BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Seven Number Three August 2021 ISSN: 1350-7265

CONTENTS

(continued)

BERNOULLI

Official Journal of the Bernoulli Society for Mathematical Statistics and Probability

Volume Twenty Seven Number Three August 2021 ISSN: 1350-7265

CONTENTS

(continued)

The papers published in Bernoulli are indexed or abstracted in *Current Index to Statistics*, *Mathematical Reviews*, *Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z)*, and *Zentralblatt für Mathematik* (also avalaible on the *MATH via STN* database and *Compact MATH CD-ROM*). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

BERNOULLI Volume 27 Number 3 August 2021 Pages 1441–2158 ISI/BS

Volume 27 Number 3   August 2021   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 *Current Index to Statistics, Mathematical Reviews, Statistical Theory and Method Abstracts-Zentralblatt (STMA-Z), Thomson Scientific* and *Zentralblatt für Mathematik* (also available on the *MATH via STN* database and *Compact MATH CD-ROM*). A list of forthcoming papers can be found online at http://www.bernoulli-society.org/index.php/publications/bernoulli-journal/bernoulli-journal-papers

In 2021 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

Bernoulli **27**(3), 2021, 1441–1466
https://doi.org/10.3150/20-BEJ1239

Bootstrap based inference for sparse high-dimensional time series models

JONAS KRAMPE[1], JENS-PETER KREISS[2] and EFSTATHIOS PAPARODITIS[3]

[1]*University of Mannheim, Germany.*
E-mail: j.krampe@uni-mannheim.de
[2]*Technische Universität Braunschweig, Germany.*
E-mail: j.kreiss@tu-bs.de
[3]*University of Cyprus, Cyprus.*
E-mail: stathisp@ucy.ac.cy

Fitting sparse models to high-dimensional time series is an important area of statistical inference. In this paper, we consider sparse vector autoregressive models and develop appropriate bootstrap methods to infer properties of such processes. Our bootstrap methodology generates pseudo time series using a model-based bootstrap procedure which involves an estimated, sparsified version of the underlying vector autoregressive model. Inference is performed using so-called de-sparsified or de-biased estimators of the autoregressive model parameters. We derive the asymptotic distribution of such estimators in the time series context and establish asymptotic validity of the bootstrap procedure proposed for estimation and, appropriately modified, for testing purposes. In particular, we focus on testing that large groups of autoregressive coefficients equal zero. Our theoretical results are complemented by simulations which investigate the finite sample performance of the bootstrap methodology proposed. A real-life data application is also presented.

Keywords: De-sparsified estimators; testing; vector autoregressive models

References

[1] Basu, S. and Michailidis, G. (2015). Regularized estimation in sparse high-dimensional time series models. *Ann. Statist.* **43** 1535–1567. MR3357870 https://doi.org/10.1214/15-AOS1315

[2] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217. MR0630103

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

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

[5] Cai, T. and Liu, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **106** 672–684. MR2847949 https://doi.org/10.1198/jasa.2011.tm10560

[6] Cai, T., Liu, W. and Luo, X. (2011). A constrained ℓ_1 minimization approach to sparse precision matrix estimation. *J. Amer. Statist. Assoc.* **106** 594–607. MR2847973 https://doi.org/10.1198/jasa.2011.tm10155

[7] Cai, T.T., Liu, W. and Zhou, H.H. (2016). Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *Ann. Statist.* **44** 455–488. MR3476606 https://doi.org/10.1214/13-AOS1171

[8] Cai, T.T., Ren, Z. and Zhou, H.H. (2016). Estimating structured high-dimensional covariance and precision matrices: Optimal rates and adaptive estimation. *Electron. J. Stat.* **10** 1–59. MR3466172 https://doi.org/10.1214/15-EJS1081

[9] Chatterjee, A. and Lahiri, S.N. (2010). Asymptotic properties of the residual bootstrap for Lasso estimators. *Proc. Amer. Math. Soc.* **138** 4497–4509. MR2680074 https://doi.org/10.1090/S0002-9939-2010-10474-4

[10] Chatterjee, A. and Lahiri, S.N. (2011). Bootstrapping lasso estimators. *J. Amer. Statist. Assoc.* **106** 608–625. MR2847974 https://doi.org/10.1198/jasa.2011.tm10159

1350-7265

[11] Chaudhry, A., Xu, P. and Gu, Q. (2017). Uncertainty assessment and false discovery rate control in high-dimensional granger causal inference. In *International Conference on Machine Learning* 684–693.

[12] Chernozhukov, V., Härdle, W.K., Huang, C. and Wang, W. (2018). Lasso-driven inference in time and space. Preprint. Available at arXiv:1806.05081.

[13] Davis, R.A., Zang, P. and Zheng, T. (2016). Sparse vector autoregressive modeling. *J. Comput. Graph. Statist.* **25** 1077–1096. MR3572029 https://doi.org/10.1080/10618600.2015.1092978

[14] Dezeure, R., Bühlmann, P. and Zhang, C.-H. (2017). High-dimensional simultaneous inference with the bootstrap. *TEST* **26** 685–719. MR3713586 https://doi.org/10.1007/s11749-017-0554-2

[15] Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Canad. J. Statist.* **9** 139–172. MR0640014 https://doi.org/10.2307/3314608

[16] Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. *Monographs on Statistics and Applied Probability* **57**. New York: CRC Press. MR1270903 https://doi.org/10.1007/978-1-4899-4541-9

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

[18] Farmer, R.E. (2015). The stock market crash really did cause the great recession. *Oxf. Bull. Econ. Stat.* **77** 617–633.

[19] Garcia, M.G., Medeiros, M.C. and Vasconcelos, G.F. (2017). Real-time inflation forecasting with high-dimensional models: The case of Brazil. *Int. J. Forecast.* **33** 679–693.

[20] Han, F., Lu, H. and Liu, H. (2015). A direct estimation of high dimensional stationary vector autoregressions. *J. Mach. Learn. Res.* **16** 3115–3150. MR3450535

[21] Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Cambridge: Cambridge Univ. Press.

[22] Knight, K. and Fu, W. (2000). Asymptotics for lasso-type estimators. *Ann. Statist.* **28** 1356–1378. MR1805787 https://doi.org/10.1214/aos/1015957397

[23] Kock, A.B. and Callot, L. (2015). Oracle inequalities for high dimensional vector autoregressions. *J. Econometrics* **186** 325–344. MR3343790 https://doi.org/10.1016/j.jeconom.2015.02.013

[24] Krampe, J., Kreiss, J.-P. and Paparoditis, E. (2021). Supplement to "Bootstrap based inference for sparse high-dimensional time series models." https://doi.org/10.3150/20-BEJ1239SUPP

[25] Ledoit, O. and Wolf, M. (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Statist.* **40** 1024–1060. MR2985942 https://doi.org/10.1214/12-AOS989

[26] Lin, J. and Michailidis, G. (2017). Regularized estimation and testing for high-dimensional multi-block vector-autoregressive models. *J. Mach. Learn. Res.* **18** 4188–4236. MR3725456 https://doi.org/10.1631/jzus.a1500279

[27] Liu, W., Xiao, H. and Wu, W.B. (2013). Probability and moment inequalities under dependence. *Statist. Sinica* **23** 1257–1272. MR3114713

[28] Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer. MR2172368 https://doi.org/10.1007/978-3-540-27752-1

[29] McCracken, M.W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *J. Bus. Econom. Statist.* **34** 574–589. MR3547997 https://doi.org/10.1080/07350015.2015.1086655

[30] Neykov, M., Ning, Y., Liu, J.S. and Liu, H. (2018). A unified theory of confidence regions and testing for high-dimensional estimating equations. *Statist. Sci.* **33** 427–443. MR3843384 https://doi.org/10.1214/18-STS661

[31] Ning, Y. and Liu, H. (2017). A general theory of hypothesis tests and confidence regions for sparse high dimensional models. *Ann. Statist.* **45** 158–195. MR3611489 https://doi.org/10.1214/16-AOS1448

[32] Phelps, E.S. (1999). Behind this structural boom: The role of asset valuations. *Am. Econ. Rev.* **89** 63–68.

[33] R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

[34] Reinsel, G.C. (1993). *Elements of Multivariate Time Series Analysis*. *Springer Series in Statistics*. New York: Springer. MR1238940 https://doi.org/10.1007/978-1-4684-0198-1

[35] Simon, N., Friedman, J., Hastie, T. and Tibshirani, R. (2011). Regularization paths for Cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* **39** 1–13. https://doi.org/10.18637/jss.v039.i05

[36] Song, S. and Bickel, P.J. (2011). Large vector auto regressions. Preprint. Available at arXiv:1106.3915.

[37] Tsay, R.S. (2014). *Multivariate Time Series Analysis*: *With R and Financial Applications*. *Wiley Series in Probability and Statistics*. Hoboken, NJ: Wiley. MR3236787
[38] van de Geer, S., Bühlmann, P., Ritov, Y. and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann*. *Statist*. **42** 1166–1202. MR3224285 https://doi.org/10.1214/14-AOS1221
[39] Wilms, I., Basu, S., Bien, J. and Matteson, D.S. (2019). Sparse identification and estimation of large-scale vector autoregressive moving averages. Preprint. Available at arXiv:1707.09208.
[40] Wu, W.-B. and Wu, Y.N. (2016). Performance bounds for parameter estimates of high-dimensional linear models with correlated errors. *Electron*. *J*. *Stat*. **10** 352–379. MR3466186 https://doi.org/10.1214/16-EJS1108
[41] 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
[42] 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
[43] Yan, Y. and Lin, F. (2016). FinCovRegularization: Covariance matrix estimation and regularization for finance. R package version 1.1.0.
[44] Zhang, C.-H. and Zhang, S.S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **76** 217–242. MR3153940 https://doi.org/10.1111/rssb.12026
[45] 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
[46] Zhang, X. and Cheng, G. (2018). Gaussian approximation for high dimensional vector under physical dependence. *Bernoulli* **24** 2640–2675. MR3779697 https://doi.org/10.3150/17-BEJ939
[47] Zheng, L. and Raskutti, G. (2019). Testing for high-dimensional network parameters in auto-regressive models. *Electron*. *J*. *Stat*. **13** 4977–5043. MR4041701 https://doi.org/10.1214/19-EJS1646

Bernoulli **27**(3), 2021, 1467–1495
https://doi.org/10.3150/20-BEJ1280

Inference without compatibility: Using exponential weighting for inference on a parameter of a linear model

MICHAEL LAW* and YA'ACOV RITOV†

Department of Statistics, University of Michigan, Ann Arbor, USA.
E-mail: *mmylaw@umich.edu; †yritov@umich.edu

We consider hypotheses testing problems for three parameters in high-dimensional linear models with minimal sparsity assumptions of their type but without any compatibility conditions. Under this framework, we construct the first $\sqrt{n}$-consistent estimators for low-dimensional coefficients, the signal strength, and the noise level. We support our results using numerical simulations and provide comparisons with other estimators.

Keywords: Lasso; compatibility condition; exponential weighting; inference

References

[1] Bellec, P.C. (2018). The noise barrier and the large signal bias of the lasso and other convex estimators. arXiv preprint arXiv:1804.01230.
[2] Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. *Johns Hopkins Series in the Mathematical Sciences*. Baltimore, MD: Johns Hopkins Univ. Press. MR1245941
[3] Bradic, J., Claeskens, G. and Gueuning, T. (2020). Fixed effects testing in high-dimensional linear mixed models. *J. Amer. Statist. Assoc.* 1–16.
[4] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data*: *Methods*, *Theory and Applications*. *Springer Series in Statistics*. Heidelberg: Springer. MR2807761 https://doi.org/10.1007/978-3-642-20192-9
[5] Cai, T.T. and Guo, Z. (2020). Semisupervised inference for explained variance in high dimensional linear regression and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 391–419. MR4084169
[6] Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **21** C1–C68. MR3769544 https://doi.org/10.1111/ectj.12097
[7] Dezeure, R., Bühlmann, P., Meier, L. and Meinshausen, N. (2015). High-dimensional inference: Confidence intervals, p-values and R-software `hdi`. *Statist. Sci.* **30** 533–558. MR3432840 https://doi.org/10.1214/15-STS527
[8] Dicker, L.H. (2014). Variance estimation in high-dimensional linear models. *Biometrika* **101** 269–284. MR3215347 https://doi.org/10.1093/biomet/ast065
[9] Fan, J., Guo, S. and Hao, N. (2012). Variance estimation using refitted cross-validation in ultrahigh dimensional regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 37–65. MR2885839 https://doi.org/10.1111/j.1467-9868.2011.01005.x
[10] Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 849–911. MR2530322 https://doi.org/10.1111/j.1467-9868.2008.00674.x
[11] Godsil, C. and Royle, G. (2001). *Algebraic Graph Theory*. *Graduate Texts in Mathematics* **207**. New York: Springer. MR1829620 https://doi.org/10.1007/978-1-4613-0163-9
[12] 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

[13] Hsu, D., Kakade, S.M. and Zhang, T. (2012). A tail inequality for quadratic forms of subgaussian random vectors. *Electron. Commun. Probab.* **17** no. 52, 6. MR2994877 https://doi.org/10.1214/ECP.v17-2079

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

[15] Javanmard, A. and Montanari, A. (2014). Confidence intervals and hypothesis testing for high-dimensional regression. *J. Mach. Learn. Res.* **15** 2869–2909. MR3277152

[16] Javanmard, A. and Montanari, A. (2018). Debiasing the Lasso: Optimal sample size for Gaussian designs. *Ann. Statist.* **46** 2593–2622. MR3851749 https://doi.org/10.1214/17-AOS1630

[17] Law, M. and Ritov, Y. (2021). Supplement to "Inference without compatibility: Using exponential weighting for inference on a parameter of a linear model." https://doi.org/10.3150/20-BEJ1280SUPP

[18] Lee, J.D., Sun, D.L., Sun, Y. and Taylor, J.E. (2016). Exact post-selection inference, with application to the lasso. *Ann. Statist.* **44** 907–927. MR3485948 https://doi.org/10.1214/15-AOS1371

[19] Leung, G. and Barron, A.R. (2006). Information theory and mixing least-squares regressions. *IEEE Trans. Inf. Theory* **52** 3396–3410. MR2242356 https://doi.org/10.1109/TIT.2006.878172

[20] Li, S., Cai, T.T. and Li, H. (2019). Inference for high-dimensional linear mixed-effects models: A quasi-likelihood approach. arXiv preprint arXiv:1907.06116.

[21] Raskutti, G., Wainwright, M.J. and Yu, B. (2010). Restricted eigenvalue properties for correlated Gaussian designs. *J. Mach. Learn. Res.* **11** 2241–2259. MR2719855

[22] Reid, S., Tibshirani, R. and Friedman, J. (2016). A study of error variance estimation in Lasso regression. *Statist. Sinica* **26** 35–67. MR3468344

[23] Rigollet, P. and Tsybakov, A. (2011). Exponential screening and optimal rates of sparse estimation. *Ann. Statist.* **39** 731–771. MR2816337 https://doi.org/10.1214/10-AOS854

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

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

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

[27] van de Geer, S., Bühlmann, P., Ritov, Y. and Dezeure, R. (2014). On asymptotically optimal confidence regions and tests for high-dimensional models. *Ann. Statist.* **42** 1166–1202. MR3224285 https://doi.org/10.1214/14-AOS1221

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

[29] Zhang, Y., Wainwright, M.J. and Jordan, M.I. (2014). Lower bounds on the performance of polynomial-time algorithms for sparse linear regression. In *Conference on Learning Theory* 921–948.

[30] Zhang, C.-H. and Zhang, S.S. (2014). Confidence intervals for low dimensional parameters in high dimensional linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 217–242. MR3153940 https://doi.org/10.1111/rssb.12026

Bernoulli **27**(3), 2021, 1496–1528
https://doi.org/10.3150/20-BEJ1281

Flux large deviations of weakly interacting jump processes via well-posedness of an associated Hamilton–Jacobi equation

RICHARD C. KRAAIJ

Delft Institute of Applied Mathematics, Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE Delft, The Netherlands.
E-mail: r.c.kraaij@tudelft.nl

We establish uniqueness for a class of first-order Hamilton–Jacobi equations with Hamiltonians that arise from the large deviations of the empirical measure and empirical flux pair of weakly interacting Markov jump processes. As a corollary, we obtain such a large deviation principle in the context of weakly interacting processes with time-periodic rates in which the period-length converges to 0.

Keywords: Weakly interacting jump processes; empirical measure and flux; large deviations; Hamilton–Jacobi equation

References

[1] Bardi, M. and Capuzzo-Dolcetta, I. (1997). *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*. *Systems & Control*: *Foundations & Applications*. Boston, MA: Birkhäuser, Inc. MR1484411 https://doi.org/10.1007/978-0-8176-4755-1

[2] Bertini, L., Chetrite, R., Faggionato, A. and Gabrielli, D. (2018). Level 2.5 large deviations for continuous-time Markov chains with time periodic rates. *Ann. Henri Poincaré* **19** 3197–3238. MR3851785 https://doi.org/10.1007/s00023-018-0705-3

[3] Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G. and Landim, C. (2002). Macroscopic fluctuation theory for stationary non-equilibrium states. *J. Stat. Phys.* **107** 635–675. MR1898852 https://doi.org/10.1023/A:1014525911391

[4] Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G. and Landim, C. (2006). Large deviations of the empirical current in interacting particle systems. *Teor. Veroyatn. Primen.* **51** 144–170. MR2324172 https://doi.org/10.1137/S0040585X97982256

[5] Bertini, L., Faggionato, A. and Gabrielli, D. (2015). Large deviations of the empirical flow for continuous time Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 867–900. MR3365965 https://doi.org/10.1214/14-AIHP601

[6] Borkar, V.S. and Sundaresan, R. (2012). Asymptotics of the invariant measure in mean field models with jumps. *Stoch. Syst.* **2** 322–380. MR3354770 https://doi.org/10.1214/12-SSY064

[7] Collet, F. and Kraaij, R.C. (2017). Dynamical moderate deviations for the Curie–Weiss model. *Stochastic Process. Appl.* **127** 2900–2925. MR3682118 https://doi.org/10.1016/j.spa.2017.01.002

[8] Comets, F. (1987). Nucleation for a long range magnetic model. *Ann. Inst. Henri Poincaré Probab. Stat.* **23** 135–178. MR0891708

[9] Costantini, C. and Kurtz, T.G. (2015). Viscosity methods giving uniqueness for martingale problems. *Electron. J. Probab.* **20** no. 67, 27. MR3361255 https://doi.org/10.1214/EJP.v20-3624

[10] 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

[11] Dai Pra, P. and den Hollander, F. (1996). McKean–Vlasov limit for interacting random processes in random media. *J. Stat. Phys.* **84** 735–772. MR1400186 https://doi.org/10.1007/BF02179656
[12] Deng, X., Feng, J. and Liu, Y. (2011). A singular 1-D Hamilton–Jacobi equation, with application to large deviation of diffusions. *Commun. Math. Sci.* **9** 289–300. MR2836847
[13] Dupuis, P., Ramanan, K. and Wu, W. (2016). Large deviation principle for finite-state mean field interacting particle systems. Preprint. Available at arXiv:1601.06219.
[14] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658
[15] Feng, J. and Kurtz, T.G. (2006). *Large Deviations for Stochastic Processes. Mathematical Surveys and Monographs* **131**. Providence, RI: Amer. Math. Soc. MR2260560 https://doi.org/10.1090/surv/131
[16] Fleming, W.H. and Soner, H.M. (2006). *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. *Stochastic Modelling and Applied Probability* **25**. New York: Springer. MR2179357
[17] Kraaij, R. (2016). Large deviations for finite state Markov jump processes with mean-field interaction via the comparison principle for an associated Hamilton–Jacobi equation. *J. Stat. Phys.* **164** 321–345. MR3513255 https://doi.org/10.1007/s10955-016-1542-8
[18] Kraaij, R.C. (2020). The exponential resolvent of a Markov process and large deviations for Markov processes via Hamilton–Jacobi equations. *Electron. J. Probab.* **25** 134.
[19] Kraaij, R.C. (2020). A general convergence result for viscosity solutions of Hamilton–Jacobi equations and non-linear semigroups. *J. Funct. Anal.* To appear.
[20] Kraaij, R.C. and Schlottke, M.C. (2019). Comparison Principle for Generalized Hamilton–Jacobi–Bellman Equations via a Bootstrapping Procedure. Preprint. Available at arXiv:1912.06579.
[21] Kraaij, R.C. and Schlottke, M.C. (2020). A large deviation principle for Markovian slow-fast systems Preprint. Available at arXiv:2011.05686.
[22] Léonard, C. (1995). Large deviations for long range interacting particle systems with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **31** 289–323. MR1324810
[23] Maes, C., Netočný, K. and Wynants, B. (2008). On and beyond entropy production: The case of Markov jump processes. *Markov Process. Related Fields* **14** 445–464. MR2455019
[24] Patterson, R.I.A. and Renger, D.R.M. (2019). Large deviations of jump process fluxes. *Math. Phys. Anal. Geom.* **22** Paper No. 21, 32. MR4002612 https://doi.org/10.1007/s11040-019-9318-4
[25] Perko, L. (2001). *Differential Equations and Dynamical Systems. Texts in Applied Mathematics* **7**. New York: Springer. MR1801796 https://doi.org/10.1007/978-1-4613-0003-8
[26] Renger, D.R.M. (2018). Flux large deviations of independent and reacting particle systems, with implications for macroscopic fluctuation theory. *J. Stat. Phys.* **172** 1291–1326. MR3856945 https://doi.org/10.1007/s10955-018-2083-0
[27] Shwartz, A. and Weiss, A. (2005). Large deviations with diminishing rates. *Math. Oper. Res.* **30** 281–310. MR2142034 https://doi.org/10.1287/moor.1040.0121

Bernoulli **27**(3), 2021, 1529–1555
https://doi.org/10.3150/20-BEJ1282

Precise asymptotics of longest cycles in random permutations without macroscopic cycles

VOLKER BETZ[1,*], JULIAN MÜHLBAUER[1], HELGE SCHÄFER[1,†] and DIRK ZEINDLER[2]

[1]*Fachbereich Mathematik, Arbeitsgruppe Stochastik, Schlossgartenstrasse 7, 64289 Darmstadt, Germany. E-mail:* [*]*betz@mathematik.tu-darmstadt.de;* [†]*hschaefer@mathematik.tu-darmstadt.de*
[2]*Lancaster University, Mathematics and Statistics, Fylde College, Bailrigg, Lancaster LA1 4YF, United Kingdom. E-mail: d.zeindler@lancaster.ac.uk*

We consider Ewens random permutations of length n conditioned to have no cycle longer than n^β with $0 < \beta < 1$ and study the asymptotic behaviour as $n \to \infty$. We obtain very precise information on the joint distribution of the lengths of the longest cycles; in particular we prove a functional limit theorem where the cumulative number of long cycles converges to a Poisson process in the suitable scaling. Furthermore, we prove convergence of the total variation distance between joint cycle counts and suitable independent Poisson random variables up to a significantly larger maximal cycle length than previously known. Finally, we remove a superfluous assumption from a central limit theorem for the total number of cycles proved in an earlier paper.

Keywords: Random permutations; Ewens measure; long cycles; functional limit theorem; Total variation distance; cycle structure

References

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

[2] Arratia, R. and Tavaré, S. (1992). The cycle structure of random permutations. *Ann*. *Probab*. **20** 1567–1591. MR1175278

[3] Betz, V. and Schäfer, H. (2017). The number of cycles in random permutations without long cycles is asymptotically Gaussian. *ALEA Lat*. *Am*. *J*. *Probab*. *Math*. *Stat*. **14** 427–444. MR3651976

[4] Betz, V., Schäfer, H. and Zeindler, D. (2020). Random permutations without macroscopic cycles. *Ann*. *Appl*. *Probab*. **30** 1484–1505. MR4133379 https://doi.org/10.1214/19-AAP1538

[5] Betz, V. and Ueltschi, D. (2009). Spatial random permutations and infinite cycles. *Comm*. *Math*. *Phys*. **285** 469–501. MR2461985 https://doi.org/10.1007/s00220-008-0584-4

[6] Betz, V. and Ueltschi, D. (2010). Critical temperature of dilute Bose gases. *Phys*. *Rev*. *A* **81** 023611.

[7] Betz, V. and Ueltschi, D. (2011). Spatial random permutations and Poisson–Dirichlet law of cycle lengths. *Electron*. *J*. *Probab*. **16** 1173–1192. MR2820074 https://doi.org/10.1214/EJP.v16-901

[8] Betz, V., Ueltschi, D. and Velenik, Y. (2011). Random permutations with cycle weights. *Ann*. *Appl*. *Probab*. **21** 312–331. MR2759204 https://doi.org/10.1214/10-AAP697

[9] 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

[10] Bogachev, L.V. and Zeindler, D. (2015). Asymptotic statistics of cycles in surrogate-spatial permutations. *Comm*. *Math*. *Phys*. **334** 39–116. MR3304271 https://doi.org/10.1007/s00220-014-2110-1

[11] DeLaurentis, J.M. and Pittel, B.G. (1985). Random permutations and Brownian motion. *Pacific J*. *Math*. **119** 287–301. MR0803120

[12] Elboim, D. and Peled, R. (2019). Limit distributions for Euclidean random permutations. *Comm. Math. Phys.* **369** 457–522. MR3962003 https://doi.org/10.1007/s00220-019-03421-8
[13] Ercolani, N.M. and Ueltschi, D. (2014). Cycle structure of random permutations with cycle weights. *Random Structures Algorithms* **44** 109–133. MR3143592 https://doi.org/10.1002/rsa.20430
[14] Ewens, W.J. (1972). The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* **3** 87–112; erratum, ibid. 3 (1972), 240; erratum, ibid. 3 (1972), 376. MR0325177 https://doi.org/10.1016/0040-5809(72)90035-4
[15] Flajolet, P. and Sedgewick, R. (2009). *Analytic Combinatorics*. Cambridge: Cambridge Univ. Press. MR2483235 https://doi.org/10.1017/CBO9780511801655
[16] Judkovich, D. (2019). The cycle structure of permutations without long cycles.
[17] Kingman, J.F.C. (1977). The population structure associated with the Ewens sampling formula. *Theor. Popul. Biol.* **11** 274–283. MR0682238 https://doi.org/10.1016/0040-5809(77)90029-6
[18] Lees, B. and Taggi, L. (2020). Site monotonicity and uniform positivity for interacting random walks and the spin $O(N)$ model with arbitrary N. *Comm. Math. Phys.* **376** 487–520. MR4093867 https://doi.org/10.1007/s00220-019-03647-6
[19] Liggett, T.M. (2010). *Continuous Time Markov Processes: An Introduction*. *Graduate Studies in Mathematics* **113**. Providence, RI: Amer. Math. Soc. MR2574430 https://doi.org/10.1090/gsm/113
[20] Manstavičius, E. and Petuchovas, R. (2016). Local probabilities for random permutations without long cycles. *Electron. J. Combin.* **23** Paper 1.58, 25. MR3484763
[21] Pólya, G. (1937). Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen. *Acta Math.* **68** 145–254. MR1577579 https://doi.org/10.1007/BF02546665
[22] Schäfer, H. (2018). The cycle structure of random permutations without macroscopic cycles. Ph.D. thesis, TU Darmstadt.
[23] Shmidt, A.A. and Vershik, A.M. (1977). Limit measures arising in the asymptotic theory of symmetric groups. *Theory Probab. Appl.* **22** 70–85.
[24] Taggi, L. (2019). Uniformly positive correlations in the dimer model and phase transition in lattice permutations on $\mathbb{Z}^d$, $d > 2$, via reflection positivity.
[25] Ueltschi, D. (2006). Feynman cycles in the Bose gas. *J. Math. Phys.* **47** 123303, 15. MR2285149 https://doi.org/10.1063/1.2383008
[26] Yakymiv, A.L. (2009). A limit theorem for the middle members of a variational series of cycle lengths of random A-permutation. *Teor. Veroyatn. Primen.* **54** 63–79. MR2766647 https://doi.org/10.1137/S0040585X97984073
[27] Yakymiv, A.L. (2010). A limit theorem for the logarithm of the order of a random A-permutation. *Diskret. Mat.* **22** 126–149. MR2676236 https://doi.org/10.1515/DMA.2010.015
[28] Yakymiv, A.L. (2011). A generalization of the Curtiss theorem for moment generating functions. *Mat. Zametki* **90** 947–952. MR2962966 https://doi.org/10.1134/S0001434611110290

Bernoulli **27**(3), 2021, 1556–1582
https://doi.org/10.3150/20-BEJ1283

Explicit bounds for critical infection rates and expected extinction times of the contact process on finite random graphs

E. CATOR* and H. DON†

Faculty of Science, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands.
E-mail: *e.cator@math.ru.nl; †h.don@math.ru.nl

We introduce a method to prove metastability of the contact process on Erdős–Rényi graphs and on configuration model graphs. The method relies on uniformly bounding the total infection rate from below, over all sets with a fixed number of nodes. Once this bound is established, a simple comparison with a well chosen birth-and-death process will show the exponential growth of the extinction time. Our paper complements recent results on the metastability of the contact process: under a certain minimal edge density condition, we give explicit lower bounds on the infection rate needed to get metastability, and we have explicit exponentially growing lower bounds on the expected extinction time.

