BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Eight Number One February 2022 ISSN: 1350-7265

CONTENTS

(continued)

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

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Eight Number One February 2022 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI Volume 28 Number 1 February 2022 Pages 1–712 ISI/BS

Volume 28 Number 1 February 2022 ISSN 1350-7265

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Aims and Scope

BERNOULLI is the journal of the Bernoulli Society for Mathematical Statistics and Probability, issued four times per year. The journal provides a comprehensive account of important developments in the fields of statistics and probability, offering an international forum for both theoretical and applied work.

Bernoulli Society for Mathematical Statistics and Probability

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

Meetings: http://www.bernoulli-society.org/index.php/meetings

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

Executive Committee

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

---

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

©2022 International Statistical Institute/Bernoulli Society

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

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

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **28**(1), 2022, 1–32
https://doi.org/10.3150/20-BEJ1317

A general framework for SPDE-based stationary random fields

RICARDO CARRIZO VERGARA[1,4,*], DENIS ALLARD[2] and NICOLAS DESASSIS[3,1,†]

[1]*Arxitek, 4b rue Lemasson Henrion, 77250 Moret sur Loing et Orvanne, France.*
E-mail: [*]*ricardo.carrizo_vergara@arxitek.fr;* [†]*nicolas.desassis@arxitek.fr*
[2]*Biostatistics and Spatial Processes, BioSP, INRAE, 84914, Avignon, France.*
E-mail: denis.allard@inrae.fr
[3]*MINES ParisTech, PSL University, Geosciences, Geostatistics team, 35 rue St Honoré, 77300 Fontainebleau, France.*
E-mail: nicolas.desassis@mines-paristech.fr
[4]*ENSIIE, 1 place de la Résistence, 91000, Evry-Courcouronnes, France.*
E-mail: ricardo.carrizovergara@ensiie.fr; racarriz@uc.cl

This paper presents theoretical advances in the application of the Stochastic Partial Differential Equation (SPDE) approach in geostatistics. We show a general approach to construct stationary models related to a wide class of linear SPDEs, with applications to spatio-temporal models having non-trivial properties. Within the framework of Generalized Random Fields, a criterion for existence and uniqueness of stationary solutions for this class of SPDEs is proposed and proven. Their covariance are then obtained through their spectral measure. We present a result relating the covariance of the solution in the case of a White Noise source term with the covariance in a generic case through convolution. Then, we obtain a variety of SPDE-based stationary random fields. In particular, well-known results regarding the Matérn Model and Markovian models are recovered. A new relationship between the Stein model and a particular SPDE is obtained. New spatio-temporal models obtained from evolution SPDEs of arbitrary temporal derivative order are then obtained, for which properties of separability and symmetry can be controlled. We also obtain results concerning stationary solutions for physically inspired models, such as solutions to the heat equation, the advection-diffusion equation, some Langevin's equations and the wave equation.

Keywords: Evolution equation; generalized random fields; matérn model; space-time geostatistics; SPDE approach; spectral measure; symbol function

References

[1] Åberg, S. and Podgórski, K. (2011). A class of non-Gaussian second order random fields. *Extremes* **14** 187–222. MR2794324 https://doi.org/10.1007/s10687-010-0119-1
[2] Ailliot, P., Baxevani, A., Cuzol, A., Monbet, V. and Raillard, N. (2011). Space-time models for moving fields with an application to significant wave height fields. *Environmetrics* **22** 354–369. MR2843390 https://doi.org/10.1002/env.1061
[3] Anh, V.V., Angulo, J.M. and Ruiz-Medina, M.D. (1999). Possible long-range dependence in fractional random fields. *J. Statist. Plann. Inference* **80** 95–110. MR1713795 https://doi.org/10.1016/S0378-3758(98)00244-4
[4] Benoit, L., Allard, D. and Mariethoz, G. (2018). Stochastic rainfall modelling at sub-kilometer scale. *Water Resour. Res.* **54** 4108–4130.
[5] Bolin, D. (2014). Spatial Matérn fields driven by non-Gaussian noise. *Scand. J. Stat.* **41** 557–579. MR3249417 https://doi.org/10.1111/sjos.12046
[6] Bolin, D. and Kirchner, K. (2020). The rational SPDE approach for Gaussian random fields with general smoothness. *J. Comput. Graph. Statist.* **29** 274–285. MR4116041 https://doi.org/10.1080/10618600.2019.1665537

 © 2022 ISI/BS

[7] Bolin, D. and Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann. Appl. Stat.* **5** 523–550. MR2810408 https://doi.org/10.1214/10-AOAS383

[8] Brown, P.E., Kåresen, K.F., Roberts, G.O. and Tonellato, S. (2000). Blur-generated non-separable space-time models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **62** 847–860. MR1796297 https://doi.org/10.1111/1467-9868.00269

[9] Cameletti, M., Lindgren, F., Simpson, D. and Rue, H. (2011). Using the SPDE approach for air quality mapping in Piemonte region. In *Spatial*2 *Conference*: *Spatial Data Methods for Environmental and Ecological Processes*, *Foggia* (*IT*), 1–2 *September* 2011.

[10] Cameletti, M., Lindgren, F., Simpson, D. and Rue, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *AStA Adv. Stat. Anal.* **97** 109–131. MR3045763 https://doi.org/10.1007/s10182-012-0196-3

[11] Chilès, J.-P. and Delfiner, P. (2012). *Geostatistics*: *Modeling Spatial Uncertainty*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR2850475 https://doi.org/10.1002/9781118136188

[12] De Iaco, S., Myers, D.E. and Posa, D. (2001). Space-time analysis using a general product-sum model. *Statist. Probab. Lett.* **52** 21–28. MR1820046 https://doi.org/10.1016/S0167-7152(00)00200-5

[13] De Iaco, S., Myers, D.E. and Posa, D. (2002). Nonseparable space-time covariance models: Some parametric families. *Math. Geol.* **34** 23–42. MR1954124 https://doi.org/10.1023/A:1014075310344

[14] Demengel, F. and Demengel, G. (2000). *Mesures et Distributions*. *Théorie et Illustrations Par les Exemples*. Paris: Ellipses Universités.

[15] Dierolf, P. and Voigt, J. (1978). Convolution and S'-convolution of distributions. *Collect. Math.* **29** 185–196. MR0565276

[16] Dong, A. (1990). Estimation géostatistique des phénomènes régis par des équations aux dérivées partielles (Doctoral dissertation. Paris, ENMP.

[17] Donoghue, W.F. Jr. (1969). *Distributions and Fourier Transforms*. *Pure and Applied Mathematics* **32**. New York: Academic Press. MR3363413

[18] Fuglstad, G.-A., Lindgren, F., Simpson, D. and Rue, H. (2015). Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. *Statist. Sinica* **25** 115–133. MR3328806

[19] Gay, R. and Heyde, C.C. (1990). On a class of random field models which allows long range dependence. *Biometrika* **77** 401–403. MR1064814 https://doi.org/10.1093/biomet/77.2.401

[20] Gel'fand, I.M. and Vilenkin, N.Ya. (1964). *Generalized Functions*. *Vol.* 4: *Applications of Harmonic Analysis*. New York - London: Academic Press. MR0173945

[21] Gel'fand, I.M. (1955). Generalized random processes. *Dokl. Akad. Nauk SSSR* **100** 853–856. MR0068769

[22] Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97** 590–600. MR1941475 https://doi.org/10.1198/016214502760047113

[23] Gneiting, T., Genton, M.G. and Guttorp, P. (2006). Geostatistical space-time models, stationarity, separability, and full symmetry. *Monogr. Statist. Appl. Probab.* **107** 151.

[24] Gradshteyn, I.S. and Ryzhik, I.M. (1996). *Table of Integrals*, *Series*, *and Products*, 5th ed. San Diego, CA: Academic Press. MR1398882

[25] Heine, V. (1955). Models for two-dimensional stationary stochastic processes. *Biometrika* **42** 170–178. MR0071673 https://doi.org/10.1093/biomet/42.1-2.170

[26] Herrmann, L., Kirchner, K. and Schwab, C. (2020). Multilevel approximation of Gaussian random fields: Fast simulation. *Math. Models Methods Appl. Sci.* **30** 181–223. MR4059369 https://doi.org/10.1142/s0218202520500050

[27] Hörmander, L. (2007). *The Analysis of Linear Partial Differential Operators*. *III*: *Pseudo-Differential Operators*. *Classics in Mathematics*. Berlin: Springer. MR2304165 https://doi.org/10.1007/978-3-540-49938-1

[28] Hristopulos, D.T. and Tsantili, I.C. (2016). Space-time models based on random fields with local interactions. *Internat. J. Modern Phys. B* **30** 1541007, 26. MR3513397 https://doi.org/10.1142/S0217979215410076

[29] Huang, J., Malone, B.P., Minasny, B., McBratney, A.B. and Triantafilis, J. (2017). Evaluating a Bayesian modelling approach (INLA-SPDE) for environmental mapping. *Sci. Total Environ.* **609** 621–632.

[30] Itô, K. (1954). Stationary random distributions. *Mem. Coll. Sci., Univ. Kyoto*, *Ser. A*: *Math.* **28** 209–223. MR0065060 https://doi.org/10.1215/kjm/1250777359

[31] Jones, R.H. and Zhang, Y. (1997). Models for continuous stationary space-time processes. In *Modelling Longitudinal and Spatially Correlated Data* 289–298. New York, NY: Springer.
[32] Kelbert, M.Ya., Leonenko, N.N. and Ruiz-Medina, M.D. (2005). Fractional random fields associated with stochastic fractional heat equations. *Adv. in Appl. Probab.* **37** 108–133. MR2135156 https://doi.org/10.1239/aap/1113402402
[33] Lang, A. and Potthoff, J. (2011). Fast simulation of Gaussian random fields. *Monte Carlo Methods Appl.* **17** 195–214. MR2846495 https://doi.org/10.1515/MCMA.2011.009
[34] Lindgren, F., Rue, H. and Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. MR2853727 https://doi.org/10.1111/j.1467-9868.2011.00777.x
[35] Lindgren, G. (2013). *Stationary Stochastic Processes*: *Theory and Applications*. *Chapman & Hall/CRC Texts in Statistical Science Series*. Boca Raton, FL: CRC Press. MR3024784
[36] Liu, X., Guillas, S. and Lai, M.-J. (2016). Efficient spatial modeling using the SPDE approach with bivariate splines. *J. Comput. Graph. Statist.* **25** 1176–1194. MR3572035 https://doi.org/10.1080/10618600.2015.1081597
[37] Mainardi, F., Luchko, Y. and Pagnini, G. (2001). The fundamental solution of the space-time fractional diffusion equation. *Fract. Calc. Appl. Anal.* **4** 153–192. MR1829592
[38] Matheron, G. (1965). Les variables régionalisées et leur estimation: Une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson et CIE.
[39] Mena, H. and Pfurtscheller, L. (2019). An efficient SPDE approach for El Niño. *Appl. Math. Comput.* **352** 146–156. MR3910814 https://doi.org/10.1016/j.amc.2019.01.071
[40] Opitz, T. (2017). Latent Gaussian modeling and INLA: A review with focus on space-time applications. *J. SFdS* **158** 62–85. MR3720130
[41] Pereira, M. (2019). Generalized random fields on Riemannian manifolds: Theory and practice Doctoral dissertation, PSL Research University.
[42] Porcu, E., Mateu, J. and Christakos, G. (2009). Quasi-arithmetic means of covariance functions with potential applications to space-time data. *J. Multivariate Anal.* **100** 1830–1844. MR2535390 https://doi.org/10.1016/j.jmva.2009.02.013
[43] Reed, M. and Simon, B. (1980). *Methods of Modern Mathematical Physics. I. Functional Analysis*. London: Academic Press.
[44] Rozanov, Ju.A. (1977). Markovian random fields, and stochastic partial differential equations. *Math. USSR, Sb.* **32** 515–534.
[45] Rozanov, Yu.A. (1982). *Markov Random Fields*. *Applications of Mathematics*. New York-Berlin: Springer. MR0676644
[46] Rudin, W. (1987). *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill. MR0924157
[47] Ruiz-Medina, M.D., Angulo, J.M., Christakos, G. and Fernández-Pascual, R. (2016). New compactly supported spatiotemporal covariance functions from SPDEs. *Stat. Methods Appl.* **25** 125–141. MR3460490 https://doi.org/10.1007/s10260-015-0333-8
[48] Schwartz, L. (1966). *Théorie des Distributions*. Paris: Hermann.
[49] Sigrist, F., Künsch, H.R. and Stahel, W.A. (2015). Stochastic partial differential equation based modelling of large space-time data sets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 3–33. MR3299397 https://doi.org/10.1111/rssb.12061
[50] Simpson, D., Lindgren, F. and Rue, H. (2012). In order to make spatial statistics computationally feasible, we need to forget about the covariance function. *Environmetrics* **23** 65–74. MR2873784 https://doi.org/10.1002/env.1137
[51] Stein, M.L. (2005). Space-time covariance functions. *J. Amer. Statist. Assoc.* **100** 310–321. MR2156840 https://doi.org/10.1198/016214504000000854
[52] Trèves, F. (1967). *Topological Vector Spaces*, *Distributions and Kernels*. New York: Academic Press. MR0225131
[53] Vecchia, A.V. (1985). A general class of models for stationary two-dimensional random processes. *Biometrika* **72** 281–291. MR0801769 https://doi.org/10.1093/biomet/72.2.281
[54] Whittle, P. (1963). Stochastic processes in several dimensions. *Bull. Int. Stat. Inst.* **40** 974–994. MR0173287

Bernoulli **28**(1), 2022, 33–63
https://doi.org/10.3150/21-BEJ1332

Parameter estimation in branching processes with almost sure extinction

PETER BRAUNSTEINS[1,2,5,*,†], SOPHIE HAUTPHENNE[2,5,‡] and CARMEN MINUESA[3,4,5,§,¶]

[1]*Korteweg-de Vries Instituut, University of Amsterdam, Amsterdam, The Netherlands.*
E-mail: [*]*p.t.m.braunsteins@uva.nl*, 0000-0003-1864-0703
[2]*School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia.*
E-mail: [†]*braunsteins.p@unimelb.edu.au*; [‡]*sophiemh@unimelb.edu.au*, 0000-0002-8361-1901
[3]*Department of Mathematics, Autonomous University of Madrid, Madrid, Spain.*
E-mail: [§]*carmen.minuesa@uam.es*, 0000-0002-8858-3145
[4]*Department of Mathematics, University of Extremadura, Badajoz, Spain.*
E-mail: [¶]*cminuesaa@unex.es*
[5]*All authors contributed equally to this work.*

We consider population-size-dependent branching processes (PSDBPs) which eventually become extinct with probability one. For these processes, we derive maximum likelihood estimators for the mean number of offspring born to individuals when the current population size is $z \geq 1$. As is standard in branching process theory, an asymptotic analysis of the estimators requires us to condition on non-extinction up to a finite generation n and let $n \to \infty$; however, because the processes become extinct with probability one, we are able to demonstrate that our estimators do not satisfy the classical consistency property (C-consistency). This leads us to define the concept of Q-*consistency*, and we prove that our estimators are Q-consistent and asymptotically normal. To investigate the circumstances in which a C-consistent estimator is preferable to a Q-consistent estimator, we then provide two C-consistent estimators for subcritical Galton–Watson branching processes. Our results rely on a combination of linear operator theory, coupling arguments, and martingale methods.

Keywords: Branching process; population-size-dependence; almost sure extinction; inference; carrying capacity; Q-process

References

[1] Asmussen, S. and Hering, H. (1983). *Branching Processes*. *Progress in Probability and Statistics* **3**. Boston, MA: Birkhäuser. MR0701538 https://doi.org/10.1007/978-1-4615-8155-0
[2] Athreya, K.B. and Ney, P.E. (1972). *Branching Processes*. New York: Springer. MR0373040
[3] Billingsley, P. (1995). *Probability and Measure*, 3rd ed. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1324786
[4] Braunsteins, P., Hautphenne, S. and Minuesa, C. Weighted least squares estimation in population-size-dependent branching processes. In preparation.
[5] Braunsteins, P., Hautphenne, S. and Minuesa, C. (2022). Supplement to "Parameter estimation in branching processes with almost sure extinction." https://doi.org/10.3150/21-BEJ1332SUPP
[6] Dion, J.-P. (1974). Estimation of the mean and the initial probabilities of a branching process. *J. Appl. Probab.* **11** 687–694. MR0375488 https://doi.org/10.2307/3212552
[7] Elliston, E.P. (1994). The black Robin: Saving the world's most endangered bird. *BioScience* **44** 499–501.
[8] Ernst, P.A., Kendall, W.S., Roberts, G.O. and Rosenthal, J.S. (2019). MEXIT: Maximal un-coupling times for stochastic processes. *Stochastic Process. Appl.* **129** 355–380. MR3907003 https://doi.org/10.1016/j.spa.2018.03.001

 © 2022 ISI/BS

[9] González, M., Martínez, R. and del Puerto, I. (2004). Nonparametric estimation of the offspring distribution and the mean for a controlled branching process. *TEST* **13** 465–479. MR2154009 https://doi.org/10.1007/BF02595782
[10] Gosselin, F. (2001). Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology. *Ann. Appl. Probab.* **11** 261–284. MR1825466 https://doi.org/10.1214/aoap/998926993
[11] Guttorp, P. (1991). *Statistical Inference for Branching Processes*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR1254434
[12] Haccou, P., Jagers, P. and Vatutin, V.A. (2007). *Branching Processes*: *Variation*, *Growth*, *and Extinction of Populations*. *Cambridge Studies in Adaptive Dynamics* **5**. Cambridge: Cambridge Univ. Press. MR2429372
[13] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*. New York–London: Academic Press. MR0624435
[14] Hamza, K., Jagers, P. and Klebaner, F.C. (2016). On the establishment, persistence, and inevitable extinction of populations. *J. Math. Biol.* **72** 797–820. MR3459167 https://doi.org/10.1007/s00285-015-0903-2
[15] Harris, T.E. (1948). Branching processes. *Ann. Math. Stat.* **19** 474–494. MR0027465 https://doi.org/10.1214/aoms/1177730146
[16] Heyde, C.C. and Seneta, E. (1972). Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Probab.* **9** 235–256. MR0343385 https://doi.org/10.2307/3212796
[17] Heyde, C.C. and Seneta, E. (1974). Notes on "Estimation theory for growth and immigration rates in a multiplicative process" (J. Appl. Probab. **9** (1972), 235–256). *J. Appl. Probab.* **11** 572–577. MR0368196 https://doi.org/10.2307/3212702
[18] Högnäs, G. (2019). On the lifetime of a size-dependent branching process. *Stoch. Models* **35** 119–131. MR3969510 https://doi.org/10.1080/15326349.2019.1578241
[19] Jagers, P. (1975). *Branching Processes with Biological Applications*. London–New York–Sydney: Wiley. MR0488341
[20] Jagers, P. (1992). Stabilities and instabilities in population dynamics. *J. Appl. Probab.* **29** 770–780. MR1188534 https://doi.org/10.1017/s0021900200043667
[21] Jagers, P. and Klebaner, F.C. (2011). Population-size-dependent, age-structured branching processes linger around their carrying capacity. *J. Appl. Probab.* **48A** 249–260. MR2865630 https://doi.org/10.1239/jap/1318940469
[22] Jagers, P. and Zuyev, S. (2020). Populations in environments with a soft carrying capacity are eventually extinct. *J. Math. Biol.* **81** 845–851. MR4154627 https://doi.org/10.1007/s00285-020-01527-5
[23] Kimmel, M. and Axelrod, D.E. (2015). *Branching Processes in Biology*, 2nd ed. *Interdisciplinary Applied Mathematics* **19**. New York: Springer. MR3310028 https://doi.org/10.1007/978-1-4939-1559-0
[24] Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23** 1125–1138. MR1349164
[25] Newman, M. (1997). A model of mass extinction. *J. Theoret. Biol.* **189** 235–252.
[26] Pakes, A.G. (1975). Non-parametric estimation in the Galton–Watson process. *Math. Biosci.* **26** 1–18. MR0405755 https://doi.org/10.1016/0025-5564(75)90091-7
[27] Raup, D.M. (1992). *Extinction*: *Bad Genes or Bad Luck*. New York: W. W. Norton.
[28] Ross, S.M. (1996). *Stochastic Processes*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1373653
[29] Sasser, D.W. (1964). Quasi-positive operators. *Pacific J. Math.* **14** 1029–1037. MR0169067
[30] Thorisson, H. (1995). Coupling methods in probability theory. *Scand. J. Stat.* **22** 159–182. MR1339749
[31] Wei, C.Z. and Winnicki, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.* **18** 1757–1773. MR1074433 https://doi.org/10.1214/aos/1176347876
[32] Yanev, N.M. (2008). Statistical inference for branching processes. In *Records and Branching Processes* (M. Ahsanullah and G.P. Yanev, eds.) 143–168. Nova Science Publishers.

Bernoulli **28**(1), 2022, 64–86
https://doi.org/10.3150/21-BEJ1333

A note on eigenvalues estimates for one-dimensional diffusion operators

MICHEL BONNEFONT[1] and ALDÉRIC JOULIN[2]

[1] *UMR CNRS 5251, Institut de Mathématiques de Bordeaux, Université Bordeaux 1, Bordeaux, France.*
E-mail: michel.bonnefont@math.u-bordeaux.fr; url: http://www.math.u-bordeaux.fr/~mibonnef/
[2] *UMR CNRS 5219, Institut de Mathématiques de Toulouse, Université de Toulouse, Toulouse, France.*
E-mail: ajoulin@insa-toulouse.fr; url: http://perso.math.univ-toulouse.fr/joulin/

Dealing with one-dimensional diffusion operators, we obtain upper and lower variational formulae on the eigenvalues given by the max–min principle, generalizing the celebrated result of Chen and Wang on the spectral gap. Our inequalities reveal to be sharp at least when the eigenvalues considered belong to the discrete spectrum of the operator, since in this case both lower and upper bounds coincide and involve the associated eigenfunctions. Based on the intertwinings between diffusion operators and some convenient gradients with weights, our approach also allows to estimate the gap between the two first positive eigenvalues when the spectral gap belongs to the discrete spectrum.

Keywords: Diffusion operator; Schrödinger operator; intertwining; eigenvalues; spectral gap; max–min principle

References

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

[2] Bakry, D., Barthe, F., Cattiaux, P. and Guillin, A. (2008). A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.* **13** 60–66. MR2386063 https://doi.org/10.1214/ECP.v13-1352

[3] Bakry, D., Gentil, I. and Ledoux, M. (2013). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9

[4] Bobkov, S.G. (1999). Isoperimetric and analytic inequalities for log-concave probability measures. *Ann. Probab.* **27** 1903–1921. MR1742893 https://doi.org/10.1214/aop/1022874820

[5] Bonnefont, M. and Joulin, A. (2014). Intertwining relations for one-dimensional diffusions and application to functional inequalities. *Potential Anal.* **41** 1005–1031. MR3269712 https://doi.org/10.1007/s11118-014-9408-7

[6] Bonnefont, M. and Joulin, A. (2021). Intertwinings, second-order Brascamp–Lieb inequalities and spectral estimates. *Studia Math*. To appear.

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

[8] Brascamp, H.J. and Lieb, E.H. (1976). On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.* **22** 366–389. MR0450480 https://doi.org/10.1016/0022-1236(76)90004-5

[9] Chen, M.-F. and Wang, F.-Y. (1997). Estimation of spectral gap for elliptic operators. *Trans. Amer. Math. Soc.* **349** 1239–1267. MR1401516 https://doi.org/10.1090/S0002-9947-97-01812-6

[10] Helffer, B. (2002). *Semiclassical Analysis, Witten Laplacians, and Statistical Mechanics*. *Series in Partial Differential Equations and Applications* **1**. River Edge, NJ: World Scientific Co., Inc. MR1936110 https://doi.org/10.1142/9789812776891

1350-7265 © 2022 ISI/BS

[11] Johnsen, J. (2000). On the spectral properties of Witten–Laplacians, their range projections and Brascamp–Lieb's inequality. *Integral Equations Operator Theory* **36** 288–324. MR1753421 https://doi.org/10.1007/BF01213926

[12] Kannan, R., Lovász, L. and Simonovits, M. (1995). Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* **13** 541–559. MR1318794 https://doi.org/10.1007/BF02574061

[13] Karlin, S. and Studden, W.J. (1966). *Tchebycheff Systems*: *With Applications in Analysis and Statistics*. *Pure and Applied Mathematics* **XV**. New York: Interscience Publishers Wiley. MR0204922

[14] Miclo, L. (2009). Monotonicity of the extremal functions for one-dimensional inequalities of logarithmic Sobolev type. In *Séminaire de Probabilités XLII*. *Lecture Notes in Math.* **1979** 103–130. Berlin: Springer. MR2599205 https://doi.org/10.1007/978-3-642-01763-6_2

[15] Milman, E. (2018). Spectral estimates, contractions and hypercontractivity. *J. Spectr. Theory* **8** 669–714. MR3812812 https://doi.org/10.4171/JST/210

[16] Roustant, O., Barthe, F. and Iooss, B. (2017). Poincaré inequalities on intervals – Application to sensitivity analysis. *Electron. J. Stat.* **11** 3081–3119. MR3694577 https://doi.org/10.1214/17-EJS1310

[17] Wang, F.-Y. (2005). *Functional Inequalities*, *Markov Semigroups and Spectral Theory*. Beijing: Science Press.

Bernoulli **28**(1), 2022, 87–124
https://doi.org/10.3150/21-BEJ1334

Multiplier U-processes: Sharp bounds and applications

QIYANG HAN

Department of Statistics, Rutgers University, Piscataway, NJ 08854, USA.
E-mail: qh85@stat.rutgers.edu

The theory for multiplier empirical processes has been one of the central topics in the development of the classical theory of empirical processes, due to its wide applicability to various statistical problems. In this paper, we develop theory and tools for studying multiplier U-processes, a natural higher-order generalization of the multiplier empirical processes. To this end, we develop a multiplier inequality that quantifies the moduli of continuity of the multiplier U-process in terms of that of the (decoupled) symmetrized U-process. The new inequality finds a variety of applications including (i) multiplier and bootstrap central limit theorems for U-processes, (ii) general theory for bootstrap M-estimators based on U-statistics, and (iii) theory for M-estimation under general complex sampling designs, again based on U-statistics.

Keywords: U-process; multiplier inequality; bootstrap central limit theorem; bootstrap M-estimators; complex sampling design

References

[1] Arcones, M.A., Chen, Z. and Giné, E. (1994). Estimators related to U-processes with applications to multivariate medians: Asymptotic normality. *Ann. Statist.* **22** 1460–1477. MR1311984 https://doi.org/10.1214/aos/1176325637

[2] Arcones, M.A. and Giné, E. (1992). On the bootstrap of U and V statistics. *Ann. Statist.* **20** 655–674. MR1165586 https://doi.org/10.1214/aos/1176348650

[3] Arcones, M.A. and Giné, E. (1993). Limit theorems for U-processes. *Ann. Probab.* **21** 1494–1542. MR1235426

[4] Arcones, M.A. and Giné, E. (1994). U-processes indexed by Vapnik–Červonenkis classes of functions with applications to asymptotics and bootstrap of U-statistics with estimated parameters. *Stochastic Process. Appl.* **52** 17–38. MR1289166 https://doi.org/10.1016/0304-4149(94)90098-1

[5] Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *J. Theoret. Probab.* **7** 47–71. MR1256391 https://doi.org/10.1007/BF02213360

[6] Boistard, H., Lopuhaä, H.P. and Ruiz-Gazen, A. (2017). Functional central limit theorems for single-stage sampling designs. *Ann. Statist.* **45** 1728–1758. MR3670194 https://doi.org/10.1214/16-AOS1507

[7] Bose, A. and Chatterjee, S. (2001). Generalised bootstrap in non-regular M-estimation problems. *Statist. Probab. Lett.* **55** 319–328. MR1867535 https://doi.org/10.1016/S0167-7152(01)00161-4

[8] Bose, A. and Chatterjee, S. (2003). Generalized bootstrap for estimators of minimizers of convex functions. *J. Statist. Plann. Inference* **117** 225–239. MR2004657 https://doi.org/10.1016/S0378-3758(02)00386-5

[9] Breslow, N.E. and Wellner, J.A. (2007). Weighted likelihood for semiparametric models and two-phase stratified samples, with application to Cox regression. *Scand. J. Statist.* **34** 86–102. MR2325244 https://doi.org/10.1111/j.1467-9469.2006.00523.x

[10] Chatterjee, S. and Bose, A. (2005). Generalized bootstrap for estimating equations. *Ann. Statist.* **33** 414–436. MR2157808 https://doi.org/10.1214/009053604000000904

[11] Cheng, G. and Huang, J.Z. (2010). Bootstrap consistency for general semiparametric M-estimation. *Ann. Statist.* **38** 2884–2915. MR2722459 https://doi.org/10.1214/10-AOS809

[12] Clémençon, S., Bertail, P. and Papa, G. (2016). Learning from survey training samples: Rate bounds for Horvitz–Thompson risk minimizers. In *Asian Conference on Machine Learning* 142–157.

 © 2022 ISI/BS

[13] Clémençon, S., Lugosi, G. and Vayatis, N. (2008). Ranking and empirical minimization of U-statistics. *Ann. Statist.* **36** 844–874. MR2396817 https://doi.org/10.1214/009052607000000910
[14] de la Peña, V.H. and Giné, E. (1999). *Decoupling. Probability and Its Applications* (*New York*). New York: Springer. From dependence to independence, Randomly stopped processes. U-statistics and processes. Martingales and beyond. MR1666908 https://doi.org/10.1007/978-1-4612-0537-1
[15] Giné, E. and Koltchinskii, V. (2006). Concentration inequalities and asymptotic results for ratio type empirical processes. *Ann. Probab.* **34** 1143–1216. MR2243881 https://doi.org/10.1214/009117906000000070
[16] Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann. Probab.* **12** 929–998. With discussion. MR0757767
[17] Giné, E. and Zinn, J. (1986). Lectures on the central limit theorem for empirical processes. In *Probability and Banach Spaces* (*Zaragoza*, 1985). *Lecture Notes in Math.* **1221** 50–113. Berlin: Springer. MR0875007 https://doi.org/10.1007/BFb0099111
[18] Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. *Ann. Probab.* **18** 851–869. MR1055437
[19] Han, Q. (2021). Set structured global empirical risk minimizers are rate optimal in general dimensions. *Ann. Statist.*, to appear. Available at arXiv:1905.12823.
[20] Han, Q. and Wellner, J.A. (2018). Robustness of shape-restricted regression estimators: An envelope perspective. arXiv preprint. Available at arXiv:1805.02542.
[21] Han, Q. and Wellner, J.A. (2019). Convergence rates of least squares regression estimators with heavy-tailed errors. *Ann. Statist.* **47** 2286–2319. MR3953452 https://doi.org/10.1214/18-AOS1748
[22] Han, Q. and Wellner, J.A. (2021). Complex sampling designs: Uniform limit theorems and applications. *Ann. Statist.* **49** 459–485. MR4206686 https://doi.org/10.1214/20-AOS1964
[23] Hušková, M. and Janssen, P. (1993). Consistency of the generalized bootstrap for degenerate U-statistics. *Ann. Statist.* **21** 1811–1823. MR1245770 https://doi.org/10.1214/aos/1176349399
[24] Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *Ann. Statist.* **34** 2593–2656. MR2329442 https://doi.org/10.1214/009053606000001019
[25] Ledoux, M. and Talagrand, M. (1986). Conditions d'intégrabilité pour les multiplicateurs dans le TLC banachique. *Ann. Probab.* **14** 916–921. MR0841593
[26] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces. Classics in Mathematics*. Berlin: Springer. Isoperimetry and processes, Reprint of the 1991 edition. MR2814399
[27] Liu, R.Y. (1990). On a notion of data depth based on random simplices. *Ann. Statist.* **18** 405–414. MR1041400 https://doi.org/10.1214/aos/1176347507
[28] Mammen, E. and Tsybakov, A.B. (1999). Smooth discrimination analysis. *Ann. Statist.* **27** 1808–1829. MR1765618 https://doi.org/10.1214/aos/1017939240
[29] Mendelson, S. (2016). Upper bounds on product and multiplier empirical processes. *Stochastic Process. Appl.* **126** 3652–3680. MR3565471 https://doi.org/10.1016/j.spa.2016.04.019
[30] Mendelson, S. (2017). On multiplier processes under weak moment assumptions. In *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **2169** 301–318. Cham: Springer. MR3645129
[31] Nolan, D. and Pollard, D. (1987). U-processes: Rates of convergence. *Ann. Statist.* **15** 780–799. MR0888439 https://doi.org/10.1214/aos/1176350374
[32] Nolan, D. and Pollard, D. (1988). Functional limit theorems for U-processes. *Ann. Probab.* **16** 1291–1298. MR0942769
[33] Præstgaard, J. and Wellner, J.A. (1993). Exchangeably weighted bootstraps of the general empirical process. *Ann. Probab.* **21** 2053–2086. MR1245301
[34] Rubin-Bleuer, S. and Schiopu Kratina, I. (2005). On the two-phase framework for joint model and design-based inference. *Ann. Statist.* **33** 2789–2810. MR2253102 https://doi.org/10.1214/009053605000000651
[35] Stein, E.M. and Weiss, G. (1971). *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton, NJ: Princeton Univ. Press. Princeton Mathematical Series, No. 32. MR0304972
[36] Tsybakov, A.B. (2004). Optimal aggregation of classifiers in statistical learning. *Ann. Statist.* **32** 135–166. MR2051002 https://doi.org/10.1214/aos/1079120131
[37] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256

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

[39] Wellner, J.A. and Zhan, Y. (1996). Bootstrapping Z-estimators Technical Report 308, University of Washington Department of Statistics.

