered under additional conditions; and (iii) we are able to construct models with heavy-tailed empirical degrees under specific scenarios and parameter choices. We find conditions under which the spectral clustering is consistent under non-exchangeability, revealing that the application scope of classification can be broader than classic $i.i.d.$ or exchangeable assumptions. In aggregate, we explore why and under which general conditions non-exchangeable network data can be described by a stochastic block model. The new modelling framework is able to capture empirically important characteristics of network data such as sparsity combined with heavy tailed degree distribution, and add understanding as to what generative mechanisms will make them arise.

Keywords: Exchangeable arrays; nonlinear stochastic processes; statistical network analysis; stochastic block model

References

[1] Airoldi, E.M., Costa, T.B. and Chan, S.H. (2013). Stochastic blockmodel approximation of a graphon: Theory and consistent estimation. In *Adv. Neural. Inf. Process. Syst.* 692–700.
[2] Amini, A.A., Chen, A., Bickel, P.J. and Levina, E. (2013). Pseudo-likelihood methods for community detection in large sparse networks. *Ann. Statist.* **41** 2097–2122. MR3127859 https://doi.org/10.1214/13-AOS1138
[3] Amini, A.A. and Levina, E. (2018). On semidefinite relaxations for the block model. *Ann. Statist.* **46** 149–179. MR3766949 https://doi.org/10.1214/17-AOS1545
[4] Anandkumar, A., Ge, R., Hsu, D. and Kakade, S.M. (2014). A tensor approach to learning mixed membership community models. *J. Mach. Learn. Res.* **15** 2239–2312. MR3231594
[5] Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286** 509–512. MR2091634 https://doi.org/10.1126/science.286.5439.509
[6] Berkhin, P. (2006). A survey of clustering data mining techniques. In *Grouping Multidimensional Data* 25–71. Springer.
[7] Bickel, P., Choi, D., Chang, X. and Zhang, H. (2013). Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. *Ann. Statist.* **41** 1922–1943. MR3127853 https://doi.org/10.1214/13-AOS1124
[8] Bickel, P.J. and Chen, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Natl. Acad. Sci. USA* **106** 21068–21073.

1350-7265

[9] Bloem-Reddy, B. and Orbanz, P. (2017). Preferential attachment and vertex arrival times. ArXiv preprint. Available at arXiv:1710.02159.
[10] Bollobás, B. and Riordan, O. (2009). Metrics for sparse graphs. In *Surveys in Combinatorics 2009*. *London Mathematical Society Lecture Note Series* **365** 211–287. Cambridge: Cambridge Univ. Press. MR2588543
[11] Borgs, C., Chayes, J.T., Cohn, H. and Ganguly, S. (2021). Consistent nonparametric estimation for heavy-tailed sparse graphs. *Ann. Statist.* **49** 1904–1930. MR4319235 https://doi.org/10.1214/20-aos1985
[12] Borgs, C., Chayes, J.T., Cohn, H. and Holden, N. (2017). Sparse exchangeable graphs and their limits via graphon processes. *J. Mach. Learn. Res.* **18** Paper No. 210. MR3827098
[13] Borgs, C., Chayes, J.T., Cohn, H. and Veitch, V. (2019). Sampling perspectives on sparse exchangeable graphs. *Ann. Probab.* **47** 2754–2800. MR4021236 https://doi.org/10.1214/18-AOP1320
[14] Borgs, C., Chayes, J.T., Cohn, H. and Zhao, Y. (2019). An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. *Trans. Amer. Math. Soc.* **372** 3019–3062. MR3988601 https://doi.org/10.1090/tran/7543
[15] Broido, A.D. and Clauset, A. (2019). Scale-free networks are rare. *Nat. Commun.* **10** 1017.
[16] Cai, T.T. and Li, X. (2015). Robust and computationally feasible community detection in the presence of arbitrary outlier nodes. *Ann. Statist.* **43** 1027–1059. MR3346696 https://doi.org/10.1214/14-AOS1290
[17] Caron, F. and Fox, E.B. (2017). Sparse graphs using exchangeable random measures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 1295–1366. MR3731666 https://doi.org/10.1111/rssb.12233
[18] Choi, D.S., Wolfe, P.J. and Airoldi, E.M. (2012). Stochastic blockmodels with a growing number of classes. *Biometrika* **99** 273–284. MR2931253 https://doi.org/10.1093/biomet/asr053
[19] Crane, H. and Dempsey, W. (2018). Edge exchangeable models for interaction networks. *J. Amer. Statist. Assoc.* **113** 1311–1326. MR3862359 https://doi.org/10.1080/01621459.2017.1341413
[20] Davis, C. and Kahan, W.M. (1970). The rotation of eigenvectors by a perturbation. III. *SIAM J. Numer. Anal.* **7** 1–46. MR0264450 https://doi.org/10.1137/0707001
[21] Gao, C., Lu, Y., Ma, Z. and Zhou, H.H. (2016). Optimal estimation and completion of matrices with biclustering structures. *J. Mach. Learn. Res.* **17** Paper No. 161. MR3569248
[22] Gao, C., Lu, Y. and Zhou, H.H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43** 2624–2652. MR3405606 https://doi.org/10.1214/15-AOS1354
[23] Gao, C. and Ma, Z. (2021). Minimax rates in network analysis: Graphon estimation, community detection and hypothesis testing. *Statist. Sci.* **36** 16–33. MR4194201 https://doi.org/10.1214/19-STS736
[24] Gao, F. and van der Vaart, A. (2017). On the asymptotic normality of estimating the affine preferential attachment network models with random initial degrees. *Stochastic Process. Appl.* **127** 3754–3775. MR3707244 https://doi.org/10.1016/j.spa.2017.03.008
[25] Gao, J. (2007). *Nonlinear Time Series: Semiparametric and Nonparametric Methods*. *Monographs on Statistics and Applied Probability* **108**. Boca Raton, FL: CRC Press/CRC. MR2297190 https://doi.org/10.1201/9781420011210
[26] Gong, Y. (1995). Speech recognition in noisy environments: A survey. *Speech Commun.* **16** 261–291.
[27] Ji, P. and Jin, J. (2016). Coauthorship and citation networks for statisticians. *Ann. Appl. Stat.* **10** 1779–1812. MR3592033 https://doi.org/10.1214/15-AOAS896
[28] Jin, J. (2015). Fast community detection by SCORE. *Ann. Statist.* **43** 57–89. MR3285600 https://doi.org/10.1214/14-AOS1265
[29] Kallenberg, O. (2005). *Probabilistic Symmetries and Invariance Principles*. *Probability and Its Applications (New York)*. New York: Springer. MR2161313
[30] Karrer, B. and Newman, M.E.J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E (3)* **83** 016107. MR2788206 https://doi.org/10.1103/PhysRevE.83.016107
[31] Klopp, O., Tsybakov, A.B. and Verzelen, N. (2017). Oracle inequalities for network models and sparse graphon estimation. *Ann. Statist.* **45** 316–354. MR3611494 https://doi.org/10.1214/16-AOS1454
[32] Li, T., Lei, L., Bhattacharyya, S., Van den Berge, K., Sarkar, P., Bickel, P.J. and Levina, E. (2022). Hierarchical community detection by recursive partitioning. *J. Amer. Statist. Assoc.* **117** 951–968. MR4436325 https://doi.org/10.1080/01621459.2020.1833888
[33] Lovász, L. (2012). *Large Networks and Graph Limits*. American Mathematical Soc.
[34] Maier, B.F. and Brockmann, D. (2017). Cover time for random walks on arbitrary complex networks. *Phys. Rev. E (3)* **96** 042307. https://doi.org/10.1103/PhysRevE.96.042307