Keywords: Contact process; critical infection rate; exponential extinction; metastability

References

[1] Cator, E. and Van Mieghem, P. (2012). Second-order mean-field susceptible–infected–susceptible epidemic threshold. *Phys. Rev. E* **85** 056111.
[2] Chatterjee, S. and Durrett, R. (2009). Contact processes on random graphs with power law degree distributions have critical value 0. *Ann. Probab.* **37** 2332–2356. MR2573560 https://doi.org/10.1214/09-AOP471
[3] Harris, T.E. (1974). Contact interactions on a lattice. *Ann. Probab.* **2** 969–988. MR0356292 https://doi.org/10.1214/aop/1176996493
[4] Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30. MR0144363
[5] Holley, R. and Liggett, T.M. (1978). The survival of contact processes. *Ann. Probab.* **6** 198–206. MR0488379 https://doi.org/10.1214/aop/1176995567
[6] Janson, S. and Luczak, M.J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34** 197–216. MR2490288 https://doi.org/10.1002/rsa.20231
[7] Liggett, T.M. (1996). Multiple transition points for the contact process on the binary tree. *Ann. Probab.* **24** 1675–1710. MR1415225 https://doi.org/10.1214/aop/1041903202
[8] Liggett, T.M. (1999). *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **324**. Berlin: Springer. MR1717346 https://doi.org/10.1007/978-3-662-03990-8
[9] Liggett, T.M. (2005). *Interacting Particle Systems. Classics in Mathematics*. Berlin: Springer. Reprint of the 1985 original. MR2108619 https://doi.org/10.1007/b138374
[10] Mourrat, J.-C. and Valesin, D. (2016). Phase transition of the contact process on random regular graphs. *Electron. J. Probab.* **21** Paper No. 31, 17. MR3492935 https://doi.org/10.1214/16-EJP4476
[11] Norris, J.R. (1998). *Markov Chains. Cambridge Series in Statistical and Probabilistic Mathematics* **2**. Cambridge: Cambridge Univ. Press. Reprint of 1997 original. MR1600720
[12] Pemantle, R. (1992). The contact process on trees. *Ann. Probab.* **20** 2089–2116. MR1188054
[13] Bhamidi, S., Nam, D., Nguyen, O. and Sly, A. (2019). Survival and extinction of epidemics on random graphs with general degrees. Available at arXiv:1902.03263.

1350-7265

[14] Stacey, A.M. (1996). The existence of an intermediate phase for the contact process on trees. *Ann. Probab.* **24** 1711–1726. MR1415226 https://doi.org/10.1214/aop/1041903203

[15] van der Hofstad, R. (2017). *Random Graphs and Complex Networks. Vol.* 1. *Cambridge Series in Statistical and Probabilistic Mathematics* **43**. Cambridge: Cambridge Univ. Press. MR3617364 https://doi.org/10.1017/9781316779422

Bernoulli **27**(3), 2021, 1583–1615
https://doi.org/10.3150/20-BEJ1284

Weak existence and uniqueness for affine stochastic Volterra equations with L^1-kernels

EDUARDO ABI JABER

Centre d'Economie de la Sorbonne, 106, Boulevard de l'Hôpital, 75013 Paris.
E-mail: eduardo.abi-jaber@univ-paris1.fr

We provide existence, uniqueness and stability results for affine stochastic Volterra equations with L^1-kernels and jumps. Such equations arise as scaling limits of branching processes in population genetics and self-exciting Hawkes processes in mathematical finance. The strategy we adopt for the existence part is based on approximations using stochastic Volterra equations with L^2-kernels combined with a general stability result. Most importantly, we establish weak uniqueness using a duality argument on the Fourier–Laplace transform via a deterministic Riccati–Volterra integral equation. We illustrate the applicability of our results on Hawkes processes and a class of hyper-rough Volterra Heston models with a Hurst index $H \in (-1/2, 1/2]$.

Keywords: Stochastic Volterra equations; affine Volterra processes; Riccati–Volterra equations; superprocesses; Hawkes processes; rough volatility

References

[1] Abi Jaber, E. (2019). Lifting the Heston model. *Quant. Finance* **19** 1995–2013. MR4029348 https://doi.org/10.1080/14697688.2019.1615113

[2] Abi Jaber, E., Cuchiero, C., Larsson, M. and Pulido, S. (2019). A weak solution theory for stochastic Volterra equations of convolution type. Preprint. Available at arXiv:1909.01166.

[3] Abi Jaber, E. and El Euch, O. (2019). Markovian structure of the Volterra Heston model. *Statist. Probab. Lett.* **149** 63–72. MR3911660 https://doi.org/10.1016/j.spl.2019.01.024

[4] Abi Jaber, E. and El Euch, O. (2019). Multifactor approximation of rough volatility models. *SIAM J. Financial Math.* **10** 309–349. MR3934104 https://doi.org/10.1137/18M1170236

[5] Abi Jaber, E., Larsson, M. and Pulido, S. (2019). Affine Volterra processes. *Ann. Appl. Probab.* **29** 3155–3200. MR4019885 https://doi.org/10.1214/19-AAP1477

[6] 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

[7] Brezis, H. (2011). *Functional Analysis*, *Sobolev Spaces and Partial Differential Equations*. *Universitext*. New York: Springer. MR2759829

[8] Cuchiero, C. and Teichmann, J. (2020). Generalized Feller processes and Markovian lifts of stochastic Volterra processes: The affine case. *J. Evol. Equ.* 1–48.

[9] Dawson, D.A. and Fleischmann, K. (1991). Critical branching in a highly fluctuating random medium. *Probab. Theory Related Fields* **90** 241–274. MR1128072 https://doi.org/10.1007/BF01192164

[10] Dawson, D.A. and Fleischmann, K. (1994). A super-Brownian motion with a single point catalyst. *Stochastic Process. Appl.* **49** 3–40. MR1258279 https://doi.org/10.1016/0304-4149(94)90110-4

[11] Dawson, D.A., Fleischmann, K., Li, Y. and Mueller, C. (1995). Singularity of super-Brownian local time at a point catalyst. *Ann. Probab.* **23** 37–55. MR1330759

[12] El Euch, O. and Rosenbaum, M. (2019). The characteristic function of rough Heston models. *Math. Finance* **29** 3–38. MR3905737 https://doi.org/10.1111/mafi.12173

[13] Etheridge, A.M. (2000). *An Introduction to Superprocesses*. *University Lecture Series* **20**. Providence, RI: Amer. Math. Soc. MR1779100 https://doi.org/10.1090/ulect/020

1350-7265

[14] Fleischmann, K. and Le Gall, J.-F. (1995). A new approach to the single point catalytic super-Brownian motion. *Probab. Theory Related Fields* **102** 63–82. MR1351711 https://doi.org/10.1007/BF01295222

[15] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2018). Volatility is rough. *Quant. Finance* **18** 933–949. MR3805308 https://doi.org/10.1080/14697688.2017.1393551

[16] Gripenberg, G., Londen, S.-O. and Staffans, O. (1990). *Volterra Integral and Functional Equations*. *Encyclopedia of Mathematics and Its Applications* **34**. Cambridge: Cambridge Univ. Press. MR1050319 https://doi.org/10.1017/CBO9780511662805

[17] 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

[18] 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

[19] Jarrow, R.A. (2018). *Continuous-Time Asset Pricing Theory*: *A Martingale-Based Approach*. *Springer Finance Textbooks*. Cham: Springer. MR3821509 https://doi.org/10.1007/978-3-319-77821-1

[20] Jusselin, P. and Rosenbaum, M. (2020). No-arbitrage implies power-law market impact and rough volatility. *Math. Finance* **30** 1309–1336. MR4154771 https://doi.org/10.1111/mafi.12254

[21] Kallsen, J. (2006). A didactic note on affine stochastic volatility models. In *From Stochastic Calculus to Mathematical Finance* 343–368. Berlin: Springer. MR2233549 https://doi.org/10.1007/978-3-540-30788-4_18

[22] Lépingle, D. and Mémin, J. (1978). Sur l'intégrabilité uniforme des martingales exponentielles. *Z. Wahrsch. Verw. Gebiete* **42** 175–203. MR0489492 https://doi.org/10.1007/BF00641409

[23] Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. MR0242239 https://doi.org/10.1137/1010093

[24] Mytnik, L. and Salisbury, T.S. (2015). Uniqueness for Volterra-type stochastic integral equations. Preprint. Available at arXiv:1502.05513.

[25] Natanson, I.P. (2016). *Theory of Functions of a Real Variable*, New York: Dover.

[26] Perkins, E. (2002). Dawson–Watanabe superprocesses and measure-valued diffusions. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1999). *Lecture Notes in Math.* **1781** 125–324. Berlin: Springer. MR1915445

[27] Revuz, D. and Yor, M. (2013). *Continuous martingales and Brownian motion*. **293**. Berlin: Springer.

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

[29] Zähle, H. (2005). Space-time regularity of catalytic super-Brownian motion. *Math. Nachr.* **278** 942–970. MR2141969 https://doi.org/10.1002/mana.200310284

Bernoulli **27**(3), 2021, 1616–1642
https://doi.org/10.3150/20-BEJ1285

Rates of contraction of posterior distributions based on p-exponential priors

SERGIOS AGAPIOU[1], MASOUMEH DASHTI[2] and TAPIO HELIN[3]

[1]*Department of Mathematics and Statistics, University of Cyprus, Nicosia, Cyprus.*
E-mail: agapiou.sergios@ucy.ac.cy
[2]*Department of Mathematics, University of Sussex, Brighton, United Kingdom.*
E-mail: m.dashti@sussex.ac.uk
[3]*School of Engineering Science, LUT University, Lappeenranta, Finland.*
E-mail: tapio.helin@lut.fi

We consider a family of infinite dimensional product measures with tails between Gaussian and exponential, which we call p-exponential measures. We study their measure-theoretic properties and in particular their concentration. Our findings are used to develop a general contraction theory of posterior distributions on nonparametric models with p-exponential priors in separable Banach parameter spaces. Our approach builds on the general contraction theory for Gaussian process priors in (*Ann. Statist.* **36** (2008) 1435–1463), namely we use prior concentration to verify prior mass and entropy conditions sufficient for posterior contraction. However, the specific concentration properties of p-exponential priors lead to a more complex entropy bound which can influence negatively the obtained rate of contraction, depending on the topology of the parameter space. Subject to the more complex entropy bound, we show that the rate of contraction depends on the position of the true parameter relative to a certain Banach space associated to p-exponential measures and on the small ball probabilities of these measures. For example, we apply our theory in the white noise model under Besov regularity of the truth and obtain minimax rates of contraction using (rescaled) α-regular p-exponential priors. In particular, our results suggest that when interested in spatially inhomogeneous unknown functions, in terms of posterior contraction, it is preferable to use Laplace rather than Gaussian priors.

Keywords: Bayesian nonparametric inference; non-Gaussian priors; concentration of measure

References

[1] Agapiou, S., Burger, M., Dashti, M. and Helin, T. (2018). Sparsity-promoting and edge-preserving maximum *a posteriori* estimators in non-parametric Bayesian inverse problems. *Inverse Probl*. **34** 045002, 37. MR3774703 https://doi.org/10.1088/1361-6420/aaacac

[2] Agapiou, S., Dashti, M. and Helin, T. (2021). Supplement to "Rates of contraction of posterior distributions based on p-exponential priors." https://doi.org/10.3150/20-BEJ1285SUPP

[3] Agapiou, S., Papaspiliopoulos, O., Sanz-Alonso, D. and Stuart, A.M. (2017). Importance sampling: Intrinsic dimension and computational cost. *Statist*. *Sci*. **32** 405–431. MR3696003 https://doi.org/10.1214/17-STS611

[4] Arbel, J., Gayraud, G. and Rousseau, J. (2013). Bayesian optimal adaptive estimation using a sieve prior. *Scand*. *J*. *Stat*. **40** 549–570. MR3091697 https://doi.org/10.1002/sjos.12002

[5] Aurzada, F. (2007). On the lower tail probabilities of some random sequences in l_p. *J*. *Theoret*. *Probab*. **20** 843–858. MR2359058 https://doi.org/10.1007/s10959-007-0095-9

[6] Belitser, E. and Ghosal, S. (2003). Adaptive Bayesian inference on the mean of an infinite-dimensional normal distribution. *Ann*. *Statist*. **31** 536–559. Dedicated to the memory of Herbert E. Robbins. MR1983541 https://doi.org/10.1214/aos/1051027880

[7] Bogachev, V.I. (2010). *Differentiable Measures and the Malliavin Calculus*. *Mathematical Surveys and Monographs* **164**. Providence, RI: Amer. Math. Soc. MR2663405 https://doi.org/10.1090/surv/164

[8] Borell, C. (1974). Convex measures on locally convex spaces. *Ark. Mat.* **12** 239–252. MR0388475 https://doi.org/10.1007/BF02384761

[9] Borell, C. (1975). The Brunn–Minkowski inequality in Gauss space. *Invent. Math.* **30** 207–216. MR0399402 https://doi.org/10.1007/BF01425510

[10] Candes, E. and Romberg, J. (2005). l1-magic: Recovery of sparse signals via convex programming. URL: www.acm.caltech.edu/l1magic/downloads/l1magic.pdf, 4:14.

[11] Candes, E.J. and Donoho, D.L. (2000). Curvelets: A surprisingly effective nonadaptive representation for objects with edges. Technical report, Stanford University, Department of Statistics.

[12] Castillo, I. (2008). Lower bounds for posterior rates with Gaussian process priors. *Electron. J. Stat.* **2** 1281–1299. MR2471287 https://doi.org/10.1214/08-EJS273

[13] Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. MR3127856 https://doi.org/10.1214/13-AOS1133

[14] Castillo, I., Schmidt-Hieber, J. and van der Vaart, A. (2015). Bayesian linear regression with sparse priors. *Ann. Statist.* **43** 1986–2018. MR3375874 https://doi.org/10.1214/15-AOS1334

[15] Cohen, A., Daubechies, I. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.* **1** 54–81. MR1256527 https://doi.org/10.1006/acha.1993.1005

[16] Dashti, M. and Stuart, A.M. (2017). The Bayesian approach to inverse problems. In *Handbook of Uncertainty Quantification. Vol.* 1, 2, 3 311–428. Cham: Springer. MR3839555

[17] Daubechies, I., Defrise, M. and De Mol, C. (2004). An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math.* **57** 1413–1457. MR2077704 https://doi.org/10.1002/cpa.20042

[18] Donoho, D.L. and Elad, M. (2003). Optimally sparse representation in general (nonorthogonal) dictionaries via l^1 minimization. *Proc. Natl. Acad. Sci. USA* **100** 2197–2202. MR1963681 https://doi.org/10.1073/pnas.0437847100

[19] Donoho, D.L. and Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inf. Theory* **47** 2845–2862. MR1872845 https://doi.org/10.1109/18.959265

[20] Donoho, D.L. and Johnstone, I.M. (1998). Minimax estimation via wavelet shrinkage. *Ann. Statist.* **26** 879–921. MR1635414 https://doi.org/10.1214/aos/1024691081

[21] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007 https://doi.org/10.1214/aos/1016218228

[22] Ghosal, S. and van der Vaart, A. (2007). Convergence rates of posterior distributions for non-i.i.d. observations. *Ann. Statist.* **35** 192–223. MR2332274 https://doi.org/10.1214/009053606000001172

[23] Ghosal, S. and van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference. Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR3587782 https://doi.org/10.1017/9781139029834

[24] Giné, E. and Nickl, R. (2011). Rates of contraction for posterior distributions in L^r-metrics, $1 \leq r \leq \infty$. *Ann. Statist.* **39** 2883–2911. MR3012395 https://doi.org/10.1214/11-AOS924

[25] 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

[26] Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications. Lecture Notes in Statistics* **129**. New York: Springer. MR1618204 https://doi.org/10.1007/978-1-4612-2222-4

[27] Helin, T. and Burger, M. (2015). Maximum *a posteriori* probability estimates in infinite-dimensional Bayesian inverse problems. *Inverse Probl.* **31** 085009, 22. MR3377106 https://doi.org/10.1088/0266-5611/31/8/085009

[28] Johnstone, I.M. (1994). Minimax Bayes, asymptotic minimax and sparse wavelet priors. In *Statistical Decision Theory and Related Topics, V* (*West Lafayette, IN*, 1992) 303–326. New York: Springer. MR1286310

[29] Kimeldorf, G.S. and Wahba, G. (1970). A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Stat.* **41** 495–502. MR0254999 https://doi.org/10.1214/aoms/1177697089

[30] Knapik, B.T., Szabó, B.T., van der Vaart, A.W. and van Zanten, J.H. (2016). Bayes procedures for adaptive inference in inverse problems for the white noise model. *Probab. Theory Related Fields* **164** 771–813. MR3477780 https://doi.org/10.1007/s00440-015-0619-7

[31] Knapik, B.T., van der Vaart, A.W. and van Zanten, J.H. (2011). Bayesian inverse problems with Gaussian priors. *Ann. Statist.* **39** 2626–2657. MR2906881 https://doi.org/10.1214/11-AOS920

[32] Kolehmainen, V., Lassas, M., Niinimäki, K. and Siltanen, S. (2012). Sparsity-promoting Bayesian inversion. *Inverse Probl.* **28** 025005, 28. MR2876856 https://doi.org/10.1088/0266-5611/28/2/025005

[33] Lassas, M., Saksman, E. and Siltanen, S. (2009). Discretization-invariant Bayesian inversion and Besov space priors. *Inverse Probl. Imaging* **3** 87–122. MR2558305 https://doi.org/10.3934/ipi.2009.3.87

[34] Lenk, P.J. (1988). The logistic normal distribution for Bayesian, nonparametric, predictive densities. *J. Amer. Statist. Assoc.* **83** 509–516. MR0971380

[35] Meyer, Y. (1992). *Wavelets and Operators*. *Cambridge Studies in Advanced Mathematics* **37**. Cambridge: Cambridge Univ. Press. Translated from the 1990 French original by D. H. Salinger. MR1228209

[36] Papaspiliopoulos, O., Pokern, Y., Roberts, G.O. and Stuart, A.M. (2012). Nonparametric estimation of diffusions: A differential equations approach. *Biometrika* **99** 511–531. MR2966767 https://doi.org/10.1093/biomet/ass034

[37] Pokern, Y., Stuart, A.M. and van Zanten, J.H. (2013). Posterior consistency via precision operators for Bayesian nonparametric drift estimation in SDEs. *Stochastic Process. Appl.* **123** 603–628. MR3003365 https://doi.org/10.1016/j.spa.2012.08.010

[38] Ray, K. (2013). Bayesian inverse problems with non-conjugate priors. *Electron. J. Stat.* **7** 2516–2549. MR3117105 https://doi.org/10.1214/13-EJS851

[39] Roberto, C. (2010). Isoperimetry for product of probability measures: Recent results. *Markov Process. Related Fields* **16** 617–634. MR2895085

[40] Rousseau, J. and Szabo, B. (2017). Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator. *Ann. Statist.* **45** 833–865. MR3650402 https://doi.org/10.1214/16-AOS1469

[41] Schmeisser, H.-J. and Triebel, H. (1987). *Topics in Fourier Analysis and Function Spaces*. *Mathematik und Ihre Anwendungen in Physik und Technik* [*Mathematics and Its Applications in Physics and Technology*] **42**. Leipzig: Akademische Verlagsgesellschaft Geest & Portig K.-G. MR0900143

[42] Shen, X. and Wasserman, L. (2001). Rates of convergence of posterior distributions. *Ann. Statist.* **29** 687–714. MR1865337 https://doi.org/10.1214/aos/1009210686

[43] Shively, T.S., Kohn, R. and Wood, S. (1999). Variable selection and function estimation in additive nonparametric regression using a data-based prior. *J. Amer. Statist. Assoc.* **94** 777–806. With comments and a rejoinder by the authors. MR1723272 https://doi.org/10.2307/2669990

[44] Stolz, W. (1996). Some small ball probabilities for Gaussian processes under nonuniform norms. *J. Theoret. Probab.* **9** 613–630. MR1400590 https://doi.org/10.1007/BF02214078

[45] Stuart, A.M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. MR2652785 https://doi.org/10.1017/S0962492910000061

[46] Szabó, B.T., van der Vaart, A.W. and van Zanten, J.H. (2013). Empirical Bayes scaling of Gaussian priors in the white noise model. *Electron. J. Stat.* **7** 991–1018. MR3044507 https://doi.org/10.1214/13-EJS798

[47] Talagrand, M. (1994). The supremum of some canonical processes. *Amer. J. Math.* **116** 283–325. MR1269606 https://doi.org/10.2307/2374931

[48] Triebel, H. (1978). *Interpolation Theory*, *Function Spaces*, *Differential Operators*. *North-Holland Mathematical Library* **18**. Amsterdam: North-Holland. MR0503903

[49] Triebel, H. (2010). *Bases in Function Spaces*, *Sampling*, *Discrepancy*, *Numerical Integration*. *EMS Tracts in Mathematics* **11**. Zürich: European Mathematical Society (EMS). MR2667814 https://doi.org/10.4171/085

[50] van der Vaart, A. and van Zanten, H. (2007). Bayesian inference with rescaled Gaussian process priors. *Electron. J. Stat.* **1** 433–448. MR2357712 https://doi.org/10.1214/07-EJS098

[51] van der Vaart, A.W. and van Zanten, J.H. (2008). Rates of contraction of posterior distributions based on Gaussian process priors. *Ann. Statist.* **36** 1435–1463. MR2418663 https://doi.org/10.1214/009053607000000613

[52] van der Vaart, A.W. and van Zanten, J.H. (2008). Reproducing kernel Hilbert spaces of Gaussian priors. In *Pushing the Limits of Contemporary Statistics*: *Contributions in Honor of Jayanta K*. *Ghosh*. *Inst*. *Math*. *Stat*. (*IMS*) *Collect*. **3** 200–222. Beachwood, OH: IMS. MR2459226 https://doi.org/10.1214/074921708000000156
[53] van der Vaart, A.W. and van Zanten, J.H. (2009). Adaptive Bayesian estimation using a Gaussian random field with inverse gamma bandwidth. *Ann*. *Statist*. **37** 2655–2675. MR2541442 https://doi.org/10.1214/08-AOS678
[54] Zhao, L.H. (2000). Bayesian aspects of some nonparametric problems. *Ann*. *Statist*. **28** 532–552. MR1790008 https://doi.org/10.1214/aos/1016218229

Bernoulli **27**(3), 2021, 1643–1665
https://doi.org/10.3150/20-BEJ1286

Yaglom's limit for critical Galton–Watson processes in varying environment: A probabilistic approach

NATALIA CARDONA-TOBÓN[1] and SANDRA PALAU[2]

[1]*Centro de Investigación en Matemáticas. Calle Jalisco s/n. C.P. 36240, Guanajuato, México.*
E-mail: natalia.cardona@cimat.mx
[2]*IIMAS, Universidad Nacional Autónoma de México. CDMX, 04510, Ciudad de México, México.*
E-mail: sandra@sigma.iimas.unam.mx

A Galton–Watson process in varying environment is a discrete time branching process where the offspring distributions vary among generations. Based on a two-spine decomposition technique, we provide a probabilistic argument of a Yaglom-type limit for this family processes. The result states that, in the critical case, a suitable normalisation of the process conditioned on non-extinction converges in distribution to a standard exponential random variable.

Keywords: Galton–Watson processes; varying environment; Yaglom's limit; spines decompositions

References

[1] Bhattacharya, N. and Perlman, M. (2017). Time-inhomogeneous branching processes conditioned on non-extinction. Preprint. Available at arXiv:1703.00337.
[2] Dolgopyat, D., Hebbar, P., Koralov, L. and Perlman, M. (2018). Multi-type branching processes with time-dependent branching rates. *J. Appl. Probab.* **55** 701–727. MR3877306 https://doi.org/10.1017/jpr.2018.46
[3] Geiger, J. (1999). Elementary new proofs of classical limit theorems for Galton–Watson processes. *J. Appl. Probab.* **36** 301–309. MR1724856 https://doi.org/10.1239/jap/1032374454
[4] Geiger, J. (2000). A new proof of Yaglom's exponential limit law. In *Mathematics and Computer Science* (*Versailles*, 2000). *Trends Math.* 245–249. Basel: Birkhäuser. MR1798303
[5] González, M., Kersting, G., Minuesa, C. and del Puerto, I. (2019). Branching processes in varying environment with generation-dependent immigration. *Stoch. Models* **35** 148–166. MR3969512 https://doi.org/10.1080/15326349.2019.1575754
[6] Harris, S.C., Johnston, S.G.G. and Roberts, M.I. (2020). The coalescent structure of continuous-time Galton–Watson trees. *Ann. Appl. Probab.* **30** 1368–1414. MR4133376 https://doi.org/10.1214/19-AAP1532
[7] Jagers, P. (1974). Galton-Watson processes in varying environments. *J. Appl. Probab.* **11** 174–178. MR0368197 https://doi.org/10.2307/3212594
[8] Kersting, G. (2020). A unifying approach to branching processes in a varying environment. *J. Appl. Probab.* **57** 196–220. MR4094390 https://doi.org/10.1017/jpr.2019.84
[9] Kersting, G. and Vatutin, V.A. (2017). *Discrete Time Branching Processes in Random Environment*. New York: Wiley.
[10] Kolmogorov, A. (1938). Zur lösung einer biologischen aufgabe. *Comm. Math. Mech. Chebyshev Univ. Tomsk* **2** 1–12.
[11] 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
[12] Ren, Y.-X., Song, R. and Sun, Z. (2018). A 2-spine decomposition of the critical Galton–Watson tree and a probabilistic proof of Yaglom's theorem. *Electron. Commun. Probab.* **23** Paper No. 42, 12. MR3841403 https://doi.org/10.1214/18-ECP143

[13] Yaglom, A.M. (1947). Certain limit theorems of the theory of branching random processes. *Dokl*. *Akad*. *Nauk SSSR* **56** 795–798. MR0022045

Bernoulli **27**(3), 2021, 1666–1694
https://doi.org/10.3150/20-BEJ1287

Correlation bounds, mixing and m-dependence under random time-varying network distances with an application to Cox-processes

ALEXANDER KREISS

ORSTAT KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium.
E-mail: alexander.kreiss@kuleuven.be

We will consider multivariate stochastic processes indexed either by vertices or pairs of vertices of a dynamic network. Under a dynamic network, we understand a network with a fixed vertex set and an edge set which changes randomly over time. We will assume that the spatial dependence-structure of the processes conditional on the network behaves in the following way: Close vertices (or pairs of vertices) are dependent, while we assume that the dependence decreases conditionally on that the distance in the network increases. We make this intuition mathematically precise by considering three concepts based on correlation, β-mixing with time-varying β-coefficients and conditional independence. These concepts allow proving weak-dependence results, for example, an exponential inequality, which might be of independent interest. In order to demonstrate the use of these concepts in an application, we study the asymptotics (for growing networks) of a goodness of fit test in a dynamic interaction network model based on a Cox-type model for counting processes. This model is then applied to bike-sharing data.