[40] Zhang, D. (2001). Bayesian bootstraps for U-processes, hypothesis tests and convergence of Dirichlet U-processes. *Statist*. *Sinica* **11** 463–478. MR1844535

Bernoulli **28**(1), 2022, 125–149
https://doi.org/10.3150/21-BEJ1335

Mean field limits for interacting Hawkes processes in a diffusive regime

XAVIER ERNY[1,*], EVA LÖCHERBACH[2] and DASHA LOUKIANOVA[1,†]

[1]*Université Paris-Saclay, CNRS, Univ Evry, Laboratoire de Mathématiques et Modélisation d'Evry, 91037, Evry, France.*
E-mail: [*]*xavier.erny@univ-evry.fr;* [†]*dasha.loukianova@univ-evry.fr*
[2]*Statistique, Analyse et Modélisation Multidisciplinaire, Université Paris 1 Panthéon-Sorbonne, EA 4543 et FR FP2M 2036 CNRS, France.*
E-mail: eva.locherbach@univ-paris1.fr

We consider a sequence of systems of Hawkes processes having mean field interactions in a diffusive regime. The stochastic intensity of each process is a solution of a stochastic differential equation driven by N independent Poisson random measures. We show that, as the number of interacting components N tends to infinity, this intensity converges in distribution in the Skorokhod space to a CIR-type diffusion. Moreover, we prove the convergence in distribution of the Hawkes processes to the limit point process having the limit diffusion as intensity. To prove the convergence results, we use analytical technics based on the convergence of the associated infinitesimal generators and Markovian semigroups.

Keywords: Multivariate nonlinear Hawkes processes; mean field interaction; piecewise deterministic Markov processes

References

[1] Aït-Sahalia, Y., Cacho-Diaz, J. and Laeven, R.J.A. (2015). Modeling financial contagion using mutually exciting jump processes. *J. Financ. Econ.* **117** 585–606.
[2] Bacry, E. and Muzy, J.-F. (2016). First- and second-order statistics characterization of Hawkes processes and non-parametric estimation. *IEEE Trans. Inf. Theory* **62** 2184–2202. MR3480107 https://doi.org/10.1109/TIT.2016.2533397
[3] Bauwens, L. and Hautsch, N. (2009). *Modelling Financial High Frequency Data Using Point Processes*. Berlin Heidelberg: Springer.
[4] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1700749 https://doi.org/10.1002/9780470316962
[5] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24** 1563–1588. MR1411506 https://doi.org/10.1214/aop/1065725193
[6] Carmona, R., Delarue, F. and Lacker, D. (2016). Mean field games with common noise. *Ann. Probab.* **44** 3740–3803. MR3572323 https://doi.org/10.1214/15-AOP1060
[7] Costa, M., Graham, C., Marsalle, L. and Tran, V.C. (2020). Renewal in Hawkes processes with self-excitation and inhibition. *Adv. in Appl. Probab.* **52** 879–915. MR4153592 https://doi.org/10.1017/apr.2020.19
[8] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*. *Vol. I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1950431
[9] Davis, M.H.A. (1993). *Markov Models and Optimization*. *Monographs on Statistics and Applied Probability* **49**. London: CRC Press. MR1283589 https://doi.org/10.1007/978-1-4899-4483-2
[10] Dawson, D. and Vaillancourt, J. (1995). Stochastic McKean–Vlasov equations. *NoDEA Nonlinear Differential Equations Appl.* **2** 199–229. MR1328577 https://doi.org/10.1007/BF01295311

[11] Delattre, S., Fournier, N. and Hoffmann, M. (2016). Hawkes processes on large networks. *Ann. Appl. Probab.* **26** 216–261. MR3449317 https://doi.org/10.1214/14-AAP1089
[12] Ditlevsen, S. and Löcherbach, E. (2017). Multi-class oscillating systems of interacting neurons. *Stochastic Process. Appl.* **127** 1840–1869. MR3646433 https://doi.org/10.1016/j.spa.2016.09.013
[13] Ethier, S. and Kurtz, T. (2005). *Markov Processes. Characterization and Convergence. Wiley Series in Probability and Statistics*.
[14] Fournier, N. and Löcherbach, E. (2016). On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** 1844–1876. MR3573298 https://doi.org/10.1214/15-AIHP701
[15] Graham, C. (2019). Regenerative properties of the linear Hawkes process with unbounded memory. arXiv:1905.11053 [math, stat].
[16] Grün, S., Diedsmann, M. and Aertsen, A.M. (2010). *Analysis of Parallel Spike Trains. Springer Series in Computational Neurosciences*. Rotter.
[17] Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. MR0278410 https://doi.org/10.1093/biomet/58.1.83
[18] Hawkes, A.G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Probab.* **11** 493–503. MR0378093 https://doi.org/10.2307/3212693
[19] Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *J. Geophys. Res.* **107**. https://doi.org/10.1029/2001JB001580
[20] Hewlett, P. (2006). Clustering of order arrivals, price impact and trade path optimisation. In *Workshop on Financial Modeling with Jump Processes*. Ecole Polytechnique.
[21] Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, 2nd ed. North-Holland Publishing Company. MR0637061
[22] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Berlin: Springer. MR1943877 https://doi.org/10.1007/978-3-662-05265-5
[23] Kagan, Y.Y. (2009). Statistical distributions of earthquake numbers: Consequence of branching process. *Geophys. J. Int.* **180**. https://doi.org/10.1111/j.1365-246X.2009.04487.x
[24] Kallenberg, O. (1997). *Foundations of Modern Probability. Probability and Its Applications* (*New York*). New York: Springer. MR1464694
[25] Khasminskii, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Heidelberg: Springer. With contributions by G.N. Milstein and M.B. Nevelson. MR2894052 https://doi.org/10.1007/978-3-642-23280-0
[26] Kunita, H. (1986). *Lectures on Stochastic Flows and Applications. Tata Institute of Fundamental Research Lectures on Mathematics and Physics* **78**. Berlin: Springer. Published for the Tata Institute of Fundamental Research, Bombay. MR0867686
[27] Kunita, H. (1990). *Stochastic Flows and Stochastic Differential Equations. Cambridge Studies in Advanced Mathematics* **24**. Cambridge: Cambridge Univ. Press. MR1070361
[28] Kurtz, T.G. and Xiong, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Process. Appl.* **83** 103–126. MR1705602 https://doi.org/10.1016/S0304-4149(99)00024-1
[29] Lu, X. and Abergel, F. (2018). High-dimensional Hawkes processes for limit order books: Modelling, empirical analysis and numerical calibration. *Quant. Finance* **18** 249–264. MR3750734 https://doi.org/10.1080/14697688.2017.1403142
[30] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. III. Foster–Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25** 518–548. MR1234295 https://doi.org/10.2307/1427522
[31] Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Statist. Math.* **30** 243–261. MR0514494 https://doi.org/10.1007/BF02480216
[32] Ogata, Y. (1999). Seismicity analysis through point-process modeling: A review. *Pure Appl. Geophys.* **155** 471–507. https://doi.org/10.1007/s000240050275
[33] Okatan, M., Wilson, M.A. and Brown, E.N. (2005). Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural Comput.* **17** 1927–61. https://doi.org/10.1162/0899766054322973
[34] Pillow, J.W., Wilson, M.A. and Brown, E.N. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454** 995–999.

[35] Raad, M.B. (2019). Renewal time points for Hawkes processes. arXiv:1906.02036 [math].
[36] Reynaud-Bouret, P., Rivoirard, V., Grammont, F. and Tuleau-Malot, C. (2014). Goodness-of-fit tests and nonparametric adaptive estimation for spike train analysis. *J. Math. Neurosci.* **4** Art. 3, 41. MR3197017 https://doi.org/10.1186/2190-8567-4-3
[37] Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.* **38** 2781–2822. MR2722456 https://doi.org/10.1214/10-AOS806
[38] Zhou, K., Zha, H. and Song, L. (2013). Learning triggering kernels for multi-dimensional Hawkes processes. In *Proceedings of the* 30*th International Conf. on Machine Learning* (*ICML*).

Bernoulli **28**(1), 2022, 150–180
https://doi.org/10.3150/21-BEJ1337

Tree builder random walk: Recurrence, transience and ballisticity

GIULIO IACOBELLI[1,*], RODRIGO RIBEIRO[2], GLAUCO VALLE[1,†] and LEONEL ZUAZNÁBAR[3]

[1] *UFRJ – Instituto de Matemática. Caixa Postal 68530, 21945-970, Rio de Janeiro, Brasil.*
E-mail: [*] *giulio@im.ufrj.br;* [†] *glauco.valle@im.ufrj.br*
[2] *PUC Chile – Facultad de Matemáticas. Av. Vicuña Mackenna 4860, Macul, La Florida, Región Metropolitana, Chile.*
E-mail: rribeiro@impa.br
[3] *USP – Instituto de Matemática e Estatística. R. do Matão, 1010 – Butantã, São Paulo – SP, CEP: 05508-090, Brasil.*
E-mail: lzuaznabar@ime.usp.br

The Tree Builder Random Walk is a special random walk that evolves on trees whose size increases with time, randomly and depending upon the walker. After every s steps of the walker, a random number of vertices are added to the tree and attached to the current position of the walker. These processes share similarities with other important classes of Markovian and non-Markovian random walks presenting a large variety of behaviors according to parameters specifications. We show that for a large and most significant class of tree builder random walks, the process is either null recurrent or transient. If s is odd, the walker is ballistic, thus transient. If s is even, the walker's behavior can be explained from local properties of the growing tree and it can be either null recurrent or it gets trapped on some limited part of the growing tree.

Keywords: Random walks; random environment; random trees; transience; recurrence; ballisticity

References

[1] Avena, L., Güldaş, H., van der Hofstad, R. and den Hollander, F. (2018). Mixing times of random walks on dynamic configuration models. *Ann. Appl. Probab.* **28** 1977–2002. MR3843821 https://doi.org/10.1214/17-AAP1289

[2] Bouchet, É., Ramírez, A.F. and Sabot, C. (2016). Sharp ellipticity conditions for ballistic behavior of random walks in random environment. *Bernoulli* **22** 969–994. MR3449805 https://doi.org/10.3150/14-BEJ683

[3] Campos, D. and Ramírez, A.F. (2014). Ellipticity criteria for ballistic behavior of random walks in random environment. *Probab. Theory Related Fields* **160** 189–251. MR3256813 https://doi.org/10.1007/s00440-013-0527-7

[4] Davis, B. (1990). Reinforced random walk. *Probab. Theory Related Fields* **84** 203–229. MR1030727 https://doi.org/10.1007/BF01197845

[5] Dembo, A., Huang, R. and Sidoravicius, V. (2014). Walking within growing domains: Recurrence versus transience. *Electron. J. Probab.* **19** no. 106, 20. MR3275858 https://doi.org/10.1214/EJP.v19-3272

[6] Disertori, M., Sabot, C. and Tarrès, P. (2015). Transience of edge-reinforced random walk. *Comm. Math. Phys.* **339** 121–148. MR3366053 https://doi.org/10.1007/s00220-015-2392-y

[7] Durrett, R. (2007). *Random Graph Dynamics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **20**. Cambridge: Cambridge Univ. Press. MR2271734

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

 © 2022 ISI/BS

[9] Eichelsbacher, P. and Löwe, M. (2003). Moderate deviations for i.i.d. random variables. *ESAIM Probab. Stat.* **7** 209–218. MR1956079 https://doi.org/10.1051/ps:2003005
[10] Figueiredo, D., Iacobelli, G., Oliveira, R., Reed, B. and Ribeiro, R. (2021). On a random walk that grows its own tree. *Electron. J. Probab.* **26** Paper No. 6, 40. MR4216519 https://doi.org/10.1214/20-ejp574
[11] Fribergh, A. and Kious, D. (2016). Local trapping for elliptic random walks in random environments in $\mathbb{Z}^d$. *Probab. Theory Related Fields* **165** 795–834. MR3520019 https://doi.org/10.1007/s00440-015-0646-4
[12] Iacobelli, G., Figueiredo, D.R. and Neglia, G. (2019). Transient and slim versus recurrent and fat: Random walks and the trees they grow. *J. Appl. Probab.* **56** 769–786. MR4015636 https://doi.org/10.1017/jpr.2019.43
[13] Iacobelli, G., Ribeiro, R., Valle, G. and Zuaznábar, L. (2022). Supplement to "Tree builder random walk: Recurrence, transience and ballisticity." https://doi.org/10.3150/21-BEJ1337SUPP
[14] Janson, S., Łuczak, T. and Rucinski, A. (2000). *Random Graphs*. *Wiley-Interscience Series in Discrete Mathematics and Optimization*. New York: Wiley. MR1782847 https://doi.org/10.1002/9781118032718
[15] Kosygina, E. and Zerner, M.P.W. (2013). Excited random walks: Results, methods, open problems. *Bull. Inst. Math. Acad. Sin.* (*N.S.*) **8** 105–157. MR3097419
[16] Lyons, R., Pemantle, R. and Peres, Y. (1996). Biased random walks on Galton-Watson trees. *Probab. Theory Related Fields* **106** 249–264. MR1410689 https://doi.org/10.1007/s004400050064
[17] Pemantle, R. (1988). Phase transition in reinforced random walk and RWRE on trees. *Ann. Probab.* **16** 1229–1241. MR0942765
[18] Pemantle, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79. MR2282181 https://doi.org/10.1214/07-PS094
[19] Peres, Y., Stauffer, A. and Steif, J.E. (2015). Random walks on dynamical percolation: Mixing times, mean squared displacement and hitting times. *Probab. Theory Related Fields* **162** 487–530. MR3383336 https://doi.org/10.1007/s00440-014-0578-4
[20] Sznitman, A.-S. (2001). On a class of transient random walks in random environment. *Ann. Probab.* **29** 724–765. MR1849176 https://doi.org/10.1214/aop/1008956691
[21] Sznitman, A.-S. and Zerner, M. (1999). A law of large numbers for random walks in random environment. *Ann. Probab.* **27** 1851–1869. MR1742891 https://doi.org/10.1214/aop/1022874818
[22] van der Hofstad, R. (2017). *Random Graphs and Complex Networks*. *Vol.* 1. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge: Cambridge Univ. Press. MR3617364 https://doi.org/10.1017/9781316779422
[23] Zeitouni, O. (2004). Random walks in random environment. In *Lectures on Probability Theory and Statistics*. *Lecture Notes in Math.* **1837** 189–312. Berlin: Springer. MR2071631 https://doi.org/10.1007/978-3-540-39874-5_2

Bernoulli **28**(1), 2022, 181–213
https://doi.org/10.3150/21-BEJ1338

Finite sample properties of parametric MMD estimation: Robustness to misspecification and dependence

BADR-EDDINE CHÉRIEF-ABDELLATIF[1] and PIERRE ALQUIER[2]

[1]*Department of Statistics, University of Oxford.*
E-mail: badr-eddine.cherief-abdellatif@stats.ox.ac.uk
[2]*RIKEN AIP, Tokyo, Japan.*
E-mail: pierrealain.alquier@riken.jp

Many works in statistics aim at designing a universal estimation procedure, that is, an estimator that would converge to the best approximation of the (unknown) data generating distribution in a model, without any assumption on this distribution. This question is of major interest, in particular because the universality property leads to the robustness of the estimator. In this paper, we tackle the problem of universal estimation using a minimum distance estimator presented in (Briol et al. (2019)) based on the Maximum Mean Discrepancy. We show that the estimator is robust to both dependence and to the presence of outliers in the dataset. Finally, we provide a theoretical study of the stochastic gradient descent algorithm used to compute the estimator, and we support our findings with numerical simulations.

Keywords: Minimum distance estimation; kernel methods; universal estimation; robust statistics; RKHS; weak dependence

References

[1] Alon, N., Matias, Y. and Szegedy, M. (2008). The space complexity of approximating the frequency moments *J. Comput. System Sci.* **58** 137–147.
[2] Alquier, P. (2008). Density estimation with quadratic loss: A confidence intervals method. *ESAIM Probab. Stat.* **12** 438–463. MR2437718 https://doi.org/10.1051/ps:2007050
[3] Amenta, N., Bern, M., Eppstein, D. and Teng, S.-H. (2000). Regression depth and center points. *Discrete Comput. Geom.* **23** 305–323. MR1744506 https://doi.org/10.1007/PL00009502
[4] Arlot, S., Celisse, A. and Harchaoui, Z. (2019). A kernel multiple change-point algorithm via model selection. *J. Mach. Learn. Res.* **20** Paper No. 162, 56. MR4048973
[5] Baraud, Y. and Birgé, L. (2018). Rho-estimators revisited: General theory and applications. *Ann. Statist.* **46** 3767–3804. MR3852668 https://doi.org/10.1214/17-AOS1675
[6] Baraud, Y., Birgé, L. and Sart, M. (2017). A new method for estimation and model selection: ρ-estimation. *Invent. Math.* **207** 425–517. MR3595933 https://doi.org/10.1007/s00222-016-0673-5
[7] Barron, A., Schervish, M.J. and Wasserman, L. (1999). The consistency of posterior distributions in nonparametric problems. *Ann. Statist.* **27** 536–561. MR1714718 https://doi.org/10.1214/aos/1018031206
[8] Bernton, E., Jacob, P.E., Gerber, M. and Robert, C.P. (2017). Inference in generative models using the Wasserstein distance. Preprint. Available at arXiv:1701.05146.
[9] Biau, G., Cadre, B., Sangnier, M. and Tanielian, U. (2020). Some theoretical properties of GANs. *Ann. Statist.* **48** 1539–1566. MR4124334 https://doi.org/10.1214/19-AOS1858
[10] Bickel, P.J. (1976). Another look at robustness: A review of reviews and some new developments. *Scand. J. Stat.* **3** 145–168. MR0428589
[11] Birgé, L. (1983). Approximation dans les espaces métriques et théorie de l'estimation. *Z. Wahrsch. Verw. Gebiete* **65** 181–237. MR0722129 https://doi.org/10.1007/BF00532480

 © 2022 ISI/BS

[12] Birgé, L. (2006). Model selection via testing: An alternative to (penalized) maximum likelihood estimators. *Annales de l'IHP Probabilités et statistiques*. **42** 273–325.
[13] Blei, D.M., Kucukelbir, A. and McAuliffe, J.D. (2017). Variational inference: A review for statisticians. *J. Amer. Statist. Assoc*. **112** 859–877. MR3671776 https://doi.org/10.1080/01621459.2017.1285773
[14] 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
[15] Briol, F.-X., Barp, A., Duncan, A.B. and Girolami, M. (2019). Statistical inference for generative models via maximum mean discrepancy. Preprint. Available at arXiv:1906.05944.
[16] Bunea, F., Tsybakov, A.B. and Wegkamp, M.H. (2007). Sparse density estimation with ℓ_1 penalties. In *Learning Theory*. *Lecture Notes in Computer Science* **4539** 530–543. Berlin: Springer. MR2397610 https://doi.org/10.1007/978-3-540-72927-3_38
[17] Bunea, F., Tsybakov, A.B., Wegkamp, M.H. and Barbu, A. (2010). Spades and mixture models. *Ann. Statist*. **38** 2525–2558. MR2676897 https://doi.org/10.1214/09-AOS790
[18] Cai, T., Ma, Z. and Wu, Y. (2015). Optimal estimation and rank detection for sparse spiked covariance matrices. *Probab. Theory Related Fields* **161** 781–815. MR3334281 https://doi.org/10.1007/s00440-014-0562-z
[19] Huggins, J.H., Campbell, T., Kasprzak, M. and Broderick, T. (2018). Practical bounds on the error of Bayesian posterior approximations: A nonasymptotic approach. Preprint. Available at arXiv:1809.09505.
[20] Catoni, O. (2012). Challenging the empirical mean and empirical variance: A deviation study. *Ann. Inst. Henri Poincaré Probab. Stat*. **48** 1148–1185. MR3052407 https://doi.org/10.1214/11-AIHP454
[21] Catoni, O. and Giulini, I. (2017). Dimension free PAC-Bayesian bounds for the estimation of the mean of a random vector. In *NIPS*-2017 *Workshop* (*Almost*) 50 *Shades of Bayesian Learning*: *PAC-Bayesian Trends and Insights*.
[22] Chan, T.M. (2004). An optimal randomized algorithm for maximum Tukey depth. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms* 430–436. New York: ACM. MR2291081
[23] Chen, M., Gao, C. and Ren, Z. (2018). Robust covariance and scatter matrix estimation under Huber's contamination model. *Ann. Statist*. **46** 1932–1960. MR3845006 https://doi.org/10.1214/17-AOS1607
[24] Cherapanamjeri, Y., Flammarion, N. and Bartlett, P.L. (2019). Fast mean estimation with sub-Gaussian rates. *Proceedings of the Thirty-Second Conference on Learning Theory*, *PMLR* **99** 786–806.
[25] Chérief-Abdellatif, B.-E. and Alquier, P. (2018). Consistency of variational Bayes inference for estimation and model selection in mixtures. *Electron. J. Stat*. **12** 2995–3035. MR3855643 https://doi.org/10.1214/18-EJS1475
[26] Chérief-Abdellatif, B.-E. and Alquier, P. (2019). MMD-Bayes: Robust Bayesian estimation via maximum mean discrepancy. In *Proceedings of the* 2*nd Symposium on Advances in Approximate Bayesian Inference* (*AABI*), 2020. *Proceedings of Machine Learning Research* **118** 1–21.
[27] Chérief-Abdellatif, B.-E. and Alquier, P. (2022). Supplement to "Finite sample properties of parametric MMD estimation: robustness to misspecification and dependence." https://doi.org/10.3150/21-BEJ1338SUPP
[28] Chinot, G., Lecué, G. and Lerasle, M. (2020). Robust statistical learning with Lipschitz and convex loss functions. *Probab. Theory Related Fields* **176** 897–940. MR4087486 https://doi.org/10.1007/s00440-019-00931-3
[29] Collier, O. and Dalalyan, A.S. (2019). Multidimensional linear functional estimation in sparse Gaussian models and robust estimation of the mean. *Electron. J. Stat*. **13** 2830–2864. MR3998929 https://doi.org/10.1214/19-EJS1590
[30] Dalalyan, A.S. and Sebbar, M. (2018). Optimal Kullback-Leibler aggregation in mixture density estimation by maximum likelihood. *Math. Stat. Learn*. **1** 1–35. MR4049449 https://doi.org/10.4171/msl/1-1-1
[31] Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi, S. and Prieur, C. (2007). *Weak Dependence*: *With Examples and Applications*. *Lecture Notes in Statistics* **190**. New York: Springer. MR2338725
[32] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. MR0501537
[33] Depersin, J. and Lecué, G. (2019). Robust sub-Gaussian estimation of a mean vector in nearly linear time. Preprint. Available at arXiv:1906.03058.

[34] Devroye, L., Lerasle, M., Lugosi, G. and Oliveira, R.I. (2016). Sub-Gaussian mean estimators. *Ann. Statist.* **44** 2695–2725. MR3576558 https://doi.org/10.1214/16-AOS1440
[35] Devroye, L. and Lugosi, G. (2001). *Combinatorial Methods in Density Estimation*. *Springer Series in Statistics*. New York: Springer. MR1843146 https://doi.org/10.1007/978-1-4613-0125-7
[36] Diakonikolas, I., Kamath, G., Kane, D.M., Li, J., Moitra, A. and Stewart, A. (2016). Robust estimators in high dimensions without the computational intractability. In 57*th Annual IEEE Symposium on Foundations of Computer Science – FOCS* 2016 655–664. Los Alamitos, CA: IEEE Computer Soc. MR3631028 https://doi.org/10.1109/FOCS.2016.85
[37] Diakonikolas, I., Kane, D.M. and Stewart, A. (2018). List-decodable robust mean estimation and learning mixtures of spherical Gaussians. In *STOC*'18 – *Proceedings of the* 50*th Annual ACM SIGACT Symposium on Theory of Computing* 1047–1060. New York: ACM. MR3826316 https://doi.org/10.1145/3188745.3188758
[38] Diakonikolas, I., Kong, W. and Stewart, A. (2019). Efficient algorithms and lower bounds for robust linear regression. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* 2745–2754. Philadelphia, PA: SIAM. MR3909639 https://doi.org/10.1137/1.9781611975482.170
[39] Doukhan, P. (1994). *Mixing*. *Lecture Notes in Statistics*: *Properties and Examples* **85**. New York: Springer. MR1312160 https://doi.org/10.1007/978-1-4612-2642-0
[40] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** 313–342. MR1719345 https://doi.org/10.1016/S0304-4149(99)00055-1
[41] Du, S.S., Balakrishnan, S. and Singh, A. (2017). Computationally efficient robust estimation of sparse functionals.. *Proceedings of the* 2017 *Conference on Learning Theory*, *PMLR*. **65** 169–212.
[42] Dziugaite, G.K., Roy, D.M. and Ghahramani, Z. (2015). Training generative neural networks via maximum mean discrepancy optimization. In *Proc. Uncertainty in Artificial Intelligence* (*UAI*), 2015.
[43] Gao, C., Liu, J., Yao, Y. and Zhu, W. (2019). Robust estimation and generative adversarial nets. In *ICLR* 2019.
[44] Giulini, I. (2017). Robust PCA and pairs of projections in a Hilbert space. *Electron. J. Stat.* **11** 3903–3926. MR3714302 https://doi.org/10.1214/17-EJS1343
[45] Giulini, I. (2018). Robust dimension-free Gram operator estimates. *Bernoulli* **24** 3864–3923. MR3788191 https://doi.org/10.3150/17-BEJ981
[46] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* 2672–2680.
[47] Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B. and Smola, A. (2012). A kernel two-sample test. *J. Mach. Learn. Res.* **13** 723–773. MR2913716
[48] Gretton, A., Fukumizu, K., Harchaoui, Z. and Sriperumbudur, B.K. (2009). A fast, consistent kernel two-sample test. In *Advances in Neural Information Processing Systems* 673–681.
[49] Grünwald, P. and van Ommen, T. (2017). Inconsistency of Bayesian inference for misspecified linear models, and a proposal for repairing it. *Bayesian Anal.* **12** 1069–1103. MR3724979 https://doi.org/10.1214/17-BA1085
[50] Haddouche, M., Guedj, B., Rivasplata, O. and Shawe-Taylor, J. (2020). PAC-Bayes unleashed: Generalisation bounds with unbounded losses. Preprint. Available at arXiv:2006.07279.
[51] Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50** 1029–1054. MR0666123 https://doi.org/10.2307/1912775
[52] Holland, M.J. (2019). PAC-Bayes under potentially heavy tails. Preprint. Available at arXiv:1905.07900.
[53] Holland, M.J. (2019). Distribution-robust mean estimation via smoothed random perturbations. Preprint. Available at arXiv:1906.10300.
[54] Hopkins, S.B. (2020). Mean estimation with sub-Gaussian rates in polynomial time. *Ann. Statist.* **48** 1193–1213. MR4102693 https://doi.org/10.1214/19-AOS1843
[55] Hörmann, S. and Kokoszka, P. (2010). Weakly dependent functional data. *Ann. Statist.* **38** 1845–1884. MR2662361 https://doi.org/10.1214/09-AOS768
[56] Hsu, D. and Sabato, S. (2016). Loss minimization and parameter estimation with heavy tails. *J. Mach. Learn. Res.* **17** Paper No. 18, 40. MR3491112
[57] Huber, P.J. (1964). Robust estimation of a location parameter. *Ann. Math. Stat.* **35** 73–101. MR0161415 https://doi.org/10.1214/aoms/1177703732

[58] Jerrum, M.R., Valiant, L.G. and Vazirani, V.V. (1986). Random generation of combinatorial structures from a uniform distribution. *Theoret. Comput. Sci.* **43** 169–188. MR0855970 https://doi.org/10.1016/0304-3975(86)90174-X
[59] Jitkrittum, W., Xu, W., Szabó, Z., Fukumizu, K. and Gretton, A. (2017). A linear-time kernel goodness-of-fit test. In *Advances in Neural Information Processing Systems* 262–271.
[60] Kothari, P.K., Steinhardt, J. and Steurer, D. (2018). Robust moment estimation and improved clustering via sum of squares. In *STOC'*18 – *Proceedings of the* 50*th Annual ACM SIGACT Symposium on Theory of Computing* 1035–1046. New York: ACM. MR3826315
[61] Lai, K.A., Rao, A.B. and Vempala, S. (2016). Agnostic estimation of mean and covariance. In 57*th Annual IEEE Symposium on Foundations of Computer Science – FOCS* 2016 665–674. Los Alamitos, CA: IEEE Computer Soc. MR3631029
[62] Le Cam, L. (1970). On the assumptions used to prove asymptotic normality of maximum likelihood estimates. *Ann. Math. Stat.* **41** 802–828. MR0267676 https://doi.org/10.1214/aoms/1177696960
[63] Le Cam, L. (1975). On local and global properties in the theory of asymptotic normality of experiments. In *Stochastic Processes and Related Topics* (*Proc. Summer Res. Inst. Statist. Inference for Stochastic Processes*, *Indiana Univ.*, *Bloomington*, *Ind.*, 1974, *Vol.* 1; *Dedicated to Jerzy Neyman*) 13–54. MR0395005
[64] LeCam, L. (1973). Convergence of estimates under dimensionality restrictions. *Ann. Statist.* **1** 38–53. MR0334381
[65] Lerasle, M., Szabó, Z., Mathieu, T. and Lecué, G. (2020). MONK – outlier-robust mean embedding estimation by median-of-means. In *International Conference on Machine Learning*.
[66] Lecué, G., Lerasle, M. and Mathieu, T. (2020). Robust classification via MOM minimization. *Mach. Learn.* **109** 1635–1665. MR4137195 https://doi.org/10.1007/s10994-019-05863-6
[67] Lee, J. and Raginsky, M. (2018). Minimax statistical learning with Wasserstein distances. In *Advances in Neural Information Processing Systems*.
[68] Li, Y., Swersky, K. and Zemel, R. (2015). Generative moment matching networks. In *International Conference on Machine Learning* 1718–1727.
[69] Liu, J., Huang, Y., Singh, R., Vert, J.-P. and Noble, W.S. (2019). Jointly embedding multiple single-cell omics measurements. In *BioRxiv* Cold Spring Harbor Laboratory.
[70] Liu, Q., Lee, J. and Jordan, M. (2016). A kernelized Stein discrepancy for goodness-of-fit tests. In *International Conference on Machine Learning* 276–284.
[71] Louhichi, S. (1998). Théorèmes limites pour des suites positivement ou faiblement dépendantes. Thèse de doctorat de l'Université Paris-XI.
[72] Lugosi, G. and Mendelson, S. (2019). Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.* **19** 1145–1190. MR4017683 https://doi.org/10.1007/s10208-019-09427-x
[73] Lugosi, G. and Mendelson, S. (2020). Risk minimization by median-of-means tournaments. *J. Eur. Math. Soc.* (*JEMS*) **22** 925–965. MR4055993 https://doi.org/10.4171/jems/937
[74] 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
[75] Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** 2308–2335. MR3378468 https://doi.org/10.3150/14-BEJ645
[76] Muandet, K., Fukumizu, K., Sriperumbudur, B. and Schölkopf, B. (2017). Kernel mean embedding of distributions: A review and beyond. *Found. Trends Mach. Learn.* **10** 1–141.
[77] Nemirovski, A., Juditsky, A., Lan, G. and Shapiro, A. (2008). Robust stochastic approximation approach to stochastic programming. *SIAM J. Optim.* **19** 1574–1609. MR2486041 https://doi.org/10.1137/070704277
[78] Nemirovsky, A.S. and Yudin, D.B. (1983). *Problem Complexity and Method Efficiency in Optimization*. *Wiley-Interscience Series in Discrete Mathematics*. New York: Wiley. MR0702836
[79] O'Hagan, A., Murphy, T.B. and Gormley, I.C. (2012). Computational aspects of fitting mixture models via the expectation-maximization algorithm. *Comput. Statist. Data Anal.* **56** 3843–3864. MR2957835 https://doi.org/10.1016/j.csda.2012.05.011
[80] Park, M., Jitkrittum, W. and Sejdinovic, D. (2016). K2-ABC: Approximate Bayesian computation with kernel embeddings. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* **51** 398–407.