[35] Newman, M.E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46** 323–351.
[36] Olhede, S.C. and Wolfe, P.J. (2013). Degree-based network models. ArXiv preprint. Available at arXiv:1211.6537.
[37] Olhede, S.C. and Wolfe, P.J. (2014). Network histograms and universality of blockmodel approximation. *Proc. Natl. Acad. Sci. USA* **111** 14722–14727. https://doi.org/10.1073/pnas.1400374111
[38] Ouadah, S., Robin, S. and Latouche, P. (2020). Degree-based goodness-of-fit tests for heterogeneous random graph models: Independent and exchangeable cases. *Scand. J. Stat.* **47** 156–181. MR4075233 https://doi.org/10.1111/sjos.12410
[39] Pagani, G.A. and Aiello, M. (2013). The power grid as a complex network: A survey. *Phys. A* **392** 2688–2700. MR3043171 https://doi.org/10.1016/j.physa.2013.01.023
[40] Papadopoulos, F., Kitsak, M., Serrano, M.Á., Boguná, M. and Krioukov, D. (2012). Popularity versus similarity in growing networks. *Nature* **489** 537.
[41] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Berlin: Springer. MR2245368
[42] Qin, T. and Rohe, K. (2013). Regularized spectral clustering under the degree-corrected stochastic blockmodel. In *Adv. Neural. Inf. Process. Syst.* 3120–3128.
[43] Rohe, K., Chatterjee, S. and Yu, B. (2011). Spectral clustering and the high-dimensional stochastic blockmodel. *Ann. Statist.* **39** 1878–1915. MR2893856 https://doi.org/10.1214/11-AOS887
[44] Saldaña, D.F., Yu, Y. and Feng, Y. (2017). How many communities are there? *J. Comput. Graph. Statist.* **26** 171–181. MR3610418 https://doi.org/10.1080/10618600.2015.1096790
[45] Sarkar, P., Chakrabarti, D. and Jordan, M. (2014). Nonparametric link prediction in large scale dynamic networks. *Electron. J. Stat.* **8** 2022–2065. MR3273618 https://doi.org/10.1214/14-EJS943
[46] Solé, R.V. and Montoya, J.M. (2001). Complexity and fragility in ecological networks. *Proc. Biol. Sci.* **268** 2039–2045. https://doi.org/10.1098/rspb.2001.1767
[47] Truquet, L. (2019). Local stationarity and time-inhomogeneous Markov chains. *Ann. Statist.* **47** 2023–2050. MR3953443 https://doi.org/10.1214/18-AOS1739
[48] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794
[49] Varin, C., Reid, N. and Firth, D. (2011). An overview of composite likelihood methods. *Statist. Sinica* **21** 5–42. MR2796852
[50] Veitch, V. and Roy, D.M. (2015). The class of random graphs arising from exchangeable random measures. ArXiv preprint. Available at arXiv:1512.03099.
[51] Veitch, V. and Roy, D.M. (2019). Sampling and estimation for (sparse) exchangeable graphs. *Ann. Statist.* **47** 3274–3299. MR4025742 https://doi.org/10.1214/18-AOS1778
[52] von Luxburg, U. (2007). A tutorial on spectral clustering. *Stat. Comput.* **17** 395–416. MR2409803 https://doi.org/10.1007/s11222-007-9033-z
[53] Wolfe, P.J. and Olhede, S.C. (2013). Nonparametric graphon estimation. ArXiv preprint. Available at arXiv:1309.5936.
[54] Wu, W., Olhede, S.C. and Wolfe, P.J. (2024). Supplement to "Tractably modelling dependence in networks beyond exchangeability." https://doi.org/10.3150/24-BEJ1740SUPP
[55] Wu, W. and Zhou, Z. (2017). Nonparametric inference for time-varying coefficient quantile regression. *J. Bus. Econom. Statist.* **35** 98–109. MR3591532 https://doi.org/10.1080/07350015.2015.1060884
[56] 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
[57] Yuan, Y. and Qu, A. (2021). Community detection with dependent connectivity. *Ann. Statist.* **49** 2378–2428. MR4319255 https://doi.org/10.1214/20-aos2042
[58] Zhang, Y., Levina, E. and Zhu, J. (2020). Detecting overlapping communities in networks using spectral methods. *SIAM J. Math. Data Sci.* **2** 265–283. MR4081803 https://doi.org/10.1137/19M1272238
[59] Zhao, Y., Levina, E. and Zhu, J. (2012). Consistency of community detection in networks under degree-corrected stochastic block models. *Ann. Statist.* **40** 2266–2292. MR3059083 https://doi.org/10.1214/12-AOS1036
[60] 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

[61] Zhou, Z. (2014). Inference of weighted V-statistics for nonstationary time series and its applications. *Ann. Statist.* **42** 87–114. MR3161462 https://doi.org/10.1214/13-AOS1184

[62] 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 **31**(1), 2025, 609–629
https://doi.org/10.3150/24-BEJ1741

Optimal stopping of the stable process with state-dependent killing

KEES VAN SCHAIK[1,a], ALEXANDER R. WATSON[2,b] and XIN XU[3,c]

[1]*University of Manchester, Manchester, UK,* [a]*kees.vanschaik@manchester.ac.uk*
[2]*University College London, London, UK,* [b]*alexander.watson@ucl.ac.uk*
[3]*Xi'an Jiaotong-Liverpool University, Suzhou, PR China,* [c]*Xin.Xu03@xjtlu.edu.cn*

We describe the solution of an optimal stopping problem for a stable Lévy process killed at state-dependent rate. The killing rate is chosen in such a way that the killed process remains self-similar, and the solution to the optimal stopping problem is obtained by characterising a self-similar Markov process associated with the stable process. The optimal stopping strategy is to stop upon first passage into an interval, found explicitly in terms of the parameters of the model.

Keywords: Lamperti transformation; Lévy process; Markov additive process; omega-clock; optimal stopping; self-similar Markov process; stable process

References

[1] Albrecher, H., Gerber, H.U. and Shiu, E.S.W. (2011). The optimal dividend barrier in the gamma-omega model. *Eur. Actuar. J.* **1** 43–55. MR2843466 https://doi.org/10.1007/s13385-011-0006-4

[2] Albrecher, H. and Lautscham, V. (2013). From ruin to bankruptcy for compound Poisson surplus processes. *Astin Bull.* **43** 213–243. MR3388122 https://doi.org/10.1017/asb.2013.4

[3] Bichteler, K. (2002). *Stochastic Integration with Jumps*. *Encyclopedia of Mathematics and Its Applications* **89**. Cambridge: Cambridge Univ. Press. MR1906715 https://doi.org/10.1017/CBO9780511549878

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

[5] Bogdan, K., Burdzy, K. and Chen, Z.-Q. (2003). Censored stable processes. *Probab. Theory Related Fields* **127** 89–152. MR2006232 https://doi.org/10.1007/s00440-003-0275-1

[6] Budd, T. (2018). The peeling process on random planar maps coupled to an $O(n)$ loop model (with an appendix by Linxiao Chen). Preprint. Available at arXiv:1809.02012v1.

[7] Caballero, M.E. and Chaumont, L. (2006). Conditioned stable Lévy processes and the Lamperti representation. *J. Appl. Probab.* **43** 967–983. MR2274630 https://doi.org/10.1239/jap/1165505201

[8] Caballero, M.E., Pardo, J.C. and Pérez, J.L. (2011). Explicit identities for Lévy processes associated to symmetric stable processes. *Bernoulli* **17** 34–59. MR2797981 https://doi.org/10.3150/10-BEJ275

[9] Çağlar, M., Kyprianou, A. and Vardar-Acar, C. (2022). An optimal stopping problem for spectrally negative Markov additive processes. *Stochastic Process. Appl.* **150** 1109–1138. MR4440179 https://doi.org/10.1016/j.spa.2021.06.010

[10] Czarna, I., Kaszubowski, A., Li, S. and Palmowski, Z. (2020). Fluctuation identities for omega-killed spectrally negative Markov additive processes and dividend problem. *Adv. in Appl. Probab.* **52** 404–432. MR4123641 https://doi.org/10.1017/apr.2020.2

[11] Gerber, H.U., Shiu, E.S.W. and Yang, H. (2012). The Omega model: From bankruptcy to occupation times in the red. *Eur. Actuar. J.* **2** 259–272. MR3039553 https://doi.org/10.1007/s13385-012-0052-6

[12] Kaszubowski, A. (2019). Omega bankruptcy for different Lévy models. *Śląski Przegląd Stat.* **17** 31–58.

[13] Kühn, C. and van Schaik, K. (2008). Perpetual convertible bonds with credit risk. *Stochastics* **80** 585–610. MR2460249 https://doi.org/10.1080/17442500802263888

1350-7265

[14] 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
[15] Kyprianou, A.E. (2014). *Fluctuations of Lévy Processes with Applications: Introductory Lectures*, 2nd ed. *Universitext*. Heidelberg: Springer. MR3155252 https://doi.org/10.1007/978-3-642-37632-0
[16] Kyprianou, A.E. (2016). Deep factorisation of the stable process. *Electron. J. Probab.* **21** 1–28. MR3485365 https://doi.org/10.1214/16-EJP4506
[17] 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
[18] Kyprianou, A.E., Pardo, J.C. and Vidmar, M. (2021). Double hypergeometric Lévy processes and self-similarity. *J. Appl. Probab.* **58** 254–273. MR4222428 https://doi.org/10.1017/jpr.2020.86
[19] Kyprianou, A.E., Pardo, J.C. and Watson, A.R. (2014). Hitting distributions of α-stable processes via path censoring and self-similarity. *Ann. Probab.* **42** 398–430. MR3161489 https://doi.org/10.1214/12-AOP790
[20] Kyprianou, A.E., Rivero, V. and Şengül, B. (2018). Deep factorisation of the stable process II: Potentials and applications. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 343–362. MR3765892 https://doi.org/10.1214/16-AIHP806
[21] Lamperti, J. (1972). Semi-stable Markov processes. I. *Z. Wahrsch. Verw. Gebiete* **22** 205–225. MR0307358 https://doi.org/10.1007/BF00536091
[22] Li, B. and Palmowski, Z. (2018). Fluctuations of omega-killed spectrally negative Lévy processes. *Stochastic Process. Appl.* **128** 3273–3299. MR3849809 https://doi.org/10.1016/j.spa.2017.10.018
[23] Mordecki, E. (2002). Optimal stopping and perpetual options for Lévy processes. *Finance Stoch.* **6** 473–493. MR1932381 https://doi.org/10.1007/s007800200070
[24] Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary Problems*. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2256030
[25] Rodosthenous, N. and Zhang, H. (2018). Beating the omega clock: An optimal stopping problem with random time-horizon under spectrally negative Lévy models. *Ann. Appl. Probab.* **28** 2105–2140. MR3843825 https://doi.org/10.1214/17-AAP1322
[26] Rogozin, B.A. (1973). The distribution of the first hit for stable and asymptotically stable walks on an interval. *Theory Probab. Appl.* **17** 332–338.
[27] Watson, A.R. (2013). Stable processes. Ph.D. thesis, Univ. Bath.