Keywords: Dynamic networks; dependence; survival analysis; nonparametric regression; hypothesis testing

References

[1] 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

[2] Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes: A large sample study. *Ann*. *Statist*. **10** 1100–1120. MR0673646

[3] Brownlees, C., Nualart, E. and Sun, Y. (2018). Realized networks. *J. Appl*. *Econometrics* **33** 986–1006. MR3887275 https://doi.org/10.1002/jae.2642

[4] Butts, C.T. (2008). A relational event framework for social action. *Sociol*. *Method*. **38** 155–200.

[5] Chen, L.H.Y. and Shao, Q.-M. (2004). Normal approximation under local dependence. *Ann*. *Probab*. **32** 1985–2028. MR2073183 https://doi.org/10.1214/009117904000000450

[6] Cox, D.R. (1972). Regression models and life-tables. *J. Roy*. *Statist*. *Soc*. *Ser*. *B* **34** 187–220. MR0341758

[7] Cox, T.F. and Cox, M.A.A. (1994). *Multidimensional Scaling*. *Monographs on Statistics and Applied Probability* **59**. London: CRC Press. MR1335449

[8] Crane, H. and Dempsey, W. (2018). Edge exchangeable models for interaction networks. *J. Amer*. *Statist*. *Assoc*. **113** 1311–1326. MR3862359 https://doi.org/10.1080/01621459.2017.1341413

[9] 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

[10] Demirer, M., Diebold, F.X., Liu, L. and Yilmaz, K. (2018). Estimating global bank network connectedness. *J. Appl*. *Econometrics* **33** 1–15. MR3771571 https://doi.org/10.1002/jae.2585

[11] 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

[12] Fox, E.W., Short, M.B., Schoenberg, F.P., Coronges, K.D. and Bertozzi, A.L. (2016). Modeling e-mail networks and inferring leadership using self-exciting point processes. *J. Amer. Statist. Assoc.* **111** 564–584. MR3538687 https://doi.org/10.1080/01621459.2015.1135802

[13] Friedman, N., Murphy, K. and Russell, S. (1998). Learning the structure of dynamic probabilistic networks. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, *UAI'98* 139–147. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

[14] Golder, S.A., Wilkinson, D.M. and Huberman, B.A. (2007). Rhythms of social interaction: Messaging within a massive online network. In *Communities and Technologies* (C. Steinfield, B.T. Pentland, M. Ackerman and N. Contractor, eds.). London: Springer.

[15] Grzegorczyk, M., Husmeier, D., Edwards, K.D., Ghazal, P. and Millar, A.J. (2008). Modelling non-stationary gene regulatory processes with a non-homogeneous Bayesian network and the allocation sampler. *Bioinformatics* **24** 2071–2078.

[16] Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Statist.* **21** 1926–1947. MR1245774 https://doi.org/10.1214/aos/1176349403

[17] Holland, P.W., Laskey, K.B. and Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. MR0718088 https://doi.org/10.1016/0378-8733(83)90021-7

[18] Huberman, B.A., Romero, D.M. and Wu, F. (2008). Social networks that matter: Twitter under the microscope. Preprint. Available at arXiv:0812.1045.

[19] Jackson, M.O. (2008). *Social and Economic Networks*. Princeton, NJ: Princeton Univ. Press. MR2435744

[20] Kauermann, G. and Berger, U. (2003). A smooth test in proportional hazard survival models using local partial likelihood fitting. *Lifetime Data Anal.* **9** 373–393. MR2045455 https://doi.org/10.1023/B:LIDA.0000012423.68151.da

[21] Kojevnikov, D., Marmer, V. and Song, K. (2019). Limit theorems for network dependent random variables. Working paper. Available at https://economics.ubc.ca/faculty-and-staff/denis-kojevnikov/.

[22] Kolaczyk, E.D. (2009). *Statistical Analysis of Network Data*: *Methods and Models*. *Springer Series in Statistics*. New York: Springer. MR2724362 https://doi.org/10.1007/978-0-387-88146-1

[23] Kolaczyk, E.D. (2017). *Topics at the Frontier of Statistics and Network Analysis*: (*Re*)*visiting the Foundations*. *SemStat Elements*. Cambridge: Cambridge Univ. Press. MR3702038 https://doi.org/10.1017/9781108290159

[24] Kreiß, A., Mammen, E. and Polonik, W. (2019). Nonparametric inference for continuous-time event counting and link-based dynamic network models. *Electron. J. Stat.* **13** 2764–2829. MR3995010 https://doi.org/10.1214/19-EJS1588

[25] Kreiß, A. (2021). Supplement to "Correlation bounds, mixing and m-dependence under random time-varying network distances with an application to Cox-processes." https://doi.org/10.3150/20-BEJ1287SUPP

[26] Linton, O., Mammen, E., Nielsen, J.P. and Van Keilegom, I. (2011). Nonparametric regression with filtered data. *Bernoulli* **17** 60–87. MR2797982 https://doi.org/10.3150/10-BEJ260

[27] Linton, O.B., Nielsen, J.P. and van de Geer, S. (2003). Estimating multiplicative and additive hazard functions by kernel methods. *Ann. Statist.* **31** 464–492. MR1983538 https://doi.org/10.1214/aos/1051027877

[28] Mammen, E. and Nielsen, J.P. (2007). A general approach to the predictability issue in survival analysis with applications. *Biometrika* **94** 873–892. MR2416797 https://doi.org/10.1093/biomet/asm062

[29] Martinussen, T. and Scheike, T.H. (2006). *Dynamic Regression Models for Survival Data*. *Statistics for Biology and Health*. New York: Springer. MR2214443

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

[31] Newman, M.E.J. (2010). *Networks*: *An Introduction*. Oxford: Oxford Univ. Press. MR2676073 https://doi.org/10.1093/acprof:oso/9780199206650.001.0001

[32] Nielsen, J.P., Linton, O. and Bickel, P.J. (1998). On a semiparametric survival model with flexible covariate effect. *Ann. Statist.* **26** 215–241. MR1611784 https://doi.org/10.1214/aos/1030563983

[33] Nielsen, J.P. and Linton, O.B. (1995). Kernel estimation in a nonparametric marker dependent hazard model. *Ann. Statist.* **23** 1735–1748. MR1370305 https://doi.org/10.1214/aos/1176324321

[34] Nze, P.A. and Doukhan, P. (2004). Weak dependence: Models and applications to econometrics. *Econometric Theory* **20** 995–1045. MR2101950 https://doi.org/10.1017/S0266466604206016

[35] Orbanz, P. and Roy, D.M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 437–461.
[36] Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. *The Morgan Kaufmann Series in Representation and Reasoning*. San Mateo, CA: Morgan Kaufmann. MR0965765
[37] Perry, P.O. and Wolfe, P.J. (2013). Point process modelling for directed interaction networks. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 821–849. MR3124793 https://doi.org/10.1111/rssb.12013
[38] 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
[39] Schweinberger, M. and Handcock, M.S. (2015). Local dependence in random graph models: Characterization, properties and statistical inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 647–676. MR3351449 https://doi.org/10.1111/rssb.12081
[40] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359 https://doi.org/10.1007/b13794
[41] Vainora, J. (2018). Network dependence and inference. In *EEA ESEM*. European Economic Association.
[42] Viennet, G. (1997). Inequalities for absolutely regular sequences: Application to density estimation. *Probab. Theory Related Fields* **107** 467–492. MR1440142 https://doi.org/10.1007/s004400050094
[43] Vu, D., Lomi, A., Mascia, D. and Pallotti, F. (2017). Relational event models for longitudinal network data with an application to interhospital patient transfers. *Stat. Med.* **36** 2265–2287. MR3660130 https://doi.org/10.1002/sim.7247

Bernoulli **27**(3), 2021, 1695–1723
https://doi.org/10.3150/20-BEJ1288

On the law of the iterated logarithm and strong invariance principles in stochastic geometry

JOHANNES KREBS

Department of Mathematics, TU Braunschweig, 38106 Braunschweig, Germany.
E-mail: johannes.krebs@tu-braunschweig.de

We study the law of the iterated logarithm (Khinchin (1924), Kolmogorov (1929)) and related strong invariance principles for functionals in stochastic geometry. As potential applications, we think of well-known functionals defined on the k-nearest neighbors graph and important functionals in topological data analysis such as the Euler characteristic and persistent Betti numbers.

Keywords: Binomial process; Euler characteristic; law of the iterated logarithm; persistent Betti numbers; Poisson process; stochastic geometry; strong invariance principles; strong stabilization; topological data analysis

References

[1] Adler, R.J. (2008). Some new random field tools for spatial analysis. *Stoch. Environ. Res. Risk Assess.* **22** 809–822. MR2430406 https://doi.org/10.1007/s00477-008-0242-6
[2] Arcones, M.A. (1997). The law of the iterated logarithm for a triangular array of empirical processes. *Electron. J. Probab.* **2** no. 5, 39 pp. MR1475863 https://doi.org/10.1214/EJP.v2-19
[3] Arcones, M.A. and Giné, E. (1995). On the law of the iterated logarithm for canonical U-statistics and processes. *Stochastic Process. Appl.* **58** 217–245. MR1348376 https://doi.org/10.1016/0304-4149(94)00023-M
[4] Berkes, I., Liu, W. and Wu, W.B. (2014). Komlós–Major–Tusnády approximation under dependence. *Ann. Probab.* **42** 794–817. MR3178474 https://doi.org/10.1214/13-AOP850
[5] Berkes, I. and Philipp, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* **7** 29–54. MR0515811
[6] Bickel, P.J. and Breiman, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214. MR0682809
[7] Biermé, H. and Durieu, O. (2014). Invariance principles for self-similar set-indexed random fields. *Trans. Amer. Math. Soc.* **366** 5963–5989. MR3256190 https://doi.org/10.1090/S0002-9947-2014-06135-7
[8] Carlsson, G. (2009). Topology and data. *Bull. Amer. Math. Soc.* (*N.S.*) **46** 255–308. MR2476414 https://doi.org/10.1090/S0273-0979-09-01249-X
[9] Chatterjee, S. and Sen, S. (2017). Minimal spanning trees and Stein's method. *Ann. Appl. Probab.* **27** 1588–1645. MR3678480 https://doi.org/10.1214/16-AAP1239
[10] Chazal, F. and Michel, B. (2017). An introduction to topological data analysis: Fundamental and practical aspects for data scientists. Preprint. Available at arXiv:1710.04019.
[11] Crawford, L., Monod, A., Chen, A.X., Mukherjee, S. and Rabadán, R. (2020). Predicting clinical outcomes in glioblastoma: An application of topological and functional data analysis. *J. Amer. Statist. Assoc.* **115** 1139–1150. MR4143455 https://doi.org/10.1080/01621459.2019.1671198
[12] Csörgő, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. With a foreword by David Kendall. MR2743035
[13] Decreusefond, L., Ferraz, E., Randriambololona, H. and Vergne, A. (2014). Simplicial homology of random configurations. *Adv. in Appl. Probab.* **46** 325–347. MR3215536 https://doi.org/10.1239/aap/1401369697
[14] Dehling, H. and Wendler, M. (2009). Law of the iterated logarithm for U-statistics of weakly dependent observations. Preprint. Available at arXiv:0911.1200.
[15] Eberlein, E. (1986). On strong invariance principles under dependence assumptions. *Ann. Probab.* **14** 260–270. MR0815969

[16] Edelsbrunner, H. and Harer, J.L. (2010). *Computational Topology*: *An Introduction*. Providence, RI: Amer. Math. Soc. MR2572029 https://doi.org/10.1090/mbk/069
[17] Gidea, M. and Katz, Y. (2018). Topological data analysis of financial time series: Landscapes of crashes. *Phys*. *A* **491** 820–834. MR3721543 https://doi.org/10.1016/j.physa.2017.09.028
[18] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*: *Probability and Mathematical Statistics*. New York–London: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0624435
[19] Hartman, P. and Wintner, A. (1941). On the law of the iterated logarithm. *Amer*. *J*. *Math*. **63** 169–176. MR0003497 https://doi.org/10.2307/2371287
[20] Hatcher, A. (2002). *Algebraic Topology*. Cambridge: Cambridge Univ. Press. MR1867354
[21] Heyde, C.C. and Scott, D.J. (1973). Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments. *Ann*. *Probab*. **1** 428–436. MR0353403 https://doi.org/10.1214/aop/1176996937
[22] Hug, D., Last, G. and Schulte, M. (2016). Second-order properties and central limit theorems for geometric functionals of Boolean models. *Ann*. *Appl*. *Probab*. **26** 73–135. MR3449314 https://doi.org/10.1214/14-AAP1086
[23] Jiang, J. (1999). Some laws of the iterated logarithm for two parameter martingales. *J*. *Theoret*. *Probab*. **12** 49–74. MR1674964 https://doi.org/10.1023/A:1021740425864
[24] Khinchin, A. (1924). *Asymptotische Gesetze der Wahrscheinlichkeitsrechnung*. Berlin: Springer.
[25] Kolmogoroff, A. (1929). Über das Gesetz des iterierten Logarithmus. *Math*. *Ann*. **101** 126–135. MR1512520 https://doi.org/10.1007/BF01454828
[26] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I. *Z*. *Wahrsch*. *Verw*. *Gebiete* **32** 111–131. MR0375412 https://doi.org/10.1007/BF00533093
[27] Komlós, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Z*. *Wahrsch*. *Verw*. *Gebiete* **34** 33–58. MR0402883 https://doi.org/10.1007/BF00532688
[28] Krebs, J. (2021). Supplement to "On the law of the iterated logarithm and strong invariance principles in stochastic geometry." https://doi.org/10.3150/20-BEJ1288SUPP
[29] Krebs, J. and Hirsch, C. (2020). Functional central limit theorems for persistent Betti numbers on cylindrical networks. Preprint. Available at arXiv:2003.13490.
[30] Krebs, J. and Polonik, W. (2019). On the asymptotic normality of persistent Betti numbers. Preprint. Available at arXiv:1903.03280.
[31] Krengel, U. (1985). *Ergodic Theorems*, *Vol*. 1. *De Gruyter Studies in Mathematics* **6**. Berlin: de Gruyter. With a supplement by Antoine Brunel. MR0797411 https://doi.org/10.1515/9783110844641
[32] Lachièze-Rey, R., Schulte, M. and Yukich, J.E. (2019). Normal approximation for stabilizing functionals. *Ann*. *Appl*. *Probab*. **29** 931–993. MR3910021 https://doi.org/10.1214/18-AAP1405
[33] Lee, S. (1997). The central limit theorem for Euclidean minimal spanning trees. I. *Ann*. *Appl*. *Probab*. **7** 996–1020. MR1484795 https://doi.org/10.1214/aoap/1043862422
[34] Lee, S. (1999). The central limit theorem for Euclidean minimal spanning trees. II. *Adv*. *in Appl*. *Probab*. **31** 969–984. MR1747451 https://doi.org/10.1239/aap/1029955253
[35] Lee, Y., Barthel, S.D., Dłotko, P., Moosavi, S.M., Hess, K. and Smit, B. (2017). Quantifying similarity of pore-geometry in nanoporous materials. *Nat*. *Commun*. **8** 15396.
[36] Lerche, H.R. (1986). Sequential analysis and the law of the iterated logarithm. In *Adaptive Statistical Procedures and Related Topics* (*Upton*, *N*.*Y*., 1985). *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **8** 40–53. Hayward, CA: IMS. MR0898237 https://doi.org/10.1214/lnms/1215540287
[37] Li, D.L., Rao, M.B. and Wang, X.C. (1992). The law of the iterated logarithm for independent random variables with multidimensional indices. *Ann*. *Probab*. **20** 660–674. MR1159566
[38] Oodaira, H. and Yoshihara, K. (1971). The law of the iterated logarithm for stationary processes satisfying mixing conditions. *Kodai Math*. *Sem*. *Rep*. **23** 311–334. MR0307311
[39] Penrose, M. (2003). *Random Geometric Graphs*. *Oxford Studies in Probability* **5**. Oxford: Oxford Univ. Press. MR1986198 https://doi.org/10.1093/acprof:oso/9780198506263.001.0001
[40] Penrose, M.D. and Yukich, J.E. (2001). Central limit theorems for some graphs in computational geometry. *Ann*. *Appl*. *Probab*. **11** 1005–1041. MR1878288 https://doi.org/10.1214/aoap/1015345393
[41] Philipp, W. (1969). The law of the iterated logarithm for mixing stochastic processes. *Ann*. *Math*. *Stat*. **40** 1985–1991. MR0317390 https://doi.org/10.1214/aoms/1177697280

[42] Philipp, W. (1977). A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables. *Ann. Probab.* **5** 319–350. MR0443024 https://doi.org/10.1214/aop/1176995795
[43] Philipp, W. and Stout, W. (1975). Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* **2** no. 116. MR0433597 https://doi.org/10.1090/memo/0161
[44] Rio, E. (1995). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23** 1188–1203. MR1349167
[45] Robbins, H. (1970). Statistical methods related to the law of the iterated logarithm. *Ann. Math. Stat.* **41** 1397–1409. MR0277063 https://doi.org/10.1214/aoms/1177696786
[46] Schmuland, B. and Sun, W. (2004). A central limit theorem and law of the iterated logarithm for a random field with exponential decay of correlations. *Canad. J. Math.* **56** 209–224. MR2031129 https://doi.org/10.4153/CJM-2004-010-6
[47] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. *Probability and Its Applications (New York)*. Berlin: Springer. MR2455326 https://doi.org/10.1007/978-3-540-78859-1
[48] Schulte, M. (2012). A central limit theorem for the Poisson–Voronoi approximation. *Adv. in Appl. Math.* **49** 285–306. MR3017961 https://doi.org/10.1016/j.aam.2012.08.001
[49] Shao, Q.M. (1993). Almost sure invariance principles for mixing sequences of random variables. *Stochastic Process. Appl.* **48** 319–334. MR1244549 https://doi.org/10.1016/0304-4149(93)90051-5
[50] Stout, W.F. (1970). The Hartman–Wintner law of the iterated logarithm for martingales. *Ann. Math. Stat.* **41** 2158–2160.
[51] Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrsch. Verw. Gebiete* **3** 211–226. MR0175194 https://doi.org/10.1007/BF00534910
[52] Strassen, V. (1967). Almost sure behavior of sums of independent random variables and martingales. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif.*, 1965/66) 315–343. Berkeley, CA: Univ. California Press. MR0214118
[53] Thäle, C. and Yukich, J.E. (2016). Asymptotic theory for statistics of the Poisson–Voronoi approximation. *Bernoulli* **22** 2372–2400. MR3498032 https://doi.org/10.3150/15-BEJ732
[54] Thomas, A.M. and Owada, T. (2019). Functional limit theorems for the Euler characteristic process in the critical regime. Preprint. Available at arXiv:1910.00751.
[55] Wichura, M.J. (1973). Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probab.* **1** 272–296. MR0394894 https://doi.org/10.1214/aop/1176996980
[56] Wu, W.B. (2005). On the Bahadur representation of sample quantiles for dependent sequences. *Ann. Statist.* **33** 1934–1963. MR2166566 https://doi.org/10.1214/009053605000000291
[57] Wu, W.B. (2007). Strong invariance principles for dependent random variables. *Ann. Probab.* **35** 2294–2320. MR2353389 https://doi.org/10.1214/009117907000000060
[58] Yao, Y., Sun, J., Huang, X., Bowman, G.R., Singh, G., Lesnick, M., Guibas, L.J., Pande, V.S. and Carlsson, G. (2009). Topological methods for exploring low-density states in biomolecular folding pathways. *J. Chem. Phys.* **130** 04B614.
[59] Yogeshwaran, D., Subag, E. and Adler, R.J. (2017). Random geometric complexes in the thermodynamic regime. *Probab. Theory Related Fields* **167** 107–142. MR3602843 https://doi.org/10.1007/s00440-015-0678-9
[60] Zhao, O. and Woodroofe, M. (2008). Law of the iterated logarithm for stationary processes. *Ann. Probab.* **36** 127–142. MR2370600 https://doi.org/10.1214/009117907000000079
[61] Zomorodian, A. and Carlsson, G. (2005). Computing persistent homology. *Discrete Comput. Geom.* **33** 249–274. MR2121296 https://doi.org/10.1007/s00454-004-1146-y

Bernoulli **27**(3), 2021, 1724–1744
https://doi.org/10.3150/20-BEJ1289

From Poincaré inequalities to nonlinear matrix concentration

DE HUANG* and JOEL A. TROPP†

California Institute of Technology, Pasadena, USA.
E-mail: *dhuang@caltech.edu; †jtropp@cms.caltech.edu

This paper deduces exponential matrix concentration from a Poincaré inequality via a short, conceptual argument. Among other examples, this theory applies to matrix-valued functions of a uniformly log-concave random vector. The proof relies on the subadditivity of Poincaré inequalities and a chain rule inequality for the trace of the matrix Dirichlet form. It also uses a symmetrization technique to avoid difficulties associated with a direct extension of the classic scalar argument.

Keywords: Concentration inequality; functional inequality; Markov process; matrix concentration; Poincaré inequality; semigroup

References

[1] Ahlswede, R. and Winter, A. (2002). Strong converse for identification via quantum channels. *IEEE Trans. Inf. Theory* **48** 569–579. MR1889969 https://doi.org/10.1109/18.985947
[2] Aida, S. and Stroock, D. (1994). Moment estimates derived from Poincaré and logarithmic Sobolev inequalities. *Math. Res. Lett.* **1** 75–86. MR1258492 https://doi.org/10.4310/MRL.1994.v1.n1.a9
[3] Aoun, R., Banna, M. and Youssef, P. (2020). Matrix Poincaré inequalities and concentration. *Adv. Math.* **371** 107251, 33. MR4108222 https://doi.org/10.1016/j.aim.2020.107251
[4] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9
[5] Bobkov, S. and Ledoux, M. (1997). Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields* **107** 383–400. MR1440138 https://doi.org/10.1007/s004400050090
[6] Bobkov, S.G. and Tetali, P. (2006). Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* **19** 289–336. MR2283379 https://doi.org/10.1007/s10959-006-0016-3
[7] 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
[8] Buchholz, A. (2001). Operator Khintchine inequality in non-commutative probability. *Math. Ann.* **319** 1–16. MR1812816 https://doi.org/10.1007/PL00004425
[9] Caffarelli, L.A. (2000). Monotonicity properties of optimal transportation and the FKG and related inequalities. *Comm. Math. Phys.* **214** 547–563. MR1800860 https://doi.org/10.1007/s002200000257
[10] Chatterjee, S. (2007). Stein's method for concentration inequalities. *Probab. Theory Related Fields* **138** 305–321. MR2288072 https://doi.org/10.1007/s00440-006-0029-y
[11] Chen, R.Y. and Tropp, J.A. (2014). Subadditivity of matrix φ-entropy and concentration of random matrices. *Electron. J. Probab.* **19** no. 27, 30. MR3174839 https://doi.org/10.1214/ejp.v19-2964
[12] Cheng, H.-C. and Hsieh, M.-H. (2016). Characterizations of matrix and operator-valued Φ-entropies, and operator Efron–Stein inequalities. *Proc. R. Soc. A* **472** 20150563, 20. MR3488705 https://doi.org/10.1098/rspa.2015.0563
[13] Cheng, H.-C. and Hsieh, M.-H. (2019). Matrix Poincaré, Φ-Sobolev inequalities, and quantum ensembles. *J. Math. Phys.* **60** 032201, 16. MR3922774 https://doi.org/10.1063/1.5035381

1350-7265

[14] Cheng, H.-C., Hsieh, M.-H. and Tomamichel, M. (2017). Exponential decay of matrix Φ-entropies on Markov semigroups with applications to dynamical evolutions of quantum ensembles. *J. Math. Phys.* **58** 092202, 24. MR3702668 https://doi.org/10.1063/1.5000846
[15] Davies, E.B. (1969). Quantum stochastic processes. *Comm. Math. Phys.* **15** 277–304. MR0266288
[16] Garg, A., Kathuria, T. and Srivastava, N. (2020). Scalar Poincaré implies matrix Poincaré. Preprint. Available at arXiv:2006.09567 [math.PR].
[17] Garg, A., Lee, Y.T., Song, Z. and Srivastava, N. (2018). A matrix expander Chernoff bound. In *STOC'18 – Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* 1102–1114. New York: ACM. MR3826320
[18] Gromov, M. and Milman, V.D. (1983). A topological application of the isoperimetric inequality. *Amer. J. Math.* **105** 843–854. MR0708367 https://doi.org/10.2307/2374298
[19] Gross, L. (1975). Hypercontractivity and logarithmic Sobolev inequalities for the Clifford Dirichlet form. *Duke Math. J.* **42** 383–396. MR0372613
[20] Hansen, F. and Zhang, Z. (2015). Characterisation of matrix entropies. *Lett. Math. Phys.* **105** 1399–1411. MR3395224 https://doi.org/10.1007/s11005-015-0784-8
[21] Hermon, J. and Salez, J. (2019). Modified log-Sobolev inequalities for strong-Rayleigh measures.
[22] Huang, D. and Tropp, J.A. (2021). Nonlinear matrix concentration via semigroup methods. *Electron. J. Probab.* **26** 1–31. https://doi.org/10.1214/20-EJP578
[23] Junge, M. and Zeng, Q. (2015). Noncommutative martingale deviation and Poincaré type inequalities with applications. *Probab. Theory Related Fields* **161** 449–507. MR3334274 https://doi.org/10.1007/s00440-014-0552-1
[24] Kastoryano, M.J. and Temme, K. (2013). Quantum logarithmic Sobolev inequalities and rapid mixing. *J. Math. Phys.* **54** 052202, 30. MR3098923 https://doi.org/10.1063/1.4804995
[25] Ledoux, M. (1995/97). On Talagrand's deviation inequalities for product measures. *ESAIM Probab. Stat.* **1** 63–87. MR1399224 https://doi.org/10.1051/ps:1997103
[26] Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347 https://doi.org/10.1090/surv/089
[27] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces: Isoperimetry and Processes. Classics in Mathematics*. Berlin: Springer. MR2814399
[28] Lindblad, G. (1976). On the generators of quantum dynamical semigroups. *Comm. Math. Phys.* **48** 119–130. MR0413878
[29] Lust-Piquard, F. (1986). Inégalités de Khintchine dans C_p $(1 < p < \infty)$. *C. R. Acad. Sci. Paris Sér. I Math.* **303** 289–292. MR0859804
[30] Mackey, L., Jordan, M.I., Chen, R.Y., Farrell, B. and Tropp, J.A. (2014). Matrix concentration inequalities via the method of exchangeable pairs. *Ann. Probab.* **42** 906–945. MR3189061 https://doi.org/10.1214/13-AOP892
[31] Majewski, A.W., Olkiewicz, R. and Zegarlinski, B. (1998). Dissipative dynamics for quantum spin systems on a lattice. *J. Phys. A* **31** 2045–2056. MR1628657 https://doi.org/10.1088/0305-4470/31/8/015
[32] Münch, F. (2019). Li–Yau inequality under $cd(0, n)$ on graphs. Preprint. Available at arXiv:1909.10242 [math.DG].
[33] Oliveira, R.I. (2009). Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges.
[34] Oliveira, R.I. (2010). Sums of random Hermitian matrices and an inequality by Rudelson. *Electron. Commun. Probab.* **15** 203–212. MR2653725 https://doi.org/10.1214/ECP.v15-1544
[35] Olkiewicz, R. and Zegarlinski, B. (1999). Hypercontractivity in noncommutative L_p spaces. *J. Funct. Anal.* **161** 246–285. MR1670230 https://doi.org/10.1006/jfan.1998.3342
[36] Paulin, D., Mackey, L. and Tropp, J.A. (2016). Efron–Stein inequalities for random matrices. *Ann. Probab.* **44** 3431–3473. MR3551202 https://doi.org/10.1214/15-AOP1054
[37] Pemantle, R. and Peres, Y. (2014). Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures. *Combin. Probab. Comput.* **23** 140–160. MR3197973 https://doi.org/10.1017/S0963548313000345
[38] Petz, D. (1994). A survey of certain trace inequalities. In *Functional Analysis and Operator Theory (Warsaw, 1992). Banach Center Publ.* **30** 287–298. Warsaw: Polish Acad. Sci. Inst. Math. MR1285615

[39] Pisier, G. (1986). Probabilistic methods in the geometry of Banach spaces. In *Probability and Analysis* (*Varenna*, 1985). *Lecture Notes in Math*. **1206** 167–241. Berlin: Springer. MR0864714 https://doi.org/10.1007/BFb0076302

[40] Pisier, G. and Xu, Q. (1997). Non-commutative martingale inequalities. *Comm*. *Math*. *Phys*. **189** 667–698. MR1482934 https://doi.org/10.1007/s002200050224

[41] Pitrik, J. and Virosztek, D. (2015). On the joint convexity of the Bregman divergence of matrices. *Lett*. *Math*. *Phys*. **105** 675–692. MR3339204 https://doi.org/10.1007/s11005-015-0757-y

[42] Rudelson, M. (1999). Random vectors in the isotropic position. *J*. *Funct*. *Anal*. **164** 60–72. MR1694526 https://doi.org/10.1006/jfan.1998.3384

[43] Stroock, D.W. (1984). *An Introduction to the Theory of Large Deviations*. *Universitext*. New York: Springer. MR0755154 https://doi.org/10.1007/978-1-4613-8514-1

[44] Tropp, J.A. (2011). Freedman's inequality for matrix martingales. *Electron*. *Commun*. *Probab*. **16** 262–270. MR2802042 https://doi.org/10.1214/ECP.v16-1624

[45] Tropp, J.A. (2012). User-friendly tail bounds for sums of random matrices. *Found*. *Comput*. *Math*. **12** 389–434. MR2946459 https://doi.org/10.1007/s10208-011-9099-z

[46] Tropp, J.A. (2015). An introduction to matrix concentration inequalities. *Found*. *Trends Mach*. *Learn*. **8** 1–230.