[81] Parr, W.C. and Schucany, W.R. (1980). Minimum distance and robust estimation. *J. Amer. Statist. Assoc.* **75** 616–624. MR0590691
[82] Rio, E. (2000). Inégalités de Hoeffding pour les fonctions lipschitziennes de suites dépendantes. *C. R. Acad. Sci. Paris Sér. I Math.* **330** 905–908. MR1771956 https://doi.org/10.1016/S0764-4442(00)00290-1
[83] Rio, E. (2013). On McDiarmid's concentration inequality. *Electron. Commun. Probab.* **18** no. 44, 11. MR3070910 https://doi.org/10.1214/ECP.v18-2659
[84] Rio, E. (2017). *Asymptotic Theory of Weakly Dependent Random Processes*. *Probability Theory and Stochastic Modelling* **80**. Berlin: Springer. MR3642873 https://doi.org/10.1007/978-3-662-54323-8
[85] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** 43–47. MR0074711 https://doi.org/10.1073/pnas.42.1.43
[86] Song, L. (2008). Learning via Hilbert Space Embedding of Distributions Ph.D. thesis, University of Sydney.
[87] Song, L., Gretton, A., Bickson, D., Low, Y. and Guestrin, C. (2011). Kernel belief propagation. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* 707–715.
[88] Sriperumbudur, B.K., Gretton, A., Fukumizu, K., Schölkopf, B. and Lanckriet, G.R.G. (2010). Hilbert space embeddings and metrics on probability measures. *J. Mach. Learn. Res.* **11** 1517–1561. MR2645460
[89] Tolstikhin, I., Sriperumbudur, B.K. and Muandet, K. (2017). Minimax estimation of kernel mean embeddings. *J. Mach. Learn. Res.* **18** Paper No. 86, 47. MR3714249
[90] Tukey, J.W. (1975). Mathematics and the picturing of data. In *Proceedings of the International Congress of Mathematicians* (*Vancouver, B. C.*, 1974), *Vol.* 2 523–531. MR0426989
[91] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256
[92] Wolfowitz, J. (1957). The minimum distance method. *Ann. Math. Stat.* **28** 75–88. MR0088126 https://doi.org/10.1214/aoms/1177707038
[93] Wu, K., Ding, G.W., Huang, R. and Yu, Y. (2020). On minimax optimality of GANs for robust mean estimation. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* 4541–4551.
[94] Yatracos, Y.G. (1985). Rates of convergence of minimum distance estimators and Kolmogorov's entropy. *Ann. Statist.* **13** 768–774. MR0790571 https://doi.org/10.1214/aos/1176349553
[95] Zhao, S., Song, J. and Ermon, S. (2017). InfoVAE: Information Maximizing Variational Autoencoders. Preprint. Available at arXiv:1706.02262.

Bernoulli **28**(1), 2022, 214–238
https://doi.org/10.3150/21-BEJ1339

Limit theorems for time-dependent averages of nonlinear stochastic heat equations

KUNWOO KIM[*] and JAEYUN YI[†]

Department of Mathematics, Pohang University of Science and Technology (POSTECH), Pohang, Gyeongbuk 37673, Republic of Korea.
E-mail: [*]*kunwoo@postech.ac.kr*; [†]*stork@postech.ac.kr*

We study limit theorems for time-dependent averages of the form $X_t := \frac{1}{2L(t)} \int_{-L(t)}^{L(t)} u(t,x)\,dx$, as $t \to \infty$, where $L(t) = \exp(\lambda t)$ and $u(t,x)$ is the solution to a stochastic heat equation on $\mathbb{R}_+ \times \mathbb{R}$ driven by space-time white noise with $u_0(x) = 1$ for all $x \in \mathbb{R}$. We show that for X_t

(i) the weak law of large numbers holds when $\lambda > \lambda_1$,
(ii) the strong law of large numbers holds when $\lambda > \lambda_2$,
(iii) the central limit theorem holds when $\lambda > \lambda_3$, but fails when $\lambda < \lambda_4 \le \lambda_3$,
(iv) the quantitative central limit theorem holds when $\lambda > \lambda_5$,

where λ_i's are positive constants depending on the moment Lyapunov exponents of $u(t,x)$.

Keywords: Stochastic heat equation; weak law of large numbers; strong law of large numbers; central limit theorem

References

[1] Ben Arous, G., Molchanov, S. and Ramírez, A.F. (2005). Transition from the annealed to the quenched asymptotics for a random walk on random obstacles. *Ann. Probab.* **33** 2149–2187. MR2184094 https://doi.org/10.1214/009117905000000404
[2] Carmona, R.A. and Molchanov, S.A. (1994). Parabolic Anderson problem and intermittency. *Mem. Amer. Math. Soc.* **108** viii+125. MR1185878 https://doi.org/10.1090/memo/0518
[3] Chen, L., Khoshnevisan, D., Nualart, D. and Pu, F. (2019). Poincaré inequality, and central limit theorems for parabolic stochastic partial differential equations. Preprint. Available at arXiv:1912.01482.
[4] Chen, L., Khoshnevisan, D., Nualart, D. and Pu, F. (2019). Spatial ergodicity for SPDEs via Poincaré-type inequalities. Preprint. Available at arXiv:1912.01482.
[5] Chen, L., Khoshnevisan, D., Nualart, D. and Pu, F. (2020). Central limit theorems for spatial averges of the stochastic heat equation via Malliavin–Stein's method. Preprint. Available at arXiv:2008.02408.
[6] Chen, L. and Kim, K. (2017). On comparison principle and strict positivity of solutions to the nonlinear stochastic fractional heat equations. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 358–388. MR3606745 https://doi.org/10.1214/15-AIHP719
[7] Chen, L. and Kim, K. (2020). Stochastic comparisons for stochastic heat equation. *Electron. J. Probab.* **25** Paper No. 140, 1–38. MR4186259 https://doi.org/10.1214/20-ejp541
[8] Chen, X. (2015). Precise intermittency for the parabolic Anderson equation with an $(1+1)$-dimensional time-space white noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1486–1499. MR3414455 https://doi.org/10.1214/15-AIHP673
[9] Conus, D., Joseph, M. and Khoshnevisan, D. (2013). On the chaotic character of the stochastic heat equation, before the onset of intermitttency. *Ann. Probab.* **41** 2225–2260. MR3098071 https://doi.org/10.1214/11-AOP717
[10] Cranston, M. and Molchanov, S. (2007). Quenched to annealed transition in the parabolic Anderson problem. *Probab. Theory Related Fields* **138** 177–193. MR2288068 https://doi.org/10.1007/s00440-006-0020-7

 © 2022 ISI/BS

[11] Cranston, M., Mountford, T.S. and Shiga, T. (2002). Lyapunov exponents for the parabolic Anderson model. *Acta Math. Univ. Comenian.* (*N.S.*) **71** 163–188. MR1980378

[12] Dalang, R.C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** no. 6, 1–29. MR1684157 https://doi.org/10.1214/EJP.v4-43

[13] Das, S. and Tsai, L.C. (2019). Fractional moments of the stochastic heat equation. Preprint. Available at arXiv:1910.09271.

[14] Dharmadhikari, S.W. and Jogdeo, K. (1969). Bounds on moments of certain random variables. *Ann. Math. Stat.* **40** 1506–1509. MR0243587 https://doi.org/10.1214/aoms/1177697526

[15] Foondun, M. and Khoshnevisan, D. (2009). Intermittence and nonlinear parabolic stochastic partial differential equations. *Electron. J. Probab.* **14** 548–568. MR2480553 https://doi.org/10.1214/EJP.v14-614

[16] Gärtner, J. and Schnitzler, A. (2015). Stable limit laws for the parabolic Anderson model between quenched and annealed behaviour. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 194–206. MR3300968 https://doi.org/10.1214/13-AIHP574

[17] Ghosal, P. and Lin, Y. (2020). Lyapunov exponents of the SHE for general initial data. Preprint. Available at arXiv:2007.06505.

[18] Huang, J., Nualart, D. and Viitasaari, L. (2020). A central limit theorem for the stochastic heat equation. *Stochastic Process. Appl.* **130** 7170–7184. MR4167203 https://doi.org/10.1016/j.spa.2020.07.010

[19] Joseph, M., Khoshnevisan, D. and Mueller, C. (2017). Strong invariance and noise-comparison principles for some parabolic stochastic PDEs. *Ann. Probab.* **45** 377–403. MR3601652 https://doi.org/10.1214/15-AOP1009

[20] 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

[21] Khoshnevisan, D., Kim, K. and Xiao, Y. (2017). Intermittency and multifractality: A case study via parabolic stochastic PDEs. *Ann. Probab.* **45** 3697–3751. MR3729613 https://doi.org/10.1214/16-AOP1147

[22] Khoshnevisan, D., Kim, K. and Xiao, Y. (2018). A macroscopic multifractal analysis of parabolic stochastic PDEs. *Comm. Math. Phys.* **360** 307–346. MR3795193 https://doi.org/10.1007/s00220-018-3136-6

[23] Mueller, C. (1991). On the support of solutions to the heat equation with noise. *Stoch. Stoch. Rep.* **37** 225–245. MR1149348 https://doi.org/10.1080/17442509108833738

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

[25] von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Stat.* **36** 299–303. MR0170407 https://doi.org/10.1214/aoms/1177700291

[26] 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

Bernoulli **28**(1), 2022, 239–254
https://doi.org/10.3150/21-BEJ1340

Geometrical smeariness – a new phenomenon of Fréchet means

BENJAMIN ELTZNER

Felix Bernstein Institute for Mathematical Statistics in the Biosciences, Georg-August-University of Goettingen, Goettingen, Germany.
E-mail: benjamin.eltzner@mathematik.uni-goettingen.de

In the past decades, the central limit theorem (CLT) has been generalized to non-Euclidean data spaces. Some years ago, it was found that for some random variables on the circle, the sample Fréchet mean fluctuates around the population mean asymptotically at a scale $n^{-\tau}$ with exponent $\tau < 1/2$ with a non-normal distribution if the probability density at the antipodal point of the mean is $\frac{1}{2\pi}$. The author and his collaborator recently discovered that $\tau = 1/6$ for some random variables on higher dimensional spheres. In this article we show that, even more surprisingly, the phenomenon on spheres of higher dimension is qualitatively different from that on the circle, as it depends purely on geometrical properties of the space, namely its curvature, and not on the density at the antipodal point. This gives rise to the new concept of geometrical smeariness. In consequence, the sphere can be deformed, say, by removing a neighborhood of the antipodal point of the mean and gluing a flat space there, with a smooth transition piece. This yields smeariness on a manifold, which is diffeomorphic to Euclidean space. We give an example family of random variables with 2-smeary mean, that is, with $\tau = 1/6$, whose range has a hole containing the cut locus of the mean. The hole size exhibits a curse of dimensionality as it can increase with dimension, converging to the whole hemisphere opposite a local Fréchet mean. We observe smeariness in simulated landmark shapes on Kendall pre-shape space and in real data of geomagnetic north pole positions on the two-dimensional sphere.

Keywords: Fréchet means; asymptotics on manifolds; lower asymptotic rate; directional data; landmark shapes

References

[1] Afsari, B. (2009). *Means and Averaging on Riemannian Manifolds*. Ann Arbor, MI: ProQuest LLC. MR2718072
[2] 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
[3] Ahidar-Coutrix, A., Le Gouic, T. and Paris, Q. (2020). Convergence rates for empirical barycenters in metric spaces: Curvature, convexity and extendable geodesics. *Probab. Theory Related Fields* **177** 323–368. MR4095017 https://doi.org/10.1007/s00440-019-00950-0
[4] Arnaudon, M. and Miclo, L. (2014). Means in complete manifolds: Uniqueness and approximation. *ESAIM Probab. Stat.* **18** 185–206. MR3230874 https://doi.org/10.1051/ps/2013033
[5] Barden, D., Le, H. and Owen, M. (2013). Central limit theorems for Fréchet means in the space of phylogenetic trees. *Electron. J. Probab.* **18** no. 25, 25. MR3035753 https://doi.org/10.1214/EJP.v18-2201
[6] Barden, D., Le, H. and Owen, M. (2018). Limiting behaviour of Fréchet means in the space of phylogenetic trees. *Ann. Inst. Statist. Math.* **70** 99–129. MR3742820 https://doi.org/10.1007/s10463-016-0582-9
[7] Bhattacharya, A. and Bhattacharya, R. (2008). Statistics on Riemannian manifolds: Asymptotic distribution and curvature. *Proc. Amer. Math. Soc.* **136** 2959–2967. MR2399064 https://doi.org/10.1090/S0002-9939-08-09445-8
[8] Bhattacharya, A. and Bhattacharya, R. (2012). *Nonparametric Inference on Manifolds. Institute of Mathematical Statistics* (*IMS*) *Monographs* **2**. Cambridge: Cambridge Univ. Press. MR2934285 https://doi.org/10.1017/CBO9781139094764

1350-7265 © 2022 ISI/BS

[9] 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

[10] 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

[11] 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

[12] Chen, X. and Christensen, T.M. (2015). Optimal uniform convergence rates and asymptotic normality for series estimators under weak dependence and weak conditions. *J. Econometrics* **188** 447–465. MR3383220 https://doi.org/10.1016/j.jeconom.2015.03.010

[13] Clement, B. and Kent, D. (1985). A comparison of two sequential geomagnetic polarity transitions (upper Olduvai and lower Jaramillo) from the Southern Hemisphere. *Physics of the Earth and Planetary Interiors* (*Special Issue Irregularities in the Secular Variation and Geodynamic Implications*) **39** 301–313.

[14] Clement, B. and Kent, D. (1987). Geomagnetic polarity transition records from five hydraulic piston core sites in the North Atlantic. *Initial Reports of the Deep Sea Drilling Project* **94** 831–852.

[15] Clement, B., Kent, D. and Opdyke, N.D. (1982). Brunhes-matuyama polarity transition in three deep-sea sediment cores. *Philosophical Transactions of the Royal Society of London* **306** 113–119.

[16] Clement, B.M., Sierra, L.L., Smith, E. and Rodda, P. (1996). A record of the upper Cochiti polarity transition from the Southern Hemisphere (FIJI). *Surveys in Geophysics* **17** 189–196.

[17] Eltzner, B. (2022). Supplement to "Geometrical smeariness – a new phenomenon of Fréchet means." https://doi.org/10.3150/21-BEJ1340SUPP

[18] Eltzner, B., Galaz-Garcia, F., Huckemann, S.F. and Tuschmann, W. (2019). Stability of the cut locus and a central limit theorem for Fréchet means of Riemannian manifolds. arXiv:1909.00410 math.DG.

[19] Eltzner, B. and Huckemann, S.F. (2019). A smeary central limit theorem for manifolds with application to high-dimensional spheres. *Ann. Statist.* **47** 3360–3381. MR4025745 https://doi.org/10.1214/18-AOS1781

[20] Groisser, D. (2005). On the convergence of some Procrustean averaging algorithms. *Stochastics* **77** 31–60. MR2138772 https://doi.org/10.1080/17442500512331341059

[21] Gurariy, G.Z. (1988). Geomagnetic field during reversals in the late Cenozoic (Nauka, Moscow). pp. 207. USSR: Academy of Sciences of the USSR.

[22] Holt, J.W. and Kirschvink, J.L. (1995). The upper Olduvai geomagnetic field reversal from Death Valley, California: A fold test of transitional directions. *Earth Planet. Sci. Lett.* **133** 475–491.

[23] Hotz, T. and Huckemann, S. (2015). Intrinsic means on the circle: Uniqueness, locus and asymptotics. *Ann. Inst. Statist. Math.* **67** 177–193. MR3297863 https://doi.org/10.1007/s10463-013-0444-7

[24] Hotz, T., Huckemann, S., Le, H., Marron, J.S., Mattingly, J.C., Miller, E., Nolen, J., Owen, M., Patrangenaru, V. and Skwerer, S. (2013). Sticky central limit theorems on open books. *Ann. Appl. Probab.* **23** 2238–2258. MR3127934 https://doi.org/10.1214/12-AAP899

[25] Huckemann, S. (2011). Inference on 3D Procrustes means: Tree bole growth, rank deficient diffusion tensors and perturbation models. *Scand. J. Stat.* **38** 424–446. MR2833839 https://doi.org/10.1111/j.1467-9469.2010.00724.x

[26] Huckemann, S., Mattingly, J.C., Miller, E. and Nolen, J. (2015). Sticky central limit theorems at isolated hyperbolic planar singularities. *Electron. J. Probab.* **20** no. 78, 34. MR3371437 https://doi.org/10.1214/EJP.v20-3887

[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] Hundrieser, S.M.A. (2017). Exploring Smeary Limit Theorems. Bachelor Thesis, University of Göttingen.

[29] Hundrieser, S.M.A., Eltzner, B. and Huckemann, S.F. (2020). Finite sample smeariness of Fréchet means and application to climate. arXiv:2005.02321 stat.ME.

[30] Karcher, H. (1977). Riemannian center of mass and mollifier smoothing. *Comm. Pure Appl. Math.* **30** 509–541. MR0442975 https://doi.org/10.1002/cpa.3160300502

[31] Karcher, H. (2014). Riemannian center of mass and so called karcher mean. arXiv:1407.2087 math.HO.

[32] Kawai, N., Nakajima, T., Hirooka, K. and Kobayashi, K. (1973). The transition of field at the Brunhes and Jaramillo boundaries in the Matuyama Geomagnetic Epoch. *Proc. Jpn. Acad.* **49** 820–824.

[33] Kendall, W.S. (1990). Probability, convexity, and harmonic maps with small image. I. Uniqueness and fine existence. *Proc. Lond. Math. Soc.* (3) **61** 371–406. MR1063050 https://doi.org/10.1112/plms/s3-61.2.371

[34] Le, H. (2001). Locating Fréchet means with application to shape spaces. *Adv. in Appl. Probab.* **33** 324–338. MR1842295 https://doi.org/10.1239/aap/999188316
[35] Le, H. and Barden, D. (2014). On the measure of the cut locus of a Fréchet mean. *Bull. Lond. Math. Soc.* **46** 698–708. MR3239607 https://doi.org/10.1112/blms/bdu025
[36] Le Gouic, T., Paris, Q., Rigollet, P. and Stromme, A. (2019). Fast convergence of empirical barycenters in Alexandrov spaces and the Wasserstein space. arXiv:1908.00828 math.ST.
[37] 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
[38] Linssen, J.H. (1988). Preliminary results of a study of four successive sedimentary geomagnetic reversal records from the Mediterranean (Upper Thvera, Lower and Upper Sidufjall, and Lower Nunivak). *Phys. Earth Planet. Inter.* **52** 207–231.
[39] Linssen, J.H. (1991). Properties of Pliocene Sedimentary Geomagnetic Reversal Records from the Mediterranean (Ph.D Thesis, University of Utrecht) 80. Geologica Ultraiectina.
[40] Liu, J. and Yang, X. (2012). The convergence rate and asymptotic distribution of the bootstrap quantile variance estimator for importance sampling. *Adv. in Appl. Probab.* **44** 815–841. MR3024611 https://doi.org/10.1239/aap/1346955266
[41] Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis. Probability and Mathematical Statistics*: *A Series of Monographs and Textbooks*. London–New York–Toronto, Ont.: Academic Press. MR0560319
[42] McElhinny, M.W. and Lock, J. (1996). IAGA paleomagnetic databases with access. *Surv. Geophys.* **17** 575–591.
[43] Pennec, X. (2019). Curvature effects on the empirical mean in Riemannian and affine manifolds: A non-asymptotic high concentration expansion in the small-sample regime. arXiv:1906.07418 math.DG.
[44] Rolph, T.C. (1993). The matuyama-jaramillo R-n transition recorded in a loess section near Lanzhou, P. R. China. *J. Geomagn. Geoelectr.* **45** 301–318.
[45] Schötz, C. (2019). Convergence rates for the generalized Fréchet mean via the quadruple inequality. *Electron. J. Stat.* **13** 4280–4345. MR4023955 https://doi.org/10.1214/19-EJS1618
[46] Sturm, K.-T. (2003). Probability measures on metric spaces of nonpositive curvature. In *Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces* (*Paris*, 2002). *Contemp. Math.* **338** 357–390. Providence, RI: Amer. Math. Soc. MR2039961 https://doi.org/10.1090/conm/338/06080
[47] van Zijl, J.S.V., Graham, K.W.T. and Hales, A.L. (1962). The palaeomagnetism of the stormberg lavas, II. The behaviour of the magnetic field during a reversal. *Geophys. J. R. Astron. Soc.* **7** 169–182.
[48] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256
[49] Van Hoof, A.A.M. and Langereis, C.G. (1992). The Upper Kaena sedimentary geomagnetic reversal record from southern Sicily. *J. Geophys. Res., Solid Earth* **97** 6941–6957.
[50] Wellman, P., McElhinny, M.W. and McDougall, I. (1969). On the Polar-Wander Path for Australia during the Cenozoic. *Geophys. J. Int.* **18** 371–395.

Bernoulli **28**(1), 2022, 255–276
https://doi.org/10.3150/21-BEJ1341

Nonhomogeneous Euclidean first-passage percolation and distance learning

PABLO GROISMAN[1,2], MATTHIEU JONCKHEERE[3] and FACUNDO SAPIENZA[4]

[1]*IMAS-CONICET, Departamento de Matemática, Fac. Cs. Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina.*
E-mail: pgroisma@dm.uba.ar
[2]*NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai, Shanghai, China.*
E-mail: pgroisma@nyu.edu
[3]*Instituto de Calculo-CONICET, Fac. Cs. Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina.*
E-mail: mjonckhe@dm.uba.ar
[4]*Aristas SRL, Buenos Aires, Argentina.*
E-mail: f.sapienza@aristas.com.ar

Consider an i.i.d. sample from an unknown density function supported on an unknown manifold embedded in a high dimensional Euclidean space. We tackle the problem of learning a distance between points, able to capture both the geometry of the manifold and the underlying density. We define such a sample distance and prove the convergence, as the sample size goes to infinity, to a macroscopic one that we call *Fermat distance* as it minimizes a path functional, resembling Fermat principle in optics. The proof boils down to the study of geodesics in Euclidean first-passage percolation for nonhomogeneous Poisson point processes.

Keywords: Distance learning; Euclidean first-passage percolation; nonhomogeneous point processes

References

[1] Aggarwal, C.C., Hinneburg, A. and Keim, D.A. (2001). On the surprising behavior of distance metrics in high dimensional space. In *Database Theory*, *ICDT* 2001. Berlin, Heidelberg: Springer.
[2] Alamgir, M. and Von Luxburg, U. (2012). Shortest path distance in random k-nearest neighbor graphs. In *ICML*.
[3] Bernstein, M., de Silva, V., Langford, J.C. and Tenenbaum, J.B. (2000). Graph approximations to geodesics on embedded manifolds. Technical report, Stanford Univ.
[4] Bijral, A.S., Ratliff, N. and Srebro, N. (2011). Semi-supervised learning with density based distances. In *UAI* 11 43–50. AUAI Press.
[5] Carpio, A., Bonilla, L.L., Mathews, J.C. and Tannenbaum, A.R. (2019). Fingerprints of cancer by persistent homology. *bioRxiv*.
[6] Chang, H. and Yeung, D.Y. (2005). Robust path-based spectral clustering with application to image segmentation. In *Tenth IEEE International Conference on Computer Vision* (*ICCV* '05) **1** 278–285. IEEE.
[7] Costa, J.A. and Hero, A.O. (2004). Geodesic entropic graphs for dimension and entropy estimation in manifold learning. *IEEE Trans. Signal Process.* **52** 2210–2221. MR2085582 https://doi.org/10.1109/TSP.2004.831130
[8] Costa, J.A. and Hero, A.O. III (2006). Determining intrinsic dimension and entropy of high-dimensional shape spaces. In *Statistics and Analysis of Shapes. Model. Simul. Sci. Eng. Technol.* 231–252. Boston, MA: Birkhäuser. MR2274197 https://doi.org/10.1007/0-8176-4481-4_9
[9] Davis, E. and Sethuraman, S. (2019). Approximating geodesics via random points. *Ann. Appl. Probab.* **29** 1446–1486. MR3914549 https://doi.org/10.1214/18-AAP1414

 © 2022 ISI/BS

[10] De Silva, V. and Tenenbaum, J.B. (2004). Sparse multidimensional scaling using landmark points. Technical report, Stanford Univ.
[11] Howard, C.D. and Newman, C.M. (1997). Euclidean models of first-passage percolation. *Probab. Theory Related Fields* **108** 153–170. MR1452554 https://doi.org/10.1007/s004400050105
[12] Howard, C.D. and Newman, C.M. (2001). Geodesics and spanning trees for Euclidean first-passage percolation. *Ann. Probab.* **29** 577–623. MR1849171 https://doi.org/10.1214/aop/1008956685
[13] Hwang, S.J., Damelin, S.B. and Hero, A.O. III (2016). Shortest path through random points. *Ann. Appl. Probab.* **26** 2791–2823. MR3563194 https://doi.org/10.1214/15-AAP1162
[14] Mckenzie, D. and Damelin, S. (2019). Power weighted shortest paths for clustering Euclidean data. *Found. Data Sci.* **1** 307.
[15] Møller, J. and Waagepetersen, R.P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. *Monographs on Statistics and Applied Probability* **100**. Boca Raton, FL: CRC Press/CRC. MR2004226
[16] Myers, S.B. (1945). Arcs and geodesics in metric spaces. *Trans. Amer. Math. Soc.* **57** 217–227. MR0011792 https://doi.org/10.2307/1990203
[17] Penrose, M.D. and Yukich, J.E. (2013). Limit theory for point processes in manifolds. *Ann. Appl. Probab.* **23** 2161–2211. MR3127932 https://doi.org/10.1214/12-AAP897
[18] Potamias, M., Bonchi, F., Castillo, C. and Gionis, A. (2009). Fast shortest path distance estimation in large networks. In *Proceedings of the* 18*th ACM Conference on Information and Knowledge Management*, *CIKM* '09 867–876. New York, NY, USA: ACM.
[19] Sajama and Orlitsky, A. (2005). Estimating and computing density based distance metrics. In *Proceedings of the* 22*nd International Conference on Machine Learning*, *ICML* '05 760–767. New York, NY, USA: Association for Computing Machinery.
[20] Sapienza, F., Groisman, P. and Jonckheere, M. (2018). Weighted geodesic distance following Fermat's principle. In *International Conference on Learning Representation*.
[21] Tenenbaum, J.B., de Silva, V. and Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science* **290** 2319–2323.

Bernoulli **28**(1), 2022, 277–306
https://doi.org/10.3150/21-BEJ1342

The Hausdorff measure of the range and level sets of Gaussian random fields with sectorial local nondeterminism

CHEUK YIN LEE

Institut de Mathématiques, École Polytechnique Fédérale de Lausanne, Station 8, CH-1015 Lausanne, Switzerland.
E-mail: cheuk.lee@epfl.ch

We determine the exact Hausdorff measure functions for the range and level sets of a class of Gaussian random fields satisfying sectorial local nondeterminism and other assumptions. We also establish a Chung-type law of the iterated logarithm. The results can be applied to the Brownian sheet, fractional Brownian sheets whose Hurst indices are the same in all directions, and systems of linear stochastic wave equations in one spatial dimension driven by space–time white noise or colored noise.