Bernoulli **31**(1), 2025, 630–648
https://doi.org/10.3150/24-BEJ1742

Tail probability of maximal displacement in critical branching Lévy process with stable branching

HAOJIE HOU[1,a], YIYANG JIANG[1,b], YAN-XIA REN[2,c] and RENMING SONG[3,d]

[1]*School of Mathematical Sciences, Peking University, Beijing, P.R. China,* [a]*houhaojie@pku.edu.cn,* [b]*jyy.0916@stu.pku.edu.cn*
[2]*LMAM School of Mathematical Sciences & Center for Statistical Science, Peking University, Beijing, P.R. China,* [c]*yxren@math.pku.edu.cn*
[3]*Department of Mathematics, University of Illinois Urbana-Champaign, Urbana, U.S.A.,* [d]*rsong@illinois.edu*

Consider a critical branching Lévy process $\{X_t, t \geq 0\}$ with branching rate $\beta > 0$, offspring distribution $\{p_k : k \geq 0\}$ and spatial motion $\{\xi_t, \mathcal{P}_x\}$. For any $t \geq 0$, let N_t be the collection of particles alive at time t, and, for any $u \in N_t$, let $X_u(t)$ be the position of u at time t. We study the tail probability of the maximal displacement $M := \sup_{t>0} \sup_{u \in N_t} X_u(t)$ under the assumption $\lim_{n\to\infty} n^\alpha \sum_{k=n}^\infty p_k = \kappa \in (0,\infty)$ for some $\alpha \in (1,2)$, $\mathcal{E}_0(\xi_1) = 0$ and $\mathcal{E}_0((\xi_1^+)^r) \in (0,\infty)$ for some $r > 2\alpha/(\alpha-1)$. Our main result is a generalization of the main result of Sawyer and Fleischman (1979) for branching Brownian motions and that of Lalley and Shao (2015) for branching random walks, both of these results are proved under the assumption $\sum_{k=0}^\infty k^3 p_k < \infty$.

Keywords: Branching Lévy process; critical branching process; Feynman-Kac representation

References

Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR1015093

Chow, Y.S. (1986). On moments of ladder height variables. *Adv. in Appl. Math.* **7** 46–54. MR0834219 https://doi.org/10.1016/0196-8858(86)90005-9

Chow, Y.S. and Lai, T.L. (1979). Moments of ladder variables for driftless random walks. *Z. Wahrsch. Verw. Gebiete* **48** 253–257. MR0537671 https://doi.org/10.1007/BF00537523

Chung, K.L. and Zhao, Z.X. (1995). *From Brownian Motion to Schrödinger's Equation. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **312**. Berlin: Springer. MR1329992 https://doi.org/10.1007/978-3-642-57856-4

Doney, R.A. (1980). Moments of ladder heights in random walks. *J. Appl. Probab.* **17** 248–252. MR0557453 https://doi.org/10.2307/3212942

Doney, R.A. and Maller, R.A. (2002). Stability of the overshoot for Lévy processes. *Ann. Probab.* **30** 188–212. MR1894105 https://doi.org/10.1214/aop/1020107765

Dynkin, E.B. (2001). Branching exit Markov systems and superprocesses. *Ann. Probab.* **29** 1833–1858. MR1880244 https://doi.org/10.1214/aop/1015345774

Kyprianou, A.E. (1999). A note on branching Lévy processes. *Stochastic Process. Appl.* **82** 1–14. MR1695066 https://doi.org/10.1016/S0304-4149(99)00010-1

Lalley, S.P. and Shao, Y. (2015). On the maximal displacement of critical branching random walk. *Probab. Theory Related Fields* **162** 71–96. MR3350041 https://doi.org/10.1007/s00440-014-0566-8

Lalley, S.P. and Shao, Y. (2016). Maximal displacement of critical branching symmetric stable processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1161–1177. MR3531704 https://doi.org/10.1214/15-AIHP677

Profeta, C. (2022). Extreme values of critical and subcritical branching stable processes with positive jumps. *ALEA Lat. Am. J. Probab. Math. Stat.* **19** 1421–1433. MR4517728

Profeta, C. (2024). Maximal displacement of spectrally negative branching Lévy processes. *Bernoulli* **30** 961–982. MR4699541 https://doi.org/10.3150/23-bej1620

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

Sawyer, S. and Fleischman, J. (1979). Maximum geographic range of a mutant allele considered as a subtype of a Brownian branching random field. *Proc. Natl. Acad. Sci. USA* **76** 872–875. https://doi.org/10.1073/pnas.76.2.872

Skorohod, A.V. (1957). Limit theorems for stochastic processes with independent increments. *Teor. Veroyatn. Primen.* **2** 145–177. MR0094842

Bernoulli **31**(1), 2025, 649–670
https://doi.org/10.3150/24-BEJ1743

Dynamic principal component analysis from a global perspective

LINGXUAN SHAO[1,a] and FANG YAO[2,b]

[1]*Department of Statistics and Data Science, School of Management, Fudan University, Shanghai, China,* [a]*shao_lingxuan@fudan.edu.cn*

[2]*Department of Probability and Statistics, Center for Statistical Science, School of Mathematical Sciences, Peking University, Beijing, China,* [b]*fyao@math.pku.edu.cn*

Principal component analysis has been one of prominent techniques in multi-dimensional statistics, and the core is to estimate the eigenvalues in descending ordering with their corresponding eigenvectors. However, when data are collected longitudinally in many scientific applications, the eigenvalues become dynamic over time, and the ordering of them may not be well-defined, for instance, does not coincide with the pointwise ordering. To deal with this issue, we propose a new framework, namely the dynamic principal component analysis. This addresses the identifiability of principal components from a global perspective, and transforms the problem into a regression model for data situated on the orthogonal matrix group space. The one-step unrolling method is exploited to solve the regression problem with a suitably constructed regular base curve. The minimax rate of the proposed estimators is established through theoretical analysis of the one-step unrolling and its connection to smoothing spline in the context of manifold-valued data.

Keywords: Eigenvalue switch; discrete data; one-step unrolling; orthogonal matrix group; smoothing spline

References

Bordin, C.J. and Bruno, M.G.S. (2020). Particle filtering on the complex Stiefel manifold with application to subspace tracking. In *ICASSP* 5485–5489.

Brillinger, D.R. (1981). *Time Series: Data Analysis and Theory*, 2nd ed. *Holden-Day Series in Time Series Analysis*. Oakland, CA: Holden-Day, Inc. MR0595684

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

Chen, E.Z. and Li, H. (2016). A two-part mixed-effects model for analyzing longitudinal microbiome compositional data. *Bioinformatics* **32** 2611–2617.

Clark, R.M. and Thompson, R. (1984). Statistical comparison of palaeomagnetic directional records from lake sediments. *Geophys. J. Int.* **76** 337–368.

do Carmo, M.P. (1992). *Riemannian Geometry*. *Mathematics: Theory & Applications*. Boston, MA: Birkhäuser, Inc. MR1138207 https://doi.org/10.1007/978-1-4757-2201-7

Gilmore, R. (2012). *Lie groups, Lie Algebras, and Some of Their Applications*. Courier Corporation.

Green, P.J. and Silverman, B.W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. *Monographs on Statistics and Applied Probability* **58**. London: CRC Press. MR1270012 https://doi.org/10.1007/978-1-4899-4473-3

Groisser, D., Jung, S. and Schwartzman, A. (2017). Geometric foundations for scaling-rotation statistics on symmetric positive definite matrices: Minimal smooth scaling-rotation curves in low dimensions. *Electron. J. Stat.* **11** 1092–1159. MR3631822 https://doi.org/10.1214/17-EJS1250

Hatcher, A. (2005). *Algebraic Topology*. Tsinghua University Press.

Hörmann, S., Kidziński, Ł. and Hallin, M. (2015). Dynamic functional principal components. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 319–348. MR3310529 https://doi.org/10.1111/rssb.12076

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

Hu, X. and Yao, F. (2024). Dynamic principal component analysis in high dimensions. *J. Amer. Statist. Assoc.* **119** 308–319. MR4713894 https://doi.org/10.1080/01621459.2022.2115917

Jakubiak, J., Silva Leite, F. and Rodrigues, R.C. (2006). A two-step algorithm of smooth spline generation on Riemannian manifolds. *J. Comput. Appl. Math.* **194** 177–191. MR2239067 https://doi.org/10.1016/j.cam.2005.07.003

Jolliffe, I.T. (2002). *Principal Component Analysis*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2036084

Jupp, P.E. and Kent, J.T. (1987). Fitting smooth paths to spherical data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **36** 34–46. MR0887825 https://doi.org/10.2307/2347843

Kim, K.-R., Dryden, I.L., Le, H. and Severn, K.E. (2021). Smoothing splines on Riemannian manifolds, with applications to 3D shape space. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **83** 108–132. MR4220986 https://doi.org/10.1111/rssb.12402

Knapp, A.W. (2013). *Lie groups Beyond an Introduction*. Springer Science & Business Media.

Le, H. (2003). Unrolling shape curves. *J. Lond. Math. Soc. (2)* **68** 511–526. MR1994697 https://doi.org/10.1112/S0024610703004393

Machado, L. and Leite, F.S. (2006). Fitting smooth paths on Riemannian manifolds. *Int. J. Appl. Math. Stat.* **4** 25–53. MR2262660

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philos. Mag.* **2** 559–572.

