BERNOULLI

*Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*

Volume Twenty Five Number Four Part A November 2019 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

*Official Journal of the Bernoulli Society for Mathematical Statistics and Probability*

Volume Twenty Five Number Four Part A November 2019 ISSN: 1350-7265


CONTENTS

*(continued)*

BERNOULLI Volume 25 Number 4A November 2019 Pages 2409–3201 ISI/BS

Volume 25 Number 4A November 2019 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**


---

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 2019 Bernoulli consists of 4 issues published in February, May, August and November.

Bernoulli Society
for Mathematical Statistics
and Probability

*Bernoulli* **25**(4A), 2019, 2409–2438
https://doi.org/10.3150/18-BEJ1058

# Scaling limit of random forests with prescribed degree sequences

TAO LEI[1]

[1]*805 Sherbrooke Street West, Montréal, Québec, Canada H3A 0B9. E-mail: tao.lei@mail.mcgill.ca*

In this paper, we consider the random plane forest uniformly drawn from all possible plane forests with a given degree sequence. Under suitable conditions on the degree sequences, we consider the limit of a sequence of such forests with the number of vertices tends to infinity in terms of Gromov–Hausdorff–Prokhorov topology. This work falls into the general framework of showing convergence of random combinatorial structures to certain Gromov–Hausdorff scaling limits, described in terms of the *Brownian Continuum Random Tree* (BCRT), pioneered by the work of Aldous (*Ann. Probab.* **19** (1991) 1–28; In *Stochastic Analysis (Durham, 1990)* (1991) 23–70 Cambridge Univ. Press; *Ann. Probab.* **21** (1993) 248–289). In fact, we identify the limiting random object as a sequence of random real trees encoded by excursions of some first passage bridges reflected at minimum. We establish such convergence by studying the associated Lukasiewicz walk of the degree sequences. In particular, our work is closely related to and uses the results from the recent work of Broutin and Marckert (*Random Structures Algorithms* **44** (2014) 290–316) on scaling limit of random trees with prescribed degree sequences, and the work of Addario-Berry (*Random Structures Algorithms* **41** (2012) 253–261) on tail bounds of the height of a random tree with prescribed degree sequence.

*Keywords:* first passage bridge; Gromov–Hausdorff–Prokhorov distance; random forests

## References

[1] Abraham, R., Delmas, J.-F. and Hoscheit, P. (2013). A note on the Gromov–Hausdorff–Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.* **18** 1–21. MR3035742
[2] Addario-Berry, L. (2012). Tail bounds for the height and width of a random tree with a given degree sequence. *Random Structures Algorithms* **41** 253–261. MR2956056
[3] Addario-Berry, L., Broutin, N. and Goldschmidt, C. (2010). Critical random graphs: Limiting constructions and distributional properties. *Electron. J. Probab.* **15** 741–775. MR2650781
[4] Aldous, D. (1991). The continuum random tree. I. *Ann. Probab.* **19** 1–28. MR1085326
[5] Aldous, D. (1991). The continuum random tree. II. An overview. In *Stochastic Analysis* (*Durham*, 1990). *London Mathematical Society Lecture Note Series* **167** 23–70. Cambridge: Cambridge Univ. Press. MR1166406
[6] Aldous, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226
[7] Aldous, D. and Pitman, J. (1998). The standard additive coalescent. *Ann. Probab.* **26** 1703–1726. MR1675063
[8] Aldous, D.J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII*—1983. *Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. MR0883646
[9] Bender, E.A. and Canfield, E.R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307. MR0505796

1350-7265 

[10] Bertoin, J., Chaumont, L.C. and Pitman, J. (2003). Path transformations of first passage bridges. *Electron. Commun. Probab.* **8** 155–166. MR2042754
[11] Bettinelli, J. (2010). Scaling limits for random quadrangulations of positive genus. *Electron. J. Probab.* **15** 1594–1644. MR2735376
[12] Bettinelli, J. (2015). Scaling limit of random planar quadrangulations with a boundary. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 432–477. MR3335010
[13] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics: Probability and Statistics*. New York: Wiley. MR1700749
[14] Bollobás, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929
[15] Broutin, N. and Marckert, J.-F. (2014). Asymptotics of trees with a prescribed degree sequence and applications. *Random Structures Algorithms* **44** 290–316. MR3188597
[16] Burago, D., Burago, Y. and Ivanov, S. (2001). *A Course in Metric Geometry*. *Graduate Studies in Mathematics* **33**. Providence, RI: Amer. Math. Soc. MR1835418
[17] Chassaing, P. and Louchard, G. (2002). Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms* **21** 76–119. MR1913079
[18] Duquesne, T. and Le Gall, J.-F. (2002). Random trees, Lévy processes and spatial branching processes. *Astérisque* **281** vi+147. MR1954248
[19] Evans, S.N. (2008). *Probability and Real Trees*. *Lecture Notes in Math.* **1920**. Berlin: Springer. MR2351587
[20] Fitzsimmons, P., Pitman, J. and Yor, M. (1993). Markovian bridges: Construction, Palm interpretation, and splicing. In *Seminar on Stochastic Processes*, 1992 (*Seattle*, *WA*, 1992). *Progress in Probability* **33** 101–134. Birkhäuser, Boston, MA. MR1278079
[21] Janson, S. and Stefánsson, S.Ö. (2015). Scaling limits of random planar maps with a unique large face. *Ann. Probab.* **43** 1045–1081. MR3342658
[22] Le Gall, J.-F. (2005). Random trees and applications. *Probab. Surv.* **2** 245–311. MR2203728
[23] Lei, T. (2018). Supplement to "Scaling limit of random forests with prescribed degree sequences." DOI:10.3150/18-BEJ1058SUPP.
[24] Marckert, J.-F. and Miermont, G. (2011). The CRT is the scaling limit of unordered binary trees. *Random Structures Algorithms* **38** 467–501. MR2829313
[25] McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*. *Algorithms Combin.* **16** 195–248. Berlin: Springer. MR1678578
[26] Miermont, G. (2008). Invariance principles for spatial multitype Galton–Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** 1128–1161. MR2469338
[27] Mörters, P. and Peres, Y. (2010). *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics* **30**. Cambridge: Cambridge Univ. Press. MR2604525
[28] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Berlin: Springer. MR2245368

*Bernoulli* **25**(4A), 2019, 2439–2478
https://doi.org/10.3150/18-BEJ1059

# Stationary distributions and convergence for Walsh diffusions

TOMOYUKI ICHIBA[1] and ANDREY SARANTSEV[2]

[1]*Department of Statistics and Applied Probability, University of California, Santa Barbara, South Hall 5607A, Santa Barbara, CA, 93106, USA. E-mail: ichiba@pstat.ucsb.edu*
[2]*Department of Mathematics and Statistics, University of Nevada, Reno, 1664 N Virginia St, Reno, NV, 89557, USA. E-mail: asarantsev@unr.edu*

A Walsh diffusion on Euclidean space moves along each ray from the origin, as a solution to a stochastic differential equation with certain drift and diffusion coefficients, as long as it stays away from the origin. As it hits the origin, it instantaneously chooses a new direction according to a given probability law, called the spinning measure. A special example is a real-valued diffusion with skew reflections at the origin. This process continuously (in the weak sense) depends on the spinning measure. We determine a stationary measure for such process, explore long-term convergence to this distribution and establish an explicit rate of exponential convergence.

*Keywords:* ergodic process; invariant measure; Lyapunov function; reflected diffusion; stationary distribution; stochastic differential equation; Walsh Brownian motion; Walsh diffusion

## References

[1] Bakry, D., Cattiaux, P. and Guillin, A. (2008). Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.* **254** 727–759. MR2381160
[2] Barlow, M., Pitman, J. and Yor, M. (1989). On Walsh's Brownian motions. In *Séminaire de Probabilités, XXIII. Lecture Notes in Math.* **1372** 275–293. Berlin: Springer. MR1022917
[3] 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
[4] Chen, Z.-Q. and Fukushima, M. (2015). One-point reflection. *Stochastic Process. Appl.* **125** 1368–1393. MR3310351
[5] Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli* **21** 505–536. MR3322329
[6] Davies, P.L. (1986). Rates of convergence to the stationary distribution for $k$-dimensional diffusion processes. *J. Appl. Probab.* **23** 370–384. MR0839992
[7] Douc, R., Fort, G. and Guillin, A. (2009). Subgeometric rates of convergence of $f$-ergodic strong Markov processes. *Stochastic Process. Appl.* **119** 897–923. MR2499863
[8] Down, D., Meyn, S.P. and Tweedie, R.L. (1995). Exponential and uniform ergodicity of Markov processes. *Ann. Probab.* **23** 1671–1691. MR1379163
[9] Evans, S.N. and Sowers, R.B. (2003). Pinching and twisting Markov processes. *Ann. Probab.* **31** 486–527. MR1959800
[10] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications. Vol. I.* New York–London–Sydney: Wiley. MR0228020

[11] Fitzsimmons, P.J. and Kuter, K.E. (2014). Harmonic functions on Walsh's Brownian motion. *Stochastic Process. Appl.* **124** 2228–2248. MR3188354

[12] 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

[13] Freidlin, M.I. and Wentzell, A.D. (1993). Diffusion processes on graphs and the averaging principle. *Ann. Probab.* **21** 2215–2245. MR1245308

[14] Hajri, H. (2011). Stochastic flows related to Walsh Brownian motion. *Electron. J. Probab.* **16** 1563–1599. MR2835247

[15] Hajri, H. and Touhami, W. (2014). Itô's formula for Walsh's Brownian motion and applications. *Statist. Probab. Lett.* **87** 48–53. MR3168934

[16] Ichiba, T., Karatzas, I., Prokaj, V. and Yan, M. (2018). Stochastic integral equations for Walsh semimartingales. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** 726–756. MR3795064

[17] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940

[18] Karatzas, I. and Yan, M. (2016). Semimartingales on rays, Walsh diffusions, and related problems of control and stopping. Available at arXiv:1603.03124.

[19] Lund, R.B., Meyn, S.P. and Tweedie, R.L. (1996). Computable exponential convergence rates for stochastically ordered Markov processes. *Ann. Appl. Probab.* **6** 218–237. MR1389838

[20] 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

[21] 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

[22] Meyn, S.P. and Tweedie, R.L. (1994). Computable bounds for geometric convergence rates of Markov chains. *Ann. Appl. Probab.* **4** 981–1011. MR1304770

[23] Picard, J. (2005). Stochastic calculus and martingales on trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** 631–683. MR2144228

[24] Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. Chichester: Wiley. MR1105086

[25] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **293**. Berlin: Springer. MR1725357

[26] Roberts, G.O. and Rosenthal, J.S. (1996). Quantitative bounds for convergence rates of continuous time Markov processes. *Electron. J. Probab.* **1** no. 9, approx. 21 pp. MR1423462

[27] Roberts, G.O. and Tweedie, R.L. (1999). Bounds on regeneration times and convergence rates for Markov chains. *Stochastic Process. Appl.* **80** 211–229. MR1682243

[28] Roberts, G.O. and Tweedie, R.L. (2000). Rates of convergence of stochastically monotone and continuous time Markov models. *J. Appl. Probab.* **37** 359–373. MR1780996

[29] Sarantsev, A. (2016). Explicit rates of exponential convergence for reflected jump-diffusions on the half-line. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 1069–1093. MR3580089

[30] Sarantsev, A. (2017). Penalty method for obliquely reflected diffusions. Available at arXiv:1509.01777.