[47] Tropp, J.A. (2016). The expected norm of a sum of independent random matrices: An elementary approach. In *High Dimensional Probability VII*. *Progress in Probability* **71** 173–202. Cham: Springer. MR3565264 https://doi.org/10.1007/978-3-319-40519-3_8

[48] Van Handel, R. (2016). Probability in High Dimension. APC 550 lecture notes, Princeton Univ.

[49] Varopoulos, N.T. (1985). Hardy–Littlewood theory for semigroups. *J*. *Funct*. *Anal*. **63** 240–260. MR0803094 https://doi.org/10.1016/0022-1236(85)90087-4

Bernoulli **27**(3), 2021, 1745–1763
https://doi.org/10.3150/20-BEJ1290

Contact process under heavy-tailed renewals on finite graphs

LUIZ RENATO FONTES[1], PABLO ALMEIDA GOMES[2,*] and REMY SANCHIS[2,†]

[1] *Universidade de São Paulo, São Paulo, Brazil.*
E-mail: lrfontes@usp.br
[2] *Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.*
E-mail: [*] pabloag@ufmg.br; [†] rsanchis@mat.ufmg.br

We investigate a non-Markovian analogue of the Harris contact process in a finite connected graph $G = (V, E)$: an individual is attached to each site $x \in V$, and it can be infected or healthy; the infection propagates to healthy neighbors just as in the usual contact process, according to independent exponential times with a fixed rate $\lambda > 0$; however, the recovery times for an individual are given by the points of a renewal process attached to its timeline, whose waiting times have distribution μ such that $\mu(t, \infty) = t^{-\alpha} L(t)$, where $1/2 < \alpha < 1$ and $L(\cdot)$ is a slowly varying function; the renewal processes are assumed to be independent for different sites. We show that, starting with a single infected individual, if $|V| < 2 + (2\alpha - 1)/[(1 - \alpha)(2 - \alpha)]$, then the infection does not survive for any λ; and if $|V| > 1/(1 - \alpha)$, then, for every λ, the infection has positive probability to survive.

Keywords: Contact process; percolation; phase transition

References

[1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR0898871 https://doi.org/10.1017/CBO9780511721434
[2] Caravenna, F. and Doney, R. (2019). Local large deviations and the strong renewal theorem. *Electron. J. Probab.* **24** Paper No. 72, 48. MR3978222 https://doi.org/10.1214/19-EJP319
[3] Erickson, K.B. (1970). Strong renewal theorems with infinite mean. *Trans. Amer. Math. Soc.* **151** 263–291. MR0268976 https://doi.org/10.2307/1995628
[4] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd ed. New York: Wiley. MR0270403
[5] Fontes, L.R.G., Marchetti, D.H.U., Mountford, T.S. and Vares, M.E. (2019). Contact process under renewals I. *Stochastic Process. Appl.* **129** 2903–2911. MR3980148 https://doi.org/10.1016/j.spa.2018.08.007
[6] Fontes, L.R., Mountford, T.S. and Vares, M.E. (2020). Contact process under renewals II. *Stochastic Process. Appl.* **130** 1103–1118. MR4046531 https://doi.org/10.1016/j.spa.2019.04.008
[7] Garsia, A. and Lamperti, J. (1962/63). A discrete renewal theorem with infinite mean. *Comment. Math. Helv.* **37** 221–234. MR0148121 https://doi.org/10.1007/BF02566974
[8] Harris, T.E. (1974). Contact interactions on a lattice. *Ann. Probab.* **2** 969–988. MR0356292 https://doi.org/10.1214/aop/1176996493

Bernoulli **27**(3), 2021, 1764–1788
https://doi.org/10.3150/20-BEJ1291

Limit theorems for integral functionals of Hermite-driven processes

VALENTIN GARINO[1,*], IVAN NOURDIN[1,†], DAVID NUALART[2] and MAJID SALAMAT[3]

[1] *Département de mathématiques, Université du Luxembourg, Esch-sur-Alzette, Luxembourg.*
E-mail: [*] *valentin.garino@uni.lu;* [†] *ivan.nourdin@uni.lu*
[2] *Kansas University, Lawrence, KS, USA.*
E-mail: nualart@ku.edu
[3] *Department of Mathematical Sciences, Isfahan University of Technology, Isfahan, Iran.*
E-mail: msalamat@iut.ac.ir

Consider a moving average process X of the form $X(t) = \int_{-\infty}^{t} \varphi(t-u)\, dZ_u$, $t \geq 0$, where Z is a (non Gaussian) Hermite process of order $q \geq 2$ and $\varphi : \mathbb{R}_+ \to \mathbb{R}$ is sufficiently integrable. This paper investigates the fluctuations, as $T \to \infty$, of integral functionals of the form $t \mapsto \int_0^{Tt} P(X(s))\, ds$, in the case where P is any given polynomial function. It extends a study initiated in (*Stoch. Dyn.* **18** (2018) 1850028, 18), where only the quadratic case $P(x) = x^2$ and the convergence in the sense of finite-dimensional distributions were considered.

Keywords: Hermite processes; chaotic decomposition; fractional Brownian motion (fBm); multiple Wiener–Itô integrals

References

[1] Beckner, W. (1995). Geometric inequalities in Fourier anaylsis. In *Essays on Fourier Analysis in Honor of Elias M. Stein* (*Princeton, NJ*, 1991). *Princeton Math. Ser.* **42** 36–68. Princeton, NJ: Princeton Univ. Press. MR1315541

[2] Bell, D. and Nualart, D. (2017). Noncentral limit theorem for the generalized Hermite process. *Electron. Commun. Probab.* **22** Paper No. 66, 13. MR3734105 https://doi.org/10.1214/17-ECP99

[3] Biagini, F., Hu, Y., Øksendal, B. and Zhang, T. (2008). *Stochastic Calculus for Fractional Brownian Motion and Applications*. *Probability and Its Applications* (*New York*). London: Springer. MR2387368 https://doi.org/10.1007/978-1-84628-797-8

[4] Breuer, P. and Major, P. (1983). Central limit theorems for nonlinear functionals of Gaussian fields. *J. Multivariate Anal.* **13** 425–441. MR0716933 https://doi.org/10.1016/0047-259X(83)90019-2

[5] Campese, S., Nourdin, I. and Nualart, D. (2020). Continuous Breuer-Major theorem: Tightness and nonstationarity. *Ann. Probab.* **48** 147–177. MR4079433 https://doi.org/10.1214/19-AOP1357

[6] Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein–Uhlenbeck processes. *Electron. J. Probab.* **8** no. 3, 14. MR1961165 https://doi.org/10.1214/EJP.v8-125

[7] Diu Tran, T.T. (2018). Non-central limit theorems for quadratic functionals of Hermite-driven long memory moving average processes. *Stoch. Dyn.* **18** 1850028, 18. MR3842251 https://doi.org/10.1142/S0219493718500284

[8] Dobrushin, R.L. and Major, P. (1979). Non-central limit theorems for nonlinear functionals of Gaussian fields. *Z. Wahrsch. Verw. Gebiete* **50** 27–52. MR0550122 https://doi.org/10.1007/BF00535673

[9] Maejima, M. and Tudor, C.A. (2007). Wiener integrals with respect to the Hermite process and a non-central limit theorem. *Stoch. Anal. Appl.* **25** 1043–1056. MR2352951 https://doi.org/10.1080/07362990701540519

[10] Nourdin, I., Nualart, D. and Peccati, G. (2016). Strong asymptotic independence on Wiener chaos. *Proc. Amer. Math. Soc.* **144** 875–886. MR3430861 https://doi.org/10.1090/proc12769

1350-7265

[11] 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

[12] Nourdin, I. and Tran, T.T.D. (2019). Statistical inference for Vasicek-type model driven by Hermite processes. *Stochastic Process*. *Appl*. **129** 3774–3791. MR3997661 https://doi.org/10.1016/j.spa.2018.10.005

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

[14] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos*: *Moments*, *Cumulants and Diagrams*: *A Survey with Computer Implementation*. *Bocconi & Springer Series* **1**. Milan: Springer; Milan: Bocconi Univ. Press. MR2791919 https://doi.org/10.1007/978-88-470-1679-8

[15] Pipiras, V. and Taqqu, M.S. (2017). *Long-Range Dependence and Self-Similarity*. *Cambridge Series in Statistical and Probabilistic Mathematics*, [45]. Cambridge: Cambridge Univ. Press. MR3729426

[16] Slaoui, M. and Tudor, C.A. (2018). Limit behavior of the Rosenblatt Ornstein–Uhlenbeck process with respect to the Hurst index. *Teor. Ĭmovīr. Mat*. *Stat*. **98** 173–187. MR3824686 https://doi.org/10.1090/tpms/1070

[17] Stein, E.M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series*, *No*. 30. Princeton, NJ: Princeton Univ. Press. MR0290095

[18] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch*. *Verw. Gebiete* **50** 53–83. MR0550123 https://doi.org/10.1007/BF00535674

[19] Terrin, N. and Taqqu, M.S. (1991). Power counting theorem in Euclidean space. In *Random Walks*, *Brownian Motion*, *and Interacting Particle Systems*. *Progress in Probability* **28** 425–440. Boston, MA: Birkhäuser. MR1146462

Bernoulli **27**(3), 2021, 1789–1823
https://doi.org/10.3150/20-BEJ1292

Invariance and attraction properties of Galton–Watson trees

YEVGENIY KOVCHEGOV[1] and ILYA ZALIAPIN[2]

[1]*Department of Mathematics, Oregon State University, Corvallis, OR 97331-4605, USA.*
E-mail: kovchegy@math.oregonstate.edu; url: https://math.oregonstate.edu/people/view/kovchegy
[2]*Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557-0084, USA.*
E-mail: zal@unr.edu

We give a description of invariants and attractors of the critical and subcritical Galton–Watson tree measures under the operation of Horton pruning (cutting tree leaves with subsequent series reduction). Under a regularity condition, the class of invariant measures consists of the critical binary Galton–Watson tree and a one-parameter family of critical Galton–Watson trees with offspring distribution $\{q_k\}$ that has a power tail $q_k \sim Ck^{-(1+1/q_0)}$, where $q_0 \in (1/2, 1)$. Each invariant measure has a non-empty domain of attraction under consecutive Horton pruning, specified by the tail behavior of the initial Galton–Watson offspring distribution. The invariant measures satisfy the Toeplitz property for the Tokunaga coefficients and obey the Horton law with exponent $R = (1 - q_0)^{-1/q_0}$.

Keywords: Galton–Watson processes; self-similar trees; Horton–Strahler order; invariant measures; attractor

References

[1] Athreya, K.B. and Ney, P.E. (2004). *Branching Processes*. Mineola, NY: Dover. MR2047480
[2] Burd, G.A., Waymire, E.C. and Winn, R.D. (2000). A self-similar invariance of critical binary Galton–Watson trees. *Bernoulli* **6** 1–21. MR1781179 https://doi.org/10.2307/3318630
[3] Devroye, L. and Kruszewski, P. (1995). A note on the Horton–Strahler number for random trees. *Inform. Process. Lett.* **56** 95–99. MR1359172 https://doi.org/10.1016/0020-0190(95)00114-R
[4] Duquesne, T. and Winkel, M. (2007). Growth of Lévy trees. *Probab. Theory Related Fields* **139** 313–371. MR2322700 https://doi.org/10.1007/s00440-007-0064-3
[5] Duquesne, T. and Winkel, M. (2019). Hereditary tree growth and Lévy forests. *Stochastic Process. Appl.* **129** 3690–3747. MR3997659 https://doi.org/10.1016/j.spa.2018.10.007
[6] Evans, S.N. (2008). *Probability and Real Trees*. *Lecture Notes in Math.* **1920**. Berlin: Springer. MR2351587 https://doi.org/10.1007/978-3-540-74798-7
[7] Evans, S.N., Pitman, J. and Winter, A. (2006). Rayleigh processes, real trees, and root growth with re-grafting. *Probab. Theory Related Fields* **134** 81–126. MR2221786 https://doi.org/10.1007/s00440-004-0411-6
[8] He, H. and Winkel, M. Invariance principles for pruning processes of Galton–Watson trees. Preprint. Available at arXiv:1409.1014.
[9] Horton, R.E. (1945). Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology. *Geol. Soc. Amer. Bull.* **56** 275–370.
[10] Kesten, H. (1986). Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** 425–487. MR0871905
[11] Kovchegov, Y. and Zaliapin, I. (2016). Horton law in self-similar trees. *Fractals* **24** 1650017, 10. MR3507614 https://doi.org/10.1142/S0218348X16500171
[12] Kovchegov, Y. and Zaliapin, I. (2020). Dynamical pruning of rooted trees with applications to 1-d ballistic annihilation. *J. Stat. Phys.* **181** 618–672. MR4143639 https://doi.org/10.1007/s10955-020-02593-1
[13] Kovchegov, Y. and Zaliapin, I. (2020). Random self-similar trees: A mathematical theory of Horton laws. *Probab. Surv.* **17** 1–213. MR4067815 https://doi.org/10.1214/19-PS331

1350-7265

[14] Le Jan, Y. (1991). Superprocesses and projective limits of branching Markov process. *Ann. Inst. Henri Poincaré Probab. Stat.* **27** 91–106. MR1098565

[15] McConnell, M. and Gupta, V.K. (2008). A proof of the Horton law of stream numbers for the Tokunaga model of river networks. *Fractals* **16** 227–233. MR2451617 https://doi.org/10.1142/S0218348X08003958

[16] Neveu, J. (1986). Erasing a branching tree. *Adv. in Appl. Probab.* **1** 101–108. MR0868511

[17] Neveu, J. and Pitman, J. (1989). Renewal property of the extrema and tree property of the excursion of a one-dimensional Brownian motion. In *Séminaire de Probabilités, XXIII. Lecture Notes in Math.* **1372** 239–247. Berlin: Springer. MR1022914 https://doi.org/10.1007/BFb0083976

[18] Neveu, J. and Pitman, J.W. (1989). The branching process in a Brownian excursion. In *Séminaire de Probabilités, XXIII. Lecture Notes in Math.* **1372** 248–257. Berlin: Springer. MR1022915 https://doi.org/10.1007/BFb0083977

[19] Newman, W.I., Turcotte, D.L. and Gabrielov, A.M. (1997). Fractal trees with side branching. *Fractals* **5** 603–614.

[20] Peckham, S.D. (1995). New results for self-similar trees with applications to river networks. *Water Resour. Res.* **31** 1023–1029.

[21] Strahler, A.N. (1957). Quantitative analysis of watershed geomorphology. *Trans. – Am. Geophys. Union* **38** 913–920.

Bernoulli **27**(3), 2021, 1824–1850
https://doi.org/10.3150/20-BEJ1293

The coupling method in extreme value theory

BENJAMIN BOBBIA[*], CLÉMENT DOMBRY[†] and DAVIT VARRON[‡]

Université Bourgogne Franche-Comté, Laboratoire de Mathématiques de Besançon, UMR CNRS 6623, 16 route de Gray, 25030 Besançon Cedex, France.
E-mail: [*]benjamin.bobbia@univ-fcomte.fr; [†]clement.dombry@univ-fcomte.fr; [‡]davit.varron@univ-fcomte.fr

A coupling method is developed for univariate extreme value theory, providing an alternative to the use of the tail empirical/quantile processes. We emphasize the Peak-over-Threshold approach that approximates the distribution above high threshold by the Generalized Pareto Distribution (GPD) and compare the empirical distribution of exceedances to the empirical distribution associated with the limit GPD model. Sharp bounds for their Wasserstein distance in the second order Wasserstein space are provided. As an application, we recover standard results on the asymptotic behavior of the Hill estimator, the Weissman extreme quantile estimator or the probability weighted moment estimators, shedding some new light on the theory.

Keywords: Extreme value theory; coupling method; Wasserstein distance

References

[1] Balkema, A.A. and de Haan, L. (1974). Residual life time at great age. *Ann. Probab.* **2** 792–804. MR0359049 https://doi.org/10.1214/aop/1176996548

[2] Beirlant, J., Goegebeur, Y., Teugels, J. and Segers, J. (2004). *Statistics of Extremes: Theory and Applications*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. With contributions from Daniel De Waal and Chris Ferro. MR2108013 https://doi.org/10.1002/0470012382

[3] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217. MR0630103

[4] Bobbia, B., Dombry, C. and Varron, D. (2021). Supplement to "The coupling method in extreme value theory." https://doi.org/10.3150/20-BEJ1293SUPP.

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

[6] Bücher, A. and Segers, J. (2018). Maximum likelihood estimation for the Fréchet distribution based on block maxima extracted from a time series. *Bernoulli* **24** 1427–1462. MR3706798 https://doi.org/10.3150/16-BEJ903

[7] de Haan, L. (1971). A form of regular variation and its application to the domain of attraction of the double exponential distribution. *Z. Wahrsch. Verw. Gebiete* **17** 241–258. MR0283854 https://doi.org/10.1007/BF00536760

[8] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3

[9] de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Stat. Neerl.* **52** 60–70. MR1615558 https://doi.org/10.1111/1467-9574.00068

[10] de Valk, C. and Segers, J. (2020). Tails of optimal transport plans for regularly varying probability measures. Available at arXiv:1811.12061.

[11] Dombry, C. and Ferreira, A. (2019). Maximum likelihood estimators based on the block maxima method. *Bernoulli* **25** 1690–1723. MR3961227 https://doi.org/10.3150/18-BEJ1032

[12] Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Stat.* **25** 187–210. MR1614276 https://doi.org/10.1111/1467-9469.00097

[13] Drees, H., de Haan, L. and Li, D. (2003). On large deviation for extremes. *Statist. Probab. Lett.* **64** 51–62. MR1995809 https://doi.org/10.1016/S0167-7152(03)00130-5

1350-7265

[14] Einmahl, J.H.J., de Haan, L. and Zhou, C. (2016). Statistics of heteroscedastic extremes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 31–51. MR3453645 https://doi.org/10.1111/rssb.12099
[15] Ferreira, A. and de Haan, L. (2015). On the block maxima method in extreme value theory: PWM estimators. *Ann. Statist.* **43** 276–298. MR3285607 https://doi.org/10.1214/14-AOS1280
[16] Fisher, R. and Tipett, L. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proc. Camb. Philos. Soc.* **24** 180–190.
[17] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Ann. of Math.* (2) **44** 423–453. MR0008655 https://doi.org/10.2307/1968974
[18] Haeusler, E. and Teugels, J.L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation. *Ann. Statist.* **13** 743–756. MR0790569 https://doi.org/10.1214/aos/1176349551
[19] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc. Ser. B* **44** 37–42. MR0655370
[20] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3** 1163–1174. MR0378204
[21] Hosking, J. (1986). The theory of probability weighted moments. Research Report RC12210, IBM Thomas J. Watson Research center, Yorktown Heights, NY.
[22] Hosking, J.R.M. and Wallis, J.R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics* **29** 339–349. MR0906643 https://doi.org/10.2307/1269343
[23] Hosking, J.R.M., Wallis, J.R. and Wood, E.F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics* **27** 251–261. MR0797563 https://doi.org/10.2307/1269706
[24] Kantorovič, L.V. and Rubinšteĭn, G.Š. (1958). On a space of completely additive functions. *Vestn. Leningr. Univ.* **13** 52–59. MR0102006
[25] Le Gouic, T. and Loubes, J.-M. (2017). Existence and consistency of Wasserstein barycenters. *Probab. Theory Related Fields* **168** 901–917. MR3663634 https://doi.org/10.1007/s00440-016-0727-z
[26] Lindvall, T. (1992). *Lectures on the Coupling Method. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. MR1180522
[27] Panaretos, V.M. and Zemel, Y. (2019). Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** 405–431. MR3939527 https://doi.org/10.1146/annurev-statistics-030718-104938
[28] Pickands, J. III (1975). Statistical inference using extreme order statistics. *Ann. Statist.* **3** 119–131. MR0423667
[29] Reiss, R.-D. (1989). *Approximate Distributions of Order Statistics. Springer Series in Statistics*. New York: Springer. MR0988164 https://doi.org/10.1007/978-1-4613-9620-8
[30] Reiss, R.-D. (1993). *A Course on Point Processes. Springer Series in Statistics*. New York: Springer. MR1199815 https://doi.org/10.1007/978-1-4613-9308-5
[31] Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293. MR2861132 https://doi.org/10.1214/11-PS182
[32] Thorisson, H. (2000). *Coupling, Stationarity, and Regeneration. Probability and Its Applications (New York)*. New York: Springer. MR1741181 https://doi.org/10.1007/978-1-4612-1236-2
[33] 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
[34] 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
[35] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* **73** 812–815. MR0521329

Bernoulli **27**(3), 2021, 1851–1869
https://doi.org/10.3150/20-BEJ1294

Asymptotics and renewal approximation in the online selection of increasing subsequence

ALEXANDER GNEDIN* and AMIRLAN SEKSENBAYEV†

School of Mathematical Sciences, Queen Mary University of London, United Kingdom.
E-mail: *a.gnedin@qmul.ac.uk; †a.seksenbayev@qmul.ac.uk

We revisit the problem of maximising the expected length of increasing subsequence that can be selected from a marked Poisson process by an online strategy. Resorting to a natural size variable, we represent the problem in terms of a controlled piecewise deterministic Markov process with decreasing paths. We apply a comparison method to the optimality equation to obtain fairly complete asymptotic expansions for the moments of the maximal length, and, with the aid of a renewal approximation, give a novel proof to the central limit theorem for the length of selected subsequence under either the optimal strategy or a strategy sufficiently close to optimality.

Keywords: Online selection; monotone subsequence; renewal approximation; dynamic programming

References

[1] Alsmeyer, G. and Marynych, A. (2016). Renewal approximation for the absorption time of a decreasing Markov chain. *J. Appl. Probab.* **53** 765–782. MR3570093 https://doi.org/10.1017/jpr.2016.39
[2] Arlotto, A., Mossel, E. and Steele, J.M. (2016). Quickest online selection of an increasing subsequence of specified size. *Random Structures Algorithms* **49** 235–252. MR3536538 https://doi.org/10.1002/rsa.20634
[3] Arlotto, A., Nguyen, V.V. and Steele, J.M. (2015). Optimal online selection of a monotone subsequence: A central limit theorem. *Stochastic Process. Appl.* **125** 3596–3622. MR3357621 https://doi.org/10.1016/j.spa.2015.03.009
[4] Arlotto, A., Wei, Y. and Xie, X. (2018). An adaptive $O(\log n)$-optimal policy for the online selection of a monotone subsequence from a random sample. *Random Structures Algorithms* **52** 41–53. MR3731611 https://doi.org/10.1002/rsa.20728
[5] Baik, J., Deift, P. and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178. MR1682248 https://doi.org/10.1090/S0894-0347-99-00307-0
[6] Baryshnikov, Y.M. and Gnedin, A.V. (2000). Sequential selection of an increasing sequence from a multidimensional random sample. *Ann. Appl. Probab.* **10** 258–267. MR1765211 https://doi.org/10.1214/aoap/1019737672
[7] Bruss, F.T. and Delbaen, F. (2001). Optimal rules for the sequential selection of monotone subsequences of maximum expected length. *Stochastic Process. Appl.* **96** 313–342. MR1865761 https://doi.org/10.1016/S0304-4149(01)00122-3
[8] Bruss, F.T. and Delbaen, F. (2004). A central limit theorem for the optimal selection process for monotone subsequences of maximum expected length. *Stochastic Process. Appl.* **114** 287–311. MR2101246 https://doi.org/10.1016/j.spa.2004.09.002
[9] Bruss, F.T. and Robertson, J.B. (1991). "Wald's lemma" for sums of order statistics of i.i.d. random variables. *Adv. in Appl. Probab.* **23** 612–623. MR1122878 https://doi.org/10.2307/1427625
[10] Bunge, J. and Goldie, C.M. (2001). Record sequences and their applications. In *Stochastic Processes: Theory and Methods. Handbook of Statist.* **19** 277–308. Amsterdam: North-Holland. MR1861727 https://doi.org/10.1016/S0169-7161(01)19012-7
[11] Coffman, E.G. Jr., Flatto, L. and Weber, R.R. (1987). Optimal selection of stochastic intervals under a sum constraint. *Adv. in Appl. Probab.* **19** 454–473. MR0889945 https://doi.org/10.2307/1427427

[12] de Bruijn, N.G. (1981). *Asymptotic Methods in Analysis*, 3rd ed. New York: Dover. MR0671583
[13] Gnedin, A. and Seksenbayev, A. (2020). Diffusion approximations in the online increasing subsequence problem. Pre-print arXiv:2001.02249.
[14] Gnedin, A.V. (1999). Sequential selection of an increasing subsequence from a sample of random size. *J. Appl. Probab.* **36** 1074–1085. MR1742151 https://doi.org/10.1239/jap/1032374756
[15] Gnedin, A.V. (2000). A note on sequential selection from permutations. *Combin. Probab. Comput.* **9** 13–17. MR1751299 https://doi.org/10.1017/S0963548399004149
[16] Gnedin, A.V. (2000). Sequential selection of an increasing subsequence from a random sample with geometrically distributed sample-size. In *Game Theory, Optimal Stopping, Probability and Statistics. Institute of Mathematical Statistics Lecture Notes – Monograph Series* **35** 101–109. Beachwood, OH: IMS. MR1833854 https://doi.org/10.1214/lnms/1215089747
[17] Korshunov, D. (2008). The key renewal theorem for a transient Markov chain. *J. Theoret. Probab.* **21** 234–245. MR2384480 https://doi.org/10.1007/s10959-007-0132-8
[18] Molchanov, I. (2005). *Theory of Random Sets. Probability and Its Applications (New York)*. London: Springer London, Ltd. MR2132405
[19] Nadarajah, S. and Kotz, S. (2008). Exact distribution of max/min for two Gaussian random variables. *IEEE Trans. Very Large Scale Integr.* **16** 202–210.
[20] Peng, P. and Steele, J.M. (2016). Sequential selection of a monotone subsequence from a random permutation. *Proc. Amer. Math. Soc.* **144** 4973–4982. MR3544544 https://doi.org/10.1090/proc/13104
[21] Rhee, W. and Talagrand, M. (1991). A note on the selection of random variables under a sum constraint. *J. Appl. Probab.* **28** 919–923. MR1133802 https://doi.org/10.2307/3214697
[22] Romik, D. (2015). *The Surprising Mathematics of Longest Increasing Subsequences. Institute of Mathematical Statistics Textbooks* **4**. New York: Cambridge Univ. Press. MR3468738
[23] Samuels, S.M. and Steele, J.M. (1981). Optimal sequential selection of a monotone sequence from a random sample. *Ann. Probab.* **9** 937–947. MR0632967
[24] Seksenbayev, A. (2019). Asymptotic expansions and strategies in the online increasing subsequence problem. Pre-print.
[25] Van Cutsem, B. and Ycart, B. (1994). Renewal-type behavior of absorption times in Markov chains. *Adv. in Appl. Probab.* **26** 988–1005. MR1303873 https://doi.org/10.2307/1427901

Bernoulli **27**(3), 2021, 1870–1898
https://doi.org/10.3150/20-BEJ1295

A convolution formula for the local time of an Itô diffusion reflecting at 0 and a generalized Stroock–Williams equation

JACEK JAKUBOWSKI* and MACIEJ WIŚNIEWOLSKI†

Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland.
E-mail: *jakub@mimuw.edu.pl; †wisniewolski@mimuw.edu.pl

A new probabilistic insight into the structure of local time is presented. A convolution formula for the local time at 0 of Itô diffusions reflecting at 0 is obtained. A simple integro-differential equation for the cumulative distribution function of the local time is given. Finally, a probabilistic representation of a generalized Stroock–Williams equation is presented.