Keywords: Gaussian random fields; Hausdorff measure; local nondeterminism; Brownian sheet; harmonizable representation

References

[1] Adler, R.J. (1978). The uniform dimension of the level sets of a Brownian sheet. *Ann. Probab.* **6** 509–515. MR0490818
[2] Adler, R.J. (1981). *The Geometry of Random Fields*. Chichester: Wiley. MR0611857
[3] Anderson, T.W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176. MR0069229 https://doi.org/10.2307/2032333
[4] Ayache, A., Wu, D. and Xiao, Y. (2008). Joint continuity of the local times of fractional Brownian sheets. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 727–748. MR2446295 https://doi.org/10.1214/07-AIHP131
[5] Ayache, A. and Xiao, Y. (2005). Asymptotic properties and Hausdorff dimensions of fractional Brownian sheets. *J. Fourier Anal. Appl.* **11** 407–439. MR2169474 https://doi.org/10.1007/s00041-005-4048-3
[6] Baraka, D. and Mountford, T.S. (2011). The exact Hausdorff measure of the zero set of fractional Brownian motion. *J. Theoret. Probab.* **24** 271–293. MR2782718 https://doi.org/10.1007/s10959-009-0271-1
[7] Dalang, R.C. (1999). Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s. *Electron. J. Probab.* **4** 1–29. MR1684157 https://doi.org/10.1214/EJP.v4-43
[8] Dalang, R.C. (2003). Level sets and excursions of the Brownian sheet. In *Topics in Spatial Stochastic Processes* (*Martina Franca*, 2001). *Lecture Notes in Math.* **1802** 167–208. Berlin: Springer. MR1975520 https://doi.org/10.1007/978-3-540-36259-3_5
[9] Dalang, R.C. and Frangos, N.E. (1998). The stochastic wave equation in two spatial dimensions. *Ann. Probab.* **26** 187–212. MR1617046 https://doi.org/10.1214/aop/1022855416
[10] Dalang, R.C., Lee, C.Y., Mueller, C. and Xiao, Y. Multiple points of Gaussian random fields. Available at arXiv:1911.09793.
[11] Dalang, R.C., Mueller, C. and Xiao, Y. (2017). Polarity of points for Gaussian random fields. *Ann. Probab.* **45** 4700–4751. MR3737922 https://doi.org/10.1214/17-AOP1176
[12] Dalang, R.C. and Sanz-Solé, M. (2010). Criteria for hitting probabilities with applications to systems of stochastic wave equations. *Bernoulli* **16** 1343–1368. MR2759182 https://doi.org/10.3150/09-BEJ247
[13] Ehm, W. (1981). Sample function properties of multiparameter stable processes. *Z. Wahrsch. Verw. Gebiete* **56** 195–228. MR0618272 https://doi.org/10.1007/BF00535741

 © 2022 ISI/BS

[14] Falconer, K. (1990). *Fractal Geometry*: *Mathematical Foundations and Applications*. Chichester: Wiley. MR1102677
[15] Geman, D. and Horowitz, J. (1980). Occupation densities. *Ann. Probab.* **8** 1–67. MR0556414
[16] Goldman, A. (1988). Mouvement Brownien à plusieurs páramètres: Mesure de Hausdorff des trajectoires. *Astérisque* **167** 105 pp. MR1005613
[17] Khoshnevisan, D. and Shi, Z. (1999). Brownian sheet and capacity. *Ann. Probab.* **27** 1135–1159. MR1733143 https://doi.org/10.1214/aop/1022677442
[18] Khoshnevisan, D., Wu, D. and Xiao, Y. (2006). Sectorial local non-determinism and the geometry of the Brownian sheet. *Electron. J. Probab.* **11** 817–843. MR2261054 https://doi.org/10.1214/EJP.v11-353
[19] Khoshnevisan, D. and Xiao, Y. (2007). Images of the Brownian sheet. *Trans. Amer. Math. Soc.* **359** 3125–3151. MR2299449 https://doi.org/10.1090/S0002-9947-07-04073-1
[20] Latała, R. and Matlak, D. (2017). Royen's proof of the Gaussian correlation inequality. In *Geometric Aspects of Functional Analysis*. *Lecture Notes in Math.* **2169** 265–275. Cham: Springer. MR3645127
[21] Ledoux, M. (1996). Isoperimetry and Gaussian analysis. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1994). *Lecture Notes in Math.* **1648** 165–294. Berlin: Springer. MR1600888 https://doi.org/10.1007/BFb0095676
[22] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces*: *Isoperimetry and Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete* (3) [*Results in Mathematics and Related Areas* (3)] **23**. Berlin: Springer. MR1102015 https://doi.org/10.1007/978-3-642-20212-4
[23] Lee, C.Y. (2022). Supplement to "The Hausdorff measure of the range and level sets of Gaussian random fields with sectorial local nondeterminism." https://doi.org/10.3150/21-BEJ1342SUPP
[24] Lee, C.Y. and Xiao, Y. (2019). Local nondeterminism and the exact modulus of continuity for stochastic wave equation. *Electron. Commun. Probab.* **24** Paper No. 52, 8. MR4003126 https://doi.org/10.1214/19-ecp264
[25] Li, W.V. and Shao, Q.-M. (2001). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes*: *Theory and Methods*. *Handbook of Statist.* **19** 533–597. Amsterdam: North-Holland. MR1861734 https://doi.org/10.1016/S0169-7161(01)19019-X
[26] Luan, N. and Xiao, Y. (2012). Spectral conditions for strong local nondeterminism and exact Hausdorff measure of ranges of Gaussian random fields. *J. Fourier Anal. Appl.* **18** 118–145. MR2885561 https://doi.org/10.1007/s00041-011-9193-2
[27] Mountford, T.S. (1989). Uniform dimension results for the Brownian sheet. *Ann. Probab.* **17** 1454–1462. MR1048937
[28] Mountford, T.S. and Nualart, E. (2004). Level sets of multiparameter Brownian motions. *Electron. J. Probab.* **9** 594–614. MR2080611 https://doi.org/10.1214/EJP.v9-169
[29] Pitt, L.D. (1978). Local times for Gaussian vector fields. *Indiana Univ. Math. J.* **27** 309–330. MR0471055 https://doi.org/10.1512/iumj.1978.27.27024
[30] Rogers, C.A. (1970). *Hausdorff Measures*. London: Cambridge Univ. Press. MR0281862
[31] Rogers, C.A. and Taylor, S.J. (1961). Functions continuous and singular with respect to a Hausdorff measure. *Mathematika* **8** 1–31. MR0130336 https://doi.org/10.1112/S0025579300002084
[32] Royen, T. (2014). A simple proof of the Gaussian correlation conjecture extended to some multivariate gamma distributions. *Far East J. Theor. Stat.* **48** 139–145. MR3289621
[33] Talagrand, M. (1994). The small ball problem for the Brownian sheet. *Ann. Probab.* **22** 1331–1354. MR1303647
[34] Talagrand, M. (1995). Hausdorff measure of trajectories of multiparameter fractional Brownian motion. *Ann. Probab.* **23** 767–775. MR1334170
[35] Talagrand, M. (1998). Multiple points of trajectories of multiparameter fractional Brownian motion. *Probab. Theory Related Fields* **112** 545–563. MR1664704 https://doi.org/10.1007/s004400050200
[36] Taylor, S.J. and Tricot, C. (1985). Packing measure, and its evaluation for a Brownian path. *Trans. Amer. Math. Soc.* **288** 679–699. MR0776398 https://doi.org/10.2307/1999958
[37] 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
[38] Wu, D. and Xiao, Y. (2007). Geometric properties of fractional Brownian sheets. *J. Fourier Anal. Appl.* **13** 1–37. MR2296726 https://doi.org/10.1007/s00041-005-5078-y

[39] Wu, D. and Xiao, Y. (2011). On local times of anisotropic Gaussian random fields. *Commun. Stoch. Anal.* **5** 15–39. MR2808534 https://doi.org/10.31390/cosa.5.1.03

[40] Xiao, Y. (1996). Hausdorff measure of the sample paths of Gaussian random fields. *Osaka J. Math.* **33** 895–913. MR1435460

[41] Xiao, Y. (1997). Hölder conditions for the local times and the Hausdorff measure of the level sets of Gaussian random fields. *Probab. Theory Related Fields* **109** 129–157. MR1469923 https://doi.org/10.1007/s004400050128

[42] Xiao, Y. (2009). Sample path properties of anisotropic Gaussian random fields. In *A Minicourse on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1962** 145–212. Berlin: Springer. MR2508776 https://doi.org/10.1007/978-3-540-85994-9_5

[43] Xiao, Y. and Zhang, T. (2002). Local times of fractional Brownian sheets. *Probab. Theory Related Fields* **124** 204–226. MR1936017 https://doi.org/10.1007/s004400200210

[44] Zhou, X. (1993). Hausdorff measure of the level sets of multi-parameter Wiener processes in one dimension. *Acta Math. Sin. New Ser.* **9** 390–400. MR1380092 https://doi.org/10.1007/BF02560132

Bernoulli **28**(1), 2022, 307–342
https://doi.org/10.3150/21-BEJ1344

Information geometry approach to parameter estimation in hidden Markov model

MASAHITO HAYASHI[1,2,3,4,5,6,7]

[1]*Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China*
[2]*Guangdong Provincial Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China*
[3]*Shenzhen Key Laboratory of Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen, 518055, China*
[4]*Graduate School of Mathematics, Nagoya University, Nagoya, 464-8602, Japan*
[5]*Center for Advanced Intelligence Project, RIKEN, Tokyo, 103-0027, Japan*
[6]*Centre for Quantum Technologies, National University of Singapore, 117543, Singapore*
[7]*Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen, 518066, China.*
E-mail: hayashi@sustech.edu.cn; masahito@math.nagoya-u.ac.jp

We consider the estimation of the transition matrix of a hidden Markovian process by using information geometry with respect to transition matrices. In this paper, only the histogram of k-memory data is used for the estimation. To establish our method, we focus on a partial observation model with the Markovian process and we propose an efficient estimator whose asymptotic estimation error is given as the inverse of projective Fisher information of transition matrices. This estimator is applied to the estimation of the transition matrix of the hidden Markovian process. In this application, we carefully discuss the equivalence problem for hidden Markovian process on the tangent space.

Keywords: Hidden Markov; em-algorithm; projective Fisher information matrix; partial observation model

References

[1] Amari, S. (1995). Information geometry of the EM and em algorithms for neural networks. *Neural Netw.* **8** 1379–1408.
[2] Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry*. *Translations of Mathematical Monographs* **191**. Oxford: Oxford Univ. Press. MR1800071 https://doi.org/10.1090/mmono/191
[3] Amari, S.-I. (2009). α-divergence is unique, belonging to both f-divergence and Bregman divergence classes. *IEEE Trans. Inf. Theory* **55** 4925–4931. MR2596950 https://doi.org/10.1109/TIT.2009.2030485
[4] Baum, L.E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *Ann. Math. Stat.* **37** 1554–1563. MR0202264 https://doi.org/10.1214/aoms/1177699147
[5] Ben-Ari, I. and Neumann, M. (2012). Probabilistic approach to Perron root, the group inverse, and applications. *Linear Multilinear Algebra* **60** 39–63. MR2869672 https://doi.org/10.1080/03081087.2011.559167
[6] Brègman, L.M. (1967). A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.* **7** 200–217.
[7] Burgarth, D., Chiribella, G., Giovannetti, V., Perinotti, P. and Yuasa, K. (2013). Ergodic and mixing quantum channels in finite dimensions. *New J. Phys.* **15** 073045, 33. MR3094122 https://doi.org/10.1088/1367-2630/15/7/073045
[8] Burgarth, D. and Giovannetti, V. (2007). The generalized Lyapunov theorem and its application to quantum channels. *New J. Phys.* **9** 150.1–150.18. MR2326802 https://doi.org/10.1088/1367-2630/9/5/150
[9] Cappé, O., Moulines, E. and Rydén, T. (2005). *Inference in Hidden Markov Models*. *Springer Series in Statistics*. New York: Springer. MR2159833

 © 2022 ISI/BS

[10] Conn, A.R., Scheinberg, K. and Vicente, L.N. (2009). *Introduction to Derivative-Free Optimization*. *MPS/SIAM Series on Optimization* **8**. Philadelphia, PA: SIAM. MR2487816 https://doi.org/10.1137/1.9780898718768
[11] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. New York: Springer. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
[12] Durbin, R., Eddy, S., Krogh, A. and Mitchison, G. (1998). *Biological Sequence Analysis*: *Probabilistic Models of Proteins and Nucleic Acids*. Cambridge: CUP.
[13] Ekeland, I. and Téman, R. (1976). *Convex Analysis and Variational Problems*. Amsterdam: North-Holland; SIAM, Philadelphia, PA, 1999.
[14] Felsenstein, J. and Churchill, G.A. (1996). A hidden Markov model approach to variation among sites in rate of evolution. *Mol. Biol. Evol.* **13** 93–104.
[15] Fujimoto, Y. and Murata, N. (2007). A modified EM algorithm for mixture models based on Bregman divergence. *Ann. Inst. Statist. Math.* **59** 3–25. MR2405286 https://doi.org/10.1007/s10463-006-0097-x
[16] Halmos, P.R. (1974). *Finite-Dimensional Vector Spaces*, 2nd ed. New York: Springer. MR0409503
[17] Hayashi, M. (2017). *Quantum Information Theory*, 2nd ed. *Graduate Texts in Physics*. Berlin: Springer. MR3558531 https://doi.org/10.1007/978-3-662-49725-8
[18] Hayashi, M. (2019). Local equivalence problem in hidden Markov model. *Inf. Geom.* **2** 1–42. MR3984155 https://doi.org/10.1007/s41884-019-00016-z
[19] Hayashi, M. and Watanabe, S. (2016). Information geometry approach to parameter estimation in Markov chains. *Ann. Statist.* **44** 1495–1535. MR3519931 https://doi.org/10.1214/15-AOS1420
[20] Ito, H., Amari, S. and Kobayashi, K. (1992). Identifiability of hidden Markov information sources and their minimum degrees of freedom. *IEEE Trans. Inf. Theory* **38** 324–333. MR1162206 https://doi.org/10.1109/18.119690
[21] Jones, G.L. (2004). On the Markov chain central limit theorem. *Probab. Surv.* **1** 299–320. MR2068475 https://doi.org/10.1214/154957804100000051
[22] Kemeny, J.G. and Snell, J.L. (1960). *Finite Markov Chains*. *The University Series in Undergraduate Mathematics*. Princeton, N.J.-Toronto-London: D. Van Nostrand Co., Inc. MR0115196
[23] Khreich, W., Granger, E., Miri, A. and Sabourin, R. (2012). A survey of techniques for incremental learning of HMM parameters. *Inform. Sci.* **197** 105–130.
[24] Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Rev. Econ. Stud.* **65** 361–393.
[25] Kontoyiannis, I. and Meyn, S.P. (2003). Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab.* **13** 304–362. MR1952001 https://doi.org/10.1214/aoap/1042765670
[26] Lalley, S.P. (1986). Ruelle's Perron–Frobenius theorem and the central limit theorem for additive functionals of one-dimensional Gibbs states. In *Adaptive Statistical Procedures and Related Topics* (*Upton, N.Y.*, 1985). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **8** 428–446. Hayward, CA: IMS. MR0898264 https://doi.org/10.1214/lnms/1215540317
[27] Matsui, K., Kumagai, W. and Kanamori, T. (2017). Parallel distributed block coordinate descent methods based on pairwise comparison oracle. *J. Global Optim.* **69** 1–21. MR3687714 https://doi.org/10.1007/s10898-016-0465-x
[28] Meyn, S.P. and Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability*. *Communications and Control Engineering Series*. London: Springer London, Ltd. MR1287609 https://doi.org/10.1007/978-1-4471-3267-7
[29] Mitzenmacher, M. and Upfal, E. (2005). *Probability and Computing*: *Randomized Algorithms and Probabilistic Analysis*. Cambridge: Cambridge Univ. Press. MR2144605 https://doi.org/10.1017/CBO9780511813603
[30] Murata, N., Takenouchi, T., Kanamori, T. and Eguchi, S. (2004). Information geometry of U-boost and Bregman divergence. *Neural Comput.* **16** 1437–1481.
[31] Nagaoka, H. (2005). The exponential family of Markov chains and its information geometry. In *Proceedings of The 28th Symposium on Information Theory and Its Applications (SITA2005)*, Okinawa, Japan, Nov. 20–23. The exponential family of Markov chains and its information geometry.
[32] Nakagawa, K. and Kanaya, F. (1993). On the converse theorem in statistical hypothesis testing for Markov chains. *IEEE Trans. Inf. Theory* **39** 629–633. MR1224350 https://doi.org/10.1109/18.212294

[33] Nelder, J.A. and Mead, R. (1965). A simplex method for function minimization. *Comput. J.* **7** 308–313. MR3363409 https://doi.org/10.1093/comjnl/7.4.308

[34] Seneta, E. (1981). *Non-negative Matrices and Markov Chains*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR2209438

[35] Watanabe, S. and Hayashi, M. (2017). Finite-length analysis on tail probability for Markov chain and application to simple hypothesis testing. *Ann. Appl. Probab.* **27** 811–845. MR3655854 https://doi.org/10.1214/16-AAP1216

[36] Yoshida, Y. and Hayashi, M. (2020). Mixing and asymptotically decoupling properties in general probabilistic theories. *J. Phys. A: Math. Theor.* **53** 215303.

Bernoulli **28**(1), 2022, 343–369
https://doi.org/10.3150/21-BEJ1345

Strong and weak convergence rates for slow–fast stochastic differential equations driven by α-stable process

XIAOBIN SUN[*], LONGJIE XIE[†] and YINGCHAO XIE[‡]

School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou, 221000, China.
E-mail: [*] xbsun@jsnu.edu.cn; [†] xlj.98@whu.edu.cn; [‡] ycxie@jsnu.edu.cn

In this paper, we study the averaging principle for a class of stochastic differential equations driven by α-stable processes with slow and fast time-scales, where $\alpha \in (1, 2)$. We prove that the strong and weak convergence order are $1 - 1/\alpha$ and 1 respectively. We show, by a simple example, that $1 - 1/\alpha$ is the optimal strong convergence rate.

Keywords: Averaging principle; α-stable process; slow–fast system; convergence rates

References

[1] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge: Cambridge Univ. Press. MR2512800 https://doi.org/10.1017/CBO9780511809781
[2] Bakhtin, V. and Kifer, Y. (2004). Diffusion approximation for slow motion in fully coupled averaging. *Probab. Theory Related Fields* **129** 157–181. MR2063374 https://doi.org/10.1007/s00440-003-0326-7
[3] Bao, J., Yin, G. and Yuan, C. (2017). Two-time-scale stochastic partial differential equations driven by α-stable noises: Averaging principles. *Bernoulli* **23** 645–669. MR3556788 https://doi.org/10.3150/14-BEJ677
[4] Bertram, R. and Rubin, J.E. (2017). Multi-timescale systems and fast-slow analysis. *Math. Biosci.* **287** 105–121. MR3634156 https://doi.org/10.1016/j.mbs.2016.07.003
[5] Bréhier, C.-E. (2012). Strong and weak orders in averaging for SPDEs. *Stochastic Process. Appl.* **122** 2553–2593. MR2926167 https://doi.org/10.1016/j.spa.2012.04.007
[6] Bréhier, C.-E. (2013). Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. *SIAM J. Numer. Anal.* **51** 1185–1210. MR3040961 https://doi.org/10.1137/110853078
[7] Bréhier, C.-E. (2020). Orders of convergence in the averaging principle for SPDEs: The case of a stochastically forced slow component. *Stochastic Process. Appl.* **130** 3325–3368. MR4092407 https://doi.org/10.1016/j.spa.2019.09.015
[8] Cerrai, S. (2009). A Khasminskii type averaging principle for stochastic reaction-diffusion equations. *Ann. Appl. Probab.* **19** 899–948. MR2537194 https://doi.org/10.1214/08-AAP560
[9] Cerrai, S. (2011). Averaging principle for systems of reaction-diffusion equations with polynomial nonlinearities perturbed by multiplicative noise. *SIAM J. Math. Anal.* **43** 2482–2518. MR2854919 https://doi.org/10.1137/100806710
[10] Cerrai, S. and Freidlin, M. (2009). Averaging principle for a class of stochastic reaction-diffusion equations. *Probab. Theory Related Fields* **144** 137–177. MR2480788 https://doi.org/10.1007/s00440-008-0144-z
[11] Chen, Y., Shi, Y. and Sun, X. (2020). Averaging principle for slow–fast stochastic Burgers equation driven by α-stable process. *Appl. Math. Lett.* **103** 106199, 9. MR4053803 https://doi.org/10.1016/j.aml.2019.106199
[12] Dong, Z., Sun, X., Xiao, H. and Zhai, J. (2018). Averaging principle for one dimensional stochastic Burgers equation. *J. Differential Equations* **265** 4749–4797. MR3848236 https://doi.org/10.1016/j.jde.2018.06.020
[13] E, W., Liu, D. and Vanden-Eijnden, E. (2005). Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.* **58** 1544–1585. MR2165382 https://doi.org/10.1002/cpa.20088

[14] Fu, H. and Liu, J. (2011). Strong convergence in stochastic averaging principle for two time-scales stochastic partial differential equations. *J. Math. Anal. Appl.* **384** 70–86. MR2822851 https://doi.org/10.1016/j.jmaa.2011.02.076

[15] Fu, H., Wan, L. and Liu, J. (2015). Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales. *Stochastic Process. Appl.* **125** 3255–3279. MR3343294 https://doi.org/10.1016/j.spa.2015.03.004

[16] Fu, H., Wan, L., Liu, J. and Liu, X. (2018). Weak order in averaging principle for stochastic wave equation with a fast oscillation. *Stochastic Process. Appl.* **128** 2557–2580. MR3811697 https://doi.org/10.1016/j.spa.2017.09.021

[17] Gao, P. (2018). Averaging principle for stochastic Kuramoto–Sivashinsky equation with a fast oscillation. *Discrete Contin. Dyn. Syst.* **38** 5649–5684. MR3917783 https://doi.org/10.3934/dcds.2018247

[18] Gao, P. (2018). Averaging principle for the higher order nonlinear Schrödinger equation with a random fast oscillation. *J. Stat. Phys.* **171** 897–926. MR3800899 https://doi.org/10.1007/s10955-018-2048-3

[19] Gao, P. (2019). Averaging principle for multiscale stochastic Klein–Gordon-heat system. *J. Nonlinear Sci.* **29** 1701–1759. MR3993181 https://doi.org/10.1007/s00332-019-09529-4

[20] Gao, P. (2019). Averaging principle for stochastic Korteweg–de Vries equation. *J. Differential Equations* **267** 6872–6909. MR4011034 https://doi.org/10.1016/j.jde.2019.07.012

[21] Givon, D. (2007). Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems. *Multiscale Model. Simul.* **6** 577–594. MR2338495 https://doi.org/10.1137/060673345

[22] Givon, D., Kevrekidis, I.G. and Kupferman, R. (2006). Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems. *Commun. Math. Sci.* **4** 707–729. MR2264816

[23] Harvey, E., Kirk, V., Wechselberger, M. and Sneyd, J. (2011). Multiple timescales, mixed mode oscillations and canards in models of intracellular calcium dynamics. *J. Nonlinear Sci.* **21** 639–683.

[24] Has'minskiĭ, R.Z. (1968). On the principle of averaging the Itô's stochastic differential equations. *Kybernetika* (*Prague*) **4** 260–279. MR0260052

[25] Khasminskii, R.Z. and Yin, G. (2004). On averaging principles: An asymptotic expansion approach. *SIAM J. Math. Anal.* **35** 1534–1560. MR2083789 https://doi.org/10.1137/S0036141002403973

[26] Khasminskii, R.Z. and Yin, G. (2005). Limit behavior of two-time-scale diffusions revisited. *J. Differential Equations* **212** 85–113. MR2130548 https://doi.org/10.1016/j.jde.2004.08.013

[27] Kifer, Y. (2005). Another proof of the averaging principle for fully coupled dynamical systems with hyperbolic fast motions. *Discrete Contin. Dyn. Syst.* **13** 1187–1201. MR2166264 https://doi.org/10.3934/dcds.2005.13.1187

[28] Liu, D. (2010). Strong convergence of principle of averaging for multiscale stochastic dynamical systems. *Commun. Math. Sci.* **8** 999–1020. MR2744917

[29] Liu, D. (2012). Strong convergence rate of principle of averaging for jump-diffusion processes. *Front. Math. China* **7** 305–320. MR2897706 https://doi.org/10.1007/s11464-012-0193-6

[30] Liu, W., Röckner, M., Sun, X. and Xie, Y. (2020). Averaging principle for slow–fast stochastic differential equations with time dependent locally Lipschitz coefficients. *J. Differential Equations* **268** 2910–2948. MR4047972 https://doi.org/10.1016/j.jde.2019.09.047

[31] Pardoux, E. and Veretennikov, A.Y. (2001). On the Poisson equation and diffusion approximation. I. *Ann. Probab.* **29** 1061–1085. MR1872736 https://doi.org/10.1214/aop/1015345596

[32] Pardoux, È. and Veretennikov, A.Y. (2003). On Poisson equation and diffusion approximation. II. *Ann. Probab.* **31** 1166–1192. MR1988467 https://doi.org/10.1214/aop/1055425774

[33] Pavliotis, G.A. and Stuart, A.M. (2008). *Multiscale Methods*: *Averaging and Homogenization*. *Texts in Applied Mathematics* **53**. New York: Springer. MR2382139

[34] Peszat, S. and Zabczyk, J. (2007). *Stochastic Partial Differential Equations with Lévy Noise*: *An Evolution Equation Approach*. *Encyclopedia of Mathematics and Its Applications* **113**. Cambridge: Cambridge Univ. Press. MR2356959 https://doi.org/10.1017/CBO9780511721373

[35] Priola, E. and Zabczyk, J. (2011). Structural properties of semilinear SPDEs driven by cylindrical stable processes. *Probab. Theory Related Fields* **149** 97–137. MR2773026 https://doi.org/10.1007/s00440-009-0243-5

[36] Röckner, M., Sun, X. and Xie, L. Strong and weak convergence in the averaging principle for SDEs with Hölder coefficients. https://arxiv.org/abs/1907.09256.

[37] 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.
[38] Sun, X. and Zhai, J. (2020). Averaging principle for stochastic real Ginzburg–Landau equation driven by α-stable process. *Commun. Pure Appl. Anal.* **19** 1291–1319. MR4064032 https://doi.org/10.3934/cpaa.2020063
[39] Veretennikov, A.Y. (1991). On the averaging principle for systems of stochastic differential equations. *Math. USSR, Sb.* **69** 271–284. MR1046602 https://doi.org/10.1070/SM1991v069n01ABEH001237
[40] Wang, J. (2013). Exponential ergodicity and strong ergodicity for SDEs driven by symmetric α-stable processes. *Appl. Math. Lett.* **26** 654–658. MR3028071 https://doi.org/10.1016/j.aml.2013.01.004
[41] Wang, W. and Roberts, A.J. (2012). Average and deviation for slow–fast stochastic partial differential equations. *J. Differential Equations* **253** 1265–1286. MR2927381 https://doi.org/10.1016/j.jde.2012.05.011
[42] Xu, J. (2017). L^p-strong convergence of the averaging principle for slow–fast SPDEs with jumps. *J. Math. Anal. Appl.* **445** 342–373. MR3543771 https://doi.org/10.1016/j.jmaa.2016.07.058
[43] Xu, J., Miao, Y. and Liu, J. (2015). Strong averaging principle for slow–fast SPDEs with Poisson random measures. *Discrete Contin. Dyn. Syst. Ser. B* **20** 2233–2256. MR3423220 https://doi.org/10.3934/dcdsb.2015.20.2233
[44] Xu, J., Miao, Y. and Liu, J. (2016). Strong averaging principle for two-time-scale non-autonomous stochastic FitzHugh–Nagumo system with jumps. *J. Math. Phys.* **57** 092704, 21. MR3551157 https://doi.org/10.1063/1.4963173
[45] Zhang, B., Fu, H., Wan, L. and Liu, J. (2018). Weak order in averaging principle for stochastic differential equations with jumps. *Adv. Difference Equ.* Paper No. 197, 20. MR3805681 https://doi.org/10.1186/s13662-018-1638-3
[46] Zhang, Y., Cheng, Z., Zhang, X., Chen, X., Duan, J. and Li, X. (2017). Data assimilation and parameter estimation for a multiscale stochastic system with α-stable Lévy noise. *J. Stat. Mech. Theory Exp.* **11** 113401, 17. MR3737653 https://doi.org/10.1088/1742-5468/aa9343
[47] Zulfiqar, H., Yuan, S., He, Z. and Duan, J. (2019). Slow manifolds for a nonlocal fast-slow stochastic system with stable Lévy noise. *J. Math. Phys.* **60** 091501, 19. MR3998715 https://doi.org/10.1063/1.5093181

Bernoulli **28**(1), 2022, 370–394
https://doi.org/10.3150/21-BEJ1346

Applications of weak transport theory

J. BACKHOFF-VERAGUAS[1] and G. PAMMER[2]

[1] *Universität Wien, Vienna, Austria.*
E-mail: julio.backhoff@univie.ac.at
[2] *ETH, Zürich, Switzerland.*
E-mail: gudmund.pammer@math.ethz.ch

Motivated by applications to geometric inequalities, Gozlan, Roberto, Samson, and Tetali (*J. Funct. Anal.* **273** (2017) 3327–3405) introduced a transport problem for 'weak' cost functionals. Basic results of optimal transport theory can be extended to this setup in remarkable generality.

In this article, we collect several problems from different areas that can be recast in the framework of weak transport theory, namely: the Schrödinger problem, the Brenier–Strassen theorem, optimal mechanism design, linear transfers, semimartingale transport. Our viewpoint yields a unified approach and often allows to strengthen the original results.

Keywords: Schrödinger problem; Brenier–Strassen theorem; linear transfers; semimartingale transport; optimal mechanism design; weak transport problem; duality; cyclical monotonicity

References

[1] Alfonsi, A., Corbetta, J. and Jourdain, B. (2020). Sampling of probability measures in the convex order and approximation of Martingale Optimal Transport problems. *ArXiv e-prints*, Sept. 2017.
[2] Alibert, J.-J., Bouchitté, G. and Champion, T. (2019). A new class of costs for optimal transport planning. *European J. Appl. Math.* **30** 1229–1263. MR4028478 https://doi.org/10.1017/s0956792518000669
[3] Backhoff, J., Beiglböck, M., Lin, Y. and Zalashko, A. (2017). Causal transport in discrete time and applications. *SIAM J. Optim.* **27** 2528–2562. MR3738324 https://doi.org/10.1137/16M1080197
[4] Backhoff-Veraguas, J., Bartl, D., Beiglböck, M. and Eder, M. (2020). Adapted Wasserstein distances and stability in mathematical finance. *Finance Stoch.* **24** 601–632. MR4118990 https://doi.org/10.1007/s00780-020-00426-3
[5] Backhoff-Veraguas, J., Beiglböck, M., Huesmann, M. and Källblad, S. (2020). Martingale Benamou-Brenier: A probabilistic perspective. *Ann. Probab.* **48** 2258–2289. MR4152642 https://doi.org/10.1214/20-AOP1422
[6] Backhoff-Veraguas, J., Beiglböck, M. and Pammer, G. (2019). Existence, duality, and cyclical monotonicity for weak transport costs. *Calc. Var. Partial Differential Equations* **58** Paper No. 203, 28. MR4029731 https://doi.org/10.1007/s00526-019-1624-y
[7] Backhoff-Veraguas, J., Beiglböck, M. and Pammer, G. (2020). Weak monotone rearrangement on the line. *Electron. Commun. Probab.* **25** Paper No. 18, 16. MR4069738 https://doi.org/10.1214/20-ecp292
[8] Backhoff-Veraguas, J., Lacker, D. and Tangpi, L. (2020). Nonexponential Sanov and Schilder theorems on Wiener space: BSDEs, Schrödinger problems and control. *Ann. Appl. Probab.* **30** 1321–1367. MR4133375 https://doi.org/10.1214/19-AAP1531
[9] Backhoff-Veraguas, J. and Pammer, G. (2020). Stability of martingale optimal transport and weak optimal transport. Available at arXiv:1904.04171.
[10] Beiglböck, M. (2015). Cyclical monotonicity and the ergodic theorem. *Ergodic Theory Dynam. Systems* **35** 710–713. MR3334900 https://doi.org/10.1017/etds.2013.75
[11] Beiglböck, M., Cox, A.M.G. and Huesmann, M. (2017). Optimal transport and Skorokhod embedding. *Invent. Math.* **208** 327–400. MR3639595 https://doi.org/10.1007/s00222-016-0692-2
[12] Beiglböck, M. and Griessler, C. (2019). A land of monotone plenty. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **19** 109–127. MR3923841

1350-7265 © 2022 ISI/BS

[13] Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices—a mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985 https://doi.org/10.1007/s00780-013-0205-8

[14] Beiglböck, M. and Juillet, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. MR3456332 https://doi.org/10.1214/14-AOP966

[15] Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L. and Peyré, G. (2015). Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.* **37** A1111–A1138. MR3340204 https://doi.org/10.1137/141000439

[16] Bianchini, S. and Caravenna, L. (2010). On optimality of c-cyclically monotone transference plans. *C. R. Math. Acad. Sci. Paris* **348** 613–618. MR2652484 https://doi.org/10.1016/j.crma.2010.03.022

[17] Bowles, M. and Ghoussoub, N. (2018). A Theory of Transfers: Duality and convolution. Available at arXiv:1804.08563.