[31] Sarantsev, A. (2017). Reflected Brownian motion in a convex polyhedral cone: Tail estimates for the stationary distribution. *J. Theoret. Probab.* **30** 1200–1223. MR3687255

[32] Tsirelson, B. (1997). Triple points: From non-Brownian filtrations to harmonic measures. *Geom. Funct. Anal.* **7** 1096–1142. MR1487755

[33] Villani, C. (2009). *Optimal Transport. Old and New*. *A Series of Comprehensive Studies in Mathematics* **338**. Springer.

[34] Walsh, J. (1978). A diffusion with a discontinuous local time. *Astérisque* **52–53** (53) 37–45.

[35] Watanabe, S. (1999). The existence of a multiple spider martingale in the natural filtration of a certain diffusion in the plane. In *Séminaire de Probabilités, XXXIII*. *Lecture Notes in Math*. **1709** 277–290. Berlin: Springer. MR1768001

*Bernoulli* **25**(4A), 2019, 2479–2507
https://doi.org/10.3150/18-BEJ1060

# The eternal multiplicative coalescent encoding via excursions of Lévy-type processes

VLADA LIMIC

*IRMA, UMR 7501 de l'Université de Strasbourg et du CNRS, 7 rue René Descartes, 67084 Strasbourg Cedex, France. E-mail: vlada@math.unistra.fr*

The multiplicative coalescent is a mean-field Markov process in which any pair of blocks coalesces at rate proportional to the product of their masses. In Aldous and Limic (*Electron. J. Probab.* **3** (1998) Paper no. 3) each extreme eternal version of the multiplicative coalescent was described in three different ways, one of which matched its (marginal) law to that of the ordered excursion lengths above past minima of a certain Lévy-type process.

Using a modification of the breadth-first-walk construction from Aldous (*Ann. Probab.* **25** (1997) 812–854) and Aldous and Limic (*Electron. J. Probab.* **3** (1998) Paper no. 3), and some new insight from the thesis by Uribe Bravo (Markovian bridges, Brownian excursions, and stochastic fragmentation and coalescence (2007) UNAM), this work settles an open problem (3) from Aldous (*Ann. Probab.* **25** (1997) 812–854) in the more general context of Aldous and Limic (*Electron. J. Probab.* **3** (1998) Paper no. 3). Informally speaking, each eternal version is entirely encoded by its Lévy-type process, and contrary to Aldous' original intuition, the time for the multiplicative coalescent does correspond to the linear increase in the constant part of the drift of the Lévy-type process. In the "standard multiplicative coalescent" context of Aldous (*Ann. Probab.* **25** (1997) 812–854), this result was first announced by Armendáriz in 2001, while its first published proof is due to Broutin and Marckert (*Probab. Theory Related Fields* **166** (2016) 515–552), who simultaneously account for the process of excess (or surplus) edge counts.

*Keywords:* entrance law; excursion; Lévy process; multiplicative coalescent; near-critical; random graph; stochastic coalescent

## References

[1] Addario-Berry, L., Broutin, N. and Goldschmidt, C. (2012). The continuum limit of critical random graphs. *Probab. Theory Related Fields* **152** 367–406. MR2892951

[2] Addario-Berry, L., Broutin, N., Goldschmidt, C. and Miermont, G. Continuum Erdős-Rényi and Kruskal dynamics. Work in progress.

[3] Addario-Berry, L., Broutin, N., Goldschmidt, C. and Miermont, G. (2017). The scaling limit of the minimum spanning tree of the complete graph. *Ann. Probab.* **45** 3075–3144. Preprint. Available at arXiv:1301.1664. MR3706739

[4] Aldous, D. (1993). The continuum random tree. III. *Ann. Probab.* **21** 248–289. MR1207226

[5] Aldous, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854. MR1434128

[6] Aldous, D. and Limic, V. (1998). The entrance boundary of the multiplicative coalescent. *Electron. J. Probab.* **3** Paper no. 3, 59 pp. MR1491528

[7] Aldous, D. and Pitman, J. (1998). The standard additive coalescent. *Ann. Probab.* **26** 1703–1726. MR1675063

[8] Aldous, D. and Pitman, J. (2000). Inhomogeneous continuum random trees and the entrance boundary of the additive coalescent. *Probab. Theory Related Fields* **118** 455–482. MR1808372

[9] Aldous, D.J. (1999). Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli* **5** 3–48. MR1673235

[10] Aldous, D.J. and Pittel, B. (2000). On a random graph with immigrating vertices: Emergence of the giant component. *Random Structures Algorithms* **17** 79–102. MR1774745

[11] Armendáriz, I. (2001). Brownian excursions and coalescing particle systems. Ph.D. thesis, NY University. Available at http://proxy.library.nyu.edu/login?url=https://search.proquest.com/docview/275866769?accountid=12768.

[12] Barton, N.H., Etheridge, A.M. and Véber, A. (2010). A new model for evolution in a spatial continuum. *Electron. J. Probab.* **15** 162–216. MR2594876

[13] Berestycki, N. (2009). *Recent Progress in Coalescent Theory*. *Ensaios Matemáticos* [*Mathematical Surveys*] **16**. Rio de Janeiro: Sociedade Brasileira de Matemática. MR2574323

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

[15] Bertoin, J. (2000). A fragmentation process connected to Brownian motion. *Probab. Theory Related Fields* **117** 289–301. MR1771665

[16] Bertoin, J. (2001). Eternal additive coalescents and certain bridges with exchangeable increments. *Ann. Probab.* **29** 344–360. MR1825153

[17] Bertoin, J. (2006). *Random Fragmentation and Coagulation Processes*. *Cambridge Studies in Advanced Mathematics* **102**. Cambridge: Cambridge Univ. Press. MR2253162

[18] Bertoin, J. (2009). Two solvable systems of coagulation equations with limited aggregations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** 2073–2089. MR2569886

[19] Bertoin, J. and Le Gall, J.-F. (2003). Stochastic flows associated to coalescent processes. *Probab. Theory Related Fields* **126** 261–288. MR1990057

[20] Bhamidi, S., Budhiraja, A. and Wang, X. (2014). The augmented multiplicative coalescent, bounded size rules and critical dynamics of random graphs. *Probab. Theory Related Fields* **160** 733–796. MR3278920

[21] Bhamidi, S., Budhiraja, A. and Wang, X. (2015). Aggregation models with limited choice and the multiplicative coalescent. *Random Structures Algorithms* **46** 55–116. MR3291294

[22] Bhamidi, S., van der Hofstad, R. and Sen, S. (2018). The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. *Probab. Theory Related Fields* **170** 387–474. Preprint. Available at arXiv:1508.04645. MR3748328

[23] Bhamidi, S., van der Hofstad, R. and van Leeuwaarden, J.S.H. (2010). Scaling limits for critical inhomogeneous random graphs with finite third moments. *Electron. J. Probab.* **15** 1682–1703. MR2735378

[24] Bhamidi, S., van der Hofstad, R. and van Leeuwaarden, J.S.H. (2012). Novel scaling limits for critical inhomogeneous random graphs. *Ann. Probab.* **40** 2299–2361. MR3050505

[25] Bollobás, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **73**. Cambridge: Cambridge Univ. Press. MR1864966

[26] Broutin, N., Duquesne, T. and Wang, M. (2018). Limits of multiplicative inhomogeneous random graphs and Lévy trees. Preprint. Available at arXiv:1804.05871.

[27] Broutin, N. and Marckert, J.-F. (2016). A new encoding of coalescent processes: Applications to the additive and multiplicative cases. *Probab. Theory Related Fields* **166** 515–552. Preprint. Available at arXiv:1409.4266. MR3547745

[28] Chassaing, P. and Louchard, G. (2002). Phase transition for parking blocks, Brownian excursion and coalescence. *Random Structures Algorithms* **21** 76–119. MR1913079
[29] Donnelly, P. and Kurtz, T.G. (1999). Particle representations for measure-valued population models. *Ann. Probab.* **27** 166–205. MR1681126
[30] Durrett, R. (2007). *Random Graph Dynamics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **20**. Cambridge: Cambridge Univ. Press. MR2271734
[31] Dynkin, E.B. (1978). Sufficient statistics and extreme points. *Ann. Probab.* **6** 705–730. MR0518321
[32] Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Magy. Tud. Akad. Mat. Kut. Intéz. Közl.* **5** 17–61. Reprinted in *Paul Erdős: The Art of Counting*, MIT Press, 574–618 (1973); and in *Selected Papers of Alfréd Rényi*, Akadémiai Kiadó, 482–525 (1976).
[33] 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
[34] Evans, S.N. and Pitman, J. (1998). Construction of Markovian coalescents. *Ann. Inst. Henri Poincaré Probab. Stat.* **34** 339–383. MR1625867
[35] Fournier, N. (2006). Standard stochastic coalescence with sum kernels. *Electron. Commun. Probab.* **11** 141–148. MR2240707
[36] Fournier, N. (2006). On some stochastic coalescents. *Probab. Theory Related Fields* **136** 509–523. MR2257133
[37] Fournier, N. and Laurençot, P. (2006). Well-posedness of Smoluchowski's coagulation equation for a class of homogeneous kernels. *J. Funct. Anal.* **233** 351–379. MR2214581
[38] Fournier, N. and Laurençot, P. (2006). Local properties of self-similar solutions to Smoluchowski's coagulation equation with sum kernels. *Proc. Roy. Soc. Edinburgh Sect. A* **136** 485–508. MR2227805
[39] Freeman, N. (2015). The segregated $\Lambda$-coalescent. *Ann. Probab.* **43** 435–467. MR3305997
[40] Gumbel, E.J. (1958). *Statistics of Extremes*. New York: Columbia Univ. Press. MR0096342
[41] Jeon, I. (1998). Existence of gelling solutions for coagulation-fragmentation equations. *Comm. Math. Phys.* **194** 541–567. MR1631473
[42] Kallenberg, O. (1983). *Random Measures*. Berlin: Akademie-Verlag.
[43] Kingman, J.F.C. (1982). The coalescent. *Stochastic Process. Appl.* **13** 235–248. MR0671034
[44] Kingman, J.F.C. (1982). On the genealogy of large populations. Essays in statistical science. *J. Appl. Probab.* Special Vol. **19A** 27–43. MR0633178
[45] Limic, V. (2019). Supplement to "The eternal multiplicative coalescent encoding via excursions of Lévy-type processes." DOI:10.3150/18-BEJ1060SUPP.
[46] Limic, V. (1998). Properties of the multiplicative coalescent. Ph.D. thesis, University of California, Berkeley. MR2697979
[47] Limic, V. (2016). Eternal multiplicative coalescent is encoded by its Lévy-type processes. Preprint. Available at arXiv:1601.01325.
[48] Limic, V. (2017). A playful note on spanning and surplus edges. Preprint. Available at arXiv:1703.02574.
[49] Limic, V. and Sturm, A. (2006). The spatial $\Lambda$-coalescent. *Electron. J. Probab.* **11** 363–393. MR2223040
[50] Martin, J.B. and Ráth, B. (2017). Rigid representations of the multiplicative coalescent with linear deletion. *Electron. J. Probab.* **22** Paper no. 83, 47 pp. MR3718711
[51] Merle, M. and Normand, R. (2014). Self-organized criticality in a discrete model for Smoluchowski's equation. Preprint. Available at arXiv:1410.8338.
[52] Merle, M. and Normand, R. (2015). Self-organized criticality in a discrete model for Smoluchowski's equation with limited aggregations. Preprint. Available at arXiv:1509.00934.