Keywords: Itô diffusion; local time; excursions of Markov processes; Stroock–Williams equation

References

[1] Blumenthal, R.M. (1992). *Excursions of Markov Processes*. *Probability and Its Applications*. Boston, MA: Birkhäuser, Inc. MR1138461 https://doi.org/10.1007/978-1-4684-9412-9

[2] Borodin, A.N. and Salminen, P. (2002). *Handbook of Brownian Motion – Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Basel: Birkhäuser. MR1912205 https://doi.org/10.1007/978-3-0348-8163-0

[3] Cherny, A.S. and Engelbert, H.-J. (2005). *Singular Stochastic Differential Equations*. *Lecture Notes in Math*. **1858**. Berlin: Springer. MR2112227 https://doi.org/10.1007/b104187

[4] Engelbert, H.-J. and Peskir, G. (2014). Stochastic differential equations for sticky Brownian motion. *Stochastics* **86** 993–1021. MR3271518 https://doi.org/10.1080/17442508.2014.899600

[5] Fitzsimmons, P.J. and Pitman, J. (1999). Kac's moment formula and the Feynman–Kac formula for additive functionals of a Markov process. *Stochastic Process*. *Appl*. **79** 117–134. MR1670526 https://doi.org/10.1016/S0304-4149(98)00081-7

[6] Forman, J.L. and Sørensen, M. (2008). The Pearson diffusions: A class of statistically tractable diffusion processes. *Scand*. *J*. *Stat*. **35** 438–465. MR2446729 https://doi.org/10.1111/j.1467-9469.2007.00592.x

[7] Friedman, A. (1975). *Stochastic Differential Equations and Applications*. *Vol*. 1. *Probability and Mathematical Statistics* **28**. New York: Academic Press. MR0494490

[8] Getoor, R.K. (1979). Excursions of a Markov process. *Ann*. *Probab*. **7** 244–266. MR0525052

[9] Getoor, R.K. and Sharpe, M.J. (1979). Excursions of Brownian motion and Bessel processes. *Z*. *Wahrsch*. *Verw*. *Gebiete* **47** 83–106. MR0521534 https://doi.org/10.1007/BF00533253

[10] Getoor, R.K. and Sharpe, M.J. (1982). Excursions of dual processes. *Adv*. *Math*. **45** 259–309. MR0673804 https://doi.org/10.1016/S0001-8708(82)80006-6

[11] 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

[12] Integral calculator. Available at http://www.integral-calculator.com.

[13] Itô, K. (1972). Poisson point processes attached to Markov processes. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ*. *California*, *Berkeley*, *Calif*., 1970/1971), *Vol*. *III*: *Probability Theory* 225–239. MR0402949

[14] Itô, K. and McKean, H.P. Jr. (1995). *Diffusion Processes and Their Sample Paths*. *Die Grundlehren der Mathematischen Wissenschaften* **125**. Berlin: Springer. MR0345224

1350-7265

[15] Janson, S. and Tysk, J. (2006). Feynman–Kac formulas for Black–Scholes-type operators. *Bull. Lond. Math. Soc.* **38** 269–282. MR2214479 https://doi.org/10.1112/S0024609306018194
[16] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940 https://doi.org/10.1007/978-1-4612-0949-2
[17] Lew, J.S. (1972). On linear Volterra integral equations of convolution type. *Proc. Amer. Math. Soc.* **35** 450–456. MR0308699 https://doi.org/10.2307/2037627
[18] Maisonneuve, B. (1975). Exit systems. *Ann. Probab.* **3** 399–411. MR0400417 https://doi.org/10.1214/aop/1176996348
[19] Meier, C., Li, L. and Zhang, G. (2019). Markov chain approximation of one-dimensional sticky diffusions. Preprint.
[20] Mikusiński, J. (1959). *Operational Calculus*. *International Series of Monographs on Pure and Applied Mathematics* **8**. New York: Pergamon Press. MR0105594
[21] Molčanov, S.A. and Ostrovskiĭ, E. (1969). Symmetric stable processes as traces of degenerate diffusion processes. *Teor. Veroyatn. Primen.* **14** 127–130. MR0247668
[22] Naĭmark, M.A. (1959). *Normed Rings*. Groningen: P. Noordhoff N. V. Translated from the First Russian Edition by Leo F. Boron. MR0110956
[23] NIST Digital Library of Mathematical Functions. Available at http://dlmf.nist.gov/, Release 1.0.11 of 2016-06-08.
[24] Pang, H. and Stroock, D.W. (2007). A peculiar two point boundary value problem. *Ann. Probab.* **35** 1623–1641. MR2349570 https://doi.org/10.1214/009117906000000818
[25] Peskir, G. (2014). A probabilistic solution to the Stroock–Williams equation. *Ann. Probab.* **42** 2197–2206. MR3262501 https://doi.org/10.1214/13-AOP865
[26] Peskir, G. (2015). On boundary behaviour of one-dimensional diffusions: From Brown to Feller and beyond. In *Probab. Statist. Group Manchester, William Feller, Selected Papers II* 77–93. Berlin: Springer. Research Report No. 8 (2014).
[27] Revuz, D. and Yor, M. (2005). *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
[28] Rogers, L.C.G. (1989). A guided tour through excursions. *Bull. Lond. Math. Soc.* **21** 305–341. MR0998631 https://doi.org/10.1112/blms/21.4.305
[29] Rogers, L.C.G. and Williams, D. (1987). *Diffusions, Markov Processes, and Martingales. Vol.* 2: *Itô Calculus*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0921238
[30] Salminen, P., Vallois, P. and Yor, M. (2007). On the excursion theory for linear diffusions. *Jpn. J. Math.* **2** 97–127. MR2295612 https://doi.org/10.1007/s11537-007-0662-y
[31] Stroock, D.W. and Williams, D. (2005). A simple PDE and Wiener–Hopf Riccati equations. *Comm. Pure Appl. Math.* **58** 1116–1148. MR2143528 https://doi.org/10.1002/cpa.20081
[32] Stroock, D.W. and Williams, D. (2006). Further study of a simple PDE. *Illinois J. Math.* **50** 961–989. MR2247850
[33] Titchmarsh, E.C. (1926). The Zeros of Certain Integral Functions. *Proc. Lond. Math. Soc.* (2) **25** 283–302. MR1575285 https://doi.org/10.1112/plms/s2-25.1.283
[34] Van Mieghem, P. (2020). The Mittag-Leffler function. Preprint. Available at arXiv:2005.13330v3 [math.FA].
[35] Walsh, J.B. (1978). Excursions and local time. *Astérisque* **52–53** 159–192.
[36] Williams, D. (1974). Path decomposition and continuity of local time for one-dimensional diffusions. I. *Proc. Lond. Math. Soc.* (3) **28** 738–768. MR0350881 https://doi.org/10.1112/plms/s3-28.4.738
[37] Williams, D. and Andrews, S. (2006). Indefinite inner products: A simple illustrative example. *Math. Proc. Cambridge Philos. Soc.* **141** 127–159. MR2238647 https://doi.org/10.1017/S0305004106009364

Bernoulli **27**(3), 2021, 1899–1941
https://doi.org/10.3150/20-BEJ1296

Stochastic PDEs on graphs as scaling limits of discrete interacting systems

WAI-TONG LOUIS FAN[1,2]

[1]*Department of Mathematics, Indiana University, Bloomington, USA.*
E-mail: waifan@iu.edu
[2]*Center of Mathematical Sciences and Applications, Harvard University, Cambridge, USA.*

Stochastic partial differential equations (SPDE) on graphs were recently introduced by Cerrai and Freidlin (*Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 865–899). This class of stochastic equations in infinite dimensions provides a minimal framework for the study of the effective dynamics of much more complex systems. However, how they emerge from microscopic individual-based models is still poorly understood, partly due to complications near vertex singularities. In this work, motivated by the study of the dynamics and the genealogies of expanding populations in spatially structured environments, we obtain a new class of SPDE on graphs of Wright–Fisher type which have nontrivial boundary conditions on the vertex set. We show that these SPDE arise as scaling limits of suitably defined biased voter models (BVM), which extends the scaling limits of Durrett and Fan (*Ann. Appl. Probab.* **26** (2016) 3456–3490). We further obtain a convergent simulation scheme for each of these SPDE in terms of a system of Itô SDEs, which is useful when the size of the BVM is too large for stochastic simulations. These give the first rigorous connection between SPDE on graphs and more discrete models, specifically, interacting particle systems and interacting SDEs. Uniform heat kernel estimates for symmetric random walks approximating diffusions on graphs are the keys to our proofs. Some open problems are provided as further motivations of our study.

Keywords: Stochastic partial differential equation; graph; interacting particle system; numerical scheme; duality; scaling limit; population dynamics

References

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

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

[3] Athreya, S. and Tribe, R. (2000). Uniqueness for a class of one-dimensional stochastic PDEs using moment duality. *Ann. Probab.* **28** 1711–1734. MR1813840 https://doi.org/10.1214/aop/1019160504

[4] Barlow, M.T. and Perkins, E.A. (1988). Brownian motion on the Sierpiński gasket. *Probab. Theory Related Fields* **79** 543–623. MR0966175 https://doi.org/10.1007/BF00318785

[5] Blount, D. (1996). Fourier analysis applied to SPDEs. *Stochastic Process. Appl.* **62** 223–242. MR1397705 https://doi.org/10.1016/0304-4149(96)00056-7

[6] Brunet, E., Derrida, B., Mueller, A. and Munier, S. (2006). Phenomenological theory giving the full statistics of the position of fluctuating pulled fronts. *Physica E* **73** 056126.

[7] Caravenna, F., Sun, R. and Zygouras, N. (2017). Universality in marginally relevant disordered systems. *Ann. Appl. Probab.* **27** 3050–3112. MR3719953 https://doi.org/10.1214/17-AAP1276

[8] Carlen, E.A., Kusuoka, S. and Stroock, D.W. (1987). Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré Probab. Stat.* **23** 245–287. MR0898496

[9] Cerrai, S. and Freidlin, M. (2017). SPDEs on narrow domains and on graphs: An asymptotic approach. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 865–899. MR3634278 https://doi.org/10.1214/16-AIHP740

[10] Cerrai, S. and Freidlin, M. (2019). Fast flow asymptotics for stochastic incompressible viscous fluids in $\mathbb{R}^2$ and SPDEs on graphs. *Probab. Theory Related Fields* **173** 491–535. MR3916113 https://doi.org/10.1007/s00440-018-0839-8
[11] Chen, Z.-Q. and Fan, W.-T. (2017). Hydrodynamic limits and propagation of chaos for interacting random walks in domains. *Ann. Appl. Probab.* **27** 1299–1371. MR3678472 https://doi.org/10.1214/16-AAP1208
[12] Chen, Z.-Q. and Fan, W.-T.L. (2016). Fluctuation limit for interacting diffusions with partial annihilations through membranes. *J. Stat. Phys.* **164** 890–936. MR3529160 https://doi.org/10.1007/s10955-016-1567-z
[13] Chen, Z.-Q. and Fukushima, M. (2012). *Symmetric Markov Processes, Time Change, and Boundary Theory. London Mathematical Society Monographs Series* **35**. Princeton, NJ: Princeton Univ. Press. MR2849840
[14] Codling, E.A., Plank, M.J. and Benhamou, S. (2008). Random walk models in biology. *J. R. Soc. Interface* **5** 813–834.
[15] Conus, D. and Khoshnevisan, D. (2012). Weak nonmild solutions to some SPDEs. *Illinois J. Math.* **54** 1329–1341. MR2981850
[16] Cox, J.T., Durrett, R. and Perkins, E.A. (2000). Rescaled voter models converge to super-Brownian motion. *Ann. Probab.* **28** 185–234. MR1756003 https://doi.org/10.1214/aop/1019160117
[17] Crisan, D., Janjigian, C. and Kurtz, T.G. (2018). Particle representations for stochastic partial differential equations with boundary conditions. *Electron. J. Probab.* **23** Paper No. 65, 29. MR3835471 https://doi.org/10.1214/18-EJP186
[18] Croydon, D. (2008). Convergence of simple random walks on random discrete trees to Brownian motion on the continuum random tree. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 987–1019. MR2469332 https://doi.org/10.1214/07-AIHP153
[19] Croydon, D.A. (2008). Volume growth and heat kernel estimates for the continuum random tree. *Probab. Theory Related Fields* **140** 207–238. MR2357676 https://doi.org/10.1007/s00440-007-0063-4
[20] Croydon, D.A. (2010). Scaling limits for simple random walks on random ordered graph trees. *Adv. in Appl. Probab.* **42** 528–558. MR2675115 https://doi.org/10.1239/aap/1275055241
[21] Croydon, D.A. and Hambly, B.M. (2008). Local limit theorems for sequences of simple random walks on graphs. *Potential Anal.* **29** 351–389. MR2453564 https://doi.org/10.1007/s11118-008-9101-9
[22] 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
[23] Dalang, R., Khoshnevisan, D., Mueller, C., Nualart, D. and Xiao, Y. (2009). *A Minicourse on Stochastic Partial Differential Equations. Lecture Notes in Math.* **1962**. Berlin: Springer. MR1500166
[24] Dalang, R.C. and Quer-Sardanyons, L. (2011). Stochastic integrals for SPDE's: A comparison. *Expo. Math.* **29** 67–109. MR2785545 https://doi.org/10.1016/j.exmath.2010.09.005
[25] Delmotte, T. (1999). Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoam.* **15** 181–232. MR1681641 https://doi.org/10.4171/RMI/254
[26] Doering, C.R., Mueller, C. and Smereka, P. (2003). Interacting particles, the stochastic Fisher–Kolmogorov–Petrovsky–Piscounov equation, and duality. *Phys. A* **325** 243–259. MR2014157 https://doi.org/10.1016/S0378-4371(03)00203-6
[27] Donnelly, P. and Kurtz, T.G. (1999). Genealogical processes for Fleming–Viot models with selection and recombination. *Ann. Appl. Probab.* **9** 1091–1148. MR1728556 https://doi.org/10.1214/aoap/1029962866
[28] Durrett, R. (2009). Special invited paper: Coexistence in stochastic spatial models. *Ann. Appl. Probab.* **19** 477–496. MR2521876 https://doi.org/10.1214/08-AAP590
[29] Durrett, R. and Fan, W.-T. (2016). Genealogies in expanding populations. *Ann. Appl. Probab.* **26** 3456–3490. MR3582808 https://doi.org/10.1214/16-AAP1181
[30] Durrett, R. and Levin, S. (1994). The importance of being discrete (and spatial). *Theor. Popul. Biol.* **46** 363–394.
[31] Durrett, R. and Levin, S. (1994). Stochastic spatial models: A user's guide to ecological applications. *Philos. Trans. R. Soc. Lond. B* **343** 329–350.
[32] Durrett, R., Mytnik, L. and Perkins, E. (2005). Competing super-Brownian motions as limits of interacting particle systems. *Electron. J. Probab.* **10** 1147–1220. MR2164042 https://doi.org/10.1214/EJP.v10-229
[33] Edmonds, C.A., Lillie, A.S. and Cavalli-Sforza, L.L. (2004). Mutations arising in the wave front of an expanding population. *Proc. Natl. Acad. Sci. USA* **101** 975–979.

[34] Endler, E.E., Duca, K.A., Nealey, P.F., Whitesides, G.M. and Yin, J. (2003). Propagation of viruses on micropatterned host cells. *Biotechnol*. *Bioeng*. **81** 719–725.
[35] Etheridge, A.M. and Kurtz, T.G. (2019). Genealogical constructions of population models. *Ann*. *Probab*. **47** 1827–1910. MR3980910 https://doi.org/10.1214/18-AOP1266
[36] Ethier, S.N. and Kurtz, T.G. (1986). *Markov Processes*: *Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0838085 https://doi.org/10.1002/9780470316658
[37] Evans, L.C. (2010). *Partial Differential Equations*, 2nd ed. *Graduate Studies in Mathematics* **19**. Providence, RI: Amer. Math. Soc. MR2597943 https://doi.org/10.1090/gsm/019
[38] Evans, S.N. and Perkins, E.A. (1998). Collision local times, historical stochastic calculus, and competing superprocesses. *Electron*. *J*. *Probab*. **3** No. 5, 120. MR1615329 https://doi.org/10.1214/EJP.v3-27
[39] Fan, W.-T. (2016). Discrete approximations to local times for reflected diffusions. *Electron*. *Commun*. *Probab*. **21** Paper No. 16, 12. MR3485385 https://doi.org/10.1214/16-ECP4694
[40] Fan, W.-T.L., Hu, W. and Terlov, G. (2019). Wave propagation for reaction–diffusion equations on infinite random trees. Preprint. Available at arXiv:1907.12962.
[41] Freidlin, M. and Sheu, S.-J. (2000). Diffusion processes on graphs: Stochastic differential equations, large deviation principle. *Probab*. *Theory Related Fields* **116** 181–220. MR1743769 https://doi.org/10.1007/PL00008726
[42] Freidlin, M.I. and Wentzell, A.D. (1993). Diffusion processes on graphs and the averaging principle. *Ann*. *Probab*. **21** 2215–2245. MR1245308
[43] Fukushima, M., Oshima, Y. and Takeda, M. (2011). *Dirichlet Forms and Symmetric Markov Processes*, extended ed. *De Gruyter Studies in Mathematics* **19**. Berlin: de Gruyter. MR2778606
[44] Gillespie, D.T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J*. *Comput*. *Phys*. **22** 403–434. MR0503370 https://doi.org/10.1016/0021-9991(76)90041-3
[45] Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *J*. *Phys*. *Chem*. **81** 2340–2361.
[46] Gonçalves, P., Jara, M. and Sethuraman, S. (2015). A stochastic Burgers equation from a class of microscopic interactions. *Ann*. *Probab*. **43** 286–338. MR3298474 https://doi.org/10.1214/13-AOP878
[47] Grigor'yan, A. and Telcs, A. (2012). Two-sided estimates of heat kernels on metric measure spaces. *Ann*. *Probab*. **40** 1212–1284. MR2962091 https://doi.org/10.1214/11-AOP645
[48] Grimm, V. and Railsback, S.F. (2005). *Individual-Based Modeling and Ecology*. *Princeton Series in Theoretical and Computational Biology*. Princeton, NJ: Princeton Univ. Press. MR2153370 https://doi.org/10.1515/9781400850624
[49] Gubinelli, M., Imkeller, P. and Perkowski, N. (2015). Paracontrolled distributions and singular PDEs. *Forum Math*. *Pi* **3** e6, 75. MR3406823 https://doi.org/10.1017/fmp.2015.2
[50] Gyöngy, I. (2014). On stochastic finite difference schemes. *Stoch*. *Partial Differ*. *Equ*. *Anal*. *Comput*. **2** 539–583. MR3274891 https://doi.org/10.1007/s40072-014-0039-1
[51] Gyrya, P. and Saloff-Coste, L. (2011). Neumann and Dirichlet heat kernels in inner uniform domains. *Astérisque* **336** viii+144. MR2807275
[52] Hairer, M. (2013). Solving the KPZ equation. *Ann*. *of Math*. (2) **178** 559–664. MR3071506 https://doi.org/10.4007/annals.2013.178.2.4
[53] Hairer, M. (2014). A theory of regularity structures. *Invent*. *Math*. **198** 269–504. MR3274562 https://doi.org/10.1007/s00222-014-0505-4
[54] Hallatschek, O., Hersen, P., Ramanathan, S. and Nelson, D.R. (2007). Genetic drift at expanding frontiers promotes gene segregation. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **104** 19926–19930.
[55] Hallatschek, O. and Nelson, D.R. (2008). Gene surfing in expanding populations. *Theor*. *Popul*. *Biol*. **73** 158–170.
[56] Hallatschek, O. and Nelson, D.R. (2010). Life at the front of an expanding population. *Evolution* **64** 193–206.
[57] Inankur, B. (2015). Development of tools to study how virus-host cell interactions influence infection spread. PhD dissertation, Univ. Wisconsin-Madison.
[58] Kant, U., Klauss, T., Voigt, J. and Weber, M. (2009). Dirichlet forms for singular one-dimensional operators and on graphs. *J*. *Evol*. *Equ*. **9** 637–659. MR2563669 https://doi.org/10.1007/s00028-009-0027-5

[59] Kliem, S. (2011). Convergence of rescaled competing species processes to a class of SPDEs. *Electron. J. Probab.* **16** 618–657. MR2786644 https://doi.org/10.1214/EJP.v16-870

[60] Klopfstein, S., Currat, M. and Excoffier, L. (2005). The fate of mutations surfing on the wave of a range expansion. *Mol. Biol. Evol.* **23** 482–490.

[61] Korolev, K.S., Müller, M.J., Karahan, N., Murray, A.W., Hallatschek, O. and Nelson, D.R. (2012). Selective sweeps in growing microbial colonies. *Phys. Biol.* **9** 026008.

[62] Kotelenez, P.M. and Kurtz, T.G. (2010). Macroscopic limits for stochastic partial differential equations of McKean–Vlasov type. *Probab. Theory Related Fields* **146** 189–222. MR2550362 https://doi.org/10.1007/s00440-008-0188-0

[63] Kumagai, T. (2004). Heat kernel estimates and parabolic Harnack inequalities on graphs and resistance forms. *Publ. Res. Inst. Math. Sci.* **40** 793–818. MR2074701

[64] 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

[65] Lanchier, N. (2013). Stochastic spatial model of producer-consumer systems on the lattice. *Adv. in Appl. Probab.* **45** 1157–1181. MR3161301 https://doi.org/10.1239/aap/1386857862

[66] Lanchier, N. and Neuhauser, C. (2006). Stochastic spatial models of host-pathogen and host-mutualist interactions. I. *Ann. Appl. Probab.* **16** 448–474. MR2209349 https://doi.org/10.1214/105051605000000782

[67] Lanchier, N. and Neuhauser, C. (2010). Stochastic spatial models of host-pathogen and host-mutualist interactions. II. *Stoch. Models* **26** 399–430. MR2739344 https://doi.org/10.1080/15326349.2010.498317

[68] Le Gall, J.-F. (1993). A class of path-valued Markov processes and its applications to superprocesses. *Probab. Theory Related Fields* **95** 25–46. MR1207305 https://doi.org/10.1007/BF01197336

[69] Lee, Y. and Yin, J. (1996). Detection of evolving viruses. *Nat. Biotechnol.* **14** 491–493.

[70] Lehe, R., Hallatschek, O. and Peliti, L. (2012). The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Comput. Biol.* **8** e1002447, 13. MR2921273 https://doi.org/10.1371/journal.pcbi.1002447

[71] Lejay, A. (2003). Simulating a diffusion on a graph. Application to reservoir engineering. *Monte Carlo Methods Appl.* **9** 241–255. MR2009371 https://doi.org/10.1163/156939603322729003

[72] Ma, Z.M. and Röckner, M. (1992). *Introduction to the Theory of* (*Nonsymmetric*) *Dirichlet Forms*. *Universitext*. Berlin: Springer. MR1214375 https://doi.org/10.1007/978-3-642-77739-4

[73] 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

[74] Mueller, C., Mytnik, L. and Quastel, J. (2011). Effect of noise on front propagation in reaction–diffusion equations of KPP type. *Invent. Math.* **184** 405–453. MR2793860 https://doi.org/10.1007/s00222-010-0292-5

[75] Mueller, C., Mytnik, L. and Ryzhik, L. (2019). The speed of a random front for stochastic reaction–diffusion equations with strong noise. Preprint. Available at arXiv:1903.03645.

[76] Mueller, C. and Tribe, R. (1994). A phase transition for a stochastic PDE related to the contact process. *Probab. Theory Related Fields* **100** 131–156. MR1296425 https://doi.org/10.1007/BF01199262

[77] Müller, C. and Tribe, R. (1995). Stochastic p.d.e.'s arising from the long range contact and long range voter processes. *Probab. Theory Related Fields* **102** 519–545. MR1346264 https://doi.org/10.1007/BF01198848

[78] Nullmeier, J. and Hallatschek, O. (2013). The coalescent in boundary-limited range expansions. *Evolution* **67** 1307–1320.

[79] Ramirez, J.M. (2011). Multi-skewed Brownian motion and diffusion in layered media. *Proc. Amer. Math. Soc.* **139** 3739–3752. MR2813404 https://doi.org/10.1090/S0002-9939-2011-10766-4

[80] Riley, S., Eames, K., Isham, V., Mollison, D. and Trapman, P. (2015). Five challenges for spatial epidemic models. *Epidemics* **10** 68–71.