[18] Brenier, Y. (1987). Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.* **305** 805–808. MR0923203

[19] Brenier, Y. (1991). Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** 375–417. MR1100809 https://doi.org/10.1002/cpa.3160440402

[20] Colombo, M., De Pascale, L. and Di Marino, S. (2015). Multimarginal optimal transport maps for one-dimensional repulsive costs. *Canad. J. Math.* **67** 350–368. MR3314838 https://doi.org/10.4153/CJM-2014-011-x

[21] Cox, A.M. and Vidmar, M. (2019). The structure of non-linear martingale optimal transport problems. Preprint.

[22] 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

[23] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Proceedings of the 26th International Conference on Neural Information Processing Systems* **2** 2292–2300.

[24] Daskalakis, C., Deckelbaum, A. and Tzamos, C. (2017). Strong duality for a multiple-good monopolist. *Econometrica* **85** 735–767. MR3664178 https://doi.org/10.3982/ECTA12618

[25] Diestel, J. (1991). Uniform integrability: An introduction. School on measure theory and real analysis (Grado, 1991). *Rend. Istit. Mat. Univ. Trieste* **23** 41–80. MR1248651

[26] 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

[27] Fathi, M., Gozlan, N. and Prod'homme, M. (2020). A proof of the Caffarelli contraction theorem via entropic regularization. *Calc. Var. Partial Differential Equations* **59** Paper No. 96, 18. MR4098037 https://doi.org/10.1007/s00526-020-01754-0

[28] Fathi, M. and Shu, Y. (2018). Curvature and transport inequalities for Markov chains in discrete spaces. *Bernoulli* **24** 672–698. MR3706773 https://doi.org/10.3150/16-BEJ892

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

[30] Gangbo, W. and McCann, R.J. (1996). The geometry of optimal transportation. *Acta Math.* **177** 113–161. MR1440931 https://doi.org/10.1007/BF02392620

[31] Gozlan, N. and Juillet, N. (2020). On a mixture of Brenier and Strassen theorems. *Proc. Lond. Math. Soc.* (3) **120** 434–463. MR4008375 https://doi.org/10.1112/plms.12302

[32] Gozlan, N., Roberto, C., Samson, P.-M., Shu, Y. and Tetali, P. (2018). Characterization of a class of weak transport-entropy inequalities on the line. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 1667–1693. MR3825894 https://doi.org/10.1214/17-AIHP851

[33] Gozlan, N., Roberto, C., Samson, P.-M. and Tetali, P. (2017). Kantorovich duality for general transport costs and applications. *J. Funct. Anal.* **273** 3327–3405. MR3706606 https://doi.org/10.1016/j.jfa.2017.08.015

[34] Griessler, C. (2018). C-cyclical monotonicity as a sufficient criterion for optimality in the multimarginal Monge-Kantorovich problem. *Proc. Amer. Math. Soc.* **146** 4735–4740. MR3856141 https://doi.org/10.1090/proc/14129

[35] Guyon, J., Menegaux, R. and Nutz, M. (2017). Bounds for VIX futures given S&P 500 smiles. *Finance Stoch*. **21** 593–630. MR3663638 https://doi.org/10.1007/s00780-017-0334-6
[36] Karandikar, R.L. (2006). On almost sure convergence results in stochastic calculus. In *In Memoriam Paul-André Meyer*: *Séminaire de Probabilités XXXIX*. *Lecture Notes in Math*. **1874** 137–147. Berlin: Springer. MR2276893 https://doi.org/10.1007/978-3-540-35513-7_11
[37] Léonard, C. (2014). A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete Contin*. *Dyn*. *Syst*. **34** 1533–1574. MR3121631 https://doi.org/10.3934/dcds.2014.34.1533
[38] 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
[39] Marton, K. (1996). A measure concentration inequality for contracting Markov chains. *Geom*. *Funct*. *Anal*. **6** 556–571. MR1392329 https://doi.org/10.1007/BF02249263
[40] Mikami, T. and Thieullen, M. (2006). Duality theorem for the stochastic optimal control problem. *Stochastic Process*. *Appl*. **116** 1815–1835. MR2307060 https://doi.org/10.1016/j.spa.2006.04.014
[41] Nutz, M. and Stebegg, F. (2018). Canonical supermartingale couplings. *Ann*. *Probab*. **46** 3351–3398. MR3857858 https://doi.org/10.1214/17-AOP1249
[42] Pass, B. (2012). On the local structure of optimal measures in the multi-marginal optimal transportation problem. *Calc*. *Var*. *Partial Differential Equations* **43** 529–536. MR2875651 https://doi.org/10.1007/s00526-011-0421-z
[43] Rockafellar, R.T. and Wets, R.J.-B. (2009). *Variational Analysis* **317**. Berlin: Springer.
[44] Samson, P.-M. (2017). Transport-entropy inequalities on locally acting groups of permutations. *Electron*. *J*. *Probab*. **22** Paper No. 62, 33. MR3690287 https://doi.org/10.1214/17-EJP54
[45] Shu, Y. (2018). Hamilton-Jacobi equations on graph and applications. *Potential Anal*. **48** 125–157. MR3748388 https://doi.org/10.1007/s11118-017-9628-8
[46] Shu, Y. (2020). From Hopf-Lax formula to optimal weak transfer plan. *SIAM J*. *Math*. *Anal*. **52** 3052–3072. MR4117845 https://doi.org/10.1137/17M1152231
[47] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann*. *Math*. *Stat*. **36** 423–439. MR0177430 https://doi.org/10.1214/aoms/1177700153
[48] Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publ*. *Math*. *Inst*. *Hautes Études Sci*. **81** 73–205. MR1361756
[49] Talagrand, M. (1996). New concentration inequalities in product spaces. *Invent*. *Math*. **126** 505–563. MR1419006 https://doi.org/10.1007/s002220050108
[50] Tan, X. and Touzi, N. (2013). Optimal transportation under controlled stochastic dynamics. *Ann*. *Probab*. **41** 3201–3240. MR3127880 https://doi.org/10.1214/12-AOP797
[51] 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
[52] Villani, C. (2009). *Optimal Transport*: *Old and New*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Berlin: Springer. MR2459454 https://doi.org/10.1007/978-3-540-71050-9
[53] Zaev, D.A. (2015). On the Monge-Kantorovich problem with additional linear constraints. *Mat*. *Zametki* **98** 664–683. MR3438523 https://doi.org/10.4213/mzm10896
[54] Zălinescu, C. (2002). *Convex Analysis in General Vector Spaces*. River Edge, NJ: World Scientific. MR1921556 https://doi.org/10.1142/9789812777096
[55] Zhou, X. (2018). On the Fenchel Duality between Strong Convexity and Lipschitz Continuous Gradient. Preprint.

Bernoulli **28**(1), 2022, 395–423
https://doi.org/10.3150/21-BEJ1347

On bandwidth selection problems in nonparametric trend estimation under martingale difference errors

KARIM BENHENNI[1,*], DIDIER A. GIRARD[2] and SANA LOUHICHI[1,†]

[1]*Laboratoire Jean Kuntzmann, Université Grenoble Alpes, 700 Avenue Centrale, 38401 Saint-Martin-d'Hères, France.*
E-mail: [*]*karim.benhenni@univ-grenoble-alpes.fr;* [†]*sana.louhichi@univ-grenoble-alpes.fr*
[2]*CNRS, Laboratoire Jean Kuntzmann, 700 Avenue Centrale, 38401 Saint-Martin-d'Hères, France.*
E-mail: didier.girard@univ-grenoble-alpes.fr

In this paper, we are interested in the problem of smoothing parameter selection in nonparametric curve estimation under dependent errors. We focus on kernel estimation and the case when the errors form a general stationary sequence of martingale difference random variables where neither linearity assumption nor "all moments are finite" are required. We compare the behaviors of the smoothing bandwidths obtained by minimizing either the unknown average squared error, the theoretical mean average squared error, a Mallows-type criterion adapted to the dependent case and the family of criteria known as generalized cross validation (GCV) extensions of the Mallows' criterion. We prove that these three minimizers and those based on the GCV family are first-order equivalent in probability. We give also a normal asymptotic behavior of the gap between the minimizer of the average squared error and that of the Mallows-type criterion. This is extended to the GCV family. Finally, we apply our theoretical results to a specific case of martingale difference sequence, namely the Auto-Regressive Conditional Heteroscedastic (ARCH(1)) process. A Monte-Carlo simulation study, for this regression model with ARCH(1) process, is conducted.

Keywords: Nonparametric trend estimation; kernel nonparametric models; smoothing parameter selection; martingale difference sequences; average squared error; mean average squared error; Mallows criterion; cross validation; generalized cross validation; ARCH(1)

References

[1] Akaike, H. (1970). Statistical predictor identification. *Ann. Inst. Statist. Math.* **22** 203–217. MR0286233 https://doi.org/10.1007/BF02506337
[2] Andrew, D.W.K. and Pollard, D. (1994). An introduction to functional central limit theorems for dependent stochastic processes. *Int. Stat. Rev.* **62** 119–132.
[3] Benhenni, K. and Degras, D. (2014). Local polynomial estimation of the mean function and its derivatives based on functional data and regular designs. *ESAIM Probab. Stat.* **18** 881–899. MR3334018 https://doi.org/10.1051/ps/2014009
[4] Benhenni, K., Girard, D.A. and Louhichi, S. (2022). Supplement to "On bandwidth selection problems in nonparametric trend estimation under martingale difference errors." https://doi.org/10.3150/21-BEJ1347SUPP
[5] Burkholder, D.L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astérisque* **157–158** 75–94.
[6] Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions. *Numer. Math.* **31** 377–403.
[7] Doukhan, P. and Louhichi, S. (1999). A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** 313–342. MR1719345 https://doi.org/10.1016/S0304-4149(99)00055-1

[8] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773
[9] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. *Monographs on Statistics and Applied Probability* **66**. London: CRC Press. MR1383587
[10] Francisco-Fernandez, M., Opsomer, J. and Vilar-Fernandez, J.M. (2004). Plug-in bandwidth selector for local polynomial regression estimator with correlated errors. *J. Nonparametr. Stat.* **16** 127–151.
[11] Gasser, T. and Müller, H.-G. (1979). Kernel estimation of regression functions. In *Smoothing Techniques for Curve Estimation* (*Proc. Workshop*, *Heidelberg*, 1979). *Lecture Notes in Math.* **757** 23–68. Berlin: Springer. MR0564251
[12] Girard, D.A. (1998). Asymptotic comparison of (partial) cross-validation, GCV and randomized GCV in nonparametric regression. *Ann. Statist.* **26** 315–334. MR1608164 https://doi.org/10.1214/aos/1030563988
[13] Hall, P., Lahiri, S.N. and Polzehl, J. (1995). On bandwidth choice in nonparametric regression with both short- and long-range dependent errors. *Ann. Statist.* **23** 1921–1936. MR1389858 https://doi.org/10.1214/aos/1034713640
[14] Han, C. and Gu, C. (2008). Optimal smoothing with correlated data. *Sankhyā* **70** 38–72. MR2507467
[15] Härdle, W., Hall, P. and Marron, J.S. (1988). How far are automatically chosen regression smoothing parameters from their optimum? *J. Amer. Statist. Assoc.* **83** 86–101. MR0941001
[16] 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
[17] Lindner, A.M. (2009). Stationarity, mixing, distributional properties and moments of GARCH(p, q)-processes. In *Handbook of Financial Time Series* (T. Mikosch, J.P. Kreib, R. Davis and T. Andersen, eds.). Berlin, Heidelberg: Springer.
[18] Louhichi, S. (2000). Weak convergence for empirical processes of associated sequences. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** 547–567. MR1792655 https://doi.org/10.1016/S0246-0203(00)00140-0
[19] Mallows, C.L. (1973). Some comments on Cp. *Technometrics* **15** 661–675.
[20] McLeish, D.L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628. MR0358933 https://doi.org/10.1214/aop/1176996608
[21] Nadaraya, E.A. (1964). On estimating regression. *Theory Probab. Appl.* **9** 141–142.
[22] 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
[23] Pollard, D. (1990). *Empirical Processes*: *Theory and Applications*. *NSF-CBMS Regional Conference Series in Probability and Statistics* **2**. IMS: Hayward, CA. MR1089429
[24] Rice, J. (1984). Bandwidth choice for nonparametric regression. *Ann. Statist.* **12** 1215–1230. MR0760684 https://doi.org/10.1214/aos/1176346788
[25] Rio, E. (1993). Covariance inequalities for strongly mixing processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **29** 587–597. MR1251142
[26] Shibata, R. (1981). An optimal selection of regression variables. *Biometrika* **68** 45–54. MR0614940 https://doi.org/10.1093/biomet/68.1.45
[27] Wang, Y. (1998). Smoothing spline models with correlated random errors. *J. Amer. Statist. Assoc.* **93** 41–348.
[28] Watson, G.S. (1964). Smooth regression analysis. *Sankhyā Ser. A* **26** 359–372. MR0185765
[29] Welsh, A.H., Lin, X. and Carroll, R.J. (2002). Marginal longitudinal nonparametric regression: Locality and efficiency of spline and kernel methods. *J. Amer. Statist. Assoc.* **97** 482–493. MR1941465 https://doi.org/10.1198/016214502760047014

Bernoulli **28**(1), 2022, 424–452
https://doi.org/10.3150/21-BEJ1350

New approach to greedy vector quantization

RANCY EL NMEIR[1,2,*], HARALD LUSCHGY[3] and GILLES PAGÈS[1,†]

[1]*Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Université, Campus Pierre et Marie Curie, case 158, 4, pl. Jussieu, F-75252 Paris Cedex 5, France.*
E-mail: [*]*rancynmeir@gmail.com;* [†]*gilles.pages@sorbonne-universite.fr*
[2]*Laboratoire de Mathématiques et Applications, Université Saint-Joseph de Beyrouth, B.P. 11-514 Riad El Solh Beyrouth 1107 2050, Liban*
[3]*FB IV-Mathematik, Universität Trier, D-54286 Trier, Germany.*
E-mail: luschgy@uni-trier.de

We extend some rate of convergence results of greedy quantization sequences already investigated in 2015. We show, for a more general class of distributions satisfying a certain control, that the quantization error of these sequences has an optimal rate of convergence and that the distortion mismatch property is satisfied. We will give some non-asymptotic Pierce type estimates. The recursive character of greedy vector quantization allows some improvements to the algorithm of computation of these sequences and the implementation of a recursive formula to quantization-based numerical integration. Furthermore, we establish further properties of sub-optimality of greedy quantization sequences.

Keywords: Greedy quantization sequence; rate optimality; Lloyd's algorithm; distortion mismatch; quantization-based numerical integration; quasi-Monte Carlo methods

References

[1] Bally, V. and Pagès, G. (2003). A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli* **9** 1003–1049. MR2046816 https://doi.org/10.3150/bj/1072215199
[2] Bally, V., Pagès, G. and Printems, J. (2001). A stochastic quantization method for nonlinear problems. *Monte Carlo Methods Appl*. **7** 21–33. MR1828193 https://doi.org/10.1515/mcma.2001.7.1-2.21
[3] Bouleau, N. and Lépingle, D. (1994). *Numerical Methods for Stochastic Processes*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. New York: Wiley. MR1274043
[4] Brancolini, A., Buttazzo, G., Santambrogio, F. and Stepanov, E. (2009). Long-term planning versus short-term planning in the asymptotical location problem. *ESAIM Control Optim. Calc. Var.* **15** 509–524. MR2542570 https://doi.org/10.1051/cocv:2008034
[5] Delattre, S., Graf, S., Luschgy, H. and Pagès, G. (2004). Quantization of probability distributions under norm-based distortion measures. *Statist. Decisions* **22** 261–282. MR2158264 https://doi.org/10.1524/stnd.22.4.261.64314
[6] Doerr, C., Gnewuch, M. and Wahlström, M. (2014). Calculation of discrepancy measures and applications. In *A Panorama of Discrepancy Theory*. *Lecture Notes in Math*. **2107** 621–678. Cham: Springer. MR3330332 https://doi.org/10.1007/978-3-319-04696-9_10
[7] El Nmeir, R. (2020). Greedy quantization: New approach and application to reflected Backward SDEs. Ph.D. thesis.
[8] El Nmeir, R., Luschgy, H. and Pagès, G. (2020). New approach to greedy vector quantization. Available at http://arxiv.org/abs/2003.14145.
[9] El Nmeir, R. and Pagès, G. (2021). Quantization-based approximation of reflected BSDEs with extended upper bounds for recursive quantization. Available at https://arxiv.org/abs/2105.07684..
[10] Gersho, A. and Gray, R.M. (1998). Special issue on quantization, I-II (A. Gersho and R.M. Gray, eds.). *IEEE Trans. Inf. Theory* **28**.
[11] 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

[12] Graf, S., Luschgy, H. and Pagès, G. (2008). Distortion mismatch in the quantization of probability measures. *ESAIM Probab. Stat.* **12** 127–153. MR2374635 https://doi.org/10.1051/ps:2007044

[13] Kuipers, L. and Niederreiter, H. (1974). *Uniform Distribution of Sequences*: *Pure and Applied Mathematics*. New York: Wiley. MR0419394

[14] Luschgy, H. and Pagès, G. (2008). Functional quantization rate and mean regularity of processes with an application to Lévy processes. *Ann. Appl. Probab.* **18** 427–469. MR2398762 https://doi.org/10.1214/07-AAP459

[15] Luschgy, H. and Pagès, G. (2015). Greedy vector quantization. *J. Approx. Theory* **198** 111–131. MR3393446 https://doi.org/10.1016/j.jat.2015.05.005

[16] Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **63**. Philadelphia, PA: SIAM. MR1172997 https://doi.org/10.1137/1.9781611970081

[17] Pagès, G. (1998). A space quantization method for numerical integration. *J. Comput. Appl. Math.* **89** 1–38. MR1625987 https://doi.org/10.1016/S0377-0427(97)00190-8

[18] Pagès, G. (2015). Introduction to vector quantization and its applications for numerics. In *CEMRACS* 2013—*Modelling and Simulation of Complex Systems*: *Stochastic and Deterministic Approaches*. *ESAIM Proc. Surveys* **48** 29–79. Les Ulis: EDP Sci. MR3415387 https://doi.org/10.1051/proc/201448002

[19] Pagès, G. (2018). *Numerical Probability*: *An Introduction with Applications to Finance*. *Universitext*. Cham: Springer. MR3823785 https://doi.org/10.1007/978-3-319-90276-0

[20] Proĭnov, P.D. (1988). Discrepancy and integration of continuous functions. *J. Approx. Theory* **52** 121–131. MR0929298 https://doi.org/10.1016/0021-9045(88)90051-2

[21] Zador, P.L. (1982). Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. Inf. Theory* **28** 139–149. MR0651809 https://doi.org/10.1109/TIT.1982.1056490

Bernoulli **28**(1), 2022, 453–480
https://doi.org/10.3150/21-BEJ1351

Empirical process theory for locally stationary processes

NATHAWUT PHANDOIDAEN[*] and STEFAN RICHTER[†]

Institut für angewandte Mathematik, Universität Heidelberg, Im Neuenheimer Feld 205, Heidelberg, Germany.
E-mail: [*] phandoidaen@math.uni-heidelberg.de; [†] stefan.richter@iwr.uni-heidelberg.de

We provide a framework for empirical process theory of locally stationary processes using the functional dependence measure. Our results extend known results for stationary Markov chains and mixing sequences by another common possibility to measure dependence and allow for additional time dependence. Our main result is a functional central limit theorem for locally stationary processes. Moreover, maximal inequalities for expectations of sums are developed. We show the applicability of our theory in some examples, for instance, we provide uniform convergence rates for nonparametric regression with locally stationary noise.

Keywords: Empirical process theory; functional dependence measure; maximal inequality; functional central limit theorem; locally stationary processes

References

[1] Adamczak, R. (2008). A tail inequality for suprema of unbounded empirical processes with applications to Markov chains. *Electron. J. Probab.* **13** 1000–1034. MR2424985 https://doi.org/10.1214/EJP.v13-521
[2] Andrews, D.W.K. and Pollard, D. (1994). An introduction to functional central limit theorems for dependent stochastic processes. *Int. Stat. Rev.* **62** 119–132.
[3] Arcones, M.A. and Yu, B. (1994). Central limit theorems for empirical and U-processes of stationary mixing sequences. *J. Theoret. Probab.* **7** 47–71. MR1256391 https://doi.org/10.1007/BF02213360
[4] Berkes, I., Hörmann, S. and Schauer, J. (2009). Asymptotic results for the empirical process of stationary sequences. *Stochastic Process. Appl.* **119** 1298–1324. MR2508575 https://doi.org/10.1016/j.spa.2008.06.010
[5] Borkar, V.S. (1993). White-noise representations in stochastic realization theory. *SIAM J. Control Optim.* **31** 1093–1102. MR1233993 https://doi.org/10.1137/0331050
[6] Borovkova, S., Burton, R. and Dehling, H. (2001). Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Trans. Amer. Math. Soc.* **353** 4261–4318. MR1851171 https://doi.org/10.1090/S0002-9947-01-02819-7
[7] Dahlhaus, R. and Polonik, W. (2009). Empirical spectral processes for locally stationary time series. *Bernoulli* **15** 1–39. MR2546797 https://doi.org/10.3150/08-BEJ137
[8] 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
[9] Dedecker, J. (2010). An empirical central limit theorem for intermittent maps. *Probab. Theory Related Fields* **148** 177–195. MR2653226 https://doi.org/10.1007/s00440-009-0227-5
[10] Dedecker, J. and Louhichi, S. (2002). Maximal inequalities and empirical central limit theorems. In *Empirical Process Techniques for Dependent Data* 137–159. Boston, MA: Birkhäuser. MR1958779
[11] Dedecker, J. and Prieur, C. (2007). An empirical central limit theorem for dependent sequences. *Stochastic Process. Appl.* **117** 121–142. MR2287106 https://doi.org/10.1016/j.spa.2006.06.003
[12] Dehling, H., Durieu, O. and Volny, D. (2009). New techniques for empirical processes of dependent data. *Stochastic Process. Appl.* **119** 3699–3718. MR2568292 https://doi.org/10.1016/j.spa.2009.07.003
[13] Donsker, M.D. (1952). Justification and extension of Doob's heuristic approach to the Komogorov–Smirnov theorems. *Ann. Math. Stat.* **23** 277–281. MR0047288

[14] 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
[15] Doukhan, P., Massart, P. and Rio, E. (1995). Invariance principles for absolutely regular empirical processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 393–427. MR1324814
[16] Doukhan, P. and Neumann, M.H. (2007). Probability and moment inequalities for sums of weakly dependent random variables, with applications. *Stochastic Process. Appl.* **117** 878–903. MR2330724 https://doi.org/10.1016/j.spa.2006.10.011
[17] Dudley, R.M. (1966). Weak convergences of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** 109–126. MR0185641
[18] Dudley, R.M. (1978). Central limit theorems for empirical measures. *Ann. Probab.* **6** 899–929. MR0512411
[19] Dudley, R.M. (2014). *Uniform Central Limit Theorems*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **142**. New York: Cambridge Univ. Press. MR3445285
[20] Durieu, O. and Tusche, M. (2014). An empirical process central limit theorem for multidimensional dependent data. *J. Theoret. Probab.* **27** 249–277. MR3174225 https://doi.org/10.1007/s10959-012-0450-3
[21] Ellis, R.S. and Wyner, A.D. (1989). Uniform large deviation property of the empirical process of a Markov chain. *Ann. Probab.* **17** 1147–1151. MR1009449
[22] 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
[23] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics*, [40]. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862
[24] 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
[25] Heinrich, L. (1992). Bounds for the absolute regularity coefficient of a stationary renewal process. *Yokohama Math. J.* **40** 25–33. MR1189998
[26] Karlsen, H.A. and Tjøstheim, D. (2001). Nonparametric estimation in null recurrent time series. *Ann. Statist.* **29** 372–416. MR1863963 https://doi.org/10.1214/aos/1009210546
[27] Kulik, R., Soulier, P. and Wintenberger, O. (2019). The tail empirical process of regularly varying functions of geometrically ergodic Markov chains. *Stochastic Process. Appl.* **129** 4209–4238. MR4013860 https://doi.org/10.1016/j.spa.2018.11.014
[28] Levental, S. (1988). Uniform limit theorems for Harris recurrent Markov chains. *Probab. Theory Related Fields* **80** 101–118. MR0970473 https://doi.org/10.1007/BF00348754
[29] Li, D., Tjøstheim, D. and Gao, J. (2016). Estimation in nonlinear regression with Harris recurrent Markov chains. *Ann. Statist.* **44** 1957–1987. MR3546440 https://doi.org/10.1214/15-AOS1379
[30] Liebscher, E. (1996). Strong convergence of sums of α-mixing random variables with applications to density estimation. *Stochastic Process. Appl.* **65** 69–80. MR1422880 https://doi.org/10.1016/S0304-4149(96)00096-8
[31] Liu, W., Xiao, H. and Wu, W.B. (2013). Probability and moment inequalities under dependence. *Statist. Sinica* **23** 1257–1272. MR3114713
[32] Mayer, U. (2019). Functional weak limit theorem for a local empirical process of non-stationary time series and its application to von Mises-statistics.
[33] Meyn, S. and Tweedie, R.L. (2009). *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge: Cambridge Univ. Press. With a prologue by Peter W. Glynn. MR2509253 https://doi.org/10.1017/CBO9780511626630
[34] Mokkadem, A. (1988). Mixing properties of ARMA processes. *Stochastic Process. Appl.* **29** 309–315. MR0958507 https://doi.org/10.1016/0304-4149(88)90045-2
[35] Ossiander, M. (1987). A central limit theorem under metric entropy with L_2 bracketing. *Ann. Probab.* **15** 897–919. MR0893905
[36] Pham, T.D. and Tran, L.T. (1985). Some mixing properties of time series models. *Stochastic Process. Appl.* **19** 297–303. MR0787587 https://doi.org/10.1016/0304-4149(85)90031-6
[37] Phandoidaen, N. and Richter, S. (2022). Supplement to "Empirical process theory for locally stationary processes." https://doi.org/10.3150/21-BEJ1351SUPP

[38] Pinelis, I. (1994). Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.* **22** 1679–1706. MR1331198
[39] Pollard, D. (1982). A central limit theorem for empirical processes. *J. Aust. Math. Soc. A* **33** 235–248. MR0668445
[40] Rio, E. (1995). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23** 1188–1203. MR1349167
[41] Rio, E. (1998). Processus empiriques absolument réguliers et entropie universelle. *Probab. Theory Related Fields* **111** 585–608. MR1641838 https://doi.org/10.1007/s004400050179
[42] Rio, E. (2017). *Asymptotic Theory of Weakly Dependent Random Processes*. *Probability Theory and Stochastic Modelling* **80**. Berlin: Springer. Translated from the 2000 French edition [MR2117923]. MR3642873 https://doi.org/10.1007/978-3-662-54323-8
[43] Samur, J.D. (2004). A regularity condition and a limit theorem for Harris ergodic Markov chains. *Stochastic Process. Appl.* **111** 207–235. MR2056537 https://doi.org/10.1016/j.spa.2004.02.005
[44] Truquet, L. (2020). A perturbation analysis of Markov chains models with time-varying parameters. *Bernoulli* **26** 2876–2906. MR4140532 https://doi.org/10.3150/20-BEJ1210
[45] van der Vaart, A.W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 https://doi.org/10.1017/CBO9780511802256
[46] 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
[47] Vogt, M. (2012). Nonparametric regression for locally stationary time series. *Ann. Statist.* **40** 2601–2633. MR3097614 https://doi.org/10.1214/12-AOS1043
[48] 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
[49] Wu, W.B. (2008). Empirical processes of stationary sequences. *Statist. Sinica* **18** 313–333. MR2384990
[50] Wu, W.B. (2011). Asymptotic theory for stationary processes. *Stat. Interface* **4** 207–226. MR2812816 https://doi.org/10.4310/SII.2011.v4.n2.a15
[51] Yu, B. (1994). Rates of convergence for empirical processes of stationary mixing sequences. *Ann. Probab.* **22** 94–116. MR1258867
[52] Zhang, D. and Wu, W.B. (2017). Gaussian approximation for high dimensional time series. *Ann. Statist.* **45** 1895–1919. MR3718156 https://doi.org/10.1214/16-AOS1512

Bernoulli **28**(1), 2022, 481–503
https://doi.org/10.3150/21-BEJ1353

Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation

ANDREA COSSO[1] and FRANCESCO RUSSO[2]

[1]*University of Bologna, Department of Mathematics, Piazza di Porta San Donato 5, 40126 Bologna, Italy.*
E-mail: andrea.cosso@unibo.it
[2]*ENSTA Paris, Institut Polytechnique de Paris, Unité de Mathématiques Appliquées, 828, bd. des Maréchaux, F-91120 Palaiseau, France.*
E-mail: francesco.russo@ensta-paris.fr

We address our interest to the development of a theory of viscosity solutions à la Crandall–Lions for path-dependent partial differential equations (PDEs), namely PDEs in the space of continuous paths $C([0,T];\mathbb{R}^d)$. Path-dependent PDEs can play a central role in the study of certain classes of optimal control problems, as for instance optimal control problems with delay. Typically, they do not admit a smooth solution satisfying the corresponding HJB equation in a classical sense, it is therefore natural to search for a weaker notion of solution. While other notions of generalized solution have been proposed in the literature, the extension of the Crandall–Lions framework to the path-dependent setting is still an open problem. The question of uniqueness of the solutions, which is the most delicate issue, will be based on early ideas from the theory of viscosity solutions and a suitable variant of Ekeland's variational principle. This latter is based on the construction of a smooth gauge-type function, where smooth is meant in the horizontal/vertical (rather than Fréchet) sense. In order to make the presentation more readable, we address the path-dependent heat equation, which in particular simplifies the smoothing of its natural "candidate" solution. Finally, concerning the existence part, we provide a functional Itô formula under general assumptions, extending earlier results in the literature.