Rodrigues, R.C., Silva Leite, F. and Jakubiak, J. (2005). A new geometric algorithm to generate smooth interpolating curves on Riemannian manifolds. *LMS J. Comput. Math.* **8** 251–266. MR2193213 https://doi.org/10.1112/S146115700000098X

Rui, C. and Leite, F.S. (2007). Interpolation problems on Riemannian manifolds – a geometric approach. *IFAC Proc. Vol.* **40** 822–827.

Samir, C., Absil, P.-A., Srivastava, A. and Klassen, E. (2012). A gradient-descent method for curve fitting on Riemannian manifolds. *Found. Comput. Math.* **12** 49–73. MR2886156 https://doi.org/10.1007/s10208-011-9091-7

Shampine, L.F., Gladwell, I. and Thompson, S. (2003). *Solving ODEs with Matlab*. Cambridge: Cambridge Univ. Press. MR1985643 https://doi.org/10.1017/CBO9780511615542

Shao, L. and Yao, F. (2024). Supplement to "Dynamic principal component analysis from a global perspective." https://doi.org/10.3150/24-BEJ1743SUPP

Srivastava, A. and Klassen, E. (2004). Bayesian and geometric subspace tracking. *Adv. in Appl. Probab.* **36** 43–56. MR2035773 https://doi.org/10.1239/aap/1077134463

Yao, F., Müller, H.-G. and Wang, J.-L. (2005). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. MR2160561 https://doi.org/10.1198/016214504000001745

Yuan, Y., Zhu, H., Lin, W. and Marron, J.S. (2012). Local polynomial regression for symmetric positive definite matrices. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 697–719. MR2965956 https://doi.org/10.1111/j.1467-9868.2011.01022.x

Zhu, H., Chen, Y., Ibrahim, J.G., Li, Y., Hall, C. and Lin, W. (2009). Intrinsic regression models for positive-definite matrices with applications to diffusion tensor imaging. *J. Amer. Statist. Assoc.* **104** 1203–1212. MR2750245 https://doi.org/10.1198/jasa.2009.tm08096

Bernoulli **31**(1), 2025, 671–691
https://doi.org/10.3150/24-BEJ1744

No eigenvalues outside the support of the limiting spectral distribution of large dimensional noncentral sample covariance matrices

ZHIDONG BAI[1,2,a], JIANG HU[1,b], JACK W. SILVERSTEIN[3,d] and HUANCHAO ZHOU[1,c]

[1]*KLASMOE and School of Mathematics and Statistics, Northeast Normal University, China,* [a]*baizd@nenu.edu.cn,* [b]*huj156@nenu.edu.cn,* [c]*zhouhc782@nenu.edu.cn*
[2]*School of Mathematics and Statistics, Xi'an Jiaotong University, China*
[3]*Department of Mathematics, North Carolina State University, USA,* [d]*jack@ncsu.edu*

Let $\mathbf{B}_n = n^{-1}(\mathbf{R}_n + \mathbf{T}_n^{1/2}\mathbf{X}_n)(\mathbf{R}_n + \mathbf{T}_n^{1/2}\mathbf{X}_n)^*$, where $\mathbf{X}_n$ is a $p \times n$ matrix with independent standardized random variables, $\mathbf{R}_n$ is a $p \times n$ non-random matrix and $\mathbf{T}_n$ is a $p \times p$ non-random, nonnegative definite Hermitian matrix. The matrix $\mathbf{B}_n$ is referred to as the information-plus-noise type matrix, where $\mathbf{R}_n$ contains the information and $\mathbf{T}_n^{1/2}\mathbf{X}_n$ is the noise matrix with the covariance matrix $\mathbf{T}_n$. It is known that, as $n \to \infty$, if p/n converges to a positive number, the empirical spectral distribution of $\mathbf{B}_n$ converges almost surely to a nonrandom limit, under some conditions. In this paper, we prove that, under certain conditions on the eigenvalues of $\mathbf{R}_n$ and $\mathbf{T}_n$, for any closed interval outside the support of the limit spectral distribution, with probability one there will be no eigenvalues falling in this interval for all n sufficiently large.

Keywords: Information-plus-noise matrix; limiting spectral distribution; random matrix; Stieltjes transform

References

[1] Bai, Z., Hu, J., Silverstein, J.W. and Zhou, H. (2024). Supplement to "No eigenvalues outside the support of the limiting spectral distribution of large dimensional noncentral sample covariance matrices." https://doi.org/10.3150/24-BEJ1744SUPP

[2] Bai, Z.D. and Silverstein, J.W. (1998). No eigenvalues outside the support of the limiting spectral distribution of large-dimensional sample covariance matrices. *Ann. Probab.* **26** 316–345. MR1617051 https://doi.org/10.1214/aop/1022855421

[3] Bai, Z. and Silverstein, J.W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2567175 https://doi.org/10.1007/978-1-4419-0661-8

[4] Bai, Z. and Silverstein, J.W. (2012). No eigenvalues outside the support of the limiting spectral distribution of information-plus-noise type matrices. *Random Matrices Theory Appl.* **1** 1150004. MR2930382 https://doi.org/10.1142/S2010326311500043

[5] Belinschi, S., Bordenave, C., Capitaine, M. and Cébron, G. (2021). Outlier eigenvalues for non-Hermitian polynomials in independent i.i.d. matrices and deterministic matrices. *Electron. J. Probab.* **26** 100. MR4283371 https://doi.org/10.1214/21-EJP666

[6] Capitaine, M. (2014). Exact separation phenomenon for the eigenvalues of large information-plus-noise type matrices, and an application to spiked models. *Indiana Univ. Math. J.* **63** 1875–1910. MR3298725 https://doi.org/10.1512/iumj.2014.63.5432

[7] Couillet, R., Debbah, M. and Silverstein, J.W. (2011). A deterministic equivalent for the analysis of correlated MIMO multiple access channels. *IEEE Trans. Inf. Theory* **57** 3493–3514. MR2817033 https://doi.org/10.1109/TIT.2011.2133151

[8] Horn, R.A. and Johnson, C.R. (1985). *Matrix Analysis*. Cambridge: Cambridge Univ. Press. MR0832183 https://doi.org/10.1017/CBO9780511810817

[9] Zhou, H., Bai, Z. and Hu, J. (2023). The limiting spectral distribution of large-dimensional general information-plus-noise-type matrices. *J. Theoret. Probab.* **36** 1203–1226. MR4591871 https://doi.org/10.1007/s10959-022-01193-x

[10] Zhou, H.C., Hu, J., Bai, Z.D. and Silverstein, J.W. (2023). Analysis of the limiting spectral distribution of large dimensional general information-plus-noise type matrices. *J. Theor. Probab.* (forthcoming).

Bernoulli **31**(1), 2025, 692–708
https://doi.org/10.3150/24-BEJ1746

Some new concentration inequalities for the Itô stochastic integral

NGUYEN TIEN DUNG[a]

Department of Mathematics, University of Science, Vietnam National University, Hanoi, 334 Nguyen Trai, Thanh Xuan, Hanoi, 084 Vietnam, [a]*dung@hus.edu.vn*

In this paper, based on the techniques of Malliavin calculus, we provide some new concentration inequalities for the running supremum of the Itô stochastic integral with unbounded integrands. Several applications and examples are provided as well.

Keywords: Concentration inequality; Itô stochastic integral; Malliavin calculus

References

Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. *Springer Monographs in Mathematics*. New York: Springer. MR2319516

Alòs, E. and Ewald, C.-O. (2008). Malliavin differentiability of the Heston volatility and applications to option pricing. *Adv. in Appl. Probab.* **40** 144–162. MR2411818 https://doi.org/10.1239/aap/1208358890

Breton, J.-C., Houdré, C. and Privault, N. (2007). Dimension free and infinite variance tail estimates on Poisson space. *Acta Appl. Math.* **95** 151–203. MR2317609 https://doi.org/10.1007/s10440-007-9084-3

Gerhold, S., Hubalek, F. and Paris, R.B. (2022). The running maximum of the Cox-Ingersoll-Ross process with some properties of the Kummer function. *J. Inequal. Spec. Funct.* **13** 1–18. MR4449100 https://doi.org/10.54379/jiasf

Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2

Latała, R. (2006). Estimates of moments and tails of Gaussian chaoses. *Ann. Probab.* **34** 2315–2331. MR2294983 https://doi.org/10.1214/009117906000000421

Major, P. (2005). Tail behaviour of multiple random integrals and U-statistics. *Probab. Surv.* **2** 448–505. MR2203678 https://doi.org/10.1214/154957805100000186

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

Üstünel, A.S. (2015). *Analysis on Wiener Space and Applications.* Available at arXiv:1003.1649v2.