[53] Möhle, M. and Sagitov, S. (2001). A classification of coalescent processes for haploid exchangeable population models. *Ann*. *Probab*. **29** 1547–1562. MR1880231
[54] Norris, J.R. (1999). Smoluchowski's coagulation equation: Uniqueness, nonuniqueness and a hydrodynamic limit for the stochastic coalescent. *Ann*. *Appl*. *Probab*. **9** 78–109. MR1682596
[55] Pitman, J. (1999). Coalescents with multiple collisions. *Ann*. *Probab*. **27** 1870–1902. MR1742892
[56] Pitman, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math*. **1875**. Berlin: Springer. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. With a foreword by Jean Picard. MR2245368
[57] Riordan, O. (2012). The phase transition in the configuration model. *Combin*. *Probab*. *Comput*. **21** 265–299. MR2900063
[58] Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *J*. *Appl*. *Probab*. **36** 1116–1125. MR1742154
[59] Schweinsberg, J. (2000). Coalescents with simultaneous multiple collisions. *Electron*. *J*. *Probab*. **5** Paper no. 12, 50 pp. MR1781024
[60] Schweinsberg, J. and Durrett, R. (2005). Random partitions approximating the coalescence of lineages during a selective sweep. *Ann*. *Appl*. *Probab*. **15** 1591–1651. MR2152239
[61] Turova, T.S. (2013). Diffusion approximation for the components in critical inhomogeneous random graphs of rank 1. *Random Structures Algorithms* **43** 486–539. MR3124693
[62] Uribe Bravo, G. Coding multiplicative coalescence by an inhomogeneous random walk. Work in progress.
[63] Uribe Bravo, G. (2007). Markovian bridges, Brownian excursions, and stochastic fragmentation and coalescence. Ph.D. thesis, UNAM. Available at http://132.248.9.195/pd2007/0621617/Index.html.

*Bernoulli* **25**(4A), 2019, 2508–2537
https://doi.org/10.3150/18-BEJ1061

# Semiparametric estimation for isotropic max-stable space-time processes

SVEN BUHL[1,*], RICHARD A. DAVIS[2,**], CLAUDIA KLÜPPELBERG[1,†] and CHRISTINA STEINKOHL[1,‡]

[1]*Center for Mathematical Sciences and TUM Institute of Advanced Study, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany. E-mail:* [*]*sven.buhl@tum.de;* [†]*cklu@tum.de;* [‡]*christina.steinkohl@gmail.com, url: http://www.statistics.ma.tum.de*

[2]*Department of Statistics, Columbia University, 1255 Amsterdam Avenue, New York, NY 10027, USA. E-mail:* [**]*rdavis@stat.columbia.edu, url: http://www.stat.columbia.edu/~rdavis*

Regularly varying space-time processes have proved useful to study extremal dependence in space-time data. We propose a semiparametric estimation procedure based on a closed form expression of the extremogram to estimate parametric models of extremal dependence functions. We establish the asymptotic properties of the resulting parameter estimates and propose subsampling procedures to obtain asymptotically correct confidence intervals. A simulation study shows that the proposed procedure works well for moderate sample sizes and is robust to small departures from the underlying model. Finally, we apply this estimation procedure to fitting a max-stable process to radar rainfall measurements in a region in Florida. Complementary results and some proofs of key results are presented together with the simulation study in the supplement [Buhl et al. (2018)].

*Keywords:* Brown–Resnick process; extremogram; max-stable process; mixing; regular variation; semiparametric estimation; space-time process; subsampling

## References

[1] Andrews, D.W.K. (1999). Estimation when a parameter is on a boundary. *Econometrica* **67** 1341–1383. MR1720781
[2] Brown, B.M. and Resnick, S.I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14** 732–739. MR0517438
[3] Buhl, S. (2017). Statistical Modelling and Estimation of Space-Time Extremes. Dissertation, Technische Universität München, München.
[4] Buhl, S., Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2018). Supplement to "Semiparametric estimation for isotropic max-stable space-time processes." DOI:10.3150/18-BEJ1061SUPP.
[5] Buhl, S. and Klüppelberg, C. (2016). Anisotropic Brown–Resnick space-time processes: Estimation and model assessment. *Extremes* **19** 627–660. MR3558348
[6] Buhl, S. and Klüppelberg, C. (2017). Generalised least squares estimation of regularly varying space-time processes based on flexible observation schemes. Available at arXiv:1862319 [math.ST].
[7] Buhl, S. and Klüppelberg, C. (2018). Limit theory for the empirical extremogram of random fields. *Stochastic Process. Appl.* **128** 2060–2082. MR3797653
[8] Cho, Y.B., Davis, R.A. and Ghosh, S. (2016). Asymptotic properties of the empirical spatial extremogram. *Scand. J. Stat.* **43** 757–773. MR3543321

1350-7265 

[9] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2013). Max-stable processes for modeling extremes observed in space and time. *J. Korean Statist. Soc.* **42** 399–414. MR3255398
[10] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2013). Statistical inference for max-stable processes in space and time. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **75** 791–819. MR3124792
[11] Davis, R.A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15** 977–1009. MR2597580
[12] de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776
[13] Dombry, C., Engelke, S. and Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. MR3509888
[14] Einmahl, J.H.J., Kiriliouk, A. and Segers, J. (2018). A continuous updating weighted least squares estimator of tail dependence in high dimensions. *Extremes* **21** 205–233. MR3800297
[15] Engelke, S., Malinowski, A., Kabluchko, Z. and Schlather, M. (2015). Estimation of Hüsler–Reiss distributions and Brown–Resnick processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **77** 239–265. MR3299407
[16] Fasen, V., Klüppelberg, C. and Schlather, M. (2010). High-level dependence in time series models. *Extremes* **13** 1–33. MR2593949
[17] Genton, M.G., Ma, Y. and Sang, H. (2011). On the likelihood function of Gaussian max-stable processes. *Biometrika* **98** 481–488. MR2806443
[18] Giné, E., Hahn, M.G. and Vatan, P. (1990). Max-infinitely divisible and max-stable sample continuous processes. *Probab. Theory Related Fields* **87** 139–165. MR1080487
[19] Hult, H. and Lindskog, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.* **115** 249–274. MR2111194
[20] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math.* (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910
[21] Huser, R. and Davison, A.C. (2014). Space-time modelling of extreme events. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 439–461. MR3164873
[22] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440
[23] Lahiri, S.N., Lee, Y. and Cressie, N. (2002). On asymptotic distribution and asymptotic efficiency of least squares estimators of spatial variogram parameters. *J. Statist. Plann. Inference* **103** 65–85. MR1896984
[24] Padoan, S.A., Ribatet, M. and Sisson, S.A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105** 263–277. MR2757202
[25] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424
[26] Schlather, M. RandomFields, contributed package on random field simulation for R. http://cran.r-project.org/web/packages/RandomFields/.
[27] Steinkohl, C. (2013). Statistical Modelling of Extremes in Space and Time using Max-Stable Processes. Dissertation, Technische Universität München, München.
[28] Thibaud, E. and Opitz, T. (2015). Efficient inference and simulation for elliptical Pareto processes. *Biometrika* **102** 855–870. MR3431558
[29] Wadsworth, J.L. and Tawn, J.A. (2014). Efficient inference for spatial extreme value processes associated to log-Gaussian random functions. *Biometrika* **101** 1–15. MR3180654

*Bernoulli* **25**(4A), 2019, 2538–2563
https://doi.org/10.3150/18-BEJ1062

# Large ball probabilities, Gaussian comparison and anti-concentration

FRIEDRICH GÖTZE[1], ALEXEY NAUMOV[2,5,*], VLADIMIR SPOKOINY[2,4,5,6] and VLADIMIR ULYANOV[2,3,**]

[1]*Faculty of Mathematics, Bielefeld University, P.O. Box 10 01 31, 33501, Bielefeld, Germany. E-mail: goetze@math.uni-bielefeld.de*
[2]*National Research University Higher School of Economics, 20 Myasnitskaya ulitsa, 101000, Moscow, Russia. E-mail: *anaumov@hse.ru; **vulyanov@hse.ru*
[3]*Moscow State University, Leninskie Gory, 1, Moscow, Russia. E-mail: vulyanov@cs.msu.ru*
[4]*Skolkovo Institute of Science and Technology (Skoltech), Skolkovo Innovation Center, Building 3, 143026, Moscow, Russia*
[5]*Institute for Information Transmission Problems RAS, Bolshoy Karetny per. 19, bld.1, 127051, Moscow, Russia*
[6]*Weierstrass Institute, Mohrenstr. 39, 10117 Berlin, Germany. E-mail: spokoiny@wias-berlin.de*

We derive tight non-asymptotic bounds for the Kolmogorov distance between the probabilities of two Gaussian elements to hit a ball in a Hilbert space. The key property of these bounds is that they are dimension-free and depend on the nuclear (Schatten-one) norm of the difference between the covariance operators of the elements and on the norm of the mean shift. The obtained bounds significantly improve the bound based on Pinsker's inequality via the Kullback–Leibler divergence. We also establish an anti-concentration bound for a squared norm of a non-centered Gaussian element in Hilbert space. The paper presents a number of examples motivating our results and applications of the obtained bounds to statistical inference and to high-dimensional CLT.