Keywords: Path-dependent partial differential equations; viscosity solutions; functional Itô formula

References

[1] Aubin, J.-P. and Haddad, G. (2002). History path dependent optimal control and portfolio valuation and management. *Positivity* **6** 331–358. Special issue of the mathematical economics.
[2] Bambi, M., Fabbri, G. and Gozzi, F. (2012). Optimal policy and consumption smoothing effects in the time-to-build AK model. *Econom. Theory* **50** 635–669. MR2959118 https://doi.org/10.1007/s00199-010-0577-3
[3] Barrasso, A. and Russo, F. (2020). Decoupled mild solutions of path-dependent PDEs and integro PDEs represented by BSDEs driven by cadlag martingales. *Potential Anal.* **53** 449–481. MR4125098 https://doi.org/10.1007/s11118-019-09775-x
[4] Bayraktar, E. and Keller, C. (2018). Path-dependent Hamilton–Jacobi equations in infinite dimensions. *J. Funct. Anal.* **275** 2096–2161. MR3841538 https://doi.org/10.1016/j.jfa.2018.07.010
[5] Bensoussan, A., Chau, M.H.M., Lai, Y. and Yam, S.C.P. (2017). Linear-quadratic mean field Stackelberg games with state and control delays. *SIAM J. Control Optim.* **55** 2748–2781. MR3691722 https://doi.org/10.1137/15M1052937
[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1700749 https://doi.org/10.1002/9780470316962
[7] Borwein, J.M. and Preiss, D. (1987). A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. *Trans. Amer. Math. Soc.* **303** 517–527. MR0902782 https://doi.org/10.2307/2000681

[8] Borwein, J.M. and Zhu, Q.J. (2005). *Techniques of Variational Analysis*. *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC* **20**. New York: Springer. MR2144010
[9] Buckdahn, R., Keller, C., Ma, J. and Zhang, J. (2020). Fully nonlinear stochastic and rough PDEs: Classical and viscosity solutions. *Probab. Uncertain. Quant. Risk* **5** 7. MR4172978 https://doi.org/10.1186/s41546-020-00049-8
[10] Carmona, R., Fouque, J.-P., Mousavi, S.M. and Sun, L.-H. (2018). Systemic risk and stochastic games with delay. *J. Optim. Theory Appl.* **179** 366–399. MR3865333 https://doi.org/10.1007/s10957-018-1267-8
[11] Chojnowska-Michalik, A. (1978). Representation theorem for general stochastic delay equations. *Bull. Acad. Pol. Sci., Sér. Sci. Math. Astron. Phys.* **26** 635–642. MR0515622
[12] Cont, R. and Fournié, D.-A. (2010). Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.* **259** 1043–1072. MR2652181 https://doi.org/10.1016/j.jfa.2010.04.017
[13] Cont, R. and Fournié, D.-A. (2013). Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.* **41** 109–133. MR3059194 https://doi.org/10.1214/11-AOP721
[14] Cont, R. and Fournié, D.A. (2016). Functional Itô calculus and functional Kolmogorov equations. In *Stochastic Integration by Parts and Functional Itô Calculus*. *Adv. Courses Math. CRM Barcelona* 115–207. Cham: Birkhäuser/Springer. MR3497715
[15] Cosso, A., Di Girolami, C. and Russo, F. (2016). Calculus via regularizations in Banach spaces and Kolmogorov-type path-dependent equations. In *Probability on Algebraic and Geometric Structures*. *Contemp. Math.* **668** 43–65. Providence, RI: Amer. Math. Soc. MR3536692 https://doi.org/10.1090/conm/668/13396
[16] Cosso, A., Federico, S., Gozzi, F., Rosestolato, M. and Touzi, N. (2018). Path-dependent equations and viscosity solutions in infinite dimension. *Ann. Probab.* **46** 126–174. MR3758728 https://doi.org/10.1214/17-AOP1181
[17] Cosso, A. and Russo, F. (2016). Functional and Banach space stochastic calculi: Path-dependent Kolmogorov equations associated with the frame of a Brownian motion. In *Stochastics of Environmental and Financial Economics – Centre of Advanced Study*, *Oslo*, *Norway*, 2014–2015. *Springer Proc. Math. Stat.* **138** 27–80. Cham: Springer. MR3451168 https://doi.org/10.1007/978-3-319-23425-0_2
[18] Cosso, A. and Russo, F. (2016). Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **19** 1650024. MR3584624 https://doi.org/10.1142/S0219025716500247
[19] Cosso, A. and Russo, F. (2019). Strong-viscosity solutions: Classical and path-dependent PDEs. *Osaka J. Math.* **56** 323–373. MR3934979
[20] Cosso, A. and Russo, F. (2019). Strong-viscosity solutions: Classical and path-dependent PDEs. *Osaka J. Math.* **56** 323–373. MR3934979
[21] Crandall, M.G., Ishii, H. and Lions, P.-L. (1992). User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.* (*N.S.*) **27** 1–67. MR1118699 https://doi.org/10.1090/S0273-0979-1992-00266-5
[22] Crandall, M.G. and Lions, P.-L. (1981). Condition d'unicité pour les solutions généralisées des équations de Hamilton–Jacobi du premier ordre. *C. R. Acad. Sci. Paris Sér. I Math.* **292** 183–186. MR0610314
[23] Crandall, M.G. and Lions, P.-L. (1983). Viscosity solutions of Hamilton–Jacobi equations. *Trans. Amer. Math. Soc.* **277** 1–42. MR0690039 https://doi.org/10.2307/1999343
[24] Crandall, M.G. and Lions, P.-L. (1985). Hamilton–Jacobi equations in infinite dimensions. I. Uniqueness of viscosity solutions. *J. Funct. Anal.* **62** 379–396. MR0794776 https://doi.org/10.1016/0022-1236(85)90011-4
[25] Da Prato, G. and Zabczyk, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge: Cambridge Univ. Press. MR1417491 https://doi.org/10.1017/CBO9780511662829
[26] Di Girolami, C., Fabbri, G. and Russo, F. (2014). The covariation for Banach space valued processes and applications. *Metrika* **77** 51–104. MR3152020 https://doi.org/10.1007/s00184-013-0472-6
[27] Di Girolami, C. and Russo, F. (2010). Infinite dimensional stochastic calculus via regularization and applications. Preprint HAL-INRIA, inria-00473947, version 1, (Unpublished).
[28] Di Girolami, C. and Russo, F. (2011). Clark–Ocone type formula for non-semimartingales with finite quadratic variation. *C. R. Math. Acad. Sci. Paris* **349** 209–214. MR2769910 https://doi.org/10.1016/j.crma.2010.11.032

[29] Di Girolami, C. and Russo, F. (2012). Generalized covariation and extended Fukushima decomposition for Banach space-valued processes. Applications to windows of Dirichlet processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **15** 1250007. MR2957134 https://doi.org/10.1142/S0219025712500075
[30] Di Girolami, C. and Russo, F. (2014). Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka J. Math.* **51** 729–783. MR3272616
[31] Di Girolami, C. and Russo, F. (2020). About classical solutions of the path-dependent heat equation. *Random Oper. Stoch. Equ.* **28** 35–62. MR4069984 https://doi.org/10.1515/rose-2020-2028
[32] Dupire, B. (2009). *Functional Itô Calculus*. Bloomberg: Portfolio Research Paper.
[33] Ekeland, I. (1974). On the variational principle. *J. Math. Anal. Appl.* **47** 324–353. MR0346619 https://doi.org/10.1016/0022-247X(74)90025-0
[34] Ekren, I., Keller, C., Touzi, N. and Zhang, J. (2014). On viscosity solutions of path dependent PDEs. *Ann. Probab.* **42** 204–236. MR3161485 https://doi.org/10.1214/12-AOP788
[35] Ekren, I., Touzi, N. and Zhang, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.* **44** 1212–1253. MR3474470 https://doi.org/10.1214/14-AOP999
[36] Ekren, I., Touzi, N. and Zhang, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.* **44** 2507–2553. MR3531674 https://doi.org/10.1214/15-AOP1027
[37] Elsanosi, I., Øksendal, B. and Sulem, A. (2000). Some solvable stochastic control problems with delay. *Stoch. Stoch. Rep.* **71** 69–89. MR1813507 https://doi.org/10.1080/17442500008834259
[38] Fabbri, G., Gozzi, F. and Święch, A. (2017). *Stochastic Optimal Control in Infinite Dimension*: *Dynamic Programming and HJB Equations*, *with a Contribution by Marco Fuhrman and Gianmario Tessitore*. *Probability Theory and Stochastic Modelling* **82**. Cham: Springer. MR3674558 https://doi.org/10.1007/978-3-319-53067-3
[39] Federico, S. (2011). A stochastic control problem with delay arising in a pension fund model. *Finance Stoch.* **15** 421–459. MR2833095 https://doi.org/10.1007/s00780-010-0146-4
[40] Federico, S., Goldys, B. and Gozzi, F. (2010). HJB equations for the optimal control of differential equations with delays and state constraints, I: Regularity of viscosity solutions. *SIAM J. Control Optim.* **48** 4910–4937. MR2735510 https://doi.org/10.1137/09076742X
[41] Federico, S., Goldys, B. and Gozzi, F. (2011). HJB equations for the optimal control of differential equations with delays and state constraints, II: Verification and optimal feedbacks. *SIAM J. Control Optim.* **49** 2378–2414. MR2854622 https://doi.org/10.1137/100804292
[42] Flandoli, F. and Zanco, G. (2016). An infinite-dimensional approach to path-dependent Kolmogorov equations. *Ann. Probab.* **44** 2643–2693. MR3531677 https://doi.org/10.1214/15-AOP1031
[43] Fournie, D.-A. (2010). *Functional Ito Calculus and Applications*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.) – Columbia University. MR2782350
[44] Fuhrman, M., Masiero, F. and Tessitore, G. (2010). Stochastic equations with delay: Optimal control via BSDEs and regular solutions of Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* **48** 4624–4651. MR2683901 https://doi.org/10.1137/080730354
[45] Göllmann, L., Kern, D. and Maurer, H. (2009). Optimal control problems with delays in state and control variables subject to mixed control-state constraints. *Optimal Control Appl. Methods* **30** 341–365. MR2553382 https://doi.org/10.1002/oca.843
[46] Göllmann, L. and Maurer, H. (2018). Optimal control problems with time delays: Two case studies in biomedicine. *Math. Biosci. Eng.* **15** 1137–1154. MR3918280 https://doi.org/10.3934/mbe.2018051
[47] Gomoyunov, M. (2020). *Private Communication*.
[48] Gozzi, F. and Marinelli, C. (2006). Stochastic optimal control of delay equations arising in advertising models. In *Stochastic Partial Differential Equations and Applications – VII*. *Lect. Notes Pure Appl. Math.* **245** 133–148. Boca Raton, FL: CRC Press/CRC. MR2227226 https://doi.org/10.1201/9781420028720.ch13
[49] Gozzi, F., Marinelli, C. and Savin, S. (2009). On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. *J. Optim. Theory Appl.* **142** 291–321. MR2525792 https://doi.org/10.1007/s10957-009-9524-5
[50] Ichikawa, A. (1982). Quadratic control of evolution equations with delays in control. *SIAM J. Control Optim.* **20** 645–668. MR0667646 https://doi.org/10.1137/0320048
[51] Itô, K. and Nisio, M. (1964). On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.* **4** 1–75. MR0177456 https://doi.org/10.1215/kjm/1250524705

[52] Ivanov, A.F., Kazmerchuk, Y.I. and Swishchuk, A.V. (2003). Theory, stochastic stability and applications of stochastic delay differential equations: A survey of results. *Differ. Equ. Dyn. Syst.* **11** 55–115. MR2065279
[53] Jensen, R., Kocan, M. and Święch, A. (2002). Good and viscosity solutions of fully nonlinear elliptic equations. *Proc. Amer. Math. Soc.* **130** 533–542. MR1862134 https://doi.org/10.1090/S0002-9939-01-06115-9
[54] Kim, A.V. (1999). *Functional Differential Equations. Mathematics and Its Applications* **479**. Dordrecht: Kluwer Academic. Application of i-smooth calculus. MR1783365 https://doi.org/10.1007/978-94-017-1630-7
[55] Kim, A.V. (2015). *i-smooth Analysis*. Salem, MA: Scrivener Publishing. Theory and applications. MR3379527 https://doi.org/10.1002/9781118998519
[56] Kim, A.V. and Ivanov, A.V. (2015). *Systems with Delays*. Hoboken, NJ: Wiley. Analysis, control, and computations. MR3496968 https://doi.org/10.1002/9781119117841.index
[57] Larssen, B. and Risebro, N.H. (2003). When are HJB-equations in stochastic control of delay systems finite dimensional? *Stoch. Anal. Appl.* **21** 643–671. MR1978238 https://doi.org/10.1081/SAP-120020430
[58] Leão, D., Ohashi, A. and Simas, A.B. (2018). A weak version of path-dependent functional Itô calculus. *Ann. Probab.* **46** 3399–3441. MR3857859 https://doi.org/10.1214/17-AOP1250
[59] Lions, P.-L. (1983). On the Hamilton–Jacobi–Bellman equations. *Acta Appl. Math.* **1** 17–41. MR0702045 https://doi.org/10.1007/BF02433840
[60] Lions, P.-L. (1983). Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations. I. The dynamic programming principle and applications. *Comm. Partial Differential Equations* **8** 1101–1174. MR0709164 https://doi.org/10.1080/03605308308820297
[61] Lions, P.-L. (1983). Optimal control of diffusion processes and Hamilton–Jacobi–Bellman equations. II. Viscosity solutions and uniqueness. *Comm. Partial Differential Equations* **8** 1229–1276. MR0709162 https://doi.org/10.1080/03605308308820301
[62] Lions, P.-L. (1988). Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. I. The case of bounded stochastic evolutions. *Acta Math.* **161** 243–278. MR0971797 https://doi.org/10.1007/BF02392299
[63] Lions, P.-L. (1989). Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimensions. II. Optimal control of Zakai's equation. In *Stochastic Partial Differential Equations and Applications*, *II* (*Trento*, 1988). *Lecture Notes in Math.* **1390** 147–170. Berlin: Springer. MR1019600 https://doi.org/10.1007/BFb0083943
[64] Lions, P.-L. (1989). Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations. *J. Funct. Anal.* **86** 1–18. MR1013931 https://doi.org/10.1016/0022-1236(89)90062-1
[65] Lukoyanov, N. (2003). Functional Hamilton–Jacobi type equations in ci-derivatives for systems with distributed delays. *Nonlinear Funct. Anal. Appl.* **8** 365–397. MR2010832
[66] Lukoyanov, N. (2003). Functional Hamilton–Jacobi type equations with ci-derivatives in control problems with hereditary information. *Nonlinear Funct. Anal. Appl.* **8** 535–555. MR2031252
[67] Lukoyanov, N.Y. (2001). A minimax solution of functional equations of Hamilton–Jacobi type for hereditary systems. *Differ. Equ.* **37** 246–256.
[68] Lukoyanov, N.Y. (2007). Viscosity solution of nonanticipating equations of Hamilton–Jacobi type. *Differ. Equ.* **43** 1715–1723.
[69] Mohammed, S.-E.A. (1998). Stochastic differential systems with memory: Theory, examples and applications. In *Stochastic Analysis and Related Topics*, *VI* (*Geilo*, 1996). *Progress in Probability* **42** 1–77. Boston, MA: Birkhäuser. MR1652338
[70] Mohammed, S.E.A. (1984). *Stochastic Functional Differential Equations. Research Notes in Mathematics* **99**. Boston, MA: Pitman (Advanced Publishing Program). MR0754561
[71] Oberhauser, H. (2016). The functional Itô formula under the family of continuous semimartingale measures. *Stoch. Dyn.* **16** 1650010. MR3494685 https://doi.org/10.1142/S0219493716500106
[72] Øksendal, B., Sulem, A. and Zhang, T. (2011). Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations. *Adv. in Appl. Probab.* **43** 572–596. MR2848391 https://doi.org/10.1239/aap/1308662493
[73] Peng, S. (2012). Note on viscosity solution of path-dependent PDE and G-martingales. Preprint. Available at arXiv:1106.1144v2.

[74] Peng, S. and Song, Y. (2015). G-expectation weighted Sobolev spaces, backward SDE and path dependent PDE. *J. Math. Soc. Japan* **67** 1725–1757. MR3417511 https://doi.org/10.2969/jmsj/06741725

[75] Peng, S. and Wang, F. (2016). BSDE, path-dependent PDE and nonlinear Feynman–Kac formula. *Sci. China Math.* **59** 19–36. MR3436993 https://doi.org/10.1007/s11425-015-5086-1

[76] Pollard, D. (1984). *Convergence of Stochastic Processes*. *Springer Series in Statistics*. New York: Springer. MR0762984 https://doi.org/10.1007/978-1-4612-5254-2

[77] Ren, Z. (2016). Viscosity solutions of fully nonlinear elliptic path dependent partial differential equations. *Ann. Appl. Probab.* **26** 3381–3414. MR3582806 https://doi.org/10.1214/16-AAP1178

[78] Ren, Z., Touzi, N. and Zhang, J. (2017). Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent PDEs. *SIAM J. Math. Anal.* **49** 4093–4116. MR3715377 https://doi.org/10.1137/16M1090338

[79] Ren, Z., Touzi, N. and Zhang, J. (2020). Comparison of viscosity solutions of semilinear path-dependent PDEs. *SIAM J. Control Optim.* **58** 277–302. MR4053859 https://doi.org/10.1137/19M1239404

[80] Rodrigues, F., Silva, C.J., Torres, D.F.M. and Maurer, H. (2018). Optimal control of a delayed HIV model. *Discrete Contin. Dyn. Syst. Ser. B* **23** 443–458. MR3721854 https://doi.org/10.3934/dcdsb.2018030

[81] Scheutzow, M. (1984). Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics* **12** 41–80. MR0738934 https://doi.org/10.1080/17442508408833294

[82] Soner, H.M. (1988). On the Hamilton–Jacobi–Bellman equations in Banach spaces. *J. Optim. Theory Appl.* **57** 429–437. MR0944588 https://doi.org/10.1007/BF02346162

[83] Subbotin, A.I. (1980). Generalization of the fundamental equation of the theory of differential games. *Dokl. Akad. Nauk SSSR* **254** 293–297. MR0587158

[84] Subbotin, A.I. and Subbotina, N.N. (1978). Necessary and sufficient conditions for the piecewise smooth value of a differential game. *Dokl. Akad. Nauk SSSR* **243** 862–865. MR0514487

[85] Święch, A. (1994). "Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces. *Comm. Partial Differential Equations* **19** 1999–2036. MR1301180 https://doi.org/10.1080/03605309408821080

[86] Tang, S. and Zhang, F. (2015). Path-dependent optimal stochastic control and viscosity solution of associated Bellman equations. *Discrete Contin. Dyn. Syst.* **35** 5521–5553. MR3392684 https://doi.org/10.3934/dcds.2015.35.5521

[87] Vinter, R.B. and Kwong, R.H. (1981). The infinite time quadratic control problem for linear systems with state and control delays: An evolution equation approach. *SIAM J. Control Optim.* **19** 139–153. MR0603086 https://doi.org/10.1137/0319011

[88] Zhang, J. (2017). *Backward Stochastic Differential Equations – from Linear to Fully Nonlinear Theory*. *Probability Theory and Stochastic Modelling* **86**. New York: Springer. MR3699487 https://doi.org/10.1007/978-1-4939-7256-2

[89] Zhou, J. Viscosity solutions to second order path-dependent Hamilton–Jacobi–Bellman equations and applications. Preprint. Available at arXiv:2005.05309v1.

Bernoulli **28**(1), 2022, 504–524
https://doi.org/10.3150/21-BEJ1354

Adaptive estimation in the linear random coefficients model when regressors have limited variation

CHRISTOPHE GAILLAC[*] and ERIC GAUTIER[†]

Toulouse School of Economics, Université Toulouse Capitole, 1 esplanade de l'université, 31000 Toulouse, France. E-mail: [*] christophe.gaillac@tse-fr.eu; [†] eric.gautier@tse-fr.eu

We consider a linear model where the coefficients – intercept and slopes – are random and independent from the regressors. The law of the coefficients is nonparametric. Without further restriction, nonparametric identification requires the regressors to have a support which is the whole space. This is hardly ever the case in practice. It is possible to handle regressors with limited variation when the coefficients can have a compact support. This is not compatible with unbounded error terms as usual in regression models. In this paper, the regressors can have a support which is a proper subset but the slopes do not have heavy-tails. Lower bounds on the minimax risk for the estimation of the joint density of the random coefficients density are obtained for a wide range of smoothness. Some allow for polynomial and nearly parametric rates of convergence. We present a minimax optimal estimator and a data-driven rule for adaptive estimation. A R package is available to implement this estimator.

Keywords: Adaptation; inverse problem; minimax; random coefficients

References

[1] Alibaud, N., Maréchal, P. and Saesor, Y. (2009). A variational approach to the inversion of truncated Fourier operators. *Inverse Probl.* **25** 045002, 20. MR2482153 https://doi.org/10.1088/0266-5611/25/4/045002

[2] Alquier, P., Gautier, E. and Stoltz, G., eds. (2011) *Inverse Problems and High-Dimensional Estimation. Lecture Notes in Statistics—Proceedings* **203**. Heidelberg: Springer. MR2867623 https://doi.org/10.1007/978-3-642-19989-9

[3] Beran, R. (1995). Prediction in random coefficient regression. *J. Statist. Plann. Inference* **43** 205–213. MR1314138 https://doi.org/10.1016/0378-3758(94)00020-V

[4] Beran, R., Feuerverger, A. and Hall, P. (1996). On nonparametric estimation of intercept and slope distributions in random coefficient regression. *Ann. Statist.* **24** 2569–2592. MR1425969 https://doi.org/10.1214/aos/1032181170

[5] Beran, R. and Millar, P.W. (1994). Minimum distance estimation in random coefficient regression models. *Ann. Statist.* **22** 1976–1992. MR1329178 https://doi.org/10.1214/aos/1176325767

[6] Bertero, M. and Boccacci, P. (1998). *Introduction to Inverse Problems in Imaging*. Bristol: Institute of Physics Publishing. MR1640759 https://doi.org/10.1887/0750304359

[7] Bissantz, N., Hohage, T., Munk, A. and Ruymgaart, F. (2007). Convergence rates of general regularization methods for statistical inverse problems and applications. *SIAM J. Numer. Anal.* **45** 2610–2636. MR2361904 https://doi.org/10.1137/060651884

[8] Bonami, A. and Karoui, A. (2014). Uniform bounds of prolate spheroidal wave functions and eigenvalues decay. *C. R. Math. Acad. Sci. Paris* **352** 229–234. MR3167571 https://doi.org/10.1016/j.crma.2014.01.004

[9] Breunig, C. and Hoderlein, S. (2018). Specification testing in random coefficient models. *Quant. Econ.* **9** 1371–1417. MR3891570 https://doi.org/10.3982/QE757

[10] Carrasco, M., Florens, J.-P. and Renault, E. (2007). Linear inverse problems in structural econometrics estimation based on spectral decomposition and regularization. In *Handbook of Econometrics* **6** 5633–5751.

 © 2022 ISI/BS

[11] Cavalier, L. (2000). Efficient estimation of a density in a problem of tomography. *Ann. Statist.* **28** 630–647. MR1790012 https://doi.org/10.1214/aos/1016218233

[12] Cavalier, L., Golubev, Y., Lepski, O. and Tsybakov, A. (2003). Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Teor. Veroyatn. Primen.* **48** 534–556. MR2141349 https://doi.org/10.1137/S0040585X97980555

[13] Chen, X. and Reiss, M. (2011). On rate optimality for ill-posed inverse problems in econometrics. *Econometric Theory* **27** 497–521. MR2806258 https://doi.org/10.1017/S0266466610000381

[14] Comte, F., Genon-Catalot, V. and Samson, A. (2013). Nonparametric estimation for stochastic differential equations with random effects. *Stochastic Process. Appl.* **123** 2522–2551. MR3054535 https://doi.org/10.1016/j.spa.2013.04.009

[15] Comte, F. and Lacour, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 601–627. MR2853732 https://doi.org/10.1111/j.1467-9868.2011.00775.x

[16] 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

[17] Dion, C. (2014). New adaptive strategies for nonparametric estimation in linear mixed models. *J. Statist. Plann. Inference* **150** 30–48. MR3206719 https://doi.org/10.1016/j.jspi.2014.03.006

[18] Dunker, F., Eckle, K., Proksch, K. and Schmidt-Hieber, J. (2019). Tests for qualitative features in the random coefficients model. *Electron. J. Stat.* **13** 2257–2306. MR3980958 https://doi.org/10.1214/19-EJS1570

[19] Frikel, J. (2013). Sparse regularization in limited angle tomography. *Appl. Comput. Harmon. Anal.* **34** 117–141. MR2981336 https://doi.org/10.1016/j.acha.2012.03.005

[20] Gaillac, C. and Gautier, E. (2019). Estimates for the SVD of the truncated fourier transform on $L^2(e^{b|x|})$ and stable analytic continuation. Preprint. Available at arXiv:1905.11338.

[21] Gaillac, C. and Gautier, E. (2019). RandomCoefficients: Adaptive estimation in the linear random coefficients model. RandomCoefficients vignette.

[22] Gaillac, C. and Gautier, E. (2022). Supplement to "Adaptive estimation in the linear random coefficients model when regressors have limited variation." https://doi.org/10.3150/21-BEJ1354SUPP

[23] Gautier, E. and Hoderlein, S. (2011). A triangular treatment effect model with random coefficients in the selection equation. Preprint. Available at arXiv:1109.0362v4.

[24] Gautier, E. and Kitamura, Y. (2013). Nonparametric estimation in random coefficients binary choice models. *Econometrica* **81** 581–607. MR3043343 https://doi.org/10.3982/ECTA8675

[25] Gautier, E. and Le Pennec, E. (2018). Adaptive estimation in the nonparametric random coefficients binary choice model by needlet thresholding. *Electron. J. Stat.* **12** 277–320. MR3763073 https://doi.org/10.1214/17-EJS1383

[26] Gautier, E. (2021). Relaxing monotonicity in endogenous selection models and application to surveys. In *Advances in Contemporary Statistics and Econometrics* Heidelberg: Springer.

[27] Goldenshluger, A. and Lepski, O. (2008). Universal pointwise selection rule in multivariate function estimation. *Bernoulli* **14** 1150–1190. MR2543590 https://doi.org/10.3150/08-BEJ144

[28] Goldenshluger, A. and Lepski, O. (2014). On adaptive minimax density estimation on R^d. *Probab. Theory Related Fields* **159** 479–543. MR3230001 https://doi.org/10.1007/s00440-013-0512-1

[29] Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* **52** 271–320. MR0735309 https://doi.org/10.2307/1911491

[30] Hoderlein, S., Holzmann, H. and Meister, A. (2017). The triangular model with random coefficients. *J. Econometrics* **201** 144–169. MR3704681 https://doi.org/10.1016/j.jeconom.2017.05.018

[31] Hoderlein, S., Klemelä, J. and Mammen, E. (2010). Analyzing the random coefficient model nonparametrically. *Econometric Theory* **26** 804–837. MR2646480 https://doi.org/10.1017/S0266466609990119

[32] Hohmann, D. and Holzmann, H. (2016). Weighted angle Radon transform: Convergence rates and efficient estimation. *Statist. Sinica* **26** 157–175. MR3468349

[33] Holzmann, H. and Meister, A. (2020). Rate-optimal nonparametric estimation for random coefficient regression models. *Bernoulli* **26** 2790–2814. MR4140529 https://doi.org/10.3150/20-BEJ1207

[34] Jähnisch, M. and Nussbaum, M. (2003). Asymptotic equivalence for a model of independent non identically distributed observations. *Statist. Decisions* **21** 197–218. MR2045142 https://doi.org/10.1524/stnd.21.3.197.23430
[35] Kerkyacharian, G. and Picard, D. (2004). Regression in random design and warped wavelets. *Bernoulli* **10** 1053–1105. MR2108043 https://doi.org/10.3150/bj/1106314850
[36] Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Math. Acad. Sci. Paris* **342** 877–882. MR2224640 https://doi.org/10.1016/j.crma.2006.04.006
[37] Lacour, C. and Massart, P. (2016). Minimal penalty for Goldenshluger-Lepski method. *Stochastic Process. Appl.* **126** 3774–3789. MR3565477 https://doi.org/10.1016/j.spa.2016.04.015
[38] Lindberg, J. (2012). Mathematical concepts of optical superresolution. *J. Opt.* **14** 083001.
[39] Mair, B.A. and Ruymgaart, F.H. (1996). Statistical inverse estimation in Hilbert scales. *SIAM J. Appl. Math.* **56** 1424–1444. MR1409127 https://doi.org/10.1137/S0036139994264476
[40] Masten, M.A. (2018). Random coefficients on endogenous variables in simultaneous equations models. *Rev. Econ. Stud.* **85** 1193–1250. MR3782604 https://doi.org/10.1093/restud/rdx047
[41] Meister, A. (2007). Deconvolving compactly supported densities. *Math. Methods Statist.* **16** 63–76. MR2319471 https://doi.org/10.3103/S106653070701005X
[42] Meister, A. (2011). Asymptotic equivalence of functional linear regression and a white noise inverse problem. *Ann. Statist.* **39** 1471–1495. MR2850209 https://doi.org/10.1214/10-AOS872
[43] Morrison, J. (1962). On the commutation of finite integral operators, with difference kernels, and linear self-adjoint differential operators. *Notices Amer. Math. Soc.* **9** 119.
[44] Osipov, A., Rokhlin, V. and Xiao, H. (2013). *Prolate Spheroidal Wave Functions of Order Zero: Mathematical Tools for Bandlimited Approximation. Applied Mathematical Sciences* **187**. New York: Springer. MR3136430 https://doi.org/10.1007/978-1-4614-8259-8
[45] Reiß, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Statist.* **36** 1957–1982. MR2435461 https://doi.org/10.1214/07-AOS525
[46] Tsybakov, A. (2000). On the best rate of adaptive estimation in some inverse problems. *C. R. Acad. Sci. Paris Sér. I Math.* **330** 835–840. MR1769957 https://doi.org/10.1016/S0764-4442(00)00278-0
[47] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation. Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794
[48] Wand, M.P. and Jones, M.C. (1994). Multivariate plug-in bandwidth selection. *Comput. Statist.* **9** 97–116. MR1280754
[49] Widom, H. (1964). Asymptotic behavior of the eigenvalues of certain integral equations. II. *Arch. Ration. Mech. Anal.* **17** 215–229. MR0169015 https://doi.org/10.1007/BF00282438

Bernoulli **28**(1), 2022, 525–547
https://doi.org/10.3150/21-BEJ1355

Testing and estimation for clustered signals

HONGYUAN CAO[1,2] and WEI BIAO WU[3]

[1]*School of Mathematics, Jilin University, 2699 Qianjing Street, Changchun, 130012, China*
[2]*Department of Statistics, Florida State University, 117 N. Woodward Avenue, Tallahassee, FL, 32306, USA.*
E-mail: hongyuancao@gmail.com
[3]*Department of Statistics, University of Chicago, 5747 South Ellis Avenue, Chicago, IL, 60637, USA.*
E-mail: wbwuchicago@gmail.com

We propose a change-point detection method for large scale multiple testing problems with data having clustered signals. Unlike the classic change-point setup, the signals can vary in size within a cluster. The clustering structure on the signals enables us to effectively delineate the boundaries between signal and non-signal segments. New test statistics are proposed for observations from one and/or multiple realizations. Their asymptotic distributions are derived. We also study the associated variance estimation problem. We allow the variances to be heteroscedastic in the multiple realization case, which substantially expands the applicability of the proposed method. Simulation studies demonstrate that the proposed approach has a favorable performance. Our procedure is applied to an array based Comparative Genomic Hybridization (aCGH) dataset.