Bernoulli **31**(1), 2025, 709–730
https://doi.org/10.3150/24-BEJ1747

A log-linear model for non-stationary time series of counts

ANNE LEUCHT[1,a] and MICHAEL H. NEUMANN[2,b]

[1] *Universität Bamberg, Institut für Statistik, Feldkirchenstraße 21, D – 96052 Bamberg, Germany,*
[a] *anne.leucht@uni-bamberg.de*
[2] *Friedrich-Schiller-Universität Jena, Institut für Mathematik, Ernst-Abbe-Platz 2, D – 07743 Jena, Germany,*
[b] *michael.neumann@uni-jena.de*

We propose a new model for non-stationary integer-valued time series which is particularly suitable for data with a strong trend. In contrast to popular Poisson-INGARCH models, but in line with classical GARCH models, we propose to pick the conditional distributions from nearly scale invariant families where the mean absolute value and the standard deviation are of the same order of magnitude. As an important prerequisite for applications in statistics, we prove absolute regularity of the count process with exponentially decaying coefficients.

Keywords: Absolute regularity; count process; log-linear model; mixing; nonstationary process

References

Aagard, K., Lyons, J.E. and Thogmartin, W.E. (2018). Accounting for surveyor effort in large-scale monitoring programs. *J. Fish Wildl. Manag.* **9** 459–466.

Boussama, F. (1998). Ergodicité, mélange et estimation dans les modèles GARCH Ph.D. thesis, Université Paris 7.

Cameron, A.C. and Trivedi, P.K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *J. Appl. Econometrics* **1** 29–53.

Carrasco, M. and Chen, X. (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18** 17–39. MR1885348 https://doi.org/10.1017/S0266466602181023

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., Leucht, A. and Neumann, M.H. (2022). Mixing properties of non-stationary INGARCH(1, 1) processes. *Bernoulli* **28** 663–688. MR4337720 https://doi.org/10.3150/21-bej1362

Doukhan, P. and Neumann, M.H. (2019). Absolute regularity of semi-contractive GARCH-type processes. *J. Appl. Probab.* **56** 91–115. MR3981148 https://doi.org/10.1017/jpr.2019.8

Fokianos, K. and Tjøstheim, D. (2011). Log-linear Poisson autoregression. *J. Multivariate Anal.* **102** 563–578. MR2755016 https://doi.org/10.1016/j.jmva.2010.11.002

Francq, C. and Zakoïan, J.-M. (2006). Mixing properties of a general class of GARCH(1,1) models without moment assumptions on the observed process. *Econometric Theory* **22** 815–834. MR2291218 https://doi.org/10.1017/S0266466606060373

Kremer, C., Torneri, A., Boesmans, S. et al. (2021). Quantifying superspreading for COVID-19 using Poisson mixture distributions. *Sci. Rep.* **11** 14107.

Lindvall, T. (1992). *Lectures on the Coupling Method*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR1180522

Neumann, M.H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17** 1268–1284. MR2854772 https://doi.org/10.3150/10-BEJ313

Neumann, M.H. (2021). Bootstrap for integer-valued GARCH(p,q) processes. *Stat. Neerl.* **75** 343–363. MR4281657 https://doi.org/10.1111/stan.12238

Powell, H., Krall, J.R., Wang, Y., Bell, M.L. and Peng, R.D. (2015). Ambient coarse particulate matter and hospital admissions in the medicare cohort air pollution study, 1999-2010. *Environ. Health Perspect.* **123** 11.

1350-7265

Renshaw, A. and Haberman, S. (2003). Lee-Carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections. *J. R. Stat. Soc., Ser. C* **52** 119–137. MR1959085 https://doi.org/10.1111/1467-9876.00393

Rio, E. (1995). About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Stat.* **1** 35–61. MR1382517 https://doi.org/10.1051/ps:1997102

Shao, X. (2010). The dependent wild bootstrap. *J. Amer. Statist. Assoc.* **105** 218–235. MR2656050 https://doi.org/10.1198/jasa.2009.tm08744

Weiss, C.H. (2018). *An Introduction to Discrete-Valued Time Series*. Hoboken, NJ: Wiley.

Xu, H.-Y., Xie, M., Goh, T.N. and Fu, X. (2012). A model for integer-valued time series with conditional overdispersion. *Comput. Statist. Data Anal.* **56** 4229–4242. MR2957867 https://doi.org/10.1016/j.csda.2012.04.011

Bernoulli **31**(1), 2025, 731–758
https://doi.org/10.3150/24-BEJ1748

Irregular nonparametric autoregression

H. GRUBER[1,a] and M. JIRAK[2,b]

[1]*TU Braunschweig, Braunschweig, Germany, Universitätsplatz 2,* [a]*h.gruber@tu-bs.de*
[2]*Vienna University, Vienna, Austria, Oskar-Morgensternplatz 1, Wien,* [b]*moritz.jirak@univie.ac.at*

Both locally stationary processes and irregular models have had a long story of success in statistics and time series analysis. We combine both concepts and consider a nonparametric, first-order autoregressive model with irregular, positive innovations, where we assume that the coefficient function is Hölder continuous and positive. To estimate this function, we use a quasi-maximum likelihood based approach. A precise control of this method demands a delicate analysis of extremes of certain weakly dependent processes, our main result being a concentration inequality for such quantities. Based on our analysis, upper and matching minimax lower bounds are derived, showing the optimality of our estimators. Unlike the regular case, the information theoretic complexity depends both on the smoothness and an additional shape parameter, characterizing the irregularity of the underlying distribution. The results and ideas for the proofs are very different from classical and more recent methods in connection with locally stationary processes.

Keywords: Irregular models; local stationarity; nonparametric autoregression

References

Aigner, D., Lovell, C.A.K. and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *J. Econometrics* **6** 21–37. MR0448782 https://doi.org/10.1016/0304-4076(77)90052-5

Balakrishna, N. (2021). *Non-Gaussian Autoregressive-type Time Series*. Singapore: Springer. MR4397509 https://doi.org/10.1007/978-981-16-8162-2

Barndorff-Nielsen, O.E. and Shephard, N. (2001). Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 167–241. MR1841412 https://doi.org/10.1111/1467-9868.00282

Bibinger, M., Jirak, M. and Reiss, M. (2016). Volatility estimation under one-sided errors with applications to limit order books. *Ann. Appl. Probab.* **26** 2754–2790. MR3563193 https://doi.org/10.1214/15-AAP1161

Bondon, P. (2009). Estimation of autoregressive models with epsilon-skew-normal innovations. *J. Multivariate Anal.* **100** 1761–1776. MR2535385 https://doi.org/10.1016/j.jmva.2009.02.006

Brockwell, P.J. and Davis, R.A. (2006). *Time Series: Theory and Methods*. *Springer Series in Statistics*. New York: Springer. Reprint of the second (1991) edition. MR2839251

Dahlhaus, R. (2012). Locally stationary processes. In *Handbook of Statistics* **30** 351. Amsterdam: Elsevier.

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

Daouia, A., Florens, J.-P. and Simar, L. (2021). Robustified expected maximum production frontiers. *Econometric Theory* **37** 346–387. MR4246220 https://doi.org/10.1017/S0266466620000171

Davis, R.A., Knight, K. and Liu, J. (1992). M-estimation for autoregressions with infinite variance. *Stochastic Process. Appl.* **40** 145–180. MR1145464 https://doi.org/10.1016/0304-4149(92)90142-D

Davis, R.A. and McCormick, W.P. (1989). Estimation for first-order autoregressive processes with positive or bounded innovations. *Stochastic Process. Appl.* **31** 237–250. MR0998115 https://doi.org/10.1016/0304-4149(89)90090-2

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

Delleur, J.W., Gupta, V.K., Bhattacharya, R.N., Klemeš, V., Smith, R.L., Todorovic, P., Deutsch, S.J., Ramos, J.A., Nguyen, V.T.V., Pickard, D.K., Tory, E.M., Smith, J.A., Karr, A.F. and Woolford, S.W. (1984). Stochastic hydrology. *Adv. in Appl. Probab.* **16** 17–23.

Diaz, H.F., Hughes, M.K. and Swetnam, T.W. (2010). *Dendroclimatology: Progress and Prospects*. Netherlands: Springer.

Durbin, J. (1959). Efficient estimation of parameters in moving-average models. *Biometrika* **46** 306–316. MR0114283 https://doi.org/10.1093/biomet/46.3-4.306

Eichler, M., Motta, G. and von Sachs, R. (2011). Fitting dynamic factor models to non-stationary time series. *J. Econometrics* **163** 51–70. MR2803666 https://doi.org/10.1016/j.jeconom.2010.11.007

Eriksson, A., Preve, D.P.A. and Yu, J. (2019). Forecasting realized volatility using a nonnegative semiparametric model. *J. Financ. Risk Manag.* **12**.

Farrell, M.J. (1957). The measurement of productive efficiency. *J. Roy. Statist. Soc. Ser. A* **120** 253–290.