*Keywords:* dimension free bounds; Gaussian anti-concentration inequalities; Gaussian comparison; high-dimensional CLT; high-dimensional inference; Schatten norm

## References

[1] Ball, K. (1993). The reverse isoperimetric problem for Gaussian measure. *Discrete Comput. Geom.* **10** 411–420. MR1243336
[2] Barsov, S.S. and Ulyanov, V.V. (1987). Estimates of the proximity of Gaussian measures. *Soviet Math. Dokl.* **34** 462–466. MR866375
[3] Belitser, E. (2017). On coverage and local radial rates of credible sets. *Ann. Statist.* **45** 1124–1151. MR3662450
[4] Bentkus, V. (2003). On the dependence of the Berry–Esseen bound on dimension. *J. Statist. Plann. Inference* **113** 385–402. MR1965117
[5] Bentkus, V. (2005). A Lyapunov-type bound in $R^d$. *Theory Probab. Appl.* **49** 311–323. MR2144310
[6] Bobkov, S.G. and Chistyakov, G.P. (2014). Bounds for the maximum of the density of the sum of independent random variables. *J. Math. Sci.* **199** 100–106. MR3032208
[7] Bobkov, S.G., Chistyakov, G.P. and Götze, F. (2014). Berry–Esseen bounds in the entropic central limit theorem. *Probab. Theory Related Fields* **159** 435–478. MR3230000

1350-7265 

[8] Bogachev, V.I. (1998). *Gaussian Measures*. *Mathematical Surveys and Monographs* **62**. Providence, RI: Amer. Math. Soc. MR1642391

[9] Bontemps, D. (2011). Bernstein–von Mises theorems for Gaussian regression with increasing number of regressors. *Ann*. *Statist*. **39** 2557–2584. MR2906878

[10] Castillo, I. (2012). A semiparametric Bernstein–von Mises theorem for Gaussian process priors. *Probab*. *Theory Related Fields* **152** 53–99. MR2875753

[11] Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann*. *Statist*. **41** 1999–2028. MR3127856

[12] Christoph, G., Prokhorov, Y. and Ulyanov, V. (1996). On distribution of quadratic forms in Gaussian random variables. *Theory Probab*. *Appl*. **40** 250–260. MR1346468

[13] Chung, K.L. (2001). *A Course in Probability Theory*, 3rd ed. San Diego, CA: Academic Press. MR1796326

[14] Götze, F. and Zaitsev, A.Yu. (2018). New applications of Arak's inequalities to the Littlewood–Offord problem. *Eur*. *J*. *Math*. **4** 639–663. MR3799161

[15] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. *Vol*. 1, 2nd ed. *Wiley Series in Probability and Mathematical Statistics*: *Applied Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1299979

[16] Johnstone, I.M. (2010). High dimensional Bernstein–von Mises: Simple examples. In *Borrowing Strength*: *Theory Powering Applications—A Festschrift for Lawrence D*. *Brown*. *Inst*. *Math*. *Stat*. (*IMS*) *Collect*. **6** 87–98. Beachwood, OH: IMS. MR2798513

[17] Ledoux, M. and Talagrand, M. (2002). *Probability in Banach Spaces*. *Isoperimetry and Processes*. Berlin: Springer. MR2814399

[18] Li, W.V. and Shao, Q.-M. (2001). Gaussian processes: Inequalities, small ball probabilities and applications. In *Stochastic Processes*: *Theory and Methods*. *Handbook of Statist*. **19** 533–597. Amsterdam: North-Holland. MR1861734

[19] Lifshits, M. (2012). *Lectures on Gaussian Processes*. *SpringerBriefs in Mathematics*. Heidelberg: Springer. MR3024389

[20] Markus, A.S. (1964). The eigen- and singular values of the sum and product of linear operators. *Russian Math*. *Surveys* **19** 91–120. MR0169063

[21] Naumov, A.A., Spokoiny, V.G., Tavyrikov, Yu.E. and Ulyanov, V.V. (2018). Nonasymptotic estimates for the closeness of Gaussian measures on balls. *Dokl*. *Math*. **98**.

[22] Naumov, A.A., Spokoiny, V.G. and Ulyanov, V.V. (2018). Confidence sets for spectral projectors of covariance matrices. *Dokl*. *Math*. **98**.

[23] Panov, M. and Spokoiny, V. (2015). Finite sample Bernstein–von Mises theorem for semiparametric problems. *Bayesian Anal*. **10** 665–710. MR3420819

[24] Prokhorov, Y.V. and Ulyanov, V.V. (2013). Some approximation problems in statistics and probability. In *Limit Theorems in Probability*, *Statistics and Number Theory*. *Springer Proc*. *Math*. *Stat*. **42** 235–249. Heidelberg: Springer. MR3079145

[25] Rudelson, M. and Vershynin, R. (2008). The Littlewood–Offord problem and invertibility of random matrices. *Adv*. *Math*. **218** 600–633. MR2407948

[26] Sazonov, V.V. (1981). *Normal Approximation—Some Recent Advances*. *Lecture Notes in Math*. **879**. Berlin–New York: Springer. MR0643968

[27] Spokoiny, V. (2017). Penalized maximum likelihood estimation and effective dimension. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **53** 389–429. MR3606746

[28] Spokoiny, V. and Zhilova, M. (2015). Bootstrap confidence sets under model misspecification. *Ann*. *Statist*. **43** 2653–2675. MR3405607

[29] Tropp, J.A. (2012). User-friendly tail bounds for sums of random matrices. *Found*. *Comput*. *Math*. **12** 389–434. MR2946459

[30] Ul'yanov, V. (1987). Asymptotic expansions for distributions of sums of independent random variables in H. *Theory Probab. Appl*. **31** 25–39. MR0836950
[31] Yurinsky, V. (1995). *Sums and Gaussian Vectors*. *Lecture Notes in Math*. **1617**. Berlin: Springer. MR1442713

*Bernoulli* **25**(4A), 2019, 2564–2596
https://doi.org/10.3150/18-BEJ1063

# Limit theorems with rate of convergence under sublinear expectations

XIAO FANG[1,*], SHIGE PENG[2], QI-MAN SHAO[1,**] and YONGSHENG SONG[3]

[1]*Department of Statistics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. E-mail:* [*]*xfang@sta.cuhk.edu.hk;* [**]*qmshao@sta.cuhk.edu.hk*
[2]*Institute of Mathematics, Shandong University, Jinan 250100, China. E-mail: peng@sdu.edu.cn*
[3]*Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. E-mail: yssong@amss.ac.cn*

Under the sublinear expectation $\mathbb{E}[\cdot] := \sup_{\theta\in\Theta} E_\theta[\cdot]$ for a given set of linear expectations $\{E_\theta : \theta \in \Theta\}$, we establish a new law of large numbers and a new central limit theorem with rate of convergence. We present some interesting special cases and discuss a related statistical inference problem. We also give an approximation and a representation of the $G$-normal distribution, which was used as the limit in Peng's (Law of large numbers and central limit theorem under nonlinear expectations (2007) Preprint) central limit theorem, in a probability space.

*Keywords:* central limit theorem; $G$-normal distribution; law of large numbers; rate of convergence; Stein's method; sublinear expectation

## References

[1] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk. *Math. Finance* **9** 203–228. MR1850791
[2] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. Heidelberg: Springer. MR2732624
[3] Chen, Z. and Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica* **70** 1403–1443. MR1929974
[4] Delbaen, F., Peng, S. and Rosazza Gianin, E. (2010). Representation of the penalty term of dynamic concave utilities. *Finance Stoch.* **14** 449–472. MR2670421
[5] Denis, L., Hu, M. and Peng, S. (2011). Function spaces and capacity related to a sublinear expectation: Application to $G$-Brownian motion paths. *Potential Anal.* **34** 139–161. MR2754968
[6] Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time*, extended ed. Berlin: de Gruyter. MR2779313
[7] Huber, P.J. (1981). *Robust Statistics*. New York: Wiley. MR0606374
[8] Jin, H. and Peng, S. (2016). Optimal unbiased estimation for maximal distribution. Preprint. Available at http://arxiv.org/abs/1611.07994v1.
[9] Lieberman, G.M. (1996). *Second Order Parabolic Differential Equations*. River Edge, NJ: World Scientific. MR1465184
[10] Maccheroni, F. and Marinacci, M. (2005). A strong law of large numbers for capacities. *Ann. Probab.* **33** 1171–1178. MR2135316
[11] Marinacci, M. (1999). Limit laws for non-additive probabilities and their frequentist interpretation. *J. Econom. Theory* **84** 145–195. MR1669509

[12] Peng, S. (1997). Backward SDE and related $g$-expectation. In *Backward Stochastic Differential Equations* (*Paris*, 1995–1996). *Pitman Res*. *Notes Math*. *Ser*. **364** 141–159. Harlow: Longman. MR1752680

[13] Peng, S. (2007). Law of large numbers and central limit theorem under nonlinear expectations. Preprint. Available at https://arxiv.org/abs/math/0702358.

[14] Peng, S. (2010). Nonlinear expectations and stochastic calculus under uncertainty. Preprint. Available at http://arxiv.org/abs/1002.4546v1.

[15] Röllin, A. (2018). On quantitative bounds in the mean martingale central limit theorem. *Statist*. *Probab*. *Lett*. **138** 171–176. MR3788734

[16] Song, Y. (2017). Normal approximation by Stein's method under sublinear expectations. Preprint. Available at http://arxiv.org/abs/1711.05384.

[17] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif*., 1970/1971). *Probability Theory* **II** 583–602. Berkeley, CA: Univ. California Press.

[18] Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. *Monographs on Statistics and Applied Probability* **42**. London: Chapman & Hall. MR1145491

*Bernoulli* **25**(4A), 2019, 2597–2619
https://doi.org/10.3150/18-BEJ1064

# Functional estimation and hypothesis testing in nonparametric boundary models

MARKUS REISS* and MARTIN WAHL**

*Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. E-mail:* *mreiss@math.hu-berlin.de; **martin.wahl@math.hu-berlin.de

Consider a Poisson point process with unknown support boundary curve $g$, which forms a prototype of an irregular statistical model. We address the problem of estimating non-linear functionals of the form $\int \Phi(g(x))\,dx$. Following a nonparametric maximum-likelihood approach, we construct an estimator which is UMVU over Hölder balls and achieves the (local) minimax rate of convergence. These results hold under weak assumptions on $\Phi$ which are satisfied for $\Phi(u) = |u|^p$, $p \geq 1$. As an application, we consider the problem of estimating the $L^p$-norm and derive the minimax separation rates in the corresponding nonparametric hypothesis testing problem. Structural differences to results for regular nonparametric models are discussed.

*Keywords:* minimax hypothesis testing; non-linear functionals; Poisson point process; support estimation

## References

[1] Bergh, J. and Löfström, J. (1976). *Interpolation Spaces. An Introduction*. Berlin: Springer. Grundlehren der Mathematischen Wissenschaften, No. 223. MR0482275

[2] Drees, H., Neumeyer, N. and Selk, L. Estimation and hypotheses testing in boundary regression models. *Bernoulli*. To appear.

[3] Ibragimov, I.A., Nemirovskii, A.S. and Khas'minskii, R.Z. (1986). Some problems of nonparametric estimation in Gaussian white noise. *Teor. Veroyatn. Primen.* **31** 451–466. MR0866866

[4] Ingster, Yu.I. and Suslina, I.A. (2003). *Nonparametric Goodness-of-Fit Testing Under Gaussian Models*. *Lecture Notes in Statistics* **169**. New York: Springer. MR1991446

[5] Jirak, M., Meister, A. and Reiß, M. (2014). Adaptive function estimation in nonparametric regression with one-sided errors. *Ann. Statist.* **42** 1970–2002. MR3262474

[6] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169

[7] Karr, A.F. (1991). *Point Processes and Their Statistical Inference*, 2nd ed. *Probability*: *Pure and Applied* **7**. New York: Dekker. MR1113698

[8] Korostelëv, A.P. and Tsybakov, A.B. (1993). *Minimax Theory of Image Reconstruction*. *Lecture Notes in Statistics* **82**. New York: Springer. MR1226450

[9] Kutoyants, Yu.A. (1998). *Statistical Inference for Spatial Poisson Processes*. *Lecture Notes in Statistics* **134**. New York: Springer. MR1644620

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

[11] Lepski, O., Nemirovski, A. and Spokoiny, V. (1999). On estimation of the $L_r$ norm of a regression function. *Probab. Theory Related Fields* **113** 221–253. MR1670867

[12] Lepski, O.V. and Spokoiny, V.G. (1999). Minimax nonparametric hypothesis testing: The case of an inhomogeneous alternative. *Bernoulli* **5** 333–358. MR1681702

[13] Meister, A. and Reiß, M. (2013). Asymptotic equivalence for nonparametric regression with non-regular errors. *Probab. Theory Related Fields* **155** 201–229. MR3010397

[14] Nemirovski, A. (2000). Topics in non-parametric statistics. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1998). *Lecture Notes in Math.* **1738** 85–277. Berlin: Springer. MR1775640

[15] Reiß, M. and Selk, L. (2017). Efficient estimation of functionals in nonparametric boundary models. *Bernoulli* **23** 1022–1055. MR3606758

[16] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. MR2724359

*Bernoulli* **25**(4A), 2019, 2620–2648
https://doi.org/10.3150/18-BEJ1065

# Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance

JONATHAN WEED[1] and FRANCIS BACH[2]

[1] *Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA. E-mail: jweed@mit.edu*
[2] *INRIA, Département d'informatique de l'ENS, École normale supérieure, CNRS, PSL Research University, 75005 Paris, France. E-mail: francis.bach@inria.fr*

The Wasserstein distance between two probability measures on a metric space is a measure of closeness with applications in statistics, probability, and machine learning. In this work, we consider the fundamental question of how quickly the empirical measure obtained from $n$ independent samples from $\mu$ approaches $\mu$ in the Wasserstein distance of any order. We prove sharp asymptotic and finite-sample results for this rate of convergence for general measures on general compact metric spaces. Our finite-sample results show the existence of multi-scale behavior, where measures can exhibit radically different rates of convergence as $n$ grows.