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

[82] Saloff-Coste, L. (1995). Parabolic Harnack inequality for divergence-form second-order differential operators. *Potential Anal.* **4** 429–467. MR1354894 https://doi.org/10.1007/BF01053457

[83] Seifert, C. and Voigt, J. (2011). Dirichlet forms for singular diffusion on graphs. *Oper. Matrices* **5** 723–734. MR2906858 https://doi.org/10.7153/oam-05-51

[84] Shiga, T. (1988). Stepping stone models in population genetics and population dynamics. In *Stochastic Processes in Physics and Engineering* (*Bielefeld*, 1986). *Math*. *Appl*. **42** 345–355. Dordrecht: Reidel. MR0948717

[85] Shiga, T. (1994). Two contrasting properties of solutions for one-dimensional stochastic partial differential equations. *Canad*. *J*. *Math*. **46** 415–437. MR1271224 https://doi.org/10.4153/CJM-1994-022-8

[86] Stroock, D.W. (1988). Diffusion semigroups corresponding to uniformly elliptic divergence form operators. In *Séminaire de Probabilités*, *XXII*. *Lecture Notes in Math*. **1321** 316–347. Berlin: Springer. MR0960535 https://doi.org/10.1007/BFb0084145

[87] Sturm, A. (2003). On convergence of population processes in random environments to the stochastic heat equation with colored noise. *Electron*. *J*. *Probab*. **8** no. 6, 39. MR1986838 https://doi.org/10.1214/EJP.v8-129

[88] Su, X., Theberge, A.B., January, C.T. and Beebe, D.J. (2013). Effect of microculture on cell metabolism and biochemistry: Do cells get stressed in microchannels? *Anal*. *Chem*. **85** 1562–1570.

[89] Tindel, S. and Viens, F. (1999). On space-time regularity for the stochastic heat equation on Lie groups. *J*. *Funct*. *Anal*. **169** 559–603. MR1730556 https://doi.org/10.1006/jfan.1999.3486

[90] Tindel, S. and Viens, F. (2002). Almost sure exponential behaviour for a parabolic SPDE on a manifold. *Stochastic Process*. *Appl*. **100** 53–74. MR1919608 https://doi.org/10.1016/S0304-4149(02)00102-3

[91] Wallace, M., Feres, R. and Yablonsky, G. (2017). Reaction–diffusion on metric graphs: From 3D to 1D. *Comput*. *Math*. *Appl*. **73** 2035–2052. MR3634968 https://doi.org/10.1016/j.camwa.2017.02.033

[92] 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 **27**(3), 2021, 1942–1992
https://doi.org/10.3150/20-BEJ1297

Is there an analog of Nesterov acceleration for gradient-based MCMC?

YI-AN MA[1], NILADRI S. CHATTERJI[2], XIANG CHENG[3], NICOLAS FLAMMARION[4], PETER L. BARTLETT[3,5,*] and MICHAEL I. JORDAN[3,5,†]

[1]*Halıcıoğlu Data Science Institute, University of California San Diego.*
E-mail: yianma@ucsd.edu
[2]*Department of Physics, University of California Berkeley.*
E-mail: chatterji@berkeley.edu
[3]*Department of Electrical Engineering and Computer Sciences, University of California Berkeley.*
E-mail: x.cheng@berkeley.edu
[4]*School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne.*
E-mail: nicolas.flammarion@epfl.ch
[5]*Department of Statistics, University of California Berkeley.*
E-mail: [*]peter@berkeley.edu; [†]jordan@cs.berkeley.edu

We formulate gradient-based Markov chain Monte Carlo (MCMC) sampling as optimization on the space of probability measures, with Kullback–Leibler (KL) divergence as the objective functional. We show that an underdamped form of the Langevin algorithm performs accelerated gradient descent in this metric. To characterize the convergence of the algorithm, we construct a Lyapunov functional and exploit hypocoercivity of the underdamped Langevin algorithm. As an application, we show that accelerated rates can be obtained for a class of nonconvex functions with the Langevin algorithm.

Keywords: Markov chain Monte Carlo; Langevin Monte Carlo; accelerated gradient descent; sampling algorithms

References

[1] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. *Lectures in Mathematics ETH Zürich*. Basel: Birkhäuser. MR2401600
[2] Bernton, E. (2018). Langevin Monte Carlo and JKO splitting. In *Proceedings of the* 31*st Conference on Learning Theory* (*COLT*) 1777–1798.
[3] Bierkens, J., Fearnhead, P. and Roberts, G. (2019). The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *Ann. Statist.* **47** 1288–1320. MR3911113 https://doi.org/10.1214/18-AOS1715
[4] Bou-Rabee, N., Eberle, A. and Zimmer, R. (2020). Coupling and convergence for Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **30** 1209–1250. MR4133372 https://doi.org/10.1214/19-AAP1528
[5] Bouchard-Côté, A., Vollmer, S.J. and Doucet, A. (2018). The bouncy particle sampler: A nonreversible rejection-free Markov chain Monte Carlo method. *J. Amer. Statist. Assoc.* **113** 855–867. MR3832232 https://doi.org/10.1080/01621459.2017.1294075
[6] Calogero, S. (2012). Exponential convergence to equilibrium for kinetic Fokker–Planck equations. *Comm. Partial Differential Equations* **37** 1357–1390. MR2957543 https://doi.org/10.1080/03605302.2011.648039
[7] Carmon, Y., Hinder, O., Duchi, J.C. and Sidford, A. (2017). "Convex until proven guilty": Dimension-free acceleration of gradient descent on non-convex functions. In *Proceedings of the* 34*th International Conference on Machine Learning* (*ICML*) 654–663.
[8] Chatterji, N., Flammarion, N., Ma, Y.-A., Bartlett, P. and Jordan, M. (2018). On the theory of variance reduction for stochastic gradient Monte Carlo. In *Proceedings of the* 35*th International Conference on Machine Learning* **80** 764–773.

1350-7265

[9] Cheng, X. and Bartlett, P.L. (2018). Convergence of Langevin MCMC in KL-divergence. In *Proceedings of the* 29*th International Conference on Algorithmic Learning Theory* (*ALT*). 186–211. MR3857306

[10] Cheng, X., Chatterji, N.S., Abbasi-Yadkori, Y., Bartlett, P.L. and Jordan, M.I. (2018). Sharp convergence rates for Langevin dynamics in the nonconvex setting. Preprint. Available at arXiv:1805.01648.

[11] Cheng, X., Chatterji, N.S., Bartlett, P.L. and Jordan, M.I. (2018). Underdamped Langevin MCMC: A non-asymptotic analysis. In *Proceedings of the* 31*st Conference on Learning Theory* (*COLT*) 300–323.

[12] Dalalyan, A.S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183

[13] Dalalyan, A.S. and Karagulyan, A. (2019). User-friendly guarantees for the Langevin Monte Carlo with inaccurate gradient. *Stochastic Process. Appl.* **129** 5278–5311. MR4025705 https://doi.org/10.1016/j.spa.2019.02.016

[14] Dalalyan, A.S. and Riou-Durand, L. (2020). On sampling from a log-concave density using kinetic Langevin diffusions. *Bernoulli* **26** 1956–1988. MR4091098 https://doi.org/10.3150/19-BEJ1178

[15] Duong, M.H., Peletier, M.A. and Zimmer, J. (2014). Conservative-dissipative approximation schemes for a generalized Kramers equation. *Math. Methods Appl. Sci.* **37** 2517–2540. MR3271101 https://doi.org/10.1002/mma.2994

[16] Durmus, A., Majewski, S. and Miasojedow, B. (2019). Analysis of Langevin Monte Carlo via convex optimization. *J. Mach. Learn. Res.* **20** Paper No. 73, 46. MR3960927

[17] Durmus, A. and Moulines, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479 https://doi.org/10.1214/16-AAP1238

[18] Durmus, A. and Moulines, É. (2019). High-dimensional Bayesian inference via the unadjusted Langevin algorithm. *Bernoulli* **25** 2854–2882. MR4003567 https://doi.org/10.3150/18-BEJ1073

[19] Dwivedi, R., Chen, Y., Wainwright, M.J. and Yu, B. (2019). Log-concave sampling: Metropolis–Hastings algorithms are fast. *J. Mach. Learn. Res.* **20** Paper No. 183, 42. MR4048994

[20] Erdogdu, M.A., Mackey, L. and Shamir, O. (2018). Global non-convex optimization with discretized diffusions. In *Advances in Neural Information Processing Systems* **31** 9671–9680.

[21] Gao, X., Gurbuzbalaban, M. and Zhu, L. (2019). Breaking reversibility accelerates Langevin dynamics for global non-convex optimization. Preprint. Available at arXiv:1812.07725.

[22] Girolami, M. and Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 123–214. MR2814492 https://doi.org/10.1111/j.1467-9868.2010.00765.x

[23] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. MR0420249 https://doi.org/10.2307/2373688

[24] Huang, C. (2000). A variational principle for the Kramers equation with unbounded external forces. *J. Math. Anal. Appl.* **250** 333–367. MR1893894 https://doi.org/10.1006/jmaa.2000.7109

[25] Jordan, R., Kinderlehrer, D. and Otto, F. (1998). The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29** 1–17. MR1617171 https://doi.org/10.1137/S0036141096303359

[26] Kaiser, M., Jack, R.L. and Zimmer, J. (2017). Acceleration of convergence to equilibrium in Markov chains by breaking detailed balance. *J. Stat. Phys.* **168** 259–287. MR3667361 https://doi.org/10.1007/s10955-017-1805-z

[27] Langevin, P. (1908). On the theory of Brownian motion (sur la théorie du mouvement brownien). *C. R. Acad. Sci.* (*Paris*) **146** 530–533.

[28] Leimkuhler, B. and Shang, X. (2016). Adaptive thermostats for noisy gradient systems. *SIAM J. Sci. Comput.* **38** A712–A736. MR3465428 https://doi.org/10.1137/15M102318X

[29] Li, X., Wu, Y., Mackey, L. and Erdogdu, M.A. (2019). Stochastic Runge–Kutta accelerates Langevin Monte Carlo and beyond. In *Advances in Neural Information Processing Systems* **32** 7748–7760.

[30] Liu, C., Zhu, J. and Song, Y. (2016). Stochastic gradient geodesic MCMC methods. In *Advances in Neural Information Processing Systems* (*NIPS*) **29** 642–651.

[31] Ma, Y.-A., Chen, T. and Fox, E. (2015). A complete recipe for stochastic gradient MCMC. In *Advances in Neural Information Processing Systems* (*NIPS*) **28** 2899–2907.

[32] Ma, Y.-A., Chen, Y., Jin, C., Flammarion, N. and Jordan, M.I. (2019). Sampling can be faster than optimization. *Proc. Natl. Acad. Sci. USA* **116** 20881–20885. MR4025861 https://doi.org/10.1073/pnas.1820003116

[33] Ma, Y.-A., Fox, E.B., Chen, T. and Wu, L. (2019). Irreversible samplers from jump and continuous Markov processes. *Stat. Comput.* **29** 177–202. MR3905547 https://doi.org/10.1007/s11222-018-9802-x
[34] Mangoubi, O. and Smith, A. (2017). Rapid mixing of Hamiltonian Monte Carlo on strongly log-concave distributions. Preprint. Available at arXiv:1708.07114.
[35] Mangoubi, O. and Vishnoi, N.K. (2018). Dimensionally tight running time bounds for second-order Hamiltonian Monte Carlo. Preprint. Available at arXiv:1802.08898.
[36] Mou, W., Ma, Y.-A., Wainwright, M.J., Bartlett, P.L. and Jordan, M.I. (2019). High-order Langevin diffusion yields an accelerated MCMC algorithm.
[37] Neal, R.M. (2010). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC Handb. Mod. Stat. Methods* 113–162. Boca Raton, FL: CRC Press. MR2858447
[38] Nemirovsky, A.S. and Yudin, D.B. (1983). *Problem Complexity and Method Efficiency in Optimization. A Wiley-Interscience Publication*. New York: Wiley. MR0702836
[39] Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization: A Basic Course. Applied Optimization* **87**. Boston, MA: Kluwer Academic. MR2142598 https://doi.org/10.1007/978-1-4419-8853-9
[40] Nesterov, Y. (1983). A method of solving a convex programming problem with convergence rate $o(1/k^2)$. *Sov. Math., Dokl.* **27** 372–376.
[41] O'Donoghue, B. and Candès, E. (2015). Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.* **15** 715–732. MR3348171 https://doi.org/10.1007/s10208-013-9150-3
[42] Ohzeki, M. and Ichiki, A. (2015). Langevin dynamics neglecting balance condition. *Phys. Rev. E* (3) **92** 012105, 4. MR3504707 https://doi.org/10.1103/PhysRevE.92.012105
[43] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620 https://doi.org/10.1006/jfan.1999.3557
[44] Ottobre, M., Pillai, N.S., Pinski, F.J. and Stuart, A.M. (2016). A function space HMC algorithm with second order Langevin diffusion limit. *Bernoulli* **22** 60–106. MR3449777 https://doi.org/10.3150/14-BEJ621
[45] Peters, H.J.M. and Wakker, P.P. (1986). Convex functions on nonconvex domains. *Econom. Lett.* **22** 251–255. MR0875953 https://doi.org/10.1016/0165-1765(86)90242-9
[46] Poljak, B.T. (1963). Gradient methods for minimizing functionals. *Zh. Vychisl. Mat. Mat. Fiz.* **3** 643–653. MR0158568
[47] Rey-Bellet, L. and Spiliopoulos, K. (2016). Improving the convergence of reversible samplers. *J. Stat. Phys.* **164** 472–494. MR3519206 https://doi.org/10.1007/s10955-016-1565-1
[48] Roberts, G.O. and Stramer, O. (2002). Langevin diffusions and Metropolis–Hastings algorithms. *Methodol. Comput. Appl. Probab.* **4** 337–357. MR2002247 https://doi.org/10.1023/A:1023562417138
[49] Rossky, P.J., Doll, J.D. and Friedman, H.L. (1978). Brownian dynamics as smart Monte Carlo simulation. *J. Chem. Phys.* **69** 4628.
[50] Shi, B., Du, S.S., Su, W.J. and Jordan, M.I. (2019). Acceleration via symplectic discretization of high-resolution differential equations. Preprint. Available at arXiv:1902.03694.
[51] Su, W., Boyd, S. and Candes, E. (2014). A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights. In *Advances in Neural Information Processing Systems (NIPS)* **27** 2510–2518.
[52] Taghvaei, A. and Mehta, P. (2019). Accelerated flow for probability distributions. In *Proceedings of the* 36*th International Conference on Machine Learning (ICML)* 6076–6085.
[53] Takahashi, K. and Ohzeki, M. (2016). Conflict between fastest relaxation of a Markov process and detailed balance condition. *Phys. Rev. E* **93** 012129.
[54] Vempala, S.S. and Wibisono, A. (2019). Rapid convergence of the unadjusted Langevin algorithm: Log-Sobolev suffices. Preprint. Available at arXiv:1903.08568.
[55] Villani, C. (2009). Hypocoercivity. *Mem. Amer. Math. Soc.* **202** iv+141. MR2562709 https://doi.org/10.1090/S0065-9266-09-00567-5
[56] 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
[57] Wang, Y. and Li, W. (2019). Accelerated information gradient flow. Preprint. Available at arXiv:1909.02102.
[58] Wibisono, A. (2018). Sampling as optimization in the space of measures: The Langevin dynamics as a composite optimization problem. In *Proceedings of the* 31*st Conference on Learning Theory (COLT)* 2093–3027.

[59] Wibisono, A., Wilson, A.C. and Jordan, M.I. (2016). A variational perspective on accelerated methods in optimization. *Proc. Natl. Acad. Sci. USA* **113** E7351–E7358. MR3582442 https://doi.org/10.1073/pnas.1614734113

[60] Wilson, A., Recht, B. and Jordan, M.I. (2016). A Lyapunov analysis of momentum methods in optimization. Preprint. Available at arXiv:1611.02635.

[61] Yan, M. (2014). Extension of convex function. *J. Convex Anal.* **21** 965–987. MR3331205

[62] Zhang, J., Mokhtari, A., Sra, S. and Jadbabaie, A. (2018). Direct Runge–Kutta discretization achieves acceleration. In *Advances in Neural Information Processing Systems* (*NeuIPS*) **31** 3900–3909.

Bernoulli **27**(3), 2021, 1993–2010
https://doi.org/10.3150/20-BEJ1298

Scoring interval forecasts: Equal-tailed, shortest, and modal interval

JONAS R. BREHMER[1,2,*] and TILMANN GNEITING[2,3,†]

[1]*Institute for Mathematics, University of Mannheim, B6, 26, 68159 Mannheim, Germany*
[2]*Computational Statistics Group, Heidelberg Institute for Theoretical Studies, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany.*
E-mail: [*]*jonas.brehmer@h-its.org;* [†]*tilmann.gneiting@h-its.org*
[3]*Institute for Stochastics, Karlsruhe Institute of Technology (KIT), Englerstraße 2, 76131 Karlsruhe, Germany*

We consider different types of predictive intervals and ask whether they are elicitable, that is, are unique minimizers of a loss or scoring function in expectation. The equal-tailed interval is elicitable, with a rich class of suitable loss functions, though subject to translation invariance, or positive homogeneity and differentiability, the Winkler interval score becomes a unique choice. The modal interval also is elicitable, with a sole consistent scoring function, up to equivalence. However, the shortest interval fails to be elicitable relative to practically relevant classes of distributions. These results provide guidance in interval forecast evaluation and support recent choices of performance measures in forecast competitions.

Keywords: Elicitability; forecast evaluation; interval forecast; modal interval; predictive performance; scoring function

References

[1] Aitchison, J. and Dunsmore, I.R. (1968). Linear-loss interval estimation of location and scale parameters. *Biometrika* **55** 141–148. MR0225413 https://doi.org/10.1093/biomet/55.1.141
[2] Askanazi, R., Diebold, F.X., Schorfheide, F. and Shin, M. (2018). On the comparison of interval forecasts. *J. Time Series Anal.* **39** 953–965. MR3867133 https://doi.org/10.1111/jtsa.12426
[3] Brehmer, J.R. and Strokorb, K. (2019). Why scoring functions cannot assess tail properties. *Electron. J. Stat.* **13** 4015–4034. MR4015787 https://doi.org/10.1214/19-EJS1622
[4] Casella, G., Hwang, J.T.G. and Robert, C. (1993). A paradox in decision-theoretic interval estimation. *Statist. Sinica* **3** 141–155. MR1219296
[5] Christoffersen, P.F. (1998). Evaluating interval forecasts. *Internat. Econom. Rev.* **39** 841–862. MR1661906 https://doi.org/10.2307/2527341
[6] Clements, M.P. (2004). Evaluating the Bank of England density forecasts of inflation. *Econ. J.* **114** 844–866. https://doi.org/10.1111/j.1468-0297.2004.00246.x
[7] Czado, C., Gneiting, T. and Held, L. (2009). Predictive model assessment for count data. *Biometrics* **65** 1254–1261. MR2756513 https://doi.org/10.1111/j.1541-0420.2009.01191.x
[8] Dawid, A.P. (1986). Probability forecasting. In *Encyclopedia of Statistical Sciences* (S. Kotz, N.L. Johnson and C.B. Read, eds.) **7** 210–218. New York: John Wiley & Sons, Inc. https://doi.org/10.1002/0471667196.ess2064.pub2
[9] Dawid, A.P. and Musio, M. (2014). Theory and applications of proper scoring rules. *Metron* **72** 169–183. MR3233147 https://doi.org/10.1007/s40300-014-0039-y
[10] Dearborn, K. and Frongillo, R. (2020). On the indirect elicitability of the mode and modal interval. *Ann. Inst. Statist. Math.* **72** 1095–1108. MR4137747 https://doi.org/10.1007/s10463-019-00719-1
[11] Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **78** 505–562. MR3506792 https://doi.org/10.1111/rssb.12154

1350-7265

[12] Fissler, T., Frongillo, R., Hlavinová, J. and Rudloff, B. (2020). Forecast evaluation of quantiles, prediction intervals, and other set-valued functionals. Preprint, https://arxiv.org/abs/1910.07912v2.

[13] Fissler, T. and Ziegel, J.F. (2016). Higher order elicitability and Osband's principle. *Ann. Statist.* **44** 1680–1707. MR3519937 https://doi.org/10.1214/16-AOS1439

[14] Fissler, T. and Ziegel, J.F. (2020). Erratum: Higher order elicitability and Osband's principle. Preprint, https://arxiv.org/abs/1901.08826v2.

[15] Frongillo, R. and Kash, I.A. (2019). General truthfulness characterizations via convex analysis. Preprint, https://arxiv.org/abs/1211.3043v4.

[16] Frongillo, R. and Kash, I.A. (2020). Elicitation complexity of statistical properties. Preprint, https://arxiv.org/abs/1506.07212v3.

[17] Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. MR2847988 https://doi.org/10.1198/jasa.2011.r10138

[18] Gneiting, T. (2011). Quantiles as optimal point forecasts. *Int. J. Forecast.* **27** 197–207. https://doi.org/10.1016/j.ijforecast.2009.12.015

[19] Gneiting, T. (2017). When is the mode functional the Bayes classifier? *Stat* **6** 204–206. MR3671158 https://doi.org/10.1002/sta4.148

[20] Gneiting, T., Balabdaoui, F. and Raftery, A.E. (2007). Probabilistic forecasts, calibration and sharpness. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **69** 243–268. MR2325275 https://doi.org/10.1111/j.1467-9868.2007.00587.x

[21] Gneiting, T. and Katzfuss, M. (2014). Probabilistic forecasting. *Annu. Rev. Stat. Appl.* **1** 125–151. https://doi.org/10.1146/annurev-statistics-062713-085831

[22] Gneiting, T. and Raftery, A.E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548 https://doi.org/10.1198/016214506000001437

[23] Heinrich, C. (2014). The mode functional is not elicitable. *Biometrika* **101** 245–251. MR3180670 https://doi.org/10.1093/biomet/ast048

[24] Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A. and Hyndman, R.J. (2016). Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *Int. J. Forecast.* **32** 896–913. https://doi.org/10.1016/j.ijforecast.2016.02.001

[25] Kolassa, S. (2016). Evaluating predictive count data distributions in retail sales forecasting. *Int. J. Forecast.* **32** 788–803. https://doi.org/10.1016/j.ijforecast.2015.12.004

[26] Lambert, N.S., Pennock, D.M. and Shoham, Y. (2008). Eliciting properties of probability distributions. In *Proceedings of the* 9*th ACM Conference on Electronic Commerce*. *EC* '08 129–138. https://doi.org/10.1145/1386790.1386813

[27] Lambert, N.S. and Shoham, Y. (2009). Eliciting truthful answers to multiple-choice questions. In *Proceedings of the* 10*th ACM Conference on Electronic Commerce*. *EC* '09 109–118. https://doi.org/10.1145/1566374.1566391

[28] M Open Forecasting Center (2020). The M5 competition: Competitor's guide. Available at https://mofc.unic.ac.cy/m5-competition/.

[29] Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *Int. J. Forecast.* **36** 54–74. https://doi.org/10.1016/j.ijforecast.2019.04.014

[30] Nolde, N. and Ziegel, J.F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Ann. Appl. Stat.* **11** 1833–1874. MR3743276 https://doi.org/10.1214/17-AOAS1041

[31] Ray, E.L., Wattanachit, N., Niemi, J., Kanji, A.H., House, K., Cramer, E.Y., Bracher, J., Zheng, A., Yamana, T.K., Xiong, X., Woody, S., Wang, Y., Wang, L., Walraven, R.L., Tomar, V., Sherratt, K., Sheldon, D., Reiner, R.C., Prakash, B.A., Osthus, D., Li, M.L., Lee, E.C., Koyluoglu, U., Keskinocak, P., Gu, Y., Gu, Q., George, G.E., España, G., Corsetti, S., Chhatwal, J., Cavany, S., Biegel, H., Ben-Nun, M., Walker, J., Slayton, R., Lopez, V., Biggerstaff, M., Johansson, M.A., Reich, N.G. and COVID-19 Forecast Hub Consortium (2020). Ensemble forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S. Preprint, https://www.medrxiv.org/content/10.1101/2020.08.19.20177493v1.

[32] Steinwart, I., Pasin, C., Williamson, R. and Zhang, S. (2014). Elicitation and identification of properties. *J. Mach. Learn. Res. Workshop Conf. Proc.* **35** 1–45.

[33] Wang, R. and Wei, Y. (2020). Risk functionals with convex level sets. *Math. Finance* **30** 1337–1367. MR4154772 https://doi.org/10.1111/mafi.12270

[34] Winkler, R.L. (1972). A decision-theoretic approach to interval estimation. *J. Amer. Statist. Assoc.* **67** 187–191. MR0312611

Bernoulli **27**(3), 2021, 2011–2039
https://doi.org/10.3150/20-BEJ1299

Variable Length Memory Chains: Characterization of stationary probability measures

PEGGY CÉNAC[1], BRIGITTE CHAUVIN[2,*], CAMILLE NOÛS[3], FRÉDÉRIC PACCAUT[4] and NICOLAS POUYANNE[2,†]

[1]*Institut de Mathématiques de Bourgogne, IMB UMR 5584 CNRS, 9 rue Alain Savary - BP 47870, 21078 DIJON CEDEX, France.*
E-mail: peggy.cenac@u-bourgogne.fr
[2]*Université Paris-Saclay, UVSQ, CNRS UMR 8100, Laboratoire de Mathématiques de Versailles, 78000 Versailles, France.*
*E-mail: *brigitte.chauvin@uvsq.fr; †nicolas.pouyanne@uvsq.fr*
[3]*Cogitamus Laboratory, 78000 Versailles, France.*
E-mail: camille.nous@cogitamus.fr
[4]*Laboratoire Amiénois de Mathématique Fondamentale et Appliquée, CNRS UMR 7352, Université de Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens, France.*
E-mail: frederic.paccaut@u-picardie.fr

Variable Length Memory Chains (VLMC), which are generalizations of finite order Markov chains, are an essential tool to modelize random sequences in many domains, as well as an interesting object in contemporary probability theory. The question of existence of stationary probability measures leads us to introduce a key combinatorial structure for words produced by a VLMC: the *Longest Internal Suffix*. This notion allows us to state a necessary and sufficient condition for a VLMC to admit a unique invariant probability measure.

This condition turns out to get a much simpler form for a subclass of VLMC: the *stable* VLMC. This natural subclass, unlike the general case, enjoys a renewal property. Namely, a stable VLMC induces a semi-Markov chain on an at most countable state space. Unfortunately, this discrete time renewal process does not contain the whole information of the VLMC, preventing the study of a stable VLMC to be reduced to the study of its induced semi-Markov chain. For a subclass of stable VLMC, the convergence in distribution of a VLMC towards its stationary probability measure is established.

Finally, finite state space semi-Markov chains turn out to be very special stable VLMC, shedding some new light on their limit distributions.