Feigin, P.D., Kratz, M.F. and Resnick, S.I. (1996). Parameter estimation for moving averages with positive innovations. *Ann. Appl. Probab.* **6** 1157–1190. MR1422981 https://doi.org/10.1214/aoap/1035463327

Feigin, P.D. and Resnick, S.I. (1992). Estimation for autoregressive processes with positive innovations. *Commun. Stat., Stoch. Models* **8** 685–717. MR1182425 https://doi.org/10.1080/15326349208807235

Feigin, P.D. and Resnick, S.I. (1994). Limit distributions for linear programming time series estimators. *Stochastic Process. Appl.* **51** 135–165. MR1288286 https://doi.org/10.1016/0304-4149(94)90022-1

Gaver, D.P. and Lewis, P.A.W. (1980). First-order autoregressive gamma sequences and point processes. *Adv. in Appl. Probab.* **12** 727–745. MR0578846 https://doi.org/10.2307/1426429

Hall, P. and Van Keilegom, I. (2009). Nonparametric "regression" when errors are positioned at end-points. *Bernoulli* **15** 614–633. MR2555192 https://doi.org/10.3150/08-BEJ173

Hsiao, W.-C., Huang, H.-Y. and Ing, C.-K. (2018). Interval estimation for a first-order positive autoregressive process. *J. Time Series Anal.* **39** 447–467. MR3796529 https://doi.org/10.1111/jtsa.12297

Ing, C.-K. and Yang, C.-Y. (2014). Predictor selection for positive autoregressive processes. *J. Amer. Statist. Assoc.* **109** 243–253. MR3180560 https://doi.org/10.1080/01621459.2013.836974

Jirak, M., Meister, A. and Reiss, M. (2014). Adaptive function estimation in nonparametric regression with one-sided errors. *Ann. Statist.* **42** 1970–2002. MR3262474 https://doi.org/10.1214/14-AOS1248

Knight, K. (2001). Limiting distributions of linear programming estimators. *Extremes* **4** 87–103. MR1893873 https://doi.org/10.1023/A:1013991808181

Kreiss, J.-P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133. MR0885727 https://doi.org/10.1214/aos/1176350256

Kreiss, J.-P. and Paparoditis, E. (2015). Bootstrapping locally stationary processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 267–290. MR3299408 https://doi.org/10.1111/rssb.12068

Kumbhakar, S.C., Park, B.U., Simar, L. and Tsionas, E.G. (2007). Nonparametric stochastic frontiers: A local maximum likelihood approach. *J. Econometrics* **137** 1–27. MR2347943 https://doi.org/10.1016/j.jeconom.2006.03.006

Lawrance, A.J. and Lewis, P.A.W. (1980). The exponential autoregressive-moving average EARMA(p, q) process. *J. Roy. Statist. Soc. Ser. B* **42** 150–161. MR0583349

Leadbetter, M.R. (1974). On extreme values in stationary sequences. *Z. Wahrsch. Verw. Gebiete* **28** 289–303. MR0362465 https://doi.org/10.1007/BF00532947

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. *Springer Series in Statistics*. New York-Berlin: Springer. MR0691492

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *Internat. Econom. Rev.* **18** 435–444.

Nielsen, B. and Shephard, N. (2003). Likelihood analysis of a first-order autoregressive model with exponential innovations. *J. Time Series Anal.* **24** 337–344. MR1984600 https://doi.org/10.1111/1467-9892.00310

Park, B.U., Sickles, R.C. and Simar, L. (1998). Stochastic panel frontiers: A semiparametric approach. *J. Econometrics* **84** 273–301. MR1630206 https://doi.org/10.1016/S0304-4076(97)00087-0

Park, B.U., Sickles, R.C. and Simar, L. (2007). Semiparametric efficient estimation of dynamic panel data models. *J. Econometrics* **136** 281–301. MR2328594 https://doi.org/10.1016/j.jeconom.2006.03.004

Preve, D. (2015). Linear programming-based estimators in nonnegative autoregression. *J. Bank. Financ.* **61** 225–234. Recent Developments in Financial Econometrics and Applications.

Richter, S. and Dahlhaus, R. (2019). Cross validation for locally stationary processes. *Ann. Statist.* **47** 2145–2173. MR3953447 https://doi.org/10.1214/18-AOS1743

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

Selk, L., Tillier, C. and Marigliano, O. (2022). Multivariate boundary regression models. *Scand. J. Stat.* **49** 400–426. MR4391059 https://doi.org/10.1111/sjos.12519

Smith, R.L. (1994). Nonregular regression. *Biometrika* **81** 173–183. MR1279665 https://doi.org/10.1093/biomet/81.1.173

Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats. MR2724359 https://doi.org/10.1007/b13794

Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614 https://doi.org/10.1214/12-AOS1043

Zhou, J. and Basawa, I.V. (2005). Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors. *J. Time Series Anal.* **26** 825–842. MR2203513 https://doi.org/10.1111/j.1467-9892.2005.00440.x

Bernoulli **31**(1), 2025, 759–782
https://doi.org/10.3150/24-BEJ1749

Transportation cost inequalities for stochastic reaction diffusion equations on the whole real line

YUE LI[a], SHIJIE SHANG[b] and TUSHENG ZHANG[c]

School of Mathematical Sciences, University of Science and technology of China, Hefei, China,
[a]*liyue27@mail.ustc.edu.cn,* [b]*sjshang@ustc.edu.cn,* [c]*tusheng.zhang@manchester.ac.uk*

In this paper, we establish quadratic transportation cost inequalities for solutions of stochastic reaction diffusion equations driven by multiplicative space-time white noise on the whole real line $\mathbb{R}$. Since the space variable is defined on the unbounded domain $\mathbb{R}$, the inequalities are proved under a weighted L^2-norm and a weighted uniform metric in the so-called L^2_{tem}, C_{tem} spaces. The new moments estimates of the stochastic convolution with respect to space-time white noise play an important role. In addition, the transportation cost inequalities are also obtained for the stochastic reaction diffusion equations with random initial values.

Keywords: Concentration of measure; moment estimates for stochastic convolutions; reaction diffusion equations; stochastic partial differential equations; transportation cost inequalities

References

[1] 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

[2] Bobkov, S.G., Gentil, I. and Ledoux, M. (2001). Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)* **80** 669–696. MR1846020 https://doi.org/10.1016/S0021-7824(01)01208-9

[3] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772 https://doi.org/10.1006/jfan.1998.3326

[4] Boufoussi, B. and Hajji, S. (2018). Transportation inequalities for stochastic heat equations. *Statist. Probab. Lett.* **139** 75–83. MR3802186 https://doi.org/10.1016/j.spl.2018.03.012

[5] 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

[6] Dalang, R.C. and Quer-Sardanyons, L. (2011). Stochastic integrals for spde's: A comparison. *Expo. Math.* **29** 67–109. MR2785545 https://doi.org/10.1016/j.exmath.2010.09.005

[7] Djellout, H., Guillin, A. and Wu, L. (2004). Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.* **32** 2702–2732. MR2078555 https://doi.org/10.1214/009117904000000531

[8] Fang, S. and Shao, J. (2005). Transportation cost inequalities on path and loop groups. *J. Funct. Anal.* **218** 293–317. MR2108113 https://doi.org/10.1016/j.jfa.2004.02.002

[9] Feyel, D. and Üstünel, A.S. (2004). Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theory Related Fields* **128** 347–385. MR2036490 https://doi.org/10.1007/s00440-003-0307-x

[10] Gozlan, N., Roberto, C. and Samson, P.-M. (2011). A new characterization of Talagrand's transport-entropy inequalities and applications. *Ann. Probab.* **39** 857–880. MR2789577 https://doi.org/10.1214/10-AOP570

[11] Khoshnevisan, D. (2014). *Analysis of Stochastic Partial Differential Equations*. *CBMS Regional Conference Series in Mathematics* **119**. Providence, RI: Amer. Math. Soc.. MR3222416 https://doi.org/10.1090/cbms/119

[12] Khoshnevisan, D. and Sarantsev, A. (2019). Talagrand concentration inequalities for stochastic partial differential equations. *Stoch. Partial Differ. Equ. Anal. Comput.* **7** 679–698. MR4022285 https://doi.org/10.1007/s40072-019-00136-8

[13] Ledoux, M. (2001). *The Concentration of Measure Phenomenon*. *Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347 https://doi.org/10.1090/surv/089

[14] Ma, Y. (2010). Transportation inequalities for stochastic differential equations with jumps. *Stochastic Process. Appl.* **120** 2–21. MR2565849 https://doi.org/10.1016/j.spa.2009.09.012

[15] Marton, K. (1996). Bounding $\overline{d}$-distance by informational divergence: A method to prove measure concentration. *Ann. Probab.* **24** 857–866. MR1404531 https://doi.org/10.1214/aop/1039639365

[16] Mytnik, L., Perkins, E. and Sturm, A. (2006). On pathwise uniqueness for stochastic heat equations with non-Lipschitz coefficients. *Ann. Probab.* **34** 1910–1959. MR2271487 https://doi.org/10.1214/009117906000000331

[17] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620 https://doi.org/10.1006/jfan.1999.3557

[18] Pal, S. (2012). Concentration for multidimensional diffusions and their boundary local times. *Probab. Theory Related Fields* **154** 225–254. MR2981423 https://doi.org/10.1007/s00440-011-0368-1

[19] Saussereau, B. (2012). Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion. *Bernoulli* **18** 1–23. MR2888696 https://doi.org/10.3150/10-BEJ324

[20] Shang, S. and Wang, R. (2020). Transportation inequalities under uniform metric for a stochastic heat equation driven by time-white and space-colored noise. *Acta Appl. Math.* **170** 81–97. MR4163229 https://doi.org/10.1007/s10440-020-00325-x

[21] Shang, S. and Zhang, T. (2019). Talagrand concentration inequalities for stochastic heat-type equations under uniform distance. *Electron. J. Probab.* **24** Paper No. 129, 15. MR4040989 https://doi.org/10.1214/19-ejp388