*Keywords:* optimal transport; quantization; Wasserstein metrics

## References

[1] Ajtai, M., Komlós, J. and Tusnády, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264. MR0779885
[2] Ba, K.D., Nguyen, H.L., Nguyen, H.N. and Rubinfeld, R. (2011). Sublinear time algorithms for Earth mover's distance. *Theory Comput*. *Syst*. **48** 428–442. MR2763110
[3] Bellman, R. (1961). *Adaptive Control Processes*: *A Guided Tour*. Princeton, NJ: Princeton Univ. Press. MR0134403
[4] Berend, D. and Kontorovich, A. (2013). A sharp estimate of the binomial mean absolute deviation with applications. *Statist*. *Probab*. *Lett*. **83** 1254–1259. MR3041401
[5] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J*. *Funct*. *Anal*. **163** 1–28. MR1682772
[6] Boissard, E. (2011). Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance. *Electron*. *J*. *Probab*. **16** 2296–2333. MR2861675
[7] Boissard, E. and Le Gouic, T. (2014). On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **50** 539–563. MR3189084
[8] Bolley, F., Guillin, A. and Villani, C. (2007). Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab*. *Theory Related Fields* **137** 541–593. MR2280433
[9] Bucklew, J.A. (2013). *Introduction to Rare Event Simulation*. *Springer Series in Statistics*. New York: Springer. MR2045385

[10] Cañas, G.D. and Rosasco, L. (2012). Learning probability measures with respect to optimal transport metrics. In *Advances in Neural Information Processing Systems* 25: 26*th Annual Conference on Neural Information Processing Systems* 2012. *Proceedings of a Meeting Held December* 3–6, 2012, *Lake Tahoe*, *Nevada*, *United States* (P.L. Bartlett, F.C.N. Pereira, C.J.C. Burges, L. Bottou and K.Q. Weinberger, eds.) 2501–2509.

[11] Carlier, G., Duval, V., Peyré, G. and Schmitzer, B. (2017). Convergence of entropic schemes for optimal transport and gradient flows. *SIAM J. Math. Anal.* **49** 1385–1418. MR3635459

[12] Corlay, S. and Pagès, G. (2015). Functional quantization-based stratified sampling methods. *Monte Carlo Methods Appl.* **21** 1–32. MR3318550

[13] Cover, T.M. and Thomas, J.A. (2012). *Elements of Information Theory*. *Wiley Series in Telecommunications*. New York: Wiley. MR1122806

[14] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems* 26: 27*th Annual Conference on Neural Information Processing Systems* 2013. *Proceedings of a Meeting Held December* 5–8, 2013, *Lake Tahoe*, *Nevada*, *United States* (C.J.C. Burges, L. Bottou, Z. Ghahramani and K.Q. Weinberger, eds.) 2292–2300.

[15] David, G. (1988). Morceaux de graphes lipschitziens et intégrales singulières sur une surface. *Rev. Mat. Iberoam.* **4** 73–114. MR1009120

[16] Dereich, S., Scheutzow, M. and Schottstedt, R. (2013). Constructive quantization: Approximation by empirical measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1183–1203. MR3127919

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

[18] Dudley, R.M. (1968). The speed of mean Glivenko–Cantelli convergence. *Ann. Math. Stat.* **40** 40–50. MR0236977

[19] Falconer, K. (2003). *Fractal Geometry*: *Mathematical Foundations and Applications*, 2nd ed. Hoboken, NJ: Wiley. MR2118797

[20] Fortet, R. and Mourier, E. (1953). Convergence de la répartition empirique vers la répartition théorique. *Ann. Sci. Éc. Norm. Supér.* (3) **70** 267–285. MR0061325

[21] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341

[22] Graf, S. and Luschgy, H. (2007). *Foundations of Quantization for Probability Distributions*. *Lecture Notes in Math.* **1730**. Berlin: Springer. MR1764176

[23] Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning*: *Data Mining*, *Inference*, *and Prediction*. *Springer Series in Statistics*. New York: Springer. MR1851606

[24] Indyk, P. and Thaper, N. (2003). Fast image retrieval via embeddings. In 3*rd International Workshop on Statistical and Computational Theories of Vision*. ICCV.

[25] Kaimanovich, V.A. and Le Prince, V. (2011). Matrix random products with singular harmonic measure. *Geom. Dedicata* **150** 257–279. MR2753707

[26] Kantorovič, L.V. and Rubinšteĭn, G.Š. (1958). On a space of completely additive functions. *Vestn. Leningrad Univ.* **13** 52–59. MR0102006

[27] Kantorovitch, L. (1942). On the translocation of masses. *C. R.* (*Dokl.*) *Acad. Sci. URSS* **37** 199–201. MR0009619

[28] Kloeckner, B. (2012). Approximation by finitely supported measures. *ESAIM Control Optim. Calc. Var.* **18** 343–359. MR2954629

[29] Kusner, M.J., Sun, Y., Kolkin, N.I. and Weinberger, K.Q. (2015). From word embeddings to document distances. In *Proceedings of the* 32*nd International Conference on Machine Learning*, *ICML* 2015, *Lille*, *France*, 6–11 *July* 2015 (F.R. Bach and D.M. Blei, eds.). *JMLR Workshop and Conference Proceedings* **37** 957–966. JMLR.org.

[30] Ledoux, M. (1995). Remarks on logarithmic Sobolev constants, exponential integrability and bounds on the diameter. *J. Math. Kyoto Univ.* **35** 211–220. MR1346225

[31] Ledoux, M. (2005). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. MR1849347

[32] Ledrappier, F. (1981). Some relations between dimension and Lyapounov exponents. *Comm. Math. Phys.* **81** 229–238. MR0632758

[33] Little, A.V., Maggioni, M. and Rosasco, L. (2017). Multiscale geometric methods for data sets I: Multiscale SVD, noise and curvature. *Appl. Comput. Harmon. Anal.* **43** 504–567. MR3683673

[34] Mattila, P. (1999). *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability. Cambridge Studies in Advanced Mathematics* **44**. Cambridge: Cambridge Univ. Press. MR1333890

[35] McDiarmid, C. (1989). On the method of bounded differences. In *Surveys in Combinatorics*, 1989 (*Norwich*, 1989). *London Mathematical Society Lecture Note Series* **141** 148–188. Cambridge: Cambridge Univ. Press. MR1036755

[36] Milman, E. (2010). Isoperimetric and concentration inequalities: Equivalence under curvature lower bound. *Duke Math. J.* **154** 207–239. MR2682183

[37] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Hist. Acad. R. Sci.* **1** 666–704.

[38] Nova, D. and Estévez, P.A. (2014). A review of learning vector quantization classifiers. *Neural Comput. Appl.* **25** 511–524.

[39] Novak, E. (1988). *Deterministic and Stochastic Error Bounds in Numerical Analysis. Lecture Notes in Math.* **1349**. Berlin: Springer. MR0971255

[40] Pagès, G. (1998). A space quantization method for numerical integration. *J. Comput. Appl. Math.* **89** 1–38. MR1625987

[41] Posner, E.C., Rodemich, E.R. and Rumsey, H. Jr. (1967). Epsilon entropy of stochastic processes. *Ann. Math. Stat.* **38** 1000–1020. MR0211457

[42] Rolet, A., Cuturi, M. and Peyré, G. (2016). Fast dictionary learning with a smoothed Wasserstein loss. In *Proceedings of the* 19*th International Conference on Artificial Intelligence and Statistics*, *AISTATS* 2016, *Cadiz*, *Spain*, *May* 9–11, 2016 (A. Gretton and C.C. Robert, eds.). *JMLR Workshop and Conference Proceedings* **51** 630–638. JMLR.org.

[43] Rubner, Y., Tomasi, C. and Guibas, L.J. (2000). The Earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* **40** 99–121.

[44] Rüschendorf, L. (1991). Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals* (*Rome*, 1990). *Math. Appl.* **67** 151–187. Dordrecht: Kluwer Academic. MR1215951

[45] Rüschendorf, L. and Rachev, S.T. (1990). A characterization of random variables with minimum $L^2$-distance. *J. Multivariate Anal.* **32** 48–54. MR1035606

[46] Sandler, R. and Lindenbaum, M. (2011). Nonnegative matrix factorization with Earth mover's distance metric for image analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **33** 1590–1602.

[47] Santambrogio, F. (2015). *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Progress in Nonlinear Differential Equations and Their Applications* **87**. Cham: Birkhäuser/Springer. MR3409718

[48] Shannon, C.E. (1960). Coding theorems for a discrete source with a fidelity criterion. In *Information and Decision Processes* 93–126. New York: McGraw-Hill. MR0122612

[49] Solomon, J., de Goes, F., Peyré, G., Cuturi, M., Butscher, A., Nguyen, A., Du, T. and Guibas, L.J. (2015). Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains. *ACM Trans. Graph.* **34** 66:1–66:11. DOI:10.1145/2766963.

[50] Starck, J.-L., Murtagh, F. and Bijaoui, A. (1998). *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge: Cambridge Univ. Press. MR1637482

[51] Talagrand, M. (1992). Matching random samples in many dimensions. *Ann. Appl. Probab.* **2** 846–856. MR1189420

[52] Talagrand, M. (1995). Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. Inst. Hautes Études Sci.* **81** 73–205. MR1361756

[53] Varadarajan, V.S. (1958). On the convergence of sample probability distributions. *Sankhyā* **19** 23–26. MR0094839

[54] Vershik, A.M. (2013). Long history of the Monge–Kantorovich transportation problem. *Math. Intelligencer* **35** 1–9. MR3133757

[55] Villani, C. (2009). *Optimal Transport: Old and New. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Berlin: Springer. MR2459454

[56] Wakin, M., Donoho, D., Choi, H. and Baraniuk, R.G. (2005). The multiscale structure of non-differentiable image manifolds. In *Proc. SPIE* SPIE.