Keywords: Variable Length Memory Chains; stationary probability measure; Longest Internal Suffix; stable context trees; semi-Markov chains

References

[1] Barbu, V.S. and Limnios, N. (2008). *Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications*. *Lecture Notes in Statistics* **191**. New York: Springer. MR2452304
[2] Bejerano, G. and Yona, G. (2001). Variations on probabilistic suffix trees: Statistical modeling and prediction of protein families. *Bioinformatics* **17, 1** 23–43. https://doi.org/10.1093/bioinformatics/17.1.23
[3] Busch, J.R., Ferrari, P.A., Flesia, A.G., Fraiman, R., Grynberg, S.P. and Leonardi, F. (2009). Testing statistical hypothesis on random trees and applications to the protein classification problem. *Ann. Appl. Stat.* **3** 542–563. MR2750672 https://doi.org/10.1214/08-AOAS218
[4] Cénac, P., Chauvin, B., Herrmann, S. and Vallois, P. (2013). Persistent random walks, variable length Markov chains and piecewise deterministic Markov processes. *Markov Process. Related Fields* **19** 1–50. MR3088422

[5] Cénac, P., Chauvin, B., Noûs, C., Paccaut, F. and Pouyanne, N. (2021). Supplement to “Variable Length Memory Chains: Characterization of stationary probability measures.” https://doi.org/10.3150/20-BEJ1299SUPP
[6] Cénac, P., Chauvin, B., Paccaut, F. and Pouyanne, N. (2012). Context trees, variable length Markov chains and dynamical sources. In *Séminaire de Probabilités XLIV*. *Lecture Notes in Math*. **2046** 1–39. Heidelberg: Springer. MR2933931 https://doi.org/10.1007/978-3-642-27461-9_1
[7] Cénac, P., Chauvin, B., Paccaut, F. and Pouyanne, N. (2019). Variable length Markov chains, persistent random walks: A close encounter. Available at https://doi.org/10.1002/9781119779421.ch1
[8] Cénac, P., De Loynes, B., Offret, Y. and Rousselle, A. (2017). Recurrence of multidimensional persistent random walks. Fourier and series criteria. Available at https://doi.org/10.3150/18-BEJ1098
[9] Cénac, P., Le Ny, A., de Loynes, B. and Offret, Y. (2018). Persistent random walks. I. Recurrence versus transience. *J. Theoret. Probab.* **31** 232–243. MR3769813 https://doi.org/10.1007/s10959-016-0714-4
[10] Csiszár, I. and Talata, Z. (2006). Context tree estimation for not necessarily finite memory processes, via BIC and MDL. *IEEE Trans. Inf. Theory* **52** 1007–1016. MR2238067 https://doi.org/10.1109/TIT.2005.864431
[11] De Santis, E. and Piccioni, M. (2012). Backward coalescence times for perfect simulation of chains with infinite memory. *J. Appl. Probab.* **49** 319–337. MR2977798 https://doi.org/10.1239/jap/1339878789
[12] Doeblin, W. and Fortet, R. (1937). Sur des chaînes à liaisons complètes. *Bull. Soc. Math. France* **65** 132–148. MR1505076
[13] Fedotov, S., Tan, A. and Zubarev, A. (2015). Persistent random walk of cells involving anomalous effects and random death. *Phys. Rev. E* (3) **91** 042124, 10. MR3470981 https://doi.org/10.1103/PhysRevE.91.042124
[14] Fernández, R. and Maillard, G. (2005). Chains with complete connections: General theory, uniqueness, loss of memory and mixing properties. *J. Stat. Phys.* **118** 555–588. MR2123648 https://doi.org/10.1007/s10955-004-8821-5
[15] Ferreira, R.F., Gallo, S. and Paccaut, F. (2020). Non-regular g-measures and variable length memory chains. *Nonlinearity* **33** 6026–6052. MR4164670 https://doi.org/10.1088/1361-6544/aba0c5
[16] Furstenberg, H. (1967). Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Syst. Theory* **1** 1–49. MR0213508 https://doi.org/10.1007/BF01692494
[17] Gallo, S. (2011). Chains with unbounded variable length memory: Perfect simulation and a visible regeneration scheme. *Adv. in Appl. Probab.* **43** 735–759. MR2858219 https://doi.org/10.1239/aap/1316792668
[18] Gallo, S. and Garcia, N.L. (2013). Perfect simulation for locally continuous chains of infinite order. *Stochastic Process. Appl.* **123** 3877–3902. MR3091092 https://doi.org/10.1016/j.spa.2013.05.010
[19] Gallo, S. and Paccaut, F. (2013). On non-regular g-measures. *Nonlinearity* **26** 763–776. MR3033569 https://doi.org/10.1088/0951-7715/26/3/763
[20] Galves, A. and Leonardi, F. (2008). Exponential inequalities for empirical unbounded context trees. In *In and Out of Equilibrium*. 2. *Progress in Probability* **60** 257–269. Basel: Birkhäuser. MR2477385 https://doi.org/10.1007/978-3-7643-8786-0_12
[21] Garivier, A. and Leonardi, F. (2011). Context tree selection: A unifying view. *Stochastic Process. Appl.* **121** 2488–2506. MR2832411 https://doi.org/10.1016/j.spa.2011.06.012
[22] Harris, T.E. (1955). On chains of infinite order. *Pacific J. Math.* **5** 707–724. MR0075482
[23] Johansson, A. and Öberg, A. (2003). Square summability of variations of g-functions and uniqueness of g-measures. *Math. Res. Lett.* **10** 587–601. MR2024717 https://doi.org/10.4310/MRL.2003.v10.n5.a3
[24] Kitchens, B.P. (1998). *Symbolic Dynamics: One-Sided, Two-Sided and Countable State Markov Shifts*. *Universitext*. Berlin: Springer. MR1484730 https://doi.org/10.1007/978-3-642-58822-8
[25] Lyons, R. (1990). Random walks and percolation on trees. *Ann. Probab.* **18** 931–958. MR1062053
[26] Lyons, R. and Peres, Y. (2016). *Probability on Trees and Networks*. *Cambridge Series in Statistical and Probabilistic Mathematics* **42**. New York: Cambridge Univ. Press. MR3616205 https://doi.org/10.1017/9781316672815
[27] Rissanen, J. (1983). A universal data compression system. *IEEE Trans. Inf. Theory* **29** 656–664. MR0730903 https://doi.org/10.1109/TIT.1983.1056741

Bernoulli **27**(3), 2021, 2040–2068
https://doi.org/10.3150/20-BEJ1300

Approximation of heavy-tailed distributions via stable-driven SDEs

LU-JING HUANG[1], MATEUSZ B. MAJKA[2] and JIAN WANG[3]

[1]*College of Mathematics and Informatics, Fujian Normal University, 350007 Fuzhou, P.R. China.*
E-mail: lujingh@yeah.net
[2]*School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK.*
E-mail: m.majka@hw.ac.uk
[3]*College of Mathematics and Informatics & Fujian Key Laboratory of Mathematical Analysis and Applications (FJKLMAA) & Center for Applied Mathematics of Fujian Province (FJNU), Fujian Normal University, 350007 Fuzhou, P.R. China.*
E-mail: jianwang@fjnu.edu.cn

Constructions of numerous approximate sampling algorithms are based on the well-known fact that certain Gibbs measures are stationary distributions of ergodic stochastic differential equations (SDEs) driven by the Brownian motion. However, for some heavy-tailed distributions it can be shown that the associated SDE is not exponentially ergodic and that related sampling algorithms may perform poorly. A natural idea that has recently been explored in the machine learning literature in this context is to make use of stochastic processes with heavy tails instead of the Brownian motion. In this paper, we provide a rigorous theoretical framework for studying the problem of approximating heavy-tailed distributions via ergodic SDEs driven by symmetric (rotationally invariant) α-stable processes.

Keywords: Stochastic differential equations; symmetric α-stable processes; invariant measures; heavy-tailed distributions; approximate sampling; fractional Langevin Monte Carlo

References

[1] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9
[2] Bogdan, K., Byczkowski, T., Kulczycki, T., Ryznar, M., Song, R. and Vondraček, Z. (2009). *Potential Analysis of Stable Processes and Its Extensions*. *Lecture Notes in Math*. **1980**. Berlin: Springer. MR2569321 https://doi.org/10.1007/978-3-642-02141-1
[3] Böttcher, B., Schilling, R. and Wang, J. (2013). *Lévy-Type Processes*: *Construction*, *Approximation and Sample Path Properties*. *Lecture Notes in Math*. **2099**. Cham: Springer. Lévy Matters. III. MR3156646 https://doi.org/10.1007/978-3-319-02684-8
[4] Bucur, C. and Valdinoci, E. (2016). *Nonlocal Diffusion and Applications*. *Lecture Notes of the Unione Matematica Italiana* **20**. Bologna: Springer. MR3469920 https://doi.org/10.1007/978-3-319-28739-3
[5] Chen, M.-F. (2005). *Eigenvalues*, *Inequalities*, *and Ergodic Theory*. *Probability and Its Applications* (*New York*). London: Springer London, Ltd. MR2105651
[6] Chen, Z.-Q., Kim, P. and Song, R. (2012). Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation. *Ann*. *Probab*. **40** 2483–2538. MR3050510 https://doi.org/10.1214/11-AOP682
[7] Chen, Z.-Q., Song, R. and Zhang, X. (2018). Stochastic flows for Lévy processes with Hölder drifts. *Rev*. *Mat*. *Iberoam*. **34** 1755–1788. MR3896248 https://doi.org/10.4171/rmi/1042
[8] Chen, Z.-Q. and Zhang, X. (2018). Heat kernels for time-dependent non-symmetric stable-like operators. *J*. *Math*. *Anal*. *Appl*. **465** 1–21. MR3806688 https://doi.org/10.1016/j.jmaa.2018.03.054
[9] Cheng, X., Chatterji, N.S., Abbasi-Yadkori, Y., Bartlett, P.L. and Jordan, M.I. (2018). Sharp convergence rates for Langevin dynamics in the nonconvex setting. Available at arXiv:1805.01648.

1350-7265

[10] Dalalyan, A.S. (2017). Theoretical guarantees for approximate sampling from smooth and log-concave densities. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 651–676. MR3641401 https://doi.org/10.1111/rssb.12183

[11] Dareiotis, K., Kumar, C. and Sabanis, S. (2016). On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations. *SIAM J. Numer. Anal.* **54** 1840–1872. MR3513865 https://doi.org/10.1137/151004872

[12] Durmus, A. and Moulines, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479 https://doi.org/10.1214/16-AAP1238

[13] Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041 https://doi.org/10.1007/s00440-015-0673-1

[14] Eberle, A., Guillin, A. and Zimmer, R. (2019). Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. *Trans. Amer. Math. Soc.* **371** 7135–7173. MR3939573 https://doi.org/10.1090/tran/7576

[15] Eberle, A. and Majka, M.B. (2019). Quantitative contraction rates for Markov chains on general state spaces. *Electron. J. Probab.* **24** Paper No. 26, 36. MR3933205 https://doi.org/10.1214/19-EJP287

[16] Erdogdu, M.A. and Hosseinzadeh, R. (2020). On the convergence of Langevin Monte Carlo: The interplay between tail growth and smoothness. Available at arXiv:2005.13097.

[17] Gyöngy, I. and Krylov, N.V. (1980/81). On stochastic equations with respect to semimartingales. I. *Stochastics* **4** 1–21. MR0587426 https://doi.org/10.1080/03610918008833154

[18] Khasminskii, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Heidelberg: Springer. MR2894052 https://doi.org/10.1007/978-3-642-23280-0

[19] Kühn, F. and Schilling, R.L. (2019). Strong convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs. *Stochastic Process. Appl.* **129** 2654–2680. MR3980140 https://doi.org/10.1016/j.spa.2018.07.018

[20] Kulik, A.M. (2019). On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise. *Stochastic Process. Appl.* **129** 473–506. MR3907007 https://doi.org/10.1016/j.spa.2018.03.010

[21] Kumar, C. and Sabanis, S. (2017). On explicit approximations for Lévy driven SDEs with super-linear diffusion coefficients. *Electron. J. Probab.* **22** Paper No. 73, 19. MR3698742 https://doi.org/10.1214/17-EJP89

[22] Kwaśnicki, M. (2017). Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.* **20** 7–51. MR3613319 https://doi.org/10.1515/fca-2017-0002

[23] Liang, M., Majka, M.B. and Wang, J. (2021). Exponential ergodicity for SDEs and McKean–Vlasov processes with Lévy noise. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear. Available at arXiv:1901.11125.

[24] Liggett, T.M. (2010). *Continuous Time Markov Processes: An Introduction*. *Graduate Studies in Mathematics* **113**. Providence, RI: Amer. Math. Soc. MR2574430 https://doi.org/10.1090/gsm/113

[25] Majka, M.B. (2016). A note on existence of global solutions and invariant measures for jump SDEs with locally one-sided Lipschitz drift. *Probab. Math. Statist.* **40** 37–55. MR4170166 https://doi.org/10.37190/0208-4147.40.1.3

[26] Majka, M.B., Mijatović, A. and Szpruch, Ł. (2020). Nonasymptotic bounds for sampling algorithms without log-concavity. *Ann. Appl. Probab.* **30** 1534–1581. MR4132634 https://doi.org/10.1214/19-AAP1535

[27] Masuda, H. (2004). On multidimensional Ornstein–Uhlenbeck processes driven by a general Lévy process. *Bernoulli* **10** 97–120. MR2044595 https://doi.org/10.3150/bj/1077544605

[28] Meyn, S.P. and Tweedie, R.L. (1993). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.* **25** 487–517. MR1234294 https://doi.org/10.2307/1427521

[29] 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

[30] Mikulevičius, R. and Xu, F. (2018). On the rate of convergence of strong Euler approximation for SDEs driven by Levy processes. *Stochastics* **90** 569–604. MR3784978 https://doi.org/10.1080/17442508.2017.1381095

[31] Mou, W., Flammarion, N., Wainwright, M.J. and Bartlett, P.L. (2019). Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity. Available at arXiv:1907.11331.

[32] Nguyen, T.H., Şimşekli, U. and Richard, G. (2019). Non-asymptotic analysis of Fractional Langevin Monte Carlo for non-convex optimization. In *Proceedings of the* 36*th International Conference on Machine Learning* (K. Chaudhuri, and R. Salakhutdinov, eds.). *Proceedings of Machine Learning Research*, 97 4810–4819. Long Beach, California, USA, 09–15 Jun 2019. PMLR.

[33] Ortigueira, M.D. (2006). Riesz potential operators and inverses via fractional centred derivatives. *Int. J. Math. Math. Sci.* Art. ID 48391, 12. MR2251718 https://doi.org/10.1155/IJMMS/2006/48391

[34] Priola, E. (2012). Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.* **49** 421–447. MR2945756

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

[36] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. MR1739520

[37] Schilling, R.L., Sztonyk, P. and Wang, J. (2012). Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18** 1128–1149. MR2995789 https://doi.org/10.3150/11-BEJ375

[38] Şimşekli, U. (2017). Fractional Langevin Monte Carlo: exploring Lévy driven stochastic differential equations for Markov chain Monte Carlo. In *Proceedings of the* 34*th International Conference on Machine Learning*, (D. Precup and Y.W. Teh, eds.). *Proceedings of Machine Learning Research* 70 3200–3209. International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.

[39] Şimşekli, U., Zhu, L., Teh, Y.W. and Gürbüzbalaban, M. (2002). Fractional underdamped Langevin dynamics: Retargeting SGD with momentum under heavy-tailed gradient noise. Available at arXiv:2002.05685.

[40] Stein, E.M. (1970). *Singular Integrals and Differentiability Properties of Functions. Princeton Mathematical Series, No.* 30. Princeton, NJ: Princeton Univ. Press. MR0290095

[41] Triebel, H. (1978). *Interpolation Theory, Function Spaces, Differential Operators. North-Holland Mathematical Library* **18**. Amsterdam: North-Holland. MR0503903

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

[43] Wang, F.-Y. and Wang, J. (2015). Functional inequalities for stable-like Dirichlet forms. *J. Theoret. Probab.* **28** 423–448. MR3370660 https://doi.org/10.1007/s10959-013-0500-5

[44] Xie, L. and Zhang, X. (2014). Heat kernel estimates for critical fractional diffusion operators. *Studia Math.* **224** 221–263. MR3294616 https://doi.org/10.4064/sm224-3-3

[45] Xie, L. and Zhang, X. (2020). Ergodicity of stochastic differential equations with jumps and singular coefficients. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 175–229. MR4058986 https://doi.org/10.1214/19-AIHP959

[46] Ye, N. and Zhu, Z. (2018). Stochastic fractional Hamiltonian Monte Carlo. In *Joint Conference on Artificial Intelligence* (J. Lang, ed.). *Proceedings of the* 27*th International* 3019–3025. International Joint Conference on Artificial Intelligence, Stockholm, Sweden, July 13–19, 2018.

Bernoulli **27**(3), 2021, 2069–2090
https://doi.org/10.3150/20-BEJ1301

On multi-step estimation of delay for SDE

YURY A. KUTOYANTS[1,2]

[1]*Laboratoire Manceau de Mathematiques, Le Mans University, Le Mans, France*
[2]*Tomsk State University, Tomsk, Russia.*
E-mail: kutoyants@univ-lemans.fr

We consider the problem of delay estimation by the observations of the solutions of several SDEs. It is known that the MLEs for these models are consistent and asymptotically normal, but the likelihood ratio functions are not differentiable w.r.t. the parameter, and therefore the numerical calculation of the MLEs encounter certain difficulties. We propose One-step and Two-step MLEs, whose calculation has no such problems and provide an estimator asymptotically equivalent to the MLE. These constructions are realized in two or three steps. First, we construct preliminary estimators which are consistent and asymptotically normal, but not asymptotically efficient. Then we use these estimators and a modified Fisher-score device to obtain One-step and Two-step MLEs. We suppose that its numerical realization is much more simple. Stochastic Pantograph equation is introduced and related statistical problems are discussed.

Keywords: One-step MLE; Two-step MLE; One-step MDE; Stochastic Pantograph equation; delay estimation

References

[1] Apoyan, G.T. (1986). Parameter estimation of non differentiable trend coefficient. *Uchen. Zap. Erevan University* **1** 33–62. (In Russian).
[2] Bandrivskyy, A., Luchinsky, D.G., McClintock, P.V.E., Smelyanskiy, V.N. and Stefanovska, A. (2005). Inference of systems with delay and applications to cardiovascular dynamics. *Stoch. Dyn.* **5** 321–331. MR2147292 https://doi.org/10.1142/S0219493705001432
[3] Davies, R.B. (1985). Asymptotic inference when the amount of information is random. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer, Vol. II* (*Berkeley, Calif.*, 1983). *Wadsworth Statist./Probab. Ser.* 841–864. Belmont, CA: Wadsworth. MR0822069
[4] Frank, T.D., Friedrich, R. and Beek, P.J. (2005). Time series analysis of multivariate time-delayed systems with noise: Applications to laser physics and human movement. *Stoch. Dyn.* **5** 297–306. MR2147290 https://doi.org/10.1142/S0219493705001456
[5] Freidlin, M.I. and Wentzell, A.D. (1984). *Random Perturbations of Dynamical Systems. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **260**. New York: Springer. MR0722136 https://doi.org/10.1007/978-1-4684-0176-9
[6] Griebel, T. (2017). *The Pantograph Equation in Quantum Calculus*. Missouri Univ. Science and Technology, Masters Theses.
[7] Gushchin, A.A. and Küchler, U. (1999). Asymptotic inference for a linear stochastic differential equation with time delay. *Bernoulli* **5** 1059–1098. MR1735785 https://doi.org/10.2307/3318560
[8] Hauptmann, C., Popovych, O. and Tass, P.A. (2005). Multisite coordinated delayed feedback for an effective desynchronization of neuronal networks. *Stoch. Dyn.* **5** 307–319. MR2147291 https://doi.org/10.1142/S0219493705001420
[9] Ibragimov, I.A. and Has'minskii, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. Berlin: Springer.
[10] Kuang, Y. (1993). *Delay Differential Equations with Applications in Population Dynamics. Mathematics in Science and Engineering* **191**. Boston, MA: Academic Press. MR1218880
[11] Küchler, U. and Kutoyants, Y.A. (2000). Delay estimation for some stationary diffusion-type processes. *Scand. J. Stat.* **27** 405–414. MR1795773 https://doi.org/10.1111/1467-9469.00197
[12] Küchler, U. and Mensch, B. (1992). Langevin's stochastic differential equation extended by a time-delayed term. *Stoch. Stoch. Rep.* **40** 23–42. MR1275126 https://doi.org/10.1080/17442509208833780

[13] Kutoyants, Y. (1994). *Identification of Dynamical Systems with Small Noise*. *Mathematics and Its Applications* **300**. Dordrecht: Kluwer Academic. MR1332492 https://doi.org/10.1007/978-94-011-1020-4
[14] Kutoyants, Y.A. (1988). An example of an estimate for the parameter of a nondifferentiable drift coefficient. *Teor. Veroyatn. Primen.* **33** 188–192. MR0940007 https://doi.org/10.1137/1133024
[15] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. *Springer Series in Statistics*. London: Springer. MR2144185 https://doi.org/10.1007/978-1-4471-3866-2
[16] Kutoyants, Y.A. (2005). On delay estimation for stochastic differential equations. *Stoch. Dyn.* **5** 333–342. MR2147293 https://doi.org/10.1142/S0219493705001444
[17] Kutoyants, Y.A. (2014). Approximation of the solution of the backward stochastic differential equation. Small noise, large sample and high frequency cases. *Proc. Steklov Inst. Math.* **287** 133–154. MR3484327 https://doi.org/10.1134/S0081543814080094
[18] Kutoyants, Y.A. (2017). On the multi-step MLE-process for ergodic diffusion. *Stochastic Process. Appl.* **127** 2243–2261. MR3652412 https://doi.org/10.1016/j.spa.2016.10.007
[19] Kutoyants, Y.A. and Zhou, L. (2014). On approximation of the backward stochastic differential equation. *J. Statist. Plann. Inference* **150** 111–123. MR3206723 https://doi.org/10.1016/j.jspi.2014.03.002
[20] Le Cam, L. and Yang, G.L. (2000). *Asymptotics in Statistics*: *Some Basic Concepts*, 2nd ed. *Springer Series in Statistics*. New York: Springer. MR1784901 https://doi.org/10.1007/978-1-4612-1166-2
[21] Liptser, R.S. and Shiryaev, A.N. (2001). *Statistics of Random Processes* **1**, 2nd ed. New York: Springer.
[22] Mao, X. (2005). Delay population dynamics and environmental noise. *Stoch. Dyn.* **5** 149–162. MR2147279 https://doi.org/10.1142/S021949370500133X
[23] Myshkis, A.D. (1951). *Linear Differential Equations with Delayed Argument*. Moscow: Nauka. (In Russian).
[24] Ockendon, J.R. and Tayler, A.B. (1971). The dynamics of a current collection system for an electric locomotive. In *Proceedings of the Royal Society of London A*: *Mathematical*, *Physical and Engineering Sciences* **322** 447–468.
[25] Ohira, T. and Yamane, T. (2000). Delayed stochastic systems. *Phys. Rev.*, *E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **61** 1247–1257.
[26] Rosinberg, M.L., Tarjus, G. and Munakata, T. (2018). Influence of time delay on information exchanges between coupled linear stochastic systems. *Phys. Rev. E* (3) **98** 032130.

Bernoulli **27**(3), 2021, 2091–2116
https://doi.org/10.3150/20-BEJ1302

Context-specific independencies in stratified chain regression graphical models

FEDERICA NICOLUSSI[1] and MANUELA CAZZARO[2]

[1]*University of Milan, Via Conservatorio, 7, 20122 Milano MI, Italy.*
E-mail: federica.nicolussi@unimi.it
[2]*University of Milan Bicocca, Via Bicocca Degli Arcimboldi 8, 20126 Milano, MI, Italy.*
E-mail: manuela.cazzaro@unimib.it

Graphical models are a useful tool with increasing diffusion. In the categorical variable framework, they provide important visual support to understand the relationships among the considered variables. Besides, particular chain graphical models are suitable to represent multivariate regression models. However, the associated parameterization, such as marginal log-linear models, is often difficult to interpret when the number of variables increases because of a large number of parameters involved. On the contrary, conditional and marginal independencies reduce the number of parameters needed to represent the joint probability distribution of the variables. In compliance with the parsimonious principle, it is worthwhile to consider also the so-called context-specific independencies, which are conditional independencies holding for particular values of the variables in the conditioning set. In this work, we propose a particular chain graphical model able to represent these context-specific independencies through labeled arcs. We provide also the Markov properties able to describe marginal, conditional, and context-specific independencies from this new chain graph. Finally, we show the results in an application to a real data set.

Keywords: Graphical models; stratified Markov properties; categorical variables; multivariate regression models; marginal models

References

[1] Aitchison, J. and Silvey, S.D. (1958). Maximum-likelihood estimation of parameters subject to restraints. *Ann. Math. Stat.* **29** 813–828. MR0094873 https://doi.org/10.1214/aoms/1177706538
[2] Bartolucci, F., Colombi, R. and Forcina, A. (2007). An extended class of marginal link functions for modelling contingency tables by equality and inequality constraints. *Statist. Sinica* **17** 691–711. MR2398430
[3] Bergsma, W.P. and Rudas, T. (2002). Marginal models for categorical data. *Ann. Statist.* **30** 140–159. MR1892659 https://doi.org/10.1214/aos/1015362188
[4] Boutilier, C., Friedman, N., Goldszmidt, M. and Koller, D. (1996). Context-specific independence in Bayesian networks. In *Uncertainty in Artificial Intelligence* (*Portland*, *OR*, 1996) 115–123. San Francisco, CA: Morgan Kaufmann. MR1617129
[5] Cazzaro, M. and Colombi, R. (2008). Modelling two way contingency tables with recursive logits and odds ratios. *Stat. Methods Appl.* **17** 435–453. MR2447568 https://doi.org/10.1007/s10260-007-0068-2
[6] Cazzaro, M. and Colombi, R. (2014). Marginal nested interactions for contingency tables. *Comm. Statist. Theory Methods* **43** 2799–2814. MR3223712 https://doi.org/10.1080/03610926.2012.685550
[7] Colombi, R., Giordano, S. and Cazzaro, M. (2014). hmmm: An R package for hierarchical multinomial marginal models. *J. Stat. Softw.* **59** 1–25.
[8] Cox, D.R. and Wermuth, N. (1993). Linear dependencies represented by chain graphs. *Statist. Sci.* **8** 204–218, 247–283. MR1243593
[9] Drton, M. (2009). Discrete chain graph models. *Bernoulli* **15** 736–753. MR2555197 https://doi.org/10.3150/08-BEJ172
[10] Højsgaard, S. (2003). Split models for contingency tables. *Comput. Statist. Data Anal.* **42** 621–645. MR1967060 https://doi.org/10.1016/S0167-9473(02)00119-6

[11] Højsgaard, S. (2004). Statistical inference in context specific interaction models for contingency tables. *Scand. J. Stat.* **31** 143–158. MR2042604 https://doi.org/10.1111/j.1467-9469.2004.00378.x

[12] ISTAT (2015). Italian innovation Survey 2002–2012. Available at http://www.istat.it/en/archive/87787.

[13] La Rocca, L. and Roverato, A. (2017). *Discrete Graphical Models. Handbook of Graphical Models. Handbooks of Modern Statistical Methods*. Boca Raton, FL: CRC Press/CRC.

[14] Lauritzen, S. and Sadeghi, K. (2018). Unifying Markov properties for graphical models. *Ann. Statist.* **46** 2251–2278. MR3845017 https://doi.org/10.1214/17-AOS1618

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

[16] Marchetti, G.M. and Lupparelli, M. (2008). Parameterization and fitting of a class of discrete graphical models. *COMPSTAT* **2008** 117–128.

[17] Marchetti, G.M. and Lupparelli, M. (2011). Chain graph models of multivariate regression type for categorical data. *Bernoulli* **17** 827–844. MR2817607 https://doi.org/10.3150/10-BEJ300

[18] Nicolussi, F. (2013). Marginal parameterizations for conditional independence models and graphical models for categorical data. Ph.D. thesis, Univ. of Milan-Bicocca.

[19] Nicolussi, F. and Cazzaro, M. (2020). Context-specific independencies in hierarchical multinomial marginal models. *Stat. Methods Appl.* **29** 767–786. MR4174686 https://doi.org/10.1007/s10260-019-00503-8

[20] Nicolussi, F. and Cazzaro, M. (2021). Supplement to "Context-specific independencies in stratified chain regression graphical models". https://doi.org/10.3150/20-BEJ1302SUPP

[21] Nicolussi, F. and Colombi, R. (2017). Type II chain graph models for categorical data: A smooth subclass. *Bernoulli* **23** 863–883. MR3606753 https://doi.org/10.3150/15-BEJ762

[22] Nyman, H., Pensar, J. and Corander, J. (2016). Dependence Logic. 219–234. Springer.