[22] Shang, S. and Zhang, T. (2023). Global well-posedness to stochastic reaction-diffusion equations on the real line $\mathbb{R}$ with superlinear drifts driven by multiplicative space-time white noise. *Electron. J. Probab.* **28** Paper No. 166, 29. MR4674636 https://doi.org/10.1214/23-ejp1057

[23] Shiga, T. (1994). Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canad. J. Math.* **46** 415–437. MR1271224 https://doi.org/10.4153/CJM-1994-022-8

[24] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** 587–600. MR1392331 https://doi.org/10.1007/BF02249265

[25] Talagrand, M. (1996). New concentration inequalities in product spaces. *Invent. Math.* **126** 505–563. MR1419006 https://doi.org/10.1007/s002220050108

[26] Üstünel, A.S. (1995). *An Introduction to Analysis on Wiener Space*. *Lecture Notes in Math.* **1610**. Berlin: Springer. MR1439752 https://doi.org/10.1007/BFb0096328

[27] 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

[28] Wang, F. and Zhang, T. (2020). Talagrand inequality on free path space and application to stochastic reaction diffusion equations. *Acta Math. Appl. Sin. Engl. Ser.* **36** 253–261. MR4083440 https://doi.org/10.1007/s10255-020-0926-3

[29] Wang, F.-Y. (2004). Probability distance inequalities on Riemannian manifolds and path spaces. *J. Funct. Anal.* **206** 167–190. MR2024350 https://doi.org/10.1016/S0022-1236(02)00100-3

[30] Wu, L. (2010). Transportation inequalities for stochastic differential equations of pure jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 465–479. MR2667706 https://doi.org/10.1214/09-AIHP320

[31] Wu, L. and Zhang, Z. (2004). Talagrand's T_2-transportation inequality w.r.t. a uniform metric for diffusions. *Acta Math. Appl. Sin. Engl. Ser.* **20** 357–364. MR2086758 https://doi.org/10.1007/s10255-004-0175-x

[32] Wu, L. and Zhang, Z. (2006). Talagrand's T_2-transportation inequality and log-Sobolev inequality for dissipative SPDEs and applications to reaction-diffusion equations. *Chin. Ann. Math. Ser. B* **27** 243–262. MR2239826 https://doi.org/10.1007/s11401-005-0176-y

Bernoulli **31**(1), 2025, 783–815
https://doi.org/10.3150/24-BEJ1750

Maximum likelihood estimation for small noise multi-scale McKean-Vlasov stochastic differential equations

JIE XU[1,a], QIAO ZHENG[1,b] and JIANYONG MU[2,c]

[1]*College of Mathematics and Information Science, Henan Normal University, Xinxiang, China,*
[a]*xujiescu@163.com,* [b]*zhengqiao0329@163.com*
[2]*Department of Science, Henan Institute of Technology, Xinxiang, China,* [c]*jianyongmu@163.com*

In this paper we focus on the problem of the parameter estimation for a class of systems described by multi-scale McKean-Vlasov stochastic differential equations (MVSDEs, for short) with small noise, where the coefficients depend on the slow component, the fast component, and their own distributions. Firstly, we prove an optimal strong convergence rate of the strong averaging principle of the slow process by using the technique of Poisson equation. Secondly, we construct the maximum likelihood estimator (MLE, for short) and prove its consistency and asymptotic normality. Finally, an example is presented to illustrate the theoretical findings.

Keywords: McKean-Vlasov equation; multi-scale system; parameter estimation; small noise

References

Bao, J., Scheutzow, M. and Yuan, C. (2022). Existence of invariant probability measures for functional McKean-Vlasov SDEs. *Electron. J. Probab.* **27** Paper No. 43, 14. MR4404942 https://doi.org/10.1214/22-ejp773

Bishwal, J.P.N. (2008). *Parameter Estimation in Stochastic Differential Equations*. *Lecture Notes in Math.* **1923**. Berlin: Springer. MR2360279 https://doi.org/10.1007/978-3-540-74448-1

Butkovsky, O.A. (2014). On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations. *Theory Probab. Appl.* **58** 661–674. MR3403022 https://doi.org/10.1137/S0040585X97986825

Cardaliaguet, P. (2012). Notes on Mean Field Games.

Chaudru de Raynal, P.-E. and Frikha, N. (2022). Well-posedness for some non-linear SDEs and related PDE on the Wasserstein space. *J. Math. Pures Appl. (9)* **159** 1–167. MR4377993 https://doi.org/10.1016/j.matpur.2021.12.001

Chaudru de Raynal, P.E. (2020). Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.* **130** 79–107. MR4035024 https://doi.org/10.1016/j.spa.2019.01.006

Crisan, D. and McMurray, E. (2018). Smoothing properties of McKean-Vlasov SDEs. *Probab. Theory Related Fields* **171** 97–148. MR3800831 https://doi.org/10.1007/s00440-017-0774-0

Ding, X. and Qiao, H. (2021). Stability for stochastic McKean-Vlasov equations with non-Lipschitz coefficients. *SIAM J. Control Optim.* **59** 887–905. MR4223214 https://doi.org/10.1137/19M1289418

dos Reis, G., Salkeld, W. and Tugaut, J. (2019). Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law. *Ann. Appl. Probab.* **29** 1487–1540. MR3914550 https://doi.org/10.1214/18-AAP1416

Duan, J. and Wang, W. (2014). *Effective Dynamics of Stochastic Partial Differential Equations*. *Elsevier Insights*. Amsterdam: Elsevier. MR3289240

Eberle, A., Guillin, A. and Zimmer, R. (2019). Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.* **371** 7135–7173. MR3939573 https://doi.org/10.1090/tran/7576

Fang, K., Liu, W., Qiao, H. and Zhu, F. (2023). Asymptotic behaviors of small perturbation for multivalued McKean-Vlasov stochastic differential equations. *Appl. Math. Optim.* **88** Paper No. 22, 48. MR4587577 https://doi.org/10.1007/s00245-023-10004-6

Gailus, S. and Spiliopoulos, K. (2017). Statistical inference for perturbed multiscale dynamical systems. *Stochastic Process. Appl.* **127** 419–448. MR3583758 https://doi.org/10.1016/j.spa.2016.06.013

Gao, J., Hong, W. and Liu, W. (2023). Small noise asymptotics of multi-scale McKean-Vlasov stochastic dynamical systems. *J. Differ. Equ.* **364** 521–575. MR4574534 https://doi.org/10.1016/j.jde.2023.03.055

Hao, Z., Röckner, M. and Zhang, X. (2021). Euler scheme for density dependent stochastic differential equations. *J. Differ. Equ.* **274** 996–1014. MR4189000 https://doi.org/10.1016/j.jde.2020.11.018

Hong, W., Li, S. and Liu, W. (2022). Strong convergence rates in averaging principle for slow-fast McKean-Vlasov SPDEs. *J. Differ. Equ.* **316** 94–135. MR4374850 https://doi.org/10.1016/j.jde.2022.01.039

Huang, X. and Wang, F.-Y. (2019). Distribution dependent SDEs with singular coefficients. *Stochastic Process. Appl.* **129** 4747–4770. MR4013879 https://doi.org/10.1016/j.spa.2018.12.012

Ibragimov, I.A. and Has' Minskiĭ, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. *Applications of Mathematics* **16**. New York-Berlin: Springer. MR0620321

Krumscheid, S., Pavliotis, G.A. and Kalliadasis, S. (2013). Semiparametric drift and diffusion estimation for multiscale diffusions. *Multiscale Model. Simul.* **11** 442–473. MR3047437 https://doi.org/10.1137/110854485

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

Kutoyants, Yu. (2012). *Identification of Dynamical Systems with Small Noise*. Berlin: Springer.

Li, Y., Wu, F. and Xie, L. (2024). Poisson Equation on Wasserstein Space and Diffusion Approximations for Multiscale McKean–Vlasov Equation. *SIAM J. Math. Anal.* **56** 1495–1524. MR4711950 https://doi.org/10.1137/22M1536856

Li, Y. and Xie, L. (2023). Functional law of large numbers and central limit theorem for slow-fast McKean-Vlasov equations. *Discrete Contin. Dyn. Syst. Ser. S* **16** 846–877. MR4571162 https://doi.org/10.3934/dcdss.2022213

Liu, M. and Qiao, H. (2022). Parameter estimation of path-dependent McKean-Vlasov stochastic differential equations. *Acta Math. Sci. Ser. B Engl. Ed.* **42** 876–886. MR4411008 https://doi.org/10.1007/s10473-022-0304-8

Liu, W., Song, Y., Zhai, J. and Zhang, T. (2023). Large and moderate deviation principles for McKean-Vlasov SDEs with jumps. *Potential Anal.* **59** 1141–1190. MR4647947 https://doi.org/10.1007/s11118-022-10005-0

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

Mishura, Y. and Veretennikov, A. (2020). Existence and uniqueness theorems for solutions of McKean-Vlasov stochastic equations. *Theory Probab. Math. Statist.* **103** 59–101. MR4421344 https://doi.org/10.1090/tpms/1135

Papavasiliou, A., Pavliotis, G.A. and Stuart, A.M. (2009). Maximum likelihood drift estimation for multiscale diffusions. *Stochastic Process. Appl.* **119** 3173–3210. MR2568270 https://doi.org/10.1016/j.spa.2009.05.003

Pavliotis, G.A. and Stuart, A.M. (2007). Parameter estimation for multiscale diffusions. *J. Stat. Phys.* **127** 741–781. MR2319851 https://doi.org/10.1007/s10955-007-9300-6

Prakasa Rao, B.L.S. (1999). *Statistical Inference for Diffusion Type Processes*. *Kendall's Library of Statistics* **8**. London: Edward Arnold. MR1717690

Qiao, H. and Wei, W. (2022). Efficient filtering for multiscale McKean-Vlasov Stochastic differential equations. Preprint. Available at arXiv:2206.05037.