Keywords: Change-point inference; clustered signal; high dimension; multiple testing; signal aggregation; variance estimation

References

[1] Benjamini, Y. and Heller, R. (2007). False discovery rates for spatial signals. *J. Amer. Statist. Assoc.* **102** 1272–1281. MR2412549 https://doi.org/10.1198/016214507000000941
[2] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57** 289–300. MR1325392
[3] Bickel, P.J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095. MR0348906
[4] Cai, T.T. and Sun, W. (2009). Simultaneous testing of grouped hypotheses: Finding needles in multiple haystacks. *J. Amer. Statist. Assoc.* **104** 1467–1481. MR2597000 https://doi.org/10.1198/jasa.2009.tm08415
[5] Cao, H., Chen, J. and Zhang, X. (2021). Optimal false discovery rate control for large scale multiple testing with auxiliary information. Available at arXiv:2103.15311.
[6] Cao, H. and Kosorok, M.R. (2011). Simultaneous critical values for t-tests in very high dimensions. *Bernoulli* **17** 347–394. MR2797995 https://doi.org/10.3150/10-BEJ272
[7] Cao, H. and Wu, W.B. (2015). Changepoint estimation: Another look at multiple testing problems. *Biometrika* **102** 974–980. MR3431567 https://doi.org/10.1093/biomet/asv031
[8] Cao, H. and Wu, W.B. (2022). Supplement to "Testing and estimation for clustered signals." https://doi.org/10.3150/21-BEJ1355SUPP
[9] Chan, H.P. and Walther, G. (2013). Detection with the scan and the average likelihood ratio. *Statist. Sinica* **23** 409–428. MR3076173
[10] Chouldechova, A. (2014). *False Discovery Rate Control for Spatial Data*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–Stanford University. MR4144738
[11] Du, L. and Zhang, C. (2014). Single-index modulated multiple testing. *Ann. Statist.* **42** 30–79. MR3226157 https://doi.org/10.1214/14-AOS1222
[12] Fan, J. (1996). Test of significance based on wavelet thresholding and Neyman's truncation. *J. Amer. Statist. Assoc.* **91** 674–688. MR1395735 https://doi.org/10.2307/2291663
[13] Fan, J., Hall, P. and Yao, Q. (2007). To how many simultaneous hypothesis tests can normal, Student's t or bootstrap calibration be applied? *J. Amer. Statist. Assoc.* **102** 1282–1288. MR2372536 https://doi.org/10.1198/016214507000000969

 © 2022 ISI/BS

[14] Genovese, C.R., Roeder, K. and Wasserman, L. (2006). False discovery control with p-value weighting. *Biometrika* **93** 509–524. MR2261439 https://doi.org/10.1093/biomet/93.3.509
[15] Hall, P., Kay, J.W. and Titterington, D.M. (1990). Asymptotically optimal difference-based estimation of variance in nonparametric regression. *Biometrika* **77** 521–528. MR1087842 https://doi.org/10.1093/biomet/77.3.521
[16] Heller, R., Stanley, D., Yekutieli, D., Rubin, N. and Benjamini, Y. (2006). Cluster-based analysis of FMRI data. *NeuroImage* **33** 599–608. https://doi.org/10.1016/j.neuroimage.2006.04.233
[17] Hu, J.X., Zhao, H. and Zhou, H.H. (2010). False discovery rate control with groups. *J. Amer. Statist. Assoc.* **105** 1215–1227. MR2752616 https://doi.org/10.1198/jasa.2010.tm09329
[18] Lei, L. and Fithian, W. (2018). AdaPT: An interactive procedure for multiple testing with side information. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 649–679. MR3849338 https://doi.org/10.1111/rssb.12253
[19] Li, A. and Barber, R.F. (2017). Accumulation tests for FDR control in ordered hypothesis testing. *J. Amer. Statist. Assoc.* **112** 837–849. MR3671774 https://doi.org/10.1080/01621459.2016.1180989
[20] Li, A. and Barber, R.F. (2019). Multiple testing with the structure-adaptive Benjamini–Hochberg algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **81** 45–74. MR3904779
[21] Liu, W. and Shao, Q.-M. (2014). Phase transition and regularized bootstrap in large-scale t-tests with false discovery rate control. *Ann. Statist.* **42** 2003–2025. MR3262475 https://doi.org/10.1214/14-AOS1249
[22] Liu, W.-D. (2015). Incorporation of sparsity information in large-scale multiple two-sample t tests. ArXiv e-prints. 2015. Available at arXiv:1410.4282.
[23] Patil, G.P. and Taillie, C. (2003). Geographic and network surveillance via scan statistics for critical area detection. *Statist. Sci.* **18** 457–465. MR2109372 https://doi.org/10.1214/ss/1081443229
[24] Perone Pacifico, M., Genovese, C., Verdinelli, I. and Wasserman, L. (2004). False discovery control for random fields. *J. Amer. Statist. Assoc.* **99** 1002–1014. MR2109490 https://doi.org/10.1198/0162145000001655
[25] Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D. and Brown, P.O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23** 41–46.
[26] Pollack, J.R., Sorlie, T., Perou, C.M., Rees, C.A., Jeffrey, S.S., Lonning, P.E., Tibshirani, R., Botstein, D., Borresen-Dale, A.-L. and Brown, P.O. (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc. Natl. Acad. Sci. USA* **99** 12963–12968.
[27] Pritchard, J. (2001). Are rare variants responsible for susceptibility to complex diseases? *Am. J. Hum. Genet.* **1** 123–137.
[28] Shen, X., Huang, H.-C. and Cressie, N. (2002). Nonparametric hypothesis testing for a spatial signal. *J. Amer. Statist. Assoc.* **97** 1122–1140. MR1951265 https://doi.org/10.1198/016214502388618933
[29] Siegmund, D.O., Zhang, N.R. and Yakir, B. (2011). False discovery rate for scanning statistics. *Biometrika* **98** 979–985. MR2860337 https://doi.org/10.1093/biomet/asr057
[30] Stratton, M.R., Campbell, P.J. and Futreal, P.A. (2009). The cancer genome. *Nature* **458** 719–724.
[31] Sun, W., Reich, B.J., Cai, T.T., Guindani, M. and Schwartzman, A. (2015). False discovery control in large-scale spatial multiple testing. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 59–83. MR3299399 https://doi.org/10.1111/rssb.12064
[32] Szor, P. (2005). *The Art of Computer Virus Research and Defense*. Reading: Addison-Wesley.
[33] Tansey, W., Koyejo, O., Poldrack, R.A. and Scott, J.G. (2018). False discovery rate smoothing. *J. Amer. Statist. Assoc.* **113** 1156–1171. MR3862347 https://doi.org/10.1080/01621459.2017.1319838
[34] Tibshirani, R. and Wang, P. (2008). Spatial smoothing and hot spot detection for CGH data using the fused lasso. *Biostatistics* **9** 18–29.
[35] Wu, W. and Zhou, Z. (2018). Gradient-based structural change detection for nonstationary time series M-estimation. *Ann. Statist.* **46** 1197–1224. MR3798001 https://doi.org/10.1214/17-AOS1582
[36] Wu, W.-C. and Zhou, Z. (2020). Multiscale jump testing and estimation under complex temporal dynamics. In ArXiv E-prints, 2020. Available at arXiv:1909.06307.
[37] Yao, Q.W. (1993). Tests for change-points with epidemic alternatives. *Biometrika* **80** 179–191. MR1225223 https://doi.org/10.1093/biomet/80.1.179
[38] 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

[39] Zhang, C., Fan, J. and Yu, T. (2011). Multiple testing via FDR_L for large-scale imaging data. *Ann. Statist.* **39** 613–642. MR2797858 https://doi.org/10.1214/10-AOS848

[40] Zhou, X., Carbonetto, P. and Stephens, M. (2013). Polygenic modeling with bayesian sparse linear mixed models. *PLoS Genet.* **9** e1003264. https://doi.org/10.1371/journal.pgen.1003264

Bernoulli **28**(1), 2022, 548–578
https://doi.org/10.3150/21-BEJ1357

Adaptiveness of the empirical distribution of residuals in semi-parametric conditional location scale models

CHRISTIAN FRANCQ[*] and JEAN-MICHEL ZAKOÏAN[†]

CREST-ENSAE, 5 Avenue Henri Le Chatelier, 91120 Palaiseau, France.
E-mail: [*] *christian.francq@ensae.fr;* [†] *zakoian@ensae.fr*

This paper addresses the problem of deriving the asymptotic distribution of the empirical distribution function $\widehat{F}_n$ of the residuals in a general class of time series models, including conditional mean and conditional heteroscedasticity, whose independent and identically distributed errors have unknown distribution F. We show that, for a large class of time series models (including the standard ARMA-GARCH with symmetric innovations), the asymptotic distribution of $\sqrt{n}\{\widehat{F}_n(\cdot) - F(\cdot)\}$ is impacted by the estimation but does not depend on the model parameters. It is thus neither asymptotically estimation free, as is the case for purely linear models, nor asymptotically model dependent, as is the case for some nonlinear models. The asymptotic stochastic equicontinuity is also established. We consider an application to the estimation of the conditional Value-at-Risk.

Keywords: Adaptive estimation; GARCH; empirical distribution of residuals; asymptotic distribution of quantiles; stochastic equicontinuity; conditional VaR

References

[1] Andrews, D.W.K. (1994). Asymptotics for semiparametric econometric models via stochastic equicontinuity. *Econometrica* **62** 43–72. MR1258665 https://doi.org/10.2307/2951475
[2] Bai, J. (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.* **22** 2051–2061. MR1329182 https://doi.org/10.1214/aos/1176325771
[3] Bardet, J.-M. and Wintenberger, O. (2009). Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Ann. Statist.* **37** 2730–2759. MR2541445 https://doi.org/10.1214/08-AOS674
[4] Berkes, I. and Horváth, L. (2002). Empirical processes of residuals. In *Empirical Process Techniques for Dependent Data* 195–209. Boston, MA: Birkhäuser. MR1958782
[5] Berkes, I. and Horváth, L. (2003). Limit results for the empirical process of squared residuals in GARCH models. *Stochastic Process. Appl.* **105** 271–298. MR1978657 https://doi.org/10.1016/S0304-4149(03)00004-8
[6] Beutner, B., Heinemann, A. and Smeekes, S. (2019). A justification of conditional confidence intervals. Discussion paper. Available at arXiv:1710.00643v2.
[7] Billingsley, P. (1961). The Lindeberg–Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12** 788–792. MR0126871 https://doi.org/10.2307/2034876
[8] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396
[9] Boldin, M.V. (1982). An estimate of the distribution of the noise in an autoregressive scheme. *Theory Probab. Appl.* **27** 866–871.
[10] Boldin, M.V. (1998). On residual empirical distribution functions in ARCH models with applications to testing and estimation. *Mitt. Math. Sem. Giessen* **235** 49–66. MR1661093
[11] Burkholder, D.L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692 https://doi.org/10.1214/aop/1176997023

 © 2022 ISI/BS

[12] Drost, F.C. and Klaassen, C.A.J. (1997). Efficient estimation in semiparametric GARCH models. *J. Econometrics* **81** 193–221. MR1484585 https://doi.org/10.1016/S0304-4076(97)00042-0
[13] Engle, R.F. and Gonzalez-Rivera, G. (1991). Semiparametric ARCH models. *J. Bus. Econom. Statist.* **9** 345–359.
[14] Feigin, P.D. and Tweedie, R.L. (1985). Random coefficient autoregressive processes: A Markov chain analysis of stationarity and finiteness of moments. *J. Time Series Anal.* **6** 1–14. MR0792428 https://doi.org/10.1111/j.1467-9892.1985.tb00394.x
[15] Francq, C. and Zakoïan, J.-M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10** 605–637. MR2076065 https://doi.org/10.3150/bj/1093265632
[16] Francq, C. and Zakoïan, J.-M. (2013). Optimal predictions of powers of conditionally heteroscedastic processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 345–367. MR3021391 https://doi.org/10.1111/j.1467-9868.2012.01045.x
[17] Francq, C. and Zakoïan, J.-M. (2013). Inference in nonstationary asymmetric GARCH models. *Ann. Statist.* **41** 1970–1998. MR3127855 https://doi.org/10.1214/13-AOS1132
[18] Francq, C. and Zakoïan, J.-M. (2015). Risk-parameter estimation in volatility models. *J. Econometrics* **184** 158–173. MR3277559 https://doi.org/10.1016/j.jeconom.2014.06.019
[19] Francq, C. Zakoïan, J.-M. (2022). Supplement to "Adaptiveness of the empirical distribution of residuals in semi-parametric conditional location scale models." https://doi.org/10.3150/21-BEJ1357SUPP
[20] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*: *Probability and Mathematical Statistics*. New York: Academic Press. MR0624435
[21] Hallin, M., Paindaveine, D. and Šiman, M. (2010). Multivariate quantiles and multiple-output regression quantiles: From L_1 optimization to halfspace depth. *Ann. Statist.* **38** 635–669. MR2604670 https://doi.org/10.1214/09-AOS723
[22] Hoga, Y. (2019). Confidence intervals for conditional tail risk measures in ARMA-GARCH models. *J. Bus. Econom. Statist.* **37** 613–624. MR4016158 https://doi.org/10.1080/07350015.2017.1401543
[23] Horváth, L., Kokoszka, P. and Teyssière, G. (2001). Empirical process of the squared residuals of an ARCH sequence. *Ann. Statist.* **29** 445–469. MR1863965 https://doi.org/10.1214/aos/1009210548
[24] Koul, H.L. and Ling, S. (2006). Fitting an error distribution in some heteroscedastic time series models. *Ann. Statist.* **34** 994–1012. MR2283401 https://doi.org/10.1214/009053606000000191
[25] Kreiss, J.-P. (1987). On adaptive estimation in stationary ARMA processes. *Ann. Statist.* **15** 112–133. MR0885727 https://doi.org/10.1214/aos/1176350256
[26] Kreiss, J.-P. (1991). Estimation of the distribution function of noise in stationary processes. *Metrika* **38** 285–297. MR1142062 https://doi.org/10.1007/BF02613623
[27] Kulperger, R. and Yu, H. (2005). High moment partial sum processes of residuals in GARCH models and their applications. *Ann. Statist.* **33** 2395–2422. MR2211090 https://doi.org/10.1214/009053605000000534
[28] Le Cam, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3** 37–98. MR126903
[29] Lee, S. and Taniguchi, M. (2005). Asymptotic theory for ARCH-SM models: LAN and residual empirical processes. *Statist. Sinica* **15** 215–234. MR2125729
[30] Ling, S. (1998). Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models. *Ann. Statist.* **26** 741–754. MR1626028 https://doi.org/10.1214/aos/1028144857
[31] Ling, S. (2007). A double AR(p) model: Structure and estimation. *Statist. Sinica* **17** 161–175. MR2352507
[32] Linton, O. (1993). Adaptive estimation in ARCH models. *Econometric Theory* **9** 539–569. MR1254299 https://doi.org/10.1017/S0266466600007970
[33] Robinson, P.M. (1987). Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* **55** 875–891. MR0906567 https://doi.org/10.2307/1911033
[34] 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
[35] Stute, W. (2001). Residual analysis for ARCH(p)-time series. *TEST* **10** 393–403. MR1881147 https://doi.org/10.1007/BF02595704
[36] Stute, W. and Schumann, G. (1980). A general Glivenko–Cantelli theorem for stationary sequences of random observations. *Scand. J. Stat.* **7** 102–104. MR0574551

[37] Wang, G., Zhu, K., Li, G. and Li, W.K. (2019). Hybrid quantile estimation for asymmetric power GARCH models. Available at arXiv:1911.09343.

[38] Zheng, Y., Zhu, Q., Li, G. and Xiao, Z. (2018). Hybrid quantile regression estimation for time series models with conditional heteroscedasticity. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** 975–993. MR3874306 https://doi.org/10.1111/rssb.12277

[39] Zhu, Q. and Li, G. (2019). Quantile double autoregression. Available at SSRN 3335753.

Bernoulli **28**(1), 2022, 579–600
https://doi.org/10.3150/21-BEJ1358

Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint

JOSEPH LAM-WEIL[1], BÉATRICE LAURENT[2] and JEAN-MICHEL LOUBES[3]

[1]*Otto von Guericke Magdeburg University, Universitätspl. 2, 39106 Magdeburg, Germany.*
E-mail: joseph.lam@ovgu.de
[2]*Institut de Mathématiques de Toulouse (UMR 5219), INSA Toulouse, Université de Toulouse, 135 avenue de Rangueil, 31400 Toulouse, France.*
E-mail: laurentb@insa-toulouse.fr
[3]*Institut de Mathématiques de Toulouse (UMR 5219), Université Paul Sabatier, Université de Toulouse, 135 avenue de Rangueil, 31400 Toulouse, France.*
E-mail: loubes@math.univ-toulouse.fr

Finding anonymization mechanisms to protect personal data is at the heart of recent machine learning research. Here, we consider the consequences of local differential privacy constraints on goodness-of-fit testing, that is, the statistical problem assessing whether sample points are generated from a fixed density f_0, or not. The observations are kept hidden and replaced by a stochastic transformation satisfying the local differential privacy constraint. In this setting, we propose a testing procedure which is based on an estimation of the quadratic distance between the density f of the unobserved samples and f_0. We establish an upper bound on the separation distance associated with this test, and a matching lower bound on the minimax separation rates of testing under non-interactive privacy in the case that f_0 is uniform, in discrete and continuous settings. To the best of our knowledge, we provide the first minimax optimal test and associated private transformation under a local differential privacy constraint over Besov balls in the continuous setting, quantifying the price to pay for data privacy. We also present a test that is adaptive to the smoothness parameter of the unknown density and remains minimax optimal up to a logarithmic factor. Finally, we note that our results can be translated to the discrete case, where the treatment of probability vectors is shown to be equivalent to that of piecewise constant densities in our setting. That is why we work with a unified setting for both the continuous and the discrete cases.

Keywords: Local differential privacy; non-interactive privacy; goodness-of-fit testing; minimax separation rates; continuous and discrete distributions

References

[1] Acharya, J., Canonne, C.L., Freitag, C. and Tyagi, H. (2018). Test without trust: Optimal locally private distribution testing. Preprint. Available at arXiv:1808.02174.
[2] Acharya, J., Sun, Z. and Zhang, H. (2018). Differentially private testing of identity and closeness of discrete distributions. In *Advances in Neural Information Processing Systems* 6878–6891.
[3] Agrawal, D. and Aggarwal, C.C. (2001). On the design and quantification of privacy preserving data mining algorithms. In *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 247–255. ACM.
[4] Agrawal, R. and Srikant, R. (2000). Privacy-preserving data mining. In *ACM Sigmod Record* **29** 439–450. ACM.
[5] Agrawal, S. and Haritsa, J.R. (2005). A framework for high-accuracy privacy-preserving mining. In 21*st International Conference on Data Engineering* (*ICDE*'05) 193–204. IEEE.

[6] Aliakbarpour, M., Diakonikolas, I., Kane, D. and Rubinfeld, R. (2019). Private testing of distributions via sample permutations. In *Advances in Neural Information Processing Systems* 10878–10889.

[7] Aliakbarpour, M., Diakonikolas, I. and Rubinfeld, R. (2018). Differentially private identity and equivalence testing of discrete distributions. In *International Conference on Machine Learning* 169–178.

[8] Balakrishnan, S. and Wasserman, L. (2019). Hypothesis testing for densities and high-dimensional multinomials: Sharp local minimax rates. *Ann. Statist.* **47** 1893–1927. MR3953439 https://doi.org/10.1214/18-AOS1729

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

[10] Berrett, T.B. and Butucea, C. (2020). Locally private non-asymptotic testing of discrete distributions is faster using interactive mechanisms. Preprint. Available at arXiv:2005.12601.

[11] Bickel, P.J. and Ritov, Y. (1992). Testing for goodness of fit: A new approach. In *Nonparametric Statistics and Related Topics* (*Ottawa*, *ON*, 1991) 51–57. Amsterdam: North-Holland. MR1226715

[12] Butucea, C., Dubois, A., Kroll, M. and Saumard, A. (2020). Local differential privacy: Elbow effect in optimal density estimation and adaptation over Besov ellipsoids. *Bernoulli* **26** 1727–1764. MR4091090 https://doi.org/10.3150/19-BEJ1165

[13] Butucea, C., Rohde, A. and Steinberger, L. (2020). Interactive versus non-interactive locally, differentially private estimation: Two elbows for the quadratic functional. Preprint. Available at arXiv:2003.04773.

[14] Cai, B., Daskalakis, C. and Kamath, G. (2017). Priv'it: Private and sample efficient identity testing. Preprint. Available at arXiv:1703.10127.

[15] Canonne, C.L., Kamath, G., McMillan, A., Ullman, J. and Zakynthinou, L. (2019). Private identity testing for high-dimensional distributions. Preprint. Available at arXiv:1905.11947.

[16] Dinur, I. and Nissim, K. (2003). Revealing information while preserving privacy. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 202–210.

[17] Donoho, D.L., Johnstone, I.M., Kerkyacharian, G. and Picard, D. (1996). Density estimation by wavelet thresholding. *Ann. Statist.* **24** 508–539. MR1394974 https://doi.org/10.1214/aos/1032894451

[18] Duchi, J.C., Jordan, M.I. and Wainwright, M.J. (2013). Local privacy and statistical minimax rates. In 2013 *IEEE* 54*th Annual Symposium on Foundations of Computer Science—FOCS* 2013 429–438. Los Alamitos, CA: IEEE Computer Soc. MR3246246 https://doi.org/10.1109/FOCS.2013.53

[19] Duchi, J. C., Jordan, M. I. and Wainwright, M. J. (2013). Local privacy, data processing inequalities, and statistical minimax rates. Preprint. Available at arXiv:1302.3203.

[20] Duchi, J.C., Jordan, M.I. and Wainwright, M.J. (2018). Minimax optimal procedures for locally private estimation. *J. Amer. Statist. Assoc.* **113** 182–201. MR3803452 https://doi.org/10.1080/01621459.2017.1389735

[21] Duchi, J. C., Wainwright, M. J. and Jordan, M. I. (2013). Local privacy and minimax bounds: Sharp rates for probability estimation. In *Advances in Neural Information Processing Systems* 1529–1537.

[22] Duncan, G. T. and Lambert, D. (1986). Disclosure-limited data dissemination. *J. Amer. Statist. Assoc.* **81** 10–18.

[23] Duncan, G.T. and Lambert, D. (1989). The risk of disclosure for microdata. *J. Bus. Econom. Statist.* **7** 207–217.

[24] Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I. and Naor, M. (2006). Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology—EUROCRYPT* 2006. *Lecture Notes in Computer Science* **4004** 486–503. Berlin: Springer. MR2423560 https://doi.org/10.1007/11761679_29

[25] Dwork, C., McSherry, F., Nissim, K. and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*. *Lecture Notes in Computer Science* **3876** 265–284. Berlin: Springer. MR2241676 https://doi.org/10.1007/11681878_14

[26] Evfimievski, A., Gehrke, J. and Srikant, R. (2003). Limiting privacy breaches in privacy preserving data mining. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 211–222. ACM.

[27] Fienberg, S. E. and Steele, R. J. (1998). Disclosure limitation using perturbation and related methods for categorical data. *J. Off. Stat.* **14** 485.

[28] Fromont, M. and Laurent, B. (2006). Adaptive goodness-of-fit tests in a density model. *Ann. Statist.* **34** 680–720. MR2281881 https://doi.org/10.1214/009053606000000119

[29] Gaboardi, M., Lim, H.-W., Rogers, R.M. and Vadhan, S.P. (2016). Differentially private chi-squared hypothesis testing: Goodness of fit and independence testing. In *ICML'*16 *Proceedings of the* 33*rd International Conference on International Conference on Machine Learning* **48**, JMLR.
[30] Gaboardi, M. and Rogers, R. (2017). Local private hypothesis testing: Chi-square tests. Preprint. Available at arXiv:1709.07155.
[31] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models*. *Cambridge Series in Statistical and Probabilistic Mathematics* **40**. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862
[32] Inglot, T. and Ledwina, T. (1996). Asymptotic optimality of data-driven Neyman's tests for uniformity. *Ann. Statist.* **24** 1982–2019. MR1421157 https://doi.org/10.1214/aos/1069362306
[33] Ingster, Y. I. (1987). Minimax testing of nonparametric hypotheses on a distribution density in the L_p metrics. *Theory Probab. Appl.* **31** 333–337.
[34] Ingster, Y. I. (2000). Adaptive chi-square tests. *J. Math. Sci.* **99** 1110–1119.
[35] Ingster, Yu.I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I. *Math. Methods Statist.* **2** 85–114. MR1257978
[36] Jank, W. and Shmueli, G., eds. (2008) *Statistical Methods in E-commerce Research*. *Statistics in Practice*. Hoboken, NJ: Wiley. MR2482329 https://doi.org/10.1002/9780470315262
[37] Joseph, M., Mao, J., Neel, S. and Roth, A. (2019). The role of interactivity in local differential privacy. In 2019 *IEEE* 60*th Annual Symposium on Foundations of Computer Science* (*FOCS*) 94–105. IEEE.
[38] Kallenberg, W.C.M. and Ledwina, T. (1995). Consistency and Monte Carlo simulation of a data driven version of smooth goodness-of-fit tests. *Ann. Statist.* **23** 1594–1608. MR1370299 https://doi.org/10.1214/aos/1176324315
[39] Kasiviswanathan, S.P., Lee, H.K., Nissim, K., Raskhodnikova, S. and Smith, A. (2011). What can we learn privately? *SIAM J. Comput.* **40** 793–826. MR2823508 https://doi.org/10.1137/090756090
[40] Lam-Weil, J., Laurent, B. and Loubes, J.-M. (2020). Minimax optimal goodness-of-fit testing for densities under a local differential privacy constraint. Preprint. Available at arXiv:2002.04254.
[41] Lam-Weil, J., Laurent, B. and Loubes, J.-M. (2022). Supplement to "Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint." https://doi.org/10.3150/21-BEJ1358SUPP
[42] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. *Springer Series in Statistics*. New York: Springer. MR0856411 https://doi.org/10.1007/978-1-4612-4946-7
[43] Ledwina, T. (1994). Data-driven version of Neyman's smooth test of fit. *J. Amer. Statist. Assoc.* **89** 1000–1005. MR1294744
[44] Mishra, N. and Sandler, M. (2006). Privacy via pseudorandom sketches. In *Proceedings of the Twenty-Fifth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* 143–152. ACM.
[45] Neyman, J. (1937). Smooth test for goodness of fit. *Scand. Actuar. J.* **1937** 149–199.
[46] Paninski, L. (2008). A coincidence-based test for uniformity given very sparsely sampled discrete data. *IEEE Trans. Inf. Theory* **54** 4750–4755. MR2591136 https://doi.org/10.1109/TIT.2008.928987
[47] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. *Wiley Series in Probability and Mathematical Statistics*. New York: Wiley. MR0595165
[48] Sheffet, O. (2018). Locally private hypothesis testing. Preprint. Available at arXiv:1802.03441.
[49] Tsybakov, A.B. (2004). *Introduction à L'estimation Non-paramétrique*. *Mathématiques & Applications* (*Berlin*) [*Mathematics & Applications*] **41**. Berlin: Springer. MR2013911
[50] Valiant, G. and Valiant, P. (2017). An automatic inequality prover and instance optimal identity testing. *SIAM J. Comput.* **46** 429–455. MR3614697 https://doi.org/10.1137/151002526
[51] van den Hout, A. and van der Heijden, P.G. (2002). Randomized response, statistical disclosure control and misclassificatio: A review. *Int. Stat. Rev.* **70** 269–288.
[52] Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.* **60** 63–69.
[53] Yang, Y. and Barron, A. (1999). Information-theoretic determination of minimax rates of convergence. *Ann. Statist.* **27** 1564–1599. MR1742500 https://doi.org/10.1214/aos/1017939142
[54] Yu, B. (1997). Assouad, Fano, and Le Cam. In *Festschrift for Lucien Le Cam* 423–435. New York: Springer. MR1462963

Bernoulli **28**(1), 2022, 601–605
https://doi.org/10.3150/21-BEJ1359
Main article: https://doi.org/10.3150/19-BEJ1126

Erratum for Prediction and estimation consistency of sparse multi-class penalized optimal scoring

IRINA GAYNANOVA

Department of Statistics, Texas A&M University, MS 3143, College Station, TX 77843, USA.
E-mail: irinag@stat.tamu.edu

Correction to the proof of Theorem 3 in Gaynanova (2020) *Bernoulli* **26** 286–322.

Keywords: Classification; high-dimensional regression; sub-exponential marginals

References

[1] Gaynanova, I. (2020). Prediction and estimation consistency of sparse multi-class penalized optimal scoring. *Bernoulli* **26** 286–322. MR4036035 https://doi.org/10.3150/19-BEJ1126
[2] Gaynanova, I., Booth, J.G. and Wells, M.T. (2016). Simultaneous sparse estimation of canonical vectors in the $p \gg N$ setting. *J. Amer. Statist. Assoc.* **111** 696–706. MR3538698 https://doi.org/10.1080/01621459.2015.1034318
[3] Obozinski, G., Wainwright, M.J. and Jordan, M.I. (2011). Support union recovery in high-dimensional multivariate regression. *Ann. Statist.* **39** 1–47. MR2797839 https://doi.org/10.1214/09-AOS776
[4] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170

1350-7265 © 2022 ISI/BS

Bernoulli **28**(1), 2022, 606–637
https://doi.org/10.3150/21-BEJ1360

On eigenvalues of a high-dimensional spatial-sign covariance matrix

WEIMING LI[1], QINWEN WANG[2], JIANFENG YAO[3] and WANG ZHOU[4]

[1]*School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, China.*
E-mail: li.weiming@shufe.edu.cn
[2]*School of Data Science, Fudan University, Shanghai, China.*
E-mail: wqw@fudan.edu.cn
[3]*Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong, China.*
E-mail: jeffyao@hku.hk
[4]*Department of Statistics and Applied Probability, National University of Singapore, Singapore, 117546..*
E-mail: stazw@nus.edu.sg

This paper investigates limiting spectral properties of a high-dimensional sample spatial-sign covariance matrix when both the dimension of the observations and the sample size grow to infinity. The underlying population is general enough to include the popular independent components model and the family of elliptical distributions. The first result of the paper shows that the empirical spectral distribution of a high dimensional sample spatial-sign covariance matrix converges to a generalized Marčenko-Pastur distribution. Secondly, a new central limit theorem for a class of related linear spectral statistics is established.