[23] Nyman, H., Pensar, J., Koski, T. and Corander, J. (2016). Context-specific independence in graphical log-linear models. *Comput. Statist.* **31** 1493–1512. MR3573088 https://doi.org/10.1007/s00180-015-0606-6

[24] Pensar, J., Nyman, H., Koski, T. and Corander, J. (2015). Labeled directed acyclic graphs: A generalization of context-specific independence in directed graphical models. *Data Min. Knowl. Discov.* **29** 503–533. MR3312469 https://doi.org/10.1007/s10618-014-0355-0

[25] R Core Team (2014) R: a language and environment for statistical computing, Vienna, Austria.

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

[27] Roverato, A. (2017). *Graphical Models for Categorical Data. SemStat Elements*. Cambridge: Cambridge Univ. Press. MR3751385 https://doi.org/10.1017/9781108277495

[28] Rudas, T., Bergsma, W.P. and Németh, R. (2010). Marginal log-linear parameterization of conditional independence models. *Biometrika* **97** 1006–1012. MR2746171 https://doi.org/10.1093/biomet/asq037

[29] Sadeghi, K. (2018). Markov properties of discrete determinantal point processes. Preprint. Available at arXiv:1810.02294.

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

[31] Sadeghi, K. and Wermuth, N. (2016). Pairwise Markov properties for regression graphs. *Stat* **5** 286–294. MR3589267 https://doi.org/10.1002/sta4.122

[32] Wermuth, N. and Cox, D.R. (2004). Joint response graphs and separation induced by triangular systems. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 687–717. MR2088296 https://doi.org/10.1111/j.1467-9868.2004.b5161.x

[33] Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Chichester: Wiley. MR1112133

Bernoulli **27**(3), 2021, 2117–2138
https://doi.org/10.3150/20-BEJ1303

A new look at random projections of the cube and general product measures

ZAKHAR KABLUCHKO[1], JOSCHA PROCHNO[2] and CHRISTOPH THÄLE[3]

[1]*Institut für Mathematische Stochastik, Westfälische Wilhelms-Universität Münster, Germany. E-mail: zakhar.kabluchko@uni-muenster.de*
[2]*Institut für Mathematik & Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Austria. E-mail: joscha.prochno@uni-graz.at*
[3]*Faculty of Mathematics, Ruhr University Bochum, Germany. E-mail: christoph.thaele@rub.de*

A consequence of the celebrated Dvoretzky–Milman theorem is a strong law of large numbers for d-dimensional random projections of the n-dimensional cube. It shows that, with respect to the Hausdorff distance, a uniform random projection of the cube $[-1/\sqrt{n}, +1/\sqrt{n}]^n$ onto $\mathbb{R}^d$ converges almost surely to a centered d-dimensional Euclidean ball of radius $\sqrt{2/\pi}$, as $n \to \infty$. We start by providing an alternative proof of this strong law via the Artstein–Vitale law of large numbers for random compact sets. Then, for every point inside the ball of radius $\sqrt{2/\pi}$, we determine the asymptotic number of vertices and the volume of the part of the cube projected 'close' to this point. More generally, we study large deviations for random projections of arbitrary product measures. Let $\nu^{\otimes n}$ be the n-fold product measure of a Borel probability measure ν on $\mathbb{R}$, and let I be uniformly distributed on the Stiefel manifold of orthogonal d-frames in $\mathbb{R}^n$. It is shown that the sequence of random measures $\nu^{\otimes n} \circ (n^{-1/2} I^*)^{-1}$, $n \in \mathbb{N}$, satisfies a large deviation principle with probability 1. The rate function is explicitly identified in terms of the moment generating function of ν. At the heart of the proofs lies a transition trick which allows to replace the uniform projection by the Gaussian one. A number of concrete examples are discussed as well, including the uniform distributions on the cube $[-1, 1]^n$ and the discrete cube $\{-1, 1\}^n$ as special cases.

Keywords: Cube; Gaussian random matrices; Gaussian projections; high-dimensional probability; Hausdorff distance; large deviations; law of large numbers; random projections; Stiefel manifold

References

[1] Affentranger, F. and Schneider, R. (1992). Random projections of regular simplices. *Discrete Comput. Geom.* **7** 219–226. MR1149653 https://doi.org/10.1007/BF02187839

[2] Alonso-Gutiérrez, D., Prochno, J. and Thäle, C. (2018). Large deviations for high-dimensional random projections of ℓ_p^n-balls. *Adv. in Appl. Math.* **99** 1–35. MR3806754 https://doi.org/10.1016/j.aam.2018.04.003

[3] Alonso-Gutiérrez, D., Prochno, J. and Thäle, C. (2019). Gaussian fluctuations for high-dimensional random projections of ℓ_p^n-balls. *Bernoulli* **25** 3139–3174. MR4003577 https://doi.org/10.3150/18-BEJ1084

[4] Artstein, Z. and Vitale, R.A. (1975). A strong law of large numbers for random compact sets. *Ann. Probab.* **3** 879–882. MR0385966 https://doi.org/10.1214/aop/1176996275

[5] Artstein-Avidan, S., Giannopoulos, A. and Milman, V.D. (2015). *Asymptotic Geometric Analysis. Part I. Mathematical Surveys and Monographs* **202**. Providence, RI: Amer. Math. Soc. MR3331351 https://doi.org/10.1090/surv/202

[6] Aubrun, G. and Szarek, S.J. (2017). *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory. Mathematical Surveys and Monographs* **223**. Providence, RI: Amer. Math. Soc. MR3699754 https://doi.org/10.1090/surv/223

[7] Baryshnikov, Y.M. and Vitale, R.A. (1994). Regular simplices and Gaussian samples. *Discrete Comput. Geom.* **11** 141–147. MR1254086 https://doi.org/10.1007/BF02574000

[8] Bingham, E. and Mannila, H. (2001). Random projection in dimensionality reduction: Applications to image and text data. In *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *KDD'*01 245–250. New York, NY, USA: ACM.
[9] Böröczky, K. Jr. and Henk, M. (1999). Random projections of regular polytopes. *Arch. Math.* (*Basel*) **73** 465–473. MR1725183 https://doi.org/10.1007/s000130050424
[10] Brazitikos, S., Giannopoulos, A., Valettas, P. and Vritsiou, B.-H. (2014). *Geometry of Isotropic Convex Bodies*. *Mathematical Surveys and Monographs* **196**. Providence, RI: Amer. Math. Soc. MR3185453 https://doi.org/10.1090/surv/196
[11] Chávez-Domínguez, J.A. and Kutzarova, D. (2015). Stability of low-rank matrix recovery and its connections to Banach space geometry. *J. Math. Anal. Appl.* **427** 320–335. MR3318202 https://doi.org/10.1016/j.jmaa.2015.02.041
[12] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Berlin: Springer. Corrected reprint of the second (1998) edition. MR2571413 https://doi.org/10.1007/978-3-642-03311-7
[13] den Hollander, F. (2000). *Large Deviations*. *Fields Institute Monographs* **14**. Providence, RI: Amer. Math. Soc. MR1739680 https://doi.org/10.1007/s00440-009-0235-5
[14] Deuschel, J.-D. and Stroock, D.W. (1989). *Large Deviations*, Revised ed. *Pure and Applied Mathematics* **137**. Boston, MA: Academic Press. MR0997938
[15] Donoho, D.L. and Tanner, J. (2010). Counting the faces of randomly-projected hypercubes and orthants, with applications. *Discrete Comput. Geom.* **43** 522–541. MR2587835 https://doi.org/10.1007/s00454-009-9221-z
[16] Evans, L.C. and Gariepy, R.F. (2015). *Measure Theory and Fine Properties of Functions*, Revised ed. *Textbooks in Mathematics*. Boca Raton, FL: CRC Press. MR3409135
[17] Fern, X.Z. and Brodley, C.E. (2003). Random projection for high dimensional data clustering: A cluster ensemble approach. In *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, *ICML'*03 186–193. AAAI Press.
[18] Figiel, T., Lindenstrauss, J. and Milman, V.D. (1977). The dimension of almost spherical sections of convex bodies. *Acta Math.* **139** 53–94. MR0445274 https://doi.org/10.1007/BF02392234
[19] Foucart, S., Pajor, A., Rauhut, H. and Ullrich, T. (2010). The Gelfand widths of ℓ_p-balls for $0 < p \leq 1$. *J. Complexity* **26** 629–640. MR2735423 https://doi.org/10.1016/j.jco.2010.04.004
[20] Gantert, N., Kim, S.S. and Ramanan, K. (2017). Large deviations for random projections of ℓ^p balls. *Ann. Probab.* **45** 4419–4476. MR3737915 https://doi.org/10.1214/16-AOP1169
[21] Hinrichs, A., Krieg, D., Novak, E., Prochno, J. and Ullrich, M. (2019). On the power of random information. Available at arXiv:1903.00681.
[22] Hinrichs, A., Krieg, D., Novak, E., Prochno, J. and Ullrich, M. (2019). Random sections of ellipsoids and the power of random information. Available at arXiv:1901.06639.
[23] Hinrichs, A., Prochno, J. and Ullrich, M. (2019). The curse of dimensionality for numerical integration on general domains. *J. Complexity* **50** 25–42. MR3907362 https://doi.org/10.1016/j.jco.2018.08.003
[24] Hinrichs, A., Prochno, J. and Vybíral, J. (2017). Entropy numbers of embeddings of Schatten classes. *J. Funct. Anal.* **273** 3241–3261. MR3695893 https://doi.org/10.1016/j.jfa.2017.08.008
[25] Kabluchko, Z., Prochno, J. and Thäle, C. (2019). High-dimensional limit theorems for random vectors in ℓ_p^n-balls. *Commun. Contemp. Math.* **21** 1750092, 30. MR3904638 https://doi.org/10.1142/S0219199717500924
[26] Kabluchko, Z., Prochno, J. and Thäle, C. (2019). High-dimensional limit theorems for random vectors in ℓ_p^n-balls. II. *Commun. Contemp. Math.* To appear.
[27] Kabluchko, Z., Prochno, J. and Thäle, C. (2020). Sanov-type large deviations in Schatten classes. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** 928–953. MR4076771 https://doi.org/10.1214/19-AIHP989
[28] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[29] Kallenberg, O. (2017). *Random Measures*, *Theory and Applications*. *Probability Theory and Stochastic Modelling* **77**. Cham: Springer. MR3642325 https://doi.org/10.1007/978-3-319-41598-7
[30] Kim, S.S. and Ramanan, K. (2018). A conditional limit theorem for high-dimensional ℓ^p-spheres. *J. Appl. Probab.* **55** 1060–1077. MR3899928 https://doi.org/10.1017/jpr.2018.71
[31] Klartag, B. (2007). A central limit theorem for convex sets. *Invent. Math.* **168** 91–131. MR2285748 https://doi.org/10.1007/s00222-006-0028-8

[32] Klartag, B. (2007). Power-law estimates for the central limit theorem for convex sets. *J. Funct. Anal.* **245** 284–310. MR2311626 https://doi.org/10.1016/j.jfa.2006.12.005

[33] Klenke, A. (2014). *Probability Theory: A Comprehensive Course*, 2nd ed. *Universitext*. London: Springer. MR3112259 https://doi.org/10.1007/978-1-4471-5361-0

[34] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces: Isoperimetry and Processes*. *Classics in Mathematics*. Berlin: Springer. MR2814399

[35] Maillard, O.-A. and Munos, R. (2012). Linear regression with random projections. *J. Mach. Learn. Res.* **13** 2735–2772. MR2989913

[36] Mankiewicz, P. and Tomczak-Jaegermann, N. (2001). Geometry of families of random projections of symmetric convex bodies. *Geom. Funct. Anal.* **11** 1282–1326. MR1878321 https://doi.org/10.1007/s00039-001-8231-7

[37] Meckes, E.S. and Meckes, M.W. (2007). The central limit problem for random vectors with symmetries. *J. Theoret. Probab.* **20** 697–720. MR2359052 https://doi.org/10.1007/s10959-007-0119-5

[38] Milman, V.D. (1971). A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. *Funkcional. Anal. i Priložen.* **5** 28–37. MR0293374

[39] Milman, V.D. and Schechtman, G. (1986). *Asymptotic Theory of Finite-Dimensional Normed Spaces*. *Lecture Notes in Math.* **1200**. Berlin: Springer. MR0856576

[40] Paouris, G., Pivovarov, P. and Zinn, J. (2014). A central limit theorem for projections of the cube. *Probab. Theory Related Fields* **159** 701–719. MR3230006 https://doi.org/10.1007/s00440-013-0518-8

[41] Prochno, J., Thäle, C. and Turchi, N. (2019). Geometry of ℓ_p^n-balls: Classical results and recent developments. In *High Dimensional Probability VIII*, *the Oaxaca Volume*, *Progress in Probability* 121–150. Birkhäuser.

[42] Schechtman, G. and Schmuckenschläger, M. (1991). Another remark on the volume of the intersection of two L_p^n balls. In *Geometric Aspects of Functional Analysis* (1989–90). *Lecture Notes in Math.* **1469** 174–178. Berlin: Springer. MR1122622 https://doi.org/10.1007/BFb0089224

[43] Schmuckenschläger, M. (2001). CLT and the volume of intersections of l_p^n-balls. *Geom. Dedicata* **85** 189–195. MR1845607 https://doi.org/10.1023/A:1010353121014

Bernoulli **27**(3), 2021, 2139–2158
https://doi.org/10.3150/20-BEJ1304

Nearly optimal robust mean estimation via empirical characteristic function

SOHAIL BAHMANI

School of Electrical & Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA.
E-mail: sohail.bahmani@ece.gatech.edu

We propose an estimator for the mean of random variables in separable real Banach spaces using the empirical characteristic function. Assuming that the covariance operator of the random variable is bounded in a precise sense, we show that the proposed estimator achieves the optimal sub-Gaussian rate, except for a faster decaying mean-dependent term. Under a mild condition, an iterative refinement procedure can essentially eliminate the mean-dependent term and provide a shift-equivariant estimate. Our results particularly suggests that a certain Gaussian width that appears in the best known rate in the literature might not be necessary. Furthermore, using the boundedness of the characteristic functions, we also show that, except possibly at high signal-to-noise ratios, the proposed estimator is gracefully robust against adversarial "contamination". Our analysis is overall concise and transparent, thanks to the tractability of the characteristic functions.

Keywords: Mean estimation; robust estimation; characteristic function

References

[1] Alon, N., Matias, Y. and Szegedy, M. (1996). The space complexity of approximating the frequency moments. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing* (*Philadelphia*, *PA*, 1996) 20–29. New York: ACM. MR1427494 https://doi.org/10.1145/237814.237823

[2] 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

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

[4] 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

[5] Catoni, O. and Giulini, I. (2017). Dimension-free PAC-Bayesian bounds for matrices, vectors, and linear least squares regression. Preprint. Available at arXiv:1712.02747.

[6] Cheng, Y., Diakonikolas, I. and Ge, R. (2019). High-dimensional robust mean estimation in nearly-linear time. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* 2755–2771. Philadelphia, PA: SIAM. MR3909640 https://doi.org/10.1137/1.9781611975482.171

[7] Cherapanamjeri, Y., Flammarion, N. and Bartlett, P.L. (2019). Fast mean estimation with sub-Gaussian rates. In *Proceedings of the Thirty-Second Conference on Learning Theory* (A. Beygelzimer and D. Hsu, eds.). *Proceedings of Machine Learning Research* **99** 786–806. Phoenix, USA: PMLR.

[8] Depersin, J. and Lecué, G. (2019). Robust subgaussian estimation of a mean vector in nearly linear time.

[9] Diakonikolas, I., Kamath, G., Kane, D., Li, J., Moitra, A. and Stewart, A. (2019). Robust estimators in high-dimensions without the computational intractability. *SIAM J. Comput.* **48** 742–864. MR3945261 https://doi.org/10.1137/17M1126680

[10] Diakonikolas, I., Kamath, G., Kane, D.M., Li, J., Moitra, A. and Stewart, A. (2017). Being robust (in high dimensions) can be practical. In *Proceedings of the* 34*th International Conference on Machine Learning* (D. Precup and Y.W. Teh, eds.). *Proceedings of Machine Learning Research* **70** 999–1008. PMLR.

1350-7265

[11] Diakonikolas, I. and Kane, D.M. (2019). Recent advances in algorithmic high-dimensional robust statistics. Preprint. Available at arXiv:1911.05911.

[12] Dirksen, S. (2015). Tail bounds via generic chaining. *Electron*. *J*. *Probab*. **20** no. 53, 29. MR3354613 https://doi.org/10.1214/EJP.v20-3760

[13] van Der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*. *Springer Series in Statistics*. Springer.

[14] Giné, E. and Zinn, J. (1984). Some limit theorems for empirical processes. *Ann*. *Probab*. **12** 929–998. MR0757767

[15] Hopkins, S.B. (2020). Mean estimation with sub-Gaussian rates in polynomial time. *Ann*. *Statist*. **48** 1193–1213.

[16] Hsu, D. and Sabato, S. (2014). Heavy-tailed regression with a generalized median-of-means. In *Proceedings of the* 31*st International Conference on Machine Learning* (E.P. Xing and T. Jebara, eds.). *Proceedings of Machine Learning Research* **32** 37–45. PMLR.

[17] 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

[18] 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

[19] Ledoux, M. and Talagrand, M. (2013). *Probability in Banach Spaces*: *Isoperimetry and Processes*. Springer Science & Business Media.

[20] Lei, Z., Luh, K., Venkat, P. and Zhang, F. (2020). A fast spectral algorithm for mean estimation with sub-Gaussian rates. In *Proceedings of Machine Learning Research*. *Conference on Learning Theory* **125** 2598–2612. PMLR.

[21] Lerasle, M. and Oliveira, R.I. Robust empirical mean estimators. Preprint. Available at arXiv:1112:3914.

[22] Lugosi, G. and Mendelson, S. (2019). Sub-Gaussian estimators of the mean of a random vector. *Ann*. *Statist*. **47** 783–794. MR3909950 https://doi.org/10.1214/17-AOS1639

[23] Lugosi, G. and Mendelson, S. (2019). Near-optimal mean estimators with respect to general norms. *Probab*. *Theory Related Fields* **175** 957–973. MR4026610 https://doi.org/10.1007/s00440-019-00906-4

[24] 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

[25] Lugosi, G. and Mendelson, S. (2021). Robust multivariate mean estimation: The optimality of trimmed mean. *Ann*. *Statist*. **49** 393–410. MR4206683 https://doi.org/10.1214/20-AOS1961

[26] Mendelson, S. and Zhivotovskiy, N. (2020). Robust covariance estimation under L_4-L_2 norm equivalence. *Ann*. *Statist*. **48** 1648–1664. MR4124338 https://doi.org/10.1214/19-AOS1862

[27] Minsker, S. (2015). Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** 2308–2335. MR3378468 https://doi.org/10.3150/14-BEJ645

[28] Minsker, S. (2018). Uniform bounds for robust mean estimators. Preprint. Available at arXiv:1812.03523.

[29] Nemirovsky, A.S. and Yudin, D.B. (1983). *Problem Complexity and Method Efficiency in Optimization*. *A Wiley-Interscience Publication*. New York: Wiley. MR0702836

[30] Rodriguez, D. and Valdora, M. (2019). The breakdown point of the median of means tournament. *Statist*. *Probab*. *Lett*. **153** 108–112. MR3962785 https://doi.org/10.1016/j.spl.2019.05.012

[31] Talagrand, M. (2014). *Upper and Lower Bounds for Stochastic Processes*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete*. 3. *Folge*. *A Series of Modern Surveys in Mathematics* [*Results in Mathematics and Related Areas*. 3*rd Series*. *A Series of Modern Surveys in Mathematics*] **60**. Heidelberg: Springer. MR3184689 https://doi.org/10.1007/978-3-642-54075-2

[32] Vakhania, N.N., Tarieladze, V.I. and Chobanyan, S.A. (1987). *Probability Distributions on Banach Spaces*. *Mathematics and Its Applications* (*Soviet Series*) **14**. Dordrecht: D. Reidel Publishing Co. MR1435288 https://doi.org/10.1007/978-94-009-3873-1

[33] Vershynin, R. (2018). *High-Dimensional Probability*: *An Introduction With Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics* **47**. Cambridge: Cambridge Univ. Press. MR3837109 https://doi.org/10.1017/9781108231596

Bernoulli Forthcoming Papers

GUHA, A., HO, N. and NGUYEN, X.
On posterior contraction of parameters and interpretability in Bayesian mixture modeling

BABIĆ, S., GELBGRAS, L., HALLIN, M. and LEY, C.
Optimal tests for elliptical symmetry: Specified and unspecified location

KASPER, T.M., FUCHS, S. and TRUTSCHNIG, W.
On weak conditional convergence of bivariate Archimedean and Extreme Value copulas, and consequences to nonparametric estimation

PAGE, S. and GRÜNEWÄLDER, S.
The Goldenshluger–Lepski method for constrained least-squares estimators over RKHSs

KLEBANOV, I., SPRUNGK, B. and SULLIVAN, T.J.
The linear conditional expectation in Hilbert space

QIU, H., LUEDTKE, A. and CARONE, M.
Universal sieve-based strategies for efficient estimation using machine learning tools

JIANG, D. and BAI, Z.
Partial generalized four moment theorem revisited

PETE, G. and TIMÁR, Á.
Finite-energy infinite clusters without anchored expansion

KRELL, N. and SCHMISSER, É.
Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes

CHZHEN, E., DENIS, C. and HEBIRI, M.
Minimax semi-supervised set-valued approach to multi-class classification

GONZÁLEZ CÁZARES, J. and IVANOVS, J.
Recovering Brownian and jump parts from high-frequency observations of a Lévy process

BARBER, R.F. and SAMWORTH, R.J.
Local continuity of log-concave projection, with applications to estimation under model misspecification

GANESAN, G.
Minimum spanning trees of random geometric graphs with location dependent weights

BHUDISAKSANG, T. and CARTEA, Á.
Online drift estimation for jump-diffusion processes

CHAUDHURI, R., JAIN, V. and PILLAI, N.S.
Universality and least singular values of random matrix products: A simplified approach

Continues

Bernoulli Forthcoming Papers—*Continued*

NING, N., IONIDES, E.L. and RITOV, Y.
Scalable Monte Carlo inference and rescaled local asymptotic normality

WANG, D. and TANG, C.-F.
Testing against uniform stochastic ordering with paired observations

KOHLER, M. and KRZYŻAK, A.
Over-parametrized deep neural networks do not generalize well

BORDA, B.
Equidistribution of random walks on compact groups II. The Wasserstein metric

ČERNÝ, A. and RUF, J.
Pure-jump semimartingales

DUVAL, C. and MARIUCCI, E.
Spectral-free estimation of Lévy densities in high-frequency regime

DENIS, C., DION-BLANC, C. and MARTINEZ, M.
A ridge estimator of the drift from discrete repeated observations of the solution of a stochastic differential equation

ALTMEYER, R.
Approximation of occupation time functionals

KWAK, J., LEE, J.O. and PARK, J.
Extremal eigenvalues of sample covariance matrices with general population

XU, W.
Asymptotic results for heavy-tailed Lévy processes and their exponential functionals

LI, X.-Y., YE, Z.-S. and TANG, C.Y.
Estimating the inter-occurrence time distribution from superposed renewal processes

BRAUNSTEINS, P., HAUTPHENNE, S. and MINUESA, C.
Parameter estimation in branching processes with almost sure extinction

BONNEFONT, M. and JOULIN, A.
A note on eigenvalues estimates for one-dimensional diffusion operators

HAN, Q.
Multiplier U-processes: Sharp bounds and applications

ERNY, X., LÖCHERBACH, E. and LOUKIANOVA, D.
Mean field limits for interacting Hawkes processes in a diffusive regime

IACOBELLI, G., RIBEIRO, R., VALLE, G. and ZUAZNÁBAR, L.
Tree builder random walk: Recurrence, transience and ballisticity

CHÉRIEF-ABDELLATIF, B.-E. and ALQUIER, P.
Finite sample properties of parametric MMD estimation: Robustness to misspecification and dependence

Continues

Bernoulli Forthcoming Papers—*Continued*

KIM, K. and YI, J.
Limit theorems for time-dependent averages of nonlinear stochastic heat equations

GROISMAN, P., JONCKHEERE, M. and SAPIENZA, F.
Nonhomogeneous Euclidean first-passage percolation and distance learning

LEE, C.Y.
The Hausdorff measure of the range and level sets of Gaussian random fields with sectorial local nondeterminism

HAYASHI, M.
Information geometry approach to parameter estimation in hidden Markov model

BACKHOFF-VERAGUAS, J. and PAMMER, G.
Applications of weak transport theory

LEDOIT, O. and WOLF, M.
Quadratic shrinkage for large covariance matrices

VERGARA, R.C., ALLARD, D. and DESASSIS, N.
A general framework for SPDE-based stationary random fields

SHAO, Q.M. and ZHANG, Z.-S.
Berry–Esseen bounds for multivariate nonlinear statistics with applications to M-estimators and stochastic gradient descent algorithms

ELTZNER, B.
Geometrical smeariness – a new phenomenon of Fréchet means

MOU, W., FLAMMARION, N., WAINWRIGHT, M.J. and BARTLETT, P.J.
Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity

SUN, X., XIE, L. and XIE, Y.
Strong and weak convergence rates for slow-fast stochastic differential equations driven by α-stable process

BENHENNI, K., GIRARD, D.A. and LOUHICHI, S.
On bandwidth selection problems in nonparametric trend estimation under martingale difference errors

BELLEC, P.C. and ZHANG, C.-H.
De-biasing the Lasso with degrees-of-freedom adjustment

BURACZEWSKI, D. and DAMEK, E.
Limit theorems for supercritical branching processes in random environment

EL NAMIER, R., LUSCHGY, H. and PAGÉS, G.
New approach to greedy vector quantization

PHANDOIDAEN, N. and RICHTER, S.
Empirical process theory for locally stationary processes

Continues

Bernoulli Forthcoming Papers—*Continued*

WANG, Y. and POLITIS, D.
Model-free bootstrap for a general class of stationary time series

COSSO, A. and RUSSO, F.
Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation

GAUTIER, E.
Adaptive estimation in the linear random coefficients model when regressors have limited variation

CAO, H. and WU, W.B.
Testing and estimation for clustered signals

GRAHOVAC, D., LEONENKO, N.N. and TAQQU, M.S.
Growth and intermittency of supOU processes

FRANCQ, C. and ZAKOIAN, J.-M.
Adaptiveness of the empirical distribution of residuals in semi-parametric conditional location scale models

LAM-WEIL, J., LAURENT, B. and LOUBES, J.-M.
Minimax optimal goodness-of-fit testing for densities and multinomials under a local differential privacy constraint

GAYNANOVA, I.
Erratum for prediction and estimation consistency of sparse multi-class penalized optimal scoring

LI, W., WANG, Q., YAO, J. and ZHOU, W.
On eigenvalues of a high-dimensional spatial-sign covariance matrix

LAU, J.W. and CRIPPS, E.
Thinned completely random measures with applications in competing risks models

BHAUMIK, P., SHI, W. and GHOSAL, S.
Bayesian generalized regression in partial differential equation models

ROGERS, D. and WINKEL, M.
A Ray–Knight representation of up-down Chinese Restaurants

HUANG, Q., DUAN, J. and SONG, R.
Homogenization of nonlocal partial differential equations related to stochastic differential equations with Lévy noise