[57] Weed, J. and Bach, F. (2018). Supplement to "Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance." DOI:10.3150/18-BEJ1065SUPP.

[58] Young, L.S. (1982). Dimension, entropy and Lyapunov exponents. *Ergod. Theory Dyn. Syst.* **2** 109–124. MR0684248

[59] Zhang, M., Liu, Y., Luan, H., Sun, M., Izuha, T. and Hao, J. (2016). Building Earth mover's distance on bilingual word embeddings for machine translation. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, February* 12–17, 2016, *Phoenix, Arizona, USA* (D. Schuurmans and M.P. Wellman, eds.) 2870–2876. AAAI Press.

*Bernoulli* **25**(4A), 2019, 2649–2695
https://doi.org/10.3150/18-BEJ1066

# Uniform sampling in a structured branching population

ALINE MARGUET

*Univ. Grenoble Alpes, Inria, 38000 Grenoble, France. E-mail:* aline.marguet@inria.fr

We are interested in the dynamic of a structured branching population where the trait of each individual moves according to a Markov process. The rate of division of each individual is a function of its trait and when a branching event occurs, the trait of the descendants at birth depends on the trait of the mother and on the number of descendants. In this article, we explicitly describe the penalized Markov process, named auxiliary process, corresponding to the dynamic of the trait of a "typical" individual by giving its associated infinitesimal generator. We prove a Many-to-One formula and a Many-to-One formula for forks. Furthermore, we prove that this auxiliary process characterizes exactly the process of the trait of a uniformly sampled individual in a large population approximation. We detail three examples of growth-fragmentation models: the linear growth model, the exponential growth model and the parasite infection model.

*Keywords:* branching Markov processes; Many-to-One formulas; size-biased reproduction law

## References

[1] Athreya, K.B. (2000). Change of measures for Markov chains and the $L \log L$ theorem for branching processes. *Bernoulli* **6** 323–338. MR1748724
[2] Athreya, K.B. (2012). Coalescence in critical and subcritical Galton–Watson branching processes. *J. Appl. Probab.* **49** 627–638. MR3012088
[3] Athreya, K.B., Athreya, S.R. and Iyer, S.K. (2011). Supercritical age-dependent branching Markov processes and their scaling limits. *Bernoulli* **17** 138–154. MR2797985
[4] Bansaye, V. (2018). Ancestral lineages and limit theorems for branching Markov chains in varying environment. *J. Theoret. Probab.* To appear.
[5] Bansaye, V., Delmas, J.-F., Marsalle, L. and Tran, V.C. (2011). Limit theorems for Markov processes indexed by continuous time Galton–Watson trees. *Ann. Appl. Probab.* **21** 2263–2314. MR2895416
[6] Bansaye, V. and Méléard, S. (2015). *Stochastic Models for Structured Populations*: *Scaling Limits and Long Time Behavior*. *Mathematical Biosciences Institute Lecture Series*. *Stochastics in Biological Systems* **1**. Cham: Springer; MBI Mathematical Biosciences Institute, Ohio State Univ., Columbus, OH. MR3380810
[7] Bansaye, V. and Tran, V.C. (2011). Branching Feller diffusion for cell division with parasite infection. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** 95–127. MR2754402
[8] Bellman, R. and Harris, T. (1952). On age-dependent binary branching processes. *Ann. of Math.* (2) **55** 280–295. MR0045971
[9] Biggins, J.D. (1977). Martingale convergence in the branching random walk. *J. Appl. Probab.* **14** 25–37. MR0433619
[10] Biggins, J.D. and Kyprianou, A.E. (2004). Measure change in multitype branching. *Adv. in Appl. Probab.* **36** 544–581. MR2058149

[11] 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
[12] Campillo, F., Champagnat, N. and Fritsch, C. (2016). Links between deterministic and stochastic approaches for invasion in growth–fragmentation–death models. *J. Math. Biol.* **73** 1781–1821. MR3556850
[13] Chauvin, B. and Rouault, A. (1988). KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees. *Probab. Theory Related Fields* **80** 299–314. MR0968823
[14] Chauvin, B., Rouault, A. and Wakolbinger, A. (1991). Growing conditioned trees. *Stochastic Process. Appl.* **39** 117–130. MR1135089
[15] Cloez, B. (2017). Limit theorems for some branching measure-valued processes. *Adv. in Appl. Probab.* **49** 549–580. MR3668388
[16] Cooper, S. (2006). Distinguishing between linear and exponential cell growth during the division cycle: Single-cell studies, cell-culture studies, and the object of cell-cycle research. *Theor. Biol. Med. Model.* **3** 10.
[17] Del Moral, P. (2004). *Feynman–Kac Formulae*: *Genealogical and Interacting Particle Systems with Applications*. *Probability and Its Applications* (*New York*). New York: Springer. MR2044973
[18] Doumic, M., Hoffmann, M., Krell, N. and Robert, L. (2015). Statistical estimation of a growth–fragmentation model observed on a genealogical tree. *Bernoulli* **21** 1760–1799. MR3352060
[19] Doumic-Jauffret, M., Maia, P. and Zubelli, J. (2010). On the calibration of a size-structured population model from experimental data. *Acta Biotheor.* **58** 405–413.
[20] Durrett, R. (1978/79). The genealogy of critical branching processes. *Stochastic Process. Appl.* **8** 101–116. MR0511879
[21] 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
[22] Fournier, N. and Méléard, S. (2004). A microscopic probabilistic description of a locally regulated population and macroscopic approximations. *Ann. Appl. Probab.* **14** 1880–1919. MR2099656
[23] Georgii, H.-O. and Baake, E. (2003). Supercritical multitype branching processes: The ancestral types of typical individuals. *Adv. in Appl. Probab.* **35** 1090–1110. MR2014271
[24] Guyon, J. (2007). Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging. *Ann. Appl. Probab.* **17** 1538–1569. MR2358633
[25] Hardy, R. and Harris, S.C. (2009). A spine approach to branching diffusions with applications to $L^p$-convergence of martingales. In *Séminaire de Probabilités XLII*. *Lecture Notes in Math.* **1979** 281–330. Berlin: Springer. MR2599214
[26] Harris, S., Johnston, S. and Roberts, M. (2017). The coalescent structure of continuous-time Galton–Watson trees. Preprint. Available at arXiv:1707.07993.
[27] Harris, S.C. and Roberts, M.I. (2017). The many-to-few lemma and multiple spines. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 226–242. MR3606740
[28] Harris, S.C. and Williams, D. (1996). Large deviations and martingales for a typed branching diffusion. I. *Astérisque* **236** 133–154. Hommage à P. A. Meyer et J. Neveu. MR1417979
[29] Harris, T.E. (2002). *The Theory of Branching Processes*. *Dover Phoenix Editions*. Mineola, NY: Dover. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)]. MR1991122
[30] Hoang, V.H. (2017). Estimating the division kernel of a size-structured population. *ESAIM Probab. Stat.* **21** 275–302. MR3743915
[31] Hoffmann, M. and Olivier, A. (2016). Nonparametric estimation of the division rate of an age dependent branching process. *Stochastic Process. Appl.* **126** 1433–1471. MR3473101

[32] Hong, J.-I. (2011). *Coalescence in Bellman–Harris and Multi-Type Branching Processes*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.) – Iowa State Univ. MR2942128
[33] Kallenberg, O. (1977). Stability of critical cluster fields. *Math*. *Nachr.* **77** 7–43. MR0443078
[34] Kimmel, M. (1997). Quasistationarity in a branching model of division-within-division. In *Classical and Modern Branching Processes* (*Minneapolis*, *MN*, 1994). *IMA Vol*. *Math*. *Appl*. **84** 157–164. New York: Springer. MR1601725
[35] Kingman, J.F.C. (1975). The first birth problem for an age-dependent branching process. *Ann*. *Probab.* **3** 790–801. MR0400438
[36] Kunita, H. (1997). *Stochastic Flows and Stochastic Differential Equations*. *Cambridge Studies in Advanced Mathematics* **24**. Cambridge: Cambridge Univ. Press. Reprint of the 1990 original. MR1472487
[37] Kurtz, T., Lyons, R., Pemantle, R. and Peres, Y. (1997). A conceptual proof of the Kesten–Stigum theorem for multi-type branching processes. In *Classical and Modern Branching Processes* (*Minneapolis*, *MN*, 1994). *IMA Vol*. *Math*. *Appl*. **84** 181–185. New York: Springer. MR1601737
[38] Lambert, A. and Popovic, L. (2013). The coalescent point process of branching trees. *Ann*. *Appl*. *Probab.* **23** 99–144. MR3059232
[39] Leibler, S. and Kussell, E. (2010). Individual histories and selection in heterogeneous populations. *Proc*. *Natl*. *Acad*. *Sci*. *USA* **107** 13183–13188.
[40] 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
[41] Marguet, A. (2017). A law of large numbers for branching markov processes by the ergodicity of ancestral lineages. Preprint. Available at arXiv:1707.07993.
[42] Mischler, S. and Scher, J. (2016). Spectral analysis of semigroups and growth–fragmentation equations. *Ann*. *Inst*. *H*. *Poincaré Anal*. *Non Linéaire* **33** 849–898. MR3489637
[43] Nerman, O. and Jagers, P. (1984). The stable double infinite pedigree process of supercritical branching populations. *Z. Wahrsch*. *Verw. Gebiete* **65** 445–460. MR0731231
[44] O'Connell, N. (1995). The genealogy of branching processes and the age of our most recent common ancestor. *Adv. in Appl*. *Probab.* **27** 418–442. MR1334822
[45] Ozbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I. and Van Oudenaarden, A. (2004). Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427** 737–740.
[46] Tran, V.C. (2006). Stochastic particle models for problems of adaptive evolution and for the approximations of statistical solutions. Thesis, Univ. de Nanterre – Paris X.
[47] Williams, D. (1991). *Probability with Martingales*. *Cambridge Mathematical Textbooks*. Cambridge: Cambridge Univ. Press. MR1155402
[48] Zubkov, A.M. (1975). Limit distributions of the distance to the nearest common ancestor. *Teor. Veroyatn*. *Primen*. **20** 614–623. MR0397915

*Bernoulli* **25**(4A), 2019, 2696–2728
https://doi.org/10.3150/18-BEJ1067

# Nonparametric Bayesian posterior contraction rates for scalar diffusions with high-frequency data

KWEKU ABRAHAM

*Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK. E-mail: lkwabraham@statslab.cam.ac.uk*

We consider inference in the scalar diffusion model $\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$ with discrete data $(X_{j\Delta_n})_{0\le j\le n}$, $n \to \infty$, $\Delta_n \to 0$ and periodic coefficients. For $\sigma$ given, we prove a general theorem detailing conditions under which Bayesian posteriors will contract in $L^2$-distance around the true drift function $b_0$ at the frequentist minimax rate (up to logarithmic factors) over Besov smoothness classes. We exhibit natural nonparametric priors which satisfy our conditions. Our results show that the Bayesian method adapts both to an unknown sampling regime and to unknown smoothness.