Keywords: Central limit theorem; eigenvalue distribution; linear spectral statistics; spatial-sign covariance matrix

References

[1] Anderson, G.W. and Zeitouni, O. (2006). A CLT for a band matrix model. *Probab. Theory Related Fields* **134** 283–338. MR2222385 https://doi.org/10.1007/s00440-004-0422-3
[2] 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
[3] Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statist. Sinica* **18** 425–442. MR2411613
[4] 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
[5] Bai, Z.D. and Silverstein, J.W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845
[6] 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
[7] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. MR0233396
[8] Billingsley, P. (1979). *Probability and Measure*. *Wiley Series in Probability and Mathematical Statistics*. New York–Chichester–Brisbane: Wiley. MR0534323
[9] Chakraborty, A. and Chaudhuri, P. (2017). Tests for high-dimensional data based on means, spatial signs and spatial ranks. *Ann. Statist.* **45** 771–799. MR3650400 https://doi.org/10.1214/16-AOS1467
[10] Dürre, A., Tyler, D.E. and Vogel, D. (2016). On the eigenvalues of the spatial sign covariance matrix in more than two dimensions. *Statist. Probab. Lett.* **111** 80–85. MR3474786 https://doi.org/10.1016/j.spl.2016.01.009
[11] Dürre, A., Vogel, D. and Fried, R. (2015). Spatial sign correlation. *J. Multivariate Anal.* **135** 89–105. MR3306428 https://doi.org/10.1016/j.jmva.2014.12.002

1350-7265 © 2022 ISI/BS

[12] Feng, L. and Sun, F. (2016). Spatial-sign based high-dimensional location test. *Electron. J. Stat.* **10** 2420–2434. MR3544292 https://doi.org/10.1214/16-EJS1176
[13] Feng, L., Zou, C. and Wang, Z. (2016). Multivariate-sign-based high-dimensional tests for the two-sample location problem. *J. Amer. Statist. Assoc.* **111** 721–735. MR3538700 https://doi.org/10.1080/01621459.2015.1035380
[14] Gao, J., Han, X., Pan, G. and Yang, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 677–693. MR3641402 https://doi.org/10.1111/rssb.12189
[15] Hu, J., Li, W., Liu, Z. and Zhou, W. (2019). High-dimensional covariance matrices in elliptical distributions with application to spherical test. *Ann. Statist.* **47** 527–555. MR3909941 https://doi.org/10.1214/18-AOS1699
[16] Hu, J., Li, W. and Zhou, W. (2019). Central limit theorem for mutual information of large MIMO systems with elliptically correlated channels. *IEEE Trans. Inf. Theory* **65** 7168–7180. MR4030880 https://doi.org/10.1109/TIT.2019.2913760
[17] 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
[18] Jonsson, D. (1982). Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** 1–38. MR0650926 https://doi.org/10.1016/0047-259X(82)90080-X
[19] Li, Y., Wang, Z. and Zou, C. (2016). A simpler spatial-sign-based two-sample test for high-dimensional data. *J. Multivariate Anal.* **149** 192–198. MR3507323 https://doi.org/10.1016/j.jmva.2016.04.004
[20] Locantore, N., Marron, J.S., Simpson, D.G., Tripoli, N., Zhang, J.T. and Cohen, K.L. (1999). Robust principal component analysis for functional data. *TEST* **8** 1–73. MR1707596 https://doi.org/10.1007/BF02595862
[21] Lytova, A. and Pastur, L. (2009). Central limit theorem for linear eigenvalue statistics of the Wigner and the sample covariance random matrices. *Ann. Probab.* **37** 1778-1840.
[22] Magyar, A.F. and Tyler, D.E. (2014). The asymptotic inadmissibility of the spatial sign covariance matrix for elliptically symmetric distributions. *Biometrika* **101** 673–688. MR3254908 https://doi.org/10.1093/biomet/asu020
[23] Pastur, L. and Marčenko, V. (1967). The distribution of eigenvalues in certain sets of random matrices. *Math. USSR, Sb.* **1** 457–483.
[24] Morales-Jimenez, D., Johnstone, I., McKay, M. and Yang, J. (2021). Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. *Statist. Sinica* **31** 571–601.
[25] Pan, G.M. and Zhou, W. (2008). Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* **18** 1232–1270. MR2418244 https://doi.org/10.1214/07-AAP477
[26] Silverstein, J.W. (1995). Strong convergence of the empirical distribution of eigenvalues of large-dimensional random matrices. *J. Multivariate Anal.* **55** 331–339. MR1370408 https://doi.org/10.1006/jmva.1995.1083
[27] Sinai, Ya. and Soshnikov, A. (1998). Central limit theorem for traces of large random symmetric matrices with independent matrix elements. *Bol. Soc. Bras. Mat.* **29** 1–24. MR1620151 https://doi.org/10.1007/BF01245866
[28] Visuri, S., Koivunen, V. and Oja, H. (2000). Sign and rank covariance matrices. *J. Statist. Plann. Inference* **91** 557–575. MR1814801 https://doi.org/10.1016/S0378-3758(00)00199-3
[29] Yao, J., Zheng, S. and Bai, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics* **39**. New York: Cambridge Univ. Press. MR3468554 https://doi.org/10.1017/CBO9781107588080
[30] Yin, Y.Q. (1986). Limiting spectral distribution for a class of random matrices. *J. Multivariate Anal.* **20** 50–68. MR0862241 https://doi.org/10.1016/0047-259X(86)90019-9
[31] Zheng, S., Bai, Z. and Yao, J. (2015). Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Statist.* **43** 546–591. MR3316190 https://doi.org/10.1214/14-AOS1292

Bernoulli **28**(1), 2022, 638–662
https://doi.org/10.3150/21-BEJ1361

Thinned completely random measures with applications in competing risks models

JOHN W. LAU[*] and EDWARD CRIPPS[†]

Department of Mathematics and Statistics, University of Western Australia, Perth, Australia.
E-mail: [*] *john.w.lau@gmail.com;* [†] *edward.cripps@uwa.edu.au*

We present a posterior analysis of kernel mixtures of thinned completely random measures (CRMs) for multivariate intensities, in the context of competing risks models. The construction of the thinned CRMs is derived from a common Poisson random measure that includes the thinning probabilities in its intensity and is transferable to existing Poisson partition calculus results for the posterior analysis (James (2002; *Ann. Statist.* **33** (2005) 1771–1799)). We derive the posterior thinned CRMs, provide generalizations of both the Blackwell and MacQueen Pólya urn formula and the (weighted) Chinese restaurant process for the variates and partitions generated from the thinned CRMs, and we outline strategies for the further development of Monte Carlo simulation for estimation.

Keywords: Completely random measures; thinning; Bayesian non-parametrics

References

[1] Aldous, D.J. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal*. **11** 581–598. MR0637937 https://doi.org/10.1016/0047-259X(81)90099-3

[2] Andersen, P.K., Borgan, Ø., Gill, R.D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. *Springer Series in Statistics*. New York: Springer. MR1198884 https://doi.org/10.1007/978-1-4612-4348-9

[3] Arfé, A., Peluso, S. and Muliere, P. (2019). Reinforced urns and the subdistribution beta-Stacy process prior for competing risks analysis. *Scand. J. Stat*. **46** 706–734. MR3994167 https://doi.org/10.1111/sjos.12369

[4] Barndorff-Nielsen, O.E., Pedersen, J. and Sato, K. (2001). Multivariate subordination, self-decomposability and stability. *Adv. in Appl. Probab.* **33** 160–187. MR1825321 https://doi.org/10.1239/aap/999187902

[5] Blackwell, D. and MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist*. **1** 353–355. MR0362614

[6] Bondesson, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Gebiete* **57** 39–71. MR0623454 https://doi.org/10.1007/BF00533713

[7] Bondesson, L. (1990). Generalized gamma convolutions and complete monotonicity. *Probab. Theory Related Fields* **85** 181–194. MR1050743 https://doi.org/10.1007/BF01277981

[8] Bondesson, L. (1992). *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*. *Lecture Notes in Statistics* **76**. New York: Springer. MR1224674 https://doi.org/10.1007/978-1-4612-2948-3

[9] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. in Appl. Probab.* **31** 929–953. MR1747450 https://doi.org/10.1239/aap/1029955251

[10] Camerlenghi, F., Lijoi, A., Orbanz, P. and Prünster, I. (2019). Distribution theory for hierarchical processes. *Ann. Statist*. **47** 67–92. MR3909927 https://doi.org/10.1214/17-AOS1678

[11] Camerlenghi, F., Lijoi, A. and Prünster, I. (2021). Survival analysis via hierarchically dependent mixture hazards. *Ann. Statist*. **49** 863–884. https://doi.org/10.1214/20-AOS1982

[12] Caron, F. (2012). Bayesian nonparametric models for bipartite graphs. In *Advances in Neural Information Processing Systems* 25 (F. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger, eds.) 2051–2059. Red Hook: Curran Associates.

 © 2022 ISI/BS

[13] Chen, C., Rao, V., Buntine, W. and Teh, Y.W. (2013). Dependent normalized random measures. In *Proceedings of the* 30*th International Conference on Machine Learning* (S. Dasgupta and D. McAllester, eds.). *Proceedings of Machine Learning Research* **28** 969–977. Atlanta, Georgia, USA: PMLR. Available at http://proceedings.mlr.press/v28/chen13i.html.

[14] De Iorio, M., Müller, P., Rosner, G.L. and MacEachern, S.N. (2004). An ANOVA model for dependent random measures. *J. Amer. Statist. Assoc.* **99** 205–215. MR2054299 https://doi.org/10.1198/016214504000000205

[15] Devroye, L. and James, L. (2014). On simulation and properties of the stable law. *Stat. Methods Appl.* **23** 307–343. MR3233961 https://doi.org/10.1007/s10260-014-0260-0

[16] Doksum, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.* **2** 183–201. MR0373081 https://doi.org/10.1214/aop/1176996703

[17] Dykstra, R.L. and Laud, P. (1981). A Bayesian nonparametric approach to reliability. *Ann. Statist.* **9** 356–367. MR0606619

[18] Epifani, I. and Lijoi, A. (2010). Nonparametric priors for vectors of survival functions. *Statist. Sinica* **20** 1455–1484. MR2777332

[19] Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

[20] Ferguson, T.S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629. MR0438568

[21] Ferguson, T.S. and Phadia, E.G. (1979). Bayesian nonparametric estimation based on censored data. *Ann. Statist.* **7** 163–186. MR0515691

[22] Foti, N., Futoma, J., Rockmore, D. and Williamson, S. (2013). A unifying representation for a class of dependent random measures. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics* (C.M. Carvalho and P. Ravikumar, eds.). *Proceedings of Machine Learning Research* **31** 20–28. Scottsdale, Arizona, USA: PMLR. Available at http://proceedings.mlr.press/v31/foti13a.html.

[23] Griffin, J.E. and Leisen, F. (2017). Compound random measures and their use in Bayesian non-parametrics. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 525–545. MR3611758 https://doi.org/10.1111/rssb.12176

[24] Griffin, J.E. and Steel, M.F.J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 179–194. MR2268037 https://doi.org/10.1198/016214505000000727

[25] Griffiths, R.C. and Milne, R.K. (1978). A class of bivariate Poisson processes. *J. Multivariate Anal.* **8** 380–395. MR0512608 https://doi.org/10.1016/0047-259X(78)90061-1

[26] Hjort, N.L. (1990). Nonparametric Bayes estimators based on beta processes in models for life history data. *Ann. Statist.* **18** 1259–1294. MR1062708 https://doi.org/10.1214/aos/1176347749

[27] Ishwaran, H. and James, L.F. (2004). Computational methods for multiplicative intensity models using weighted gamma processes: Proportional hazards, marked point processes, and panel count data. *J. Amer. Statist. Assoc.* **99** 175–190. MR2054297 https://doi.org/10.1198/016214504000000179

[28] James, L.F. (2002). Poisson process partition calculus with applications to exchangeable models and Bayesian nonparametrics. Preprint. Available at arXiv:math/0205093.

[29] James, L.F. (2003). Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhyā* **65** 179–206. MR2016784

[30] James, L.F. (2005). Bayesian Poisson process partition calculus with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33** 1771–1799. MR2166562 https://doi.org/10.1214/009053605000000336

[31] James, L.F. (2017). Bayesian Poisson calculus for latent feature modeling via generalized Indian buffet process priors. *Ann. Statist.* **45** 2016–2045. MR3718160 https://doi.org/10.1214/16-AOS1517

[32] James, L.F., Orbanz, P. and Teh, Y.W. (2015). Scaled subordinators and generalizations of the Indian buffet process. Preprint. Available at arXiv:1510.07309.

[33] James, L.F., Roynette, B. and Yor, M. (2008). Generalized gamma convolutions, Dirichlet means, Thorin measures, with explicit examples. *Probab. Surv.* **5** 346–415. MR2476736 https://doi.org/10.1214/07-PS118

[34] Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR1924807 https://doi.org/10.1002/9781118032985

[35] Kingman, J.F.C. (1967). Completely random measures. *Pacific J. Math.* **21** 59–78. MR0210185

[36] Kingman, J.F.C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. New York: Oxford Univ. Press. MR1207584
[37] Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks* **7**. Cambridge: Cambridge Univ. Press. MR3791470
[38] Lau, J.W. and Cripps, E. (2015). Stick-breaking representation and computation for normalized generalized gamma processes. *Sankhya A* **77** 300–329. MR3400115 https://doi.org/10.1007/s13171-015-0070-y
[39] Lau, J.W. and Cripps, E. (2022). Supplement to "Thinned completely random measures with applications in competing risks models." https://doi.org/10.3150/21-BEJ1361SUPP
[40] Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*, 2nd ed. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley Interscience. MR1940115
[41] Lijoi, A. and Nipoti, B. (2014). A class of hazard rate mixtures for combining survival data from different experiments. *J. Amer. Statist. Assoc.* **109** 802–814. MR3223751 https://doi.org/10.1080/01621459.2013.869499
[42] Lijoi, A., Nipoti, B. and Prünster, I. (2014). Bayesian inference with dependent normalized completely random measures. *Bernoulli* **20** 1260–1291. MR3217444 https://doi.org/10.3150/13-BEJ521
[43] Lijoi, A., Nipoti, B. and Prünster, I. (2014). Dependent mixture models: Clustering and borrowing information. *Comput. Statist. Data Anal.* **71** 417–433. MR3131980 https://doi.org/10.1016/j.csda.2013.06.015
[44] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12** 351–357. MR0733519 https://doi.org/10.1214/aos/1176346412
[45] Lo, A.Y. (1992). Bayesian inference for Poisson process models with censored data. *J. Nonparametr. Stat.* **2** 71–80. MR1256374 https://doi.org/10.1080/10485259208832544
[46] Lo, A.Y. (1993). A Bayesian bootstrap for censored data. *Ann. Statist.* **21** 100–123. MR1212168 https://doi.org/10.1214/aos/1176349017
[47] Lo, A.Y., Brunner, L.J. and Chan, A.T. (1996). Weighted chinese restaurant processes and Bayesian mixture models. Research Report, Hong Kong Univ. Science and Technology, Available at http://www.utstat.utoronto.ca/~brunner/papers/wcr96.pdf.
[48] Lo, A.Y. and Weng, C.-S. (1989). On a class of Bayesian nonparametric estimates. II. Hazard rate estimates. *Ann. Inst. Statist. Math.* **41** 227–245. MR1006487 https://doi.org/10.1007/BF00049393
[49] MacEachern, S. (1999). Dependent nonparametric processes. In *ASA Proceedings of the Section on Bayesian Statistical Science* 50–55. Alexandria: Amer. Statist. Assoc.
[50] MacEachern, S. (2000). Dependent Dirichlet processes. Unpublished manuscript, Department of Statistics, The Ohio State Univ.
[51] Neal, R.M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* **9** 249–265. MR1823804 https://doi.org/10.2307/1390653
[52] Pitman, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249 https://doi.org/10.1007/BF01213386
[53] Pitman, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **30** 245–267. Hayward, CA: IMS. MR1481784 https://doi.org/10.1214/lnms/1215453576
[54] Pitman, J. (2003). Poisson–Kingman partitions. In *Statistics and Science*: *A Festschrift for Terry Speed*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **40** 1–34. Beachwood, OH: IMS. MR2004330 https://doi.org/10.1214/lnms/1215091133
[55] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Berlin: Springer. MR2245368
[56] Pitman, J. and Yor, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900. MR1434129 https://doi.org/10.1214/aop/1024404422
[57] Riva Palacio, A. and Leisen, F. (2018). Bayesian nonparametric estimation of survival functions with multiple-samples information. *Electron. J. Stat.* **12** 1330–1357. MR3797716 https://doi.org/10.1214/18-EJS1420
[58] Rodríguez, A., Dunson, D.B. and Gelfand, A.E. (2008). The nested Dirichlet process. *J. Amer. Statist. Assoc.* **103** 1131–1144. MR2528831 https://doi.org/10.1198/016214508000000553
[59] Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations. *J. Amer. Statist. Assoc.* **71** 897–902. MR0436445

[60] Teh, Y.W., Jordan, M.I., Beal, M.J. and Blei, D.M. (2006). Hierarchical Dirichlet processes. *J. Amer. Statist. Assoc*. **101** 1566–1581. MR2279480 https://doi.org/10.1198/016214506000000302

Bernoulli **28**(1), 2022, 663–688
https://doi.org/10.3150/21-BEJ1362

Mixing properties of non-stationary INGARCH(1, 1) processes

PAUL DOUKHAN[1], ANNE LEUCHT[2] and MICHAEL H. NEUMANN[3]

[1]*UMR 8088 Analyse, Géométrie et Modélisation, CY University, 2, avenue Adolphe Chauvin, 95302 Cergy-Pontoise Cedex, France.*
E-mail: doukhan@cyu.fr
[2]*Research Group of Statistics and Mathematics, Universität Bamberg, Feldkirchenstraße 21, D – 96052 Bamberg, Germany.*
E-mail: anne.leucht@uni-bamberg.de
[3]*Institut für Mathematik, Friedrich-Schiller-Universität Jena, Ernst-Abbe-Platz 2, D – 07743 Jena, Germany.*
E-mail: michael.neumann@uni-jena.de

We derive mixing properties for a broad class of Poisson count time series satisfying a certain contraction condition. Using specific coupling techniques, we prove absolute regularity at a geometric rate not only for stationary Poisson-GARCH processes but also for models with an explosive trend. We provide easily verifiable sufficient conditions for absolute regularity for a variety of models including classical (log-)linear models. Finally, we illustrate the practical use of our results for hypothesis testing.

Keywords: Absolute regularity; coupling; INGARCH; mixing

References

[1] Agosto, A., Cavaliere, G., Kristensen, D. and Rahbek, A. (2016). Modeling corporate defaults: Poisson autoregressions with exogenous covariates (PARX). *J. Empir. Finance* **38** 640–663.
[2] Berbee, H.C.P. (1979). *Random Walks with Stationary Increments and Renewal Theory*. *Mathematical Centre Tracts* **112**. Amsterdam: Mathematisch Centrum. MR0547109
[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327. MR0853051 https://doi.org/10.1016/0304-4076(86)90063-1
[4] Bradley, R.C. (2007). *Introduction to Strong Mixing Conditions*. *Vol*. 1. Heber City, UT: Kendrick Press. MR2325294
[5] Daley, D.J. and Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. *Springer Series in Statistics*. New York: Springer. MR0950166
[6] Davis, R.A., Holan, S.H., Lund, R. and Ravishanker, N., eds. (2016). *Handbook of Discrete-Valued Time Series*. *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. Boca Raton, FL: CRC Press. MR3642975
[7] 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
[8] Doukhan, P. (2018). *Stochastic Models for Time Series*. *Mathématiques & Applications* (*Berlin*) [*Mathematics & Applications*] **80**. Cham: Springer. MR3852404 https://doi.org/10.1007/978-3-319-76938-7
[9] Doukhan, P., Khan, N.M. and Neumann, M.H. (2021). Mixing properties of integer-valued GARCH processes. *ALEA Lat. Am. J. Probab. Math. Stat.* **18** 401–420. MR4219669 https://doi.org/10.30757/alea.v18-18
[10] 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
[11] Doukhan, P., Neumann, M.H. and Truquet, L. (2020). Stationarity and ergodic properties for some observation-driven models in random environments. Preprint. Available at arXiv:2007.07623.

 © 2022 ISI/BS

[12] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773
[13] Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *J. Time Series Anal.* **27** 923–942. MR2328548 https://doi.org/10.1111/j.1467-9892.2006.00496.x
[14] Fokianos, K. (2012). Count time series. In *Time Series Analysis*: *Methods and Applications* (T. Subba Rao, S. Subba Rao and C.R. Rao, eds.). *Handbook of Statistics* **30** 315–347. Amsterdam: Elsevier/North-Holland.
[15] Fokianos, K., Rahbek, A. and Tjøstheim, D. (2009). Nonlinear Poisson autoregression. *J. Amer. Statist. Assoc.* **104** 1430–1439. MR2596998 https://doi.org/10.1198/jasa.2009.tm08270
[16] 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
[17] European Centre for Disease Prevention and Control (2020-10-02). Available at https://data.europa.eu/euodp/de/data/dataset/covid-19-coronavirus-data.
[18] Francq, C. and Zakoïan, J.-M. (2010). *GARCH Models*: *Structure*, *Statistical Inference and Financial Applications*. Chichester: Wiley. MR3186556 https://doi.org/10.1002/9780470670057
[19] Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34** 1–14.
[20] McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. *Monographs on Statistics and Applied Probability*. London: CRC Press. MR3223057 https://doi.org/10.1007/978-1-4899-3242-6
[21] Neumann, M.H. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli* **17** 1268–1284. MR2854772 https://doi.org/10.3150/10-BEJ313
[22] Neumann, M.H. (2021). Bootstrap for integer-valued GARCH(p, q) processes. *Stat. Neerl.* Published online 19 February 2021. Available at https://onlinelibrary.wiley.com/doi/epdf/10.1111/stan.12238.
[23] Rio, E. (1995/97). About the Lindeberg method for strongly mixing sequences. *ESAIM Probab. Stat.* **1** 35–61. MR1382517 https://doi.org/10.1051/ps:1997102
[24] Rio, E. (2017). *Asymptotic Theory of Weakly Dependent Random Processes*. *Probability Theory and Stochastic Modelling* **80**. Berlin: Springer. MR3642873 https://doi.org/10.1007/978-3-662-54323-8
[25] Roos, B. (2003). Improvements in the Poisson approximation of mixed Poisson distributions. *J. Statist. Plann. Inference* **113** 467–483. MR1965122 https://doi.org/10.1016/S0378-3758(02)00095-2
[26] Rydberg, T.H. and Shephard, P. (2000). A modeling framework for the prices and times of trades made on the New York Stock Exchange. In *Nonlinear and Nonstationary Signal Processing* (W.J. Fitzgerald, R.L. Smith, A.T. Walden and P.C. Young, eds.) 217–246. Cambridge: Cambridge Univ. Press.
[27] Streett, S.B. (2000). Some observation driven models for time series of counts. Ph.D. Thesis, Colorado State Univ., Department of Statistics. MR2701787
[28] Weiß, C.H. (2018). *An Introduction to Discrete-Valued Time Series*. New York: Wiley.
[29] Weiß, C.H., Zhu, F. and Hoshiyar, A. (2022). Softplus INGARCH model. *Statist. Sinica* **32**. https://doi.org/10.5705/ss.202020.0353
[30] Zhu, F. (2011). A negative binomial integer-valued GARCH model. *J. Time Series Anal.* **32** 54–67. MR2790672 https://doi.org/10.1111/j.1467-9892.2010.00684.x
[31] Zhu, F. (2012). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl.* **389** 58–71. MR2876481 https://doi.org/10.1016/j.jmaa.2011.11.042

Bernoulli **28**(1), 2022, 689–712
https://doi.org/10.3150/21-BEJ1364

A Ray–Knight representation of up-down Chinese restaurants

DANE ROGERS[*] and MATTHIAS WINKEL[†]

University of Oxford, Department of Statistics, 24–29 St Giles', Oxford OX1 3LB, UK.
E-mail: [*] *danejrogers@protonmail.com;* [†] *winkel@stats.ox.ac.uk*

We study composition-valued continuous-time Markov chains that appear naturally in the framework of Chinese Restaurant Processes (CRPs). As time evolves, new customers arrive (up-step) and existing customers leave (down-step) at suitable rates derived from the ordered CRP of Pitman and Winkel (*Ann. Probab.* **37** (2009) 1999–2041). We relate such up-down CRPs to the splitting trees of Lambert (*Ann. Probab.* **38** (2010) 348–395) inducing spectrally positive Lévy processes. Conversely, we develop theorems of Ray–Knight type to recover more general up-down CRPs from the heights of Lévy processes with jumps marked by integer-valued paths. We further establish limit theorems for the Lévy process and the integer-valued paths to connect to work by Forman, Pal, Rizzolo, Shi and Winkel on interval partition diffusions and hence to some long-standing conjectures.

Keywords: Chinese Restaurant Process; composition; Ray–Knight theorem; scaling limit; squared Bessel process; stable process

References

[1] Aldous, D. (1991). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326
[2] Aldous, D. (1999). Wright–Fisher diffusions with negative mutation rate! http://www.stat.berkeley.edu/~aldous/Research/OP/fw.html.
[3] Aldous, D. and Diaconis, P. (1987). Strong uniform times and finite random walks. *Adv. in Appl. Math.* **8** 69–97. MR0876954 https://doi.org/10.1016/0196-8858(87)90006-6
[4] Asmussen, S., Avram, F. and Pistorius, M.R. (2004). Russian and American put options under exponential phase-type Lévy models. *Stochastic Process. Appl.* **109** 79–111. MR2024845 https://doi.org/10.1016/j.spa.2003.07.005
[5] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. MR1700749 https://doi.org/10.1002/9780470316962
[6] Borodin, A. and Olshanski, G. (2009). Infinite-dimensional diffusions as limits of random walks on partitions. *Probab. Theory Related Fields* **144** 281–318. MR2480792 https://doi.org/10.1007/s00440-008-0148-8
[7] Carlton, M.A. (2002). A family of densities derived from the three-parameter Dirichlet process. *J. Appl. Probab.* **39** 764–774. MR1938169 https://doi.org/10.1017/s0021900200022026
[8] Feng, S. and Sun, W. (2010). Some diffusion processes associated with two parameter Poisson-Dirichlet distribution and Dirichlet process. *Probab. Theory Related Fields* **148** 501–525. MR2678897 https://doi.org/10.1007/s00440-009-0238-2
[9] Flajolet, P. and Guillemin, F. (2000). The formal theory of birth-and-death processes, lattice path combinatorics and continued fractions. *Adv. in Appl. Probab.* **32** 750–778. MR1788094 https://doi.org/10.1239/aap/1013540243
[10] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2018). Uniform control of local times of spectrally positive stable processes. *Ann. Appl. Probab.* **28** 2592–2634. MR3843837 https://doi.org/10.1214/17-AAP1370
[11] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2018). Aldous diffusion I: a projective system of continuum k-tree evolutions. arXiv:1809.07756 [math.PR].
[12] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2020). Diffusions on a space of interval partitions: Construction from marked Lévy processes. *Electron. J. Probab.* **25** Paper No. 133, 1–46. MR4169174 https://doi.org/10.1214/20-ejp521

 © 2022 ISI/BS

[13] Forman, N., Pal, S., Rizzolo, D. and Winkel, M. (2021). Diffusions on a space of interval partitions: Poisson–Dirichlet stationary distributions. *Ann. Probab.* **49** 793–831. https://doi.org/10.1214/20-AOP1460
[14] Forman, N., Rizzolo, D., Shi, Q. and Winkel, M. (2020). Diffusions on a space of interval partitions: the two-parameter model. arXiv:2008.02823 [math.PR].
[15] Forman, N., Rizzolo, D., Shi, Q. and Winkel, M. (2020). A two-parameter family of measure-valued diffusions with Poisson–Dirichlet stationary distributions. arXiv:2007.05250 [math.PR].
[16] Forman, N., Pal, S., Rizzolo, D.and Winkel, M. (2021). Ranked masses in two-parameter Fleming–Viot diffusions. arXiv:2101.09307 [math.PR].
[17] Fulman, J. (2009). Commutation relations and Markov chains. *Probab. Theory Related Fields* **144** 99–136. MR2480787 https://doi.org/10.1007/s00440-008-0143-0
[18] Geiger, J. and Kersting, G. (1997). Depth-first search of random trees, and Poisson point processes. In *Classical and Modern Branching Processes* (*Minneapolis*, *MN*, 1994). *IMA Vol. Math. Appl.* **84** 111–126. New York: Springer. MR1601713 https://doi.org/10.1007/978-1-4612-1862-3_8
[19] Gnedin, A. and Pitman, J. (2005). Regenerative composition structures. *Ann. Probab.* **33** 445–479. MR2122798 https://doi.org/10.1214/009117904000000801
[20] Göing-Jaeschke, A. and Yor, M. (2003). A survey and some generalizations of Bessel processes. *Bernoulli* **9** 313–349. MR1997032 https://doi.org/10.3150/bj/1068128980
[21] Grimmett, G.R. and Stirzaker, D.R. (2001). *Probability and Random Processes*, 3rd ed. New York: Oxford Univ. Press. MR2059709
[22] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Berlin: Springer. MR1943877 https://doi.org/10.1007/978-3-662-05265-5
[23] Jagers, P. (1969). A general stochastic model for population development. *Skand. Aktuarietidskr.* 84–103. MR0431406 https://doi.org/10.1080/03461238.1969.10405220
[24] Jagers, P. (1975). *Branching Processes with Biological Applications*. *Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics* London–New York–Sydney: Wiley. MR0488341
[25] James, L.F. (2006). Poisson calculus for spatial neutral to the right processes. *Ann. Statist.* **34** 416–440. MR2275248 https://doi.org/10.1214/009053605000000732
[26] Joyce, P. and Tavaré, S. (1987). Cycles, permutations and the structure of the Yule process with immigration. *Stochastic Process. Appl.* **25** 309–314. MR0915145 https://doi.org/10.1016/0304-4149(87)90209-2
[27] Knight, F.B. (1963). Random walks and a sojourn density process of Brownian motion. *Trans. Amer. Math. Soc.* **109** 56–86. MR0154337 https://doi.org/10.2307/1993647
[28] Lambert, A. (2010). The contour of splitting trees is a Lévy process. *Ann. Probab.* **38** 348–395. MR2599603 https://doi.org/10.1214/09-AOP485
[29] Lorentzen, L. and Waadeland, H. (2008). *Continued Fractions*, 2nd edn. ed. Amsterdam–Paris: Atlantis Press.
[30] Pal, S. (2013). Wright-Fisher diffusion with negative mutation rates. *Ann. Probab.* **41** 503–526. MR3077518 https://doi.org/10.1214/11-AOP704
[31] Petrov, L.A. (2009). Two-parameter family of diffusion processes in the Kingman simplex. *Funct. Anal. Appl.* **43** 279–296.
[32] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Berlin: Springer. MR2245368
[33] Pitman, J. and Winkel, M. (2009). Regenerative tree growth: Binary self-similar continuum random trees and Poisson-Dirichlet compositions. *Ann. Probab.* **37** 1999–2041. MR2561439 https://doi.org/10.1214/08-AOP445
[34] Pitman, J. and Yakubovich, Y. (2018). Ordered and size-biased frequencies in GEM and Gibbs' models for species sampling. *Ann. Appl. Probab.* **28** 1793–1820. MR3809477 https://doi.org/10.1214/17-AAP1343
[35] Pitman, J. and Yor, M. (1982). A decomposition of Bessel bridges. *Z. Wahrsch. Verw. Gebiete* **59** 425–457. MR0656509 https://doi.org/10.1007/BF00532802
[36] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Berlin: Springer. MR1725357 https://doi.org/10.1007/978-3-662-06400-9

[37] Rivera-Lopez, K. and Rizzolo, D. (2020). Diffusive limits of two-parameter ordered Chinese Restaurant Process up-down chains. arXiv:2011.06577 [math.PR].
[38] Rogers, D. (2020). Up-down ordered Chinese Restaurant Processes: Representation and asymptotics DPhil thesis, University of Oxford.
[39] Shi, Q. and Winkel, M. (2020). Up-down ordered Chinese restaurant processes with two-sided immigration, emigration and diffusion limits. arXiv:2012.15758 [math.PR].
[40] Shiga, T. (1990). A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes. *J. Math. Kyoto Univ.* **30** 245–279. MR1068791 https://doi.org/10.1215/kjm/1250520071
[41] Teh, Y.W. (2006). A hierarchical Bayesian language model based on Pitman-Yor processes. In *Proceedings of the* 21*st International Conference on Computational Linguistics and the* 44*th Annual Meeting of the Association for Computational Linguistics* 985–992.