Röckner, M., Sun, X. and Xie, Y. (2021). Strong convergence order for slow-fast McKean-Vlasov stochastic differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** 547–576. MR4255184 https://doi.org/10.1214/20-aihp1087

Röckner, M. and Zhang, X. (2021). Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli* **27** 1131–1158. MR4255229 https://doi.org/10.3150/20-bej1268

Shen, G., Xiang, J. and Wu, J.-L. (2022). Averaging principle for distribution dependent stochastic differential equations driven by fractional Brownian motion and standard Brownian motion. *J. Differ. Equ.* **321** 381–414. MR4396023 https://doi.org/10.1016/j.jde.2022.03.015

Spiliopoulos, K. and Chronopoulou, A. (2013). Maximum likelihood estimation for small noise multiscale diffusions. *Stat. Inference Stoch. Process.* **16** 237–266. MR3123565 https://doi.org/10.1007/s11203-013-9088-8

Veretennikov, A.Yu. (2006). On ergodic measures for McKean-Vlasov stochastic equations. In *Monte Carlo and Quasi-Monte Carlo Methods 2004* 471–486. Berlin: Springer. MR2208726 https://doi.org/10.1007/3-540-31186-6_29

Wang, F.-Y. (2018). Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.* **128** 595–621. MR3739509 https://doi.org/10.1016/j.spa.2017.05.006

Wen, J., Wang, X., Mao, S. and Xiao, X. (2016). Maximum likelihood estimation of McKean-Vlasov stochastic differential equation and its application. *Appl. Math. Comput.* **274** 237–246. MR3433132 https://doi.org/10.1016/j.amc.2015.11.019

Xu, J., Lian, Q. and Liu, J. (2023). Strong convergence rate of the averaging principle for a class of slow-fast stochastic evolution equations. *Stochastics* **95** 581–614. MR4581949 https://doi.org/10.1080/17442508.2022.2093112

Xu, J., Liu, J., Liu, J. and Miao, Y. (2021). Strong averaging principle for two-time-scale stochastic McKean-Vlasov equations. *Appl. Math. Optim.* **84** S837–S867. MR4316803 https://doi.org/10.1007/s00245-021-09787-3

Bernoulli **31**(1), 2025, 816–841
https://doi.org/10.3150/24-BEJ1751

Strong limit theorems for empirical halfspace depth trimmed regions

ANDRII ILIENKO[1,2,a], ILYA MOLCHANOV[1,b] and RICCARDO TURIN[1,3,c]

[1] *Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, CH-3012 Bern, Switzerland,* [a]*andrii.ilienko@unibe.ch,* [b]*ilya.molchanov@unibe.ch,* [c]*Riccardo_Turin@swissre.com*
[2] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" , Beresteiskyi prosp. 37, 03056, Kyiv, Ukraine*
[3] *Swiss Re Management Ltd, Mythenquai 50/60, 8022 Zurich, Switzerland*

We study empirical variants of the halfspace (Tukey) depth of a probability measure μ, which are obtained by replacing μ with the corresponding weighted empirical measure. We prove analogues of the Marcinkiewicz–Zygmund strong law of large numbers and of the law of the iterated logarithm in terms of set inclusions and for the Hausdorff distance between the theoretical and empirical variants of depth trimmed regions. In the special case of μ being the uniform distribution on a convex body K, the depth trimmed regions are convex floating bodies of K, and we obtain strong limit theorems for their empirical estimators.

Keywords: Convex floating body; empirical measure; halfspace depth; law of the iterated logarithm; strong law of large numbers

References

[1] Alexander, K.S. and Talagrand, M. (1989). The law of the iterated logarithm for empirical processes on Vapnik-Červonenkis classes. *J. Multivariate Anal.* **30** 155–166. MR1003714 https://doi.org/10.1016/0047-259X(89)90093-6
[2] Bárány, I. (1989). Intrinsic volumes and f-vectors of random polytopes. *Math. Ann.* **285** 671–699. MR1027765 https://doi.org/10.1007/BF01452053
[3] Borovkov, A.A. (1998). *Mathematical Statistics*. Amsterdam: Gordon and Breach Science Publishers. Translated from the Russian by A. Moullagaliev and revised by the author. MR1712750
[4] Brunel, V.-E. (2019). Concentration of the empirical level sets of Tukey's halfspace depth. *Probab. Theory Related Fields* **173** 1165–1196. MR3936153 https://doi.org/10.1007/s00440-018-0850-0
[5] Csörgő, M. and Révész, P. (1978). Strong approximations of the quantile process. *Ann. Statist.* **6** 882–894. MR0501290
[6] Dutta, K., Ghosh, A. and Moran, S. (2022). Uniform brackets, containers, and combinatorial Macbeath regions. In *13th Innovations in Theoretical Computer Science Conference*. *LIPIcs. Leibniz Int. Proc. Inform.* **215** Art. No. 59, 10. Wadern: Schloss Dagstuhl. Leibniz-Zent. Inform. MR4459430
[7] Dyckerhoff, R. (2017). Convergence of depths and depth-trimmed regions. Technical Report. Available at arXiv:math:1611.08721.
[8] Fresen, D. (2013). A multivariate Gnedenko law of large numbers. *Ann. Probab.* **41** 3051–3080. MR3127874 https://doi.org/10.1214/12-AOP804
[9] Gut, A. (2013). *Probability: A Graduate Course*, 2nd ed. *Springer Texts in Statistics*. New York: Springer. MR2977961 https://doi.org/10.1007/978-1-4614-4708-5
[10] Ilienko, A. and Molchanov, I. (2018). Limit theorems for multidimensional renewal sets. *Acta Math. Hungar.* **156** 56–81. MR3856902 https://doi.org/10.1007/s10474-018-0806-y
[11] Laketa, P. and Nagy, S. (2022). Halfspace depth for general measures: The ray basis theorem and its consequences. *Statist. Papers* **63** 849–883. MR4426418 https://doi.org/10.1007/s00362-021-01259-8

1350-7265

[12] Leichtweiss, K. (1986). Über eine Formel Blaschkes zur Affinoberfläche. *Studia Sci. Math. Hungar.* **21** 453–474. MR0919390

[13] Massart, P. and Rio, E. (1998). A uniform Marcinkiewicz-Zygmund strong law of large numbers for empirical processes. In *Asymptotic Methods in Probability and Statistics (Ottawa, ON, 1997)* 199–211. Amsterdam: North-Holland. MR1661481 https://doi.org/10.1016/B978-044450083-0/50013-7

[14] Massé, J.-C. (2004). Asymptotics for the Tukey depth process, with an application to a multivariate trimmed mean. *Bernoulli* **10** 397–419. MR2061438 https://doi.org/10.3150/bj/1089206404

[15] Meyer, M. and Reisner, S. (1991). A geometric property of the boundary of symmetric convex bodies and convexity of flotation surfaces. *Geom. Dedicata* **37** 327–337. MR1094695 https://doi.org/10.1007/BF00181409

[16] Mizera, I. and Volauf, M. (2002). Continuity of halfspace depth contours and maximum depth estimators: Diagnostics of depth-related methods. *J. Multivariate Anal.* **83** 365–388. MR1945959 https://doi.org/10.1006/jmva.2001.2053

[17] Molchanov, I. (2017). *Theory of Random Sets*. *Probability Theory and Stochastic Modelling* **87**. London: Springer. Second edition of [MR2132405]. MR3751326

[18] Molchanov, I.S. (1998). A limit theorem for solutions of inequalities. *Scand. J. Stat.* **25** 235–242. MR1614288 https://doi.org/10.1111/1467-9469.00100

[19] Molina-Fructuoso, M. and Murray, R. (2022). Tukey depths and Hamilton-Jacobi differential equations. *SIAM J. Math. Data Sci.* **4** 604–633. MR4418807 https://doi.org/10.1137/21M1411998

[20] Nagy, S., Schütt, C. and Werner, E.M. (2019). Halfspace depth and floating body. *Stat. Surv.* **13** 52–118. MR3973130 https://doi.org/10.1214/19-ss123

[21] Rousseeuw, P.J. and Ruts, I. (1999). The depth function of a population distribution. *Metrika* **49** 213–244. MR1731769

[22] Schneider, R. (2014). *Convex Bodies: The Brunn-Minkowski Theory*, expanded ed. *Encyclopedia of Mathematics and Its Applications* **151**. Cambridge: Cambridge Univ. Press. MR3155183

[23] Schütt, C. and Werner, E. (1990). The convex floating body. *Math. Scand.* **66** 275–290. MR1075144 https://doi.org/10.7146/math.scand.a-12311

[24] Schütt, C. and Werner, E. (1994). Homothetic floating bodies. *Geom. Dedicata* **49** 335–348. MR1270561 https://doi.org/10.1007/BF01264033

[25] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838963

[26] Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians (Vancouver, B.C., 1974)* **2** 523–531. Montreal, QC: Canad. Math. Congr. MR0426989

[27] 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