*Keywords:* adaptive estimation; Bayesian nonparametrics; concentration inequalities; diffusion processes; discrete time observations; drift function

## References

[1] Abraham, K. (2018). Supplement to "Nonparametric Bayesian posterior contraction rates for scalar diffusions with high-frequency data". DOI:10.3150/18-BEJ1067SUPP.
[2] Baraud, Y. (2010). A Bernstein-type inequality for suprema of random processes with applications to model selection in non-Gaussian regression. *Bernoulli* **16** 1064–1085. MR2759169
[3] Bass, R.F. (2011). *Stochastic Processes*. *Cambridge Series in Statistical and Probabilistic Mathematics* **33**. Cambridge: Cambridge Univ. Press. MR2856623
[4] Bhattacharya, R., Denker, M. and Goswami, A. (1999). Speed of convergence to equilibrium and to normality for diffusions with multiple periodic scales. *Stochastic Process. Appl.* **80** 55–86. MR1670111
[5] Birgé, L. and Massart, P. (1998). Minimum contrast estimators on sieves: Exponential bounds and rates of convergence. *Bernoulli* **4** 329–375. MR1653272
[6] Castillo, I. and Nickl, R. (2013). Nonparametric Bernstein–von Mises theorems in Gaussian white noise. *Ann. Statist.* **41** 1999–2028. MR3127856
[7] Castillo, I. and Nickl, R. (2014). On the Bernstein-von Mises phenomenon for nonparametric Bayes procedures. *Ann. Statist.* **42** 1941–1969. MR3262473
[8] Chorowski, J. (2016). Statistics for diffusion processes with low and high-frequency observations. Ph.D. thesis, Mathematisch-Naturwissenschaftliche Fakultät, Humboldt-Universität zu Berlin.
[9] Comte, F., Genon-Catalot, V. and Rozenholc, Y. (2007). Penalized nonparametric mean square estimation of the coefficients of diffusion processes. *Bernoulli* **13** 514–543. MR2331262
[10] Dalalyan, A. (2005). Sharp adaptive estimation of the drift function for ergodic diffusions. *Ann. Statist.* **33** 2507–2528. MR2253093

[11] Ghosal, S., Ghosh, J.K. and van der Vaart, A.W. (2000). Convergence rates of posterior distributions. *Ann. Statist.* **28** 500–531. MR1790007

[12] 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

[13] 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

[14] Gihman, I. and Skorohod, A.V. (1972). *Stochastic Differential Equations*. New York: Springer. MR0346904

[15] 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

[16] 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

[17] Gobet, E., Hoffmann, M. and Reiß, M. (2004). Nonparametric estimation of scalar diffusions based on low frequency data. *Ann. Statist.* **32** 2223–2253. MR2102509

[18] Gugushvili, S. and Spreij, P. (2014). Nonparametric Bayesian drift estimation for multidimensional stochastic differential equations. *Lith. Math. J.* **54** 127–141. MR3212631

[19] Hoffmann, M. (1999). Adaptive estimation in diffusion processes. *Stochastic Process. Appl.* **79** 135–163. MR1670522

[20] Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scand. J. Stat.* **24** 211–229. MR1455868

[21] Kutoyants, Y.A. (2004). *Statistical Inference for Ergodic Diffusion Processes*. London: Springer. MR2144185

[22] Liptser, R.S. and Shiryayev, A.N. (1977). *Statistics of Random Processes. I. General Theory*. New York: Springer. MR0474486

[23] Nickl, R. (2019). Bernstein–von Mises theorems for statistical inverse problems I: Schrödinger equation. *J. Eur. Math. Soc.* (*JEMS*). To appear.

[24] Nickl, R. and Söhl, J. (2017). Nonparametric Bayesian posterior contraction rates for discretely observed scalar diffusions. *Ann. Statist.* **45** 1664–1693. MR3670192

[25] 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

[26] Paulin, D. (2015). Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electron. J. Probab.* **20** Article ID 79. MR3383563

[27] 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

[28] Pollard, D. (2002). *A User's Guide to Measure Theoretic Probability. Cambridge Series in Statistical and Probabilistic Mathematics* **8**. Cambridge: Cambridge Univ. Press. MR1873379

[29] Ray, K. (2013). Bayesian inverse problems with non-conjugate priors. *Electron. J. Stat.* **7** 2516–2549. MR3117105

[30] Rogers, L.C.G. and Williams, D. (2000). *Diffusions, Markov Processes, and Martingales. Vol. 2. Itô Calculus. Cambridge Mathematical Library*. Cambridge: Cambridge Univ. Press. Reprint of the second (1994) edition. MR1780932

[31] Söhl, J. and Trabs, M. (2016). Adaptive confidence bands for Markov chains and diffusions: Estimating the invariant measure and the drift. *ESAIM Probab. Stat.* **20** 432–462. MR3581829

[32] Triebel, H. (1983). *Theory of Function Spaces. Monographs in Mathematics* **78**. Basel: Birkhäuser. MR0781540

[33] van der Meulen, F., Schauer, M. and van Waaij, J. (2018). Adaptive nonparametric drift estimation for diffusion processes using Faber–Schauder expansions. *Stat*. *Inference Stoch*. *Process*. **21** 603–628. MR3846994

[34] van der Meulen, F. and van Zanten, H. (2013). Consistent nonparametric Bayesian inference for discretely observed scalar diffusions. *Bernoulli* **19** 44–63. MR3019485

[35] van Waaij, J. and van Zanten, H. (2016). Gaussian process methods for one-dimensional diffusions: Optimal rates and adaptation. *Electron*. *J*. *Stat*. **10** 628–645. MR3471991

*Bernoulli* **25**(4A), 2019, 2729–2757
https://doi.org/10.3150/18-BEJ1069

# A Benamou–Brenier formulation of martingale optimal transport

MARTIN HUESMANN[1] and DARIO TREVISAN[2]

[1] *Institut für Angewandte Mathematik, Rheinische Friedrich–Wilhelms-Universität Bonn, Endenicher Allee 60, DE-53115 DE-53115 Bonn, Germany. E-mail: huesmann@iam.uni-bonn.de*
[2] *Dipartimento di Matematica, Università degli Studi di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, Italy. E-mail: dario.trevisan@unipi.it*

We introduce a Benamou–Brenier formulation for the continuous-time martingale optimal transport problem as a weak length relaxation of its discrete-time counterpart. By the correspondence between classical martingale problems and Fokker–Planck equations, we obtain an equivalent PDE formulation for which basic properties such as existence, duality and geodesic equations can be analytically studied, yielding corresponding results for the stochastic formulation. In the one dimensional case, sufficient conditions for finiteness of the cost are also given and a link between geodesics and porous medium equations is partially investigated.

*Keywords:* Fokker–Planck equations; Martingale Optimal Transport; martingale problem; porous medium equation; Strassen's theorem

## References

[1] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, 2nd ed. Basel: Birkhäuser. MR2401600
[2] Arnaudon, M., Cruzeiro, A.B. and Fang, S. (2018). Generalized stochastic Lagrangian paths for the Navier–Stokes equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **18** 1033–1060. MR3807594
[3] Backhoff, J., Beiglboeck, M., Huesmann, M. and Källblad, S. (2017). Martingale Benamou–Brenier: A probabilistic perspective. Available at arXiv:1708.04869.
[4] Barbu, V., Benazzoli, C. and Di Persio, L. (2018). Mild solutions to the dynamic programming equation for stochastic optimal control problems. *Automatica J. IFAC* **93** 520–526. MR3810941
[5] Beiglböck, M., Cox, A.M.G. and Huesmann, M. (2017). Optimal transport and Skorokhod embedding. *Invent. Math.* **208** 327–400. MR3639595
[6] Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices – a mass transport approach. *Finance Stoch.* **17** 477–501. MR3066985
[7] Beiglböck, M., Henry-Labordère, P. and Touzi, N. (2017). Monotone martingale transport plans and Skorokhod embedding. *Stochastic Process. Appl.* **127** 3005–3013. MR3682121
[8] Beiglböck, M. and Juillet, N. (2016). On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.* **44** 42–106. MR3456332
[9] Beiglböck, M., Nutz, M. and Touzi, N. (2017). Complete duality for martingale optimal transport on the line. *Ann. Probab.* **45** 3038–3074. MR3706738
[10] Beiglboeck, M., Lim, T. and Obłój, J. (2017). Dual attainment for the martingale transport problem. Available at arXiv:1705.04273.

1350-7265 

[11] Benamou, J.-D. and Brenier, Y. (1999). A numerical method for the optimal time-continuous mass transport problem and related problems. In *Monge Ampère Equation*: *Applications to Geometry and Optimization* (*Deerfield Beach*, *FL*, 1997). *Contemp*. *Math*. **226** 1–11. Providence, RI: Amer. Math. Soc. MR1660739

[12] Benamou, J.-D. and Brenier, Y. (2000). A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math*. **84** 375–393. MR1738163

[13] Bonnans, J.F. and Tan, X. (2013). A model-free no-arbitrage price bound for variance options. *Appl*. *Math*. *Optim*. **68** 43–73. MR3072239

[14] Brenier, Y. (1999). Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. *Comm*. *Pure Appl*. *Math*. **52** 411–452. MR1658919

[15] Carlen, E.A. and Maas, J. (2014). An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker–Planck equation is gradient flow for the entropy. *Comm*. *Math*. *Phys*. **331** 887–926. MR3248053

[16] Dacorogna, B. and Moser, J. (1990). On a partial differential equation involving the Jacobian determinant. *Ann*. *Inst*. *H*. *Poincaré Anal*. *Non Linéaire* **7** 1–26. MR1046081

[17] De March, H. and Touzi, N. (2017). Irreducible convex paving for decomposition of multi-dimensional martingale transport plans. Available at arXiv:1702.08298.

[18] Dolbeault, J., Nazaret, B. and Savaré, G. (2009). A new class of transport distances between measures. *Calc*. *Var. Partial Differential Equations* **34** 193–231. MR2448650

[19] Dolinsky, Y. and Soner, H.M. (2014). Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Related Fields* **160** 391–427. MR3256817

[20] Dolinsky, Y. and Soner, H.M. (2015). Martingale optimal transport in the Skorokhod space. *Stochastic Process*. *Appl*. **125** 3893–3931. MR3373308

[21] Döring, L., Gonon, L., Prömel, D.J. and Reichmann, O. (2017). On Skorokhod embeddings and Poisson equations. Available at arXiv:1703.05673.

[22] Dupire, B. (1994). Pricing with a smile. *Risks* **7** 18–20.

[23] Erbar, M. (2014). Gradient flows of the entropy for jump processes. *Ann*. *Inst*. *Henri Poincaré Probab. Stat*. **50** 920–945. MR3224294

[24] Erbar, M. (2016). A gradient flow approach to the Boltzmann equation. Available at arXiv:1603.00540.

[25] Erbar, M. and Maas, J. (2012). Ricci curvature of finite Markov chains via convexity of the entropy. *Arch*. *Ration*. *Mech*. *Anal*. **206** 997–1038. MR2989449

[26] Figalli, A. (2008). Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J*. *Funct*. *Anal*. **254** 109–153. MR2375067

[27] 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

[28] Galichon, A., Henry-Labordère, P. and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann*. *Appl*. *Probab*. **24** 312–336. MR3161649

[29] Ghoussoub, N., Kim, Y.-H. and Lim, T. (2015). Structure of optimal martingale transport in general dimensions. Available at arXiv:1508.01806.

[30] Guo, I., Loeper, G. and Wang, S. (2017). Local volatility calibration by optimal transport. Available at arXiv:1703.05673.

[31] Hobson, D. (2011). The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris–Princeton Lectures on Mathematical Finance* 2010. *Lecture Notes in Math*. **2003** 267–318. Berlin: Springer. MR2762363

[32] Huesmann, M. and Stebegg, F. (2018). Monotonicity preserving transformations of MOT and SEP. *Stochastic Process*. *Appl*. **128** 1114–1134. MR3769657

[33] Maas, J. (2011). Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261** 2250–2292. MR2824578

[34] McCann, R.J. (1995). A convexity theory for interacting gases and equilibrium crystals. Ph.D. thesis, Princeton Univ.

[35] Mielke, A. (2011). A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity* **24** 1329–1346. MR2776123

[36] Mielke, A. (2013). Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48** 1–31. MR3090532

[37] Nutz, M., Stebegg, F. and Tan, X. (2017). Multiperiod martingale transport. Available at arXiv:1703.10588.

[38] Obłój, J. (2004). The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1** 321–390. MR2068476

[39] Obłój, J. and Siorpaes, P. (2017). Structure of martingale transports in finite dimensions. Available at arXiv:1702.08433.

[40] Otto, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations* **26** 101–174. MR1842429

[41] Revuz, D. and Yor, M. (1999). *Continuous Martingales and Brownian Motion*, 3rd ed. *Grundlehren der Mathematischen Wissenschaften* **293**. Berlin: Springer. MR1725357

[42] Rockafellar, R.T. (1970). *Convex Analysis*. *Princeton Mathematical Series* **28**. Princeton, NJ: Princeton Univ. Press. MR0274683

[43] Strassen, V. (1965). The existence of probability measures with given marginals. *Ann. Math. Stat.* **36** 423–439. MR0177430

[44] Stroock, D.W. and Varadhan, S.R.S. (2006). *Multidimensional Diffusion Processes*. Berlin: Springer. MR2190038

[45] Tan, X. and Touzi, N. (2013). Optimal transportation under controlled stochastic dynamics. *Ann. Probab.* **41** 3201–3240. MR3127880

[46] Trevisan, D. (2016). Well-posedness of multidimensional diffusion processes with weakly differentiable coefficients. *Electron. J. Probab.* **21** Paper No. 22, 41. MR3485364

[47] Vázquez, J.L. (2007). *The Porous Medium Equation*: *Mathematical Theory*. Oxford University Press, Oxford: The Clarendon Press. MR2286292

[48] Villani, C. (2009). *Optimal Transport*: *Old and New*. *Grundlehren der Mathematischen Wissenschaften* **338**. Berlin: Springer. MR2459454

*Bernoulli* **25**(4A), 2019, 2758–2792
https://doi.org/10.3150/18-BEJ1070

# Dickman approximation in simulation, summations and perpetuities

CHINMOY BHATTACHARJEE[1] and LARRY GOLDSTEIN[2]

[1]*IMSV, Universität Bern, 3012 Bern, Switzerland.*
*E-mail: chinmoy.bhattacharjee@stat.unibe.ch*
[2]*Department of Mathematics, University of Southern California, Los Angeles, CA 90089-2532, USA.*
*E-mail: larry@math.usc.edu*

The generalized Dickman distribution $\mathcal{D}_\theta$ with parameter $\theta > 0$ is the unique solution to the distributional equality $W =_d W^*$, where

$$W^* =_d U^{1/\theta}(W+1), \tag{1}$$

with $W$ non-negative with probability one, $U \sim \mathcal{U}[0,1]$ independent of $W$, and $=_d$ denoting equality in distribution. These distributions appear in number theory, stochastic geometry, perpetuities and the study of algorithms. We obtain bounds in Wasserstein type distances between $\mathcal{D}_\theta$ and the distribution of

$$W_n = \frac{1}{n}\sum_{i=1}^{n} Y_k B_k,$$

where $B_1, \dots, B_n, Y_1, \dots, Y_n$ are independent with $B_k$ distributed $\mathrm{Ber}(1/k)$ or $\mathcal{P}(\theta/k)$, $E[Y_k] = k$ and $\mathrm{Var}(Y_k) = \sigma_k^2$, and provide an application to the minimal directed spanning tree in $\mathbb{R}^2$. We also provide bounds with optimal rates for the Dickman convergence of weighted sums, arising in probabilistic number theory, of the form

$$S_n = \frac{1}{\log(p_n)}\sum_{k=1}^{n} X_k \log(p_k),$$

where $(p_k)_{k\ge 1}$ is an enumeration of the prime numbers in increasing order and $X_k$ is geometric with parameter $(1 - 1/p_k)$, Bernoulli with success probability $1/(1+p_k)$ or Poisson with mean $\lambda_k$.

Lastly, we broaden the class of generalized Dickman distributions by studying the fixed points of the transformation

$$s(W^*) =_d U^{1/\theta} s(W+1)$$

generalizing (1), that allows the use of non-identity utility functions $s(\cdot)$ in Vervaat perpetuities. We obtain distributional bounds for recursive methods that can be used to simulate from this family.

*Keywords:* delay equation; distributional approximation; primes; utility; weighted Bernoulli sums

## References

[1] Arras, B. and Houdré, C. (2017). On Stein's method for infinitely divisible laws with finite first moment. Available at https://arxiv.org/abs/1712.10051.

[2] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2016). Distances between probability distributions via characteristic functions and biasing. Available at https://arxiv.org/abs/1605.06819v1.

[3] Arras, B., Mijoule, G., Poly, G. and Swan, Y. (2017). A new approach to the Stein–Tikhomirov method: With applications to the second Wiener chaos and Dickman convergence. Available at https://arxiv.org/abs/1605.06819.

[4] Arratia, R. (2002). On the amount of dependence in the prime factorization of a uniform random integer. In *Contemporary Combinatorics*. *Bolyai Soc*. *Math*. *Stud*. **10** 29–91. Budapest: János Bolyai Math. Soc. MR1919568

[5] Arratia, R. (2017). Personal communication.

[6] Arratia, R., Barbour, A.D. and Tavaré, S. (2003). *Logarithmic Combinatorial Structures*: *A Probabilistic Approach*. Zürich: European Mathematical Society (EMS). MR2032426

[7] Barbour, A.D. and Nietlispach, B. (2011). Approximation by the Dickman distribution and quasi-logarithmic combinatorial structures. *Electron*. *J*. *Probab*. **16** 880–902. MR2793242

[8] Baxendale, P. (2017). Personal communication.

[9] Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Comentarii Academiae Scientiarum Imperiales Petropolitanae* **5** 175–192.

[10] Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica* **22** 23–36. MR0059834

[11] Bhatt, A.G. and Roy, R. (2004). On a random directed spanning tree. *Adv*. *in Appl*. *Probab*. **36** 19–42. MR2035772

[12] Bhattacharjee, C. and Goldstein, L. (2018). Supplement to "Dickman approximation in simulation, summations and perpetuities." DOI:10.3150/18-BEJ1070SUPP.

[13] Cellarosi, F. and Sinai, Y.G. (2013). Non-standard limit theorems in number theory. In *Prokhorov and Contemporary Probability Theory*. *Springer Proc*. *Math*. *Stat*. **33** 197–213. Heidelberg: Springer. MR3070473

[14] Chen, L.H.Y., Goldstein, L. and Shao, Q.-M. (2011). *Normal Approximation by Stein's Method*. Heidelberg: Springer. MR2732624

[15] Devroye, L. and Fawzi, O. (2010). Simulating the Dickman distribution. *Statist*. *Probab*. *Lett*. **80** 242–247. MR2575452

[16] Dickman, K. (1930). On the frequency of numbers containing prime factors of a certain relative magnitude. *Ark*. *Mat*. *Astron*. *Fys*. **22** 1–14.

[17] Eeckhoudt, L., Gollier, C. and Schlesinger, H. (2005). *Economic and Financial Decisions Under Risk*. Princeton: Princeton Univ. Press.

[18] Fill, J.A. and Huber, M.L. (2010). Perfect simulation of Vervaat perpetuities. *Electron*. *J*. *Probab*. **15** 96–109. MR2587562

[19] Goldstein, L. (2017). Non asymptotic distributional bounds for the Dickman approximation of the running time of the Quickselect algorithm. Available at https://arxiv.org/abs/1703.00505v1.

[20] Goldstein, L. (2018). Non asymptotic distributional bounds for the Dickman approximation of the running time of the quickselect algorithm. *Electron*. *J*. *Probab*. **23** 1–13.

[21] Goldstein, L. and Reinert, G. (1997). Stein's method and the zero bias transformation with application to simple random sampling. *Ann*. *Appl*. *Probab*. **7** 935–952. MR1484792

[22] Hwang, H.-K. and Tsai, T.-H. (2002). Quickselect and the Dickman function. *Combin*. *Probab*. *Comput*. **11** 353–371. MR1918722

[23] Peköz, E.A. and Röllin, A. (2011). New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann*. *Probab*. **39** 587–608. MR2789507

[24] Peköz, E.A., Röllin, A. and Ross, N. (2013). Degree asymptotics with rates for preferential attachment random graphs. *Ann*. *Appl*. *Probab*. **23** 1188–1218. MR3076682

[25] Penrose, M.D. and Wade, A.R. (2004). Random minimal directed spanning trees and Dickman-type distributions. *Adv. in Appl. Probab.* **36** 691–714. MR2079909

[26] Pinsky, R. (2016). A natural probabilistic model on the integers and its relation to Dickman-type distributions and Buchstab's function. Available at https://arxiv.org/abs/1606.02965.

[27] Pinsky, R. (2016). On the strange domain of attraction to generalized Dickman distributions for sums of independent random variables. Available at https://arxiv.org/abs/1611.07207.

[28] Rachev, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*. Chichester: Wiley. MR1105086

[29] Ramanujan, S. (1919). A proof of Bertrand's postulate. *J. Indian Math. Soc.* (*N.S.*) **11** 27.

[30] Rényi, A. (1962). Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.* **8** 7–13. MR0286162

[31] Ross, N. (2011). Fundamentals of Stein's method. *Probab. Surv.* **8** 210–293. MR2861132

[32] Royden, H.L. and Fitzpatrick, P. (1988). *Real Analysis*, 3rd ed. New York: Macmillan Publishing Company. MR1013117

[33] Rudin, W. (1964). *Principles of Mathematical Analysis*, *Vol.* 3. New York: McGraw-Hill. MR0166310

[34] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. 583–602. MR0402873

[35] Stein, C. (1986). *Approximate Computation of Expectations*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **7**. Hayward, CA: IMS. MR0882007

[36] Stolz, O. (1885). *Vorlesungen über Allgemeine Arithmetik*: *Nach Den Neueren Ansichten* (*Vol.* 1). Leipzig: BG Teubner.

[37] Vervaat, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. in Appl. Probab.* **11** 750–783. MR0544194

*Bernoulli* **25**(4A), 2019, 2793–2823
https://doi.org/10.3150/18-BEJ1071

# Self-normalized Cramér type moderate deviations for martingales

XIEQUAN FAN[1], ION GRAMA[2,*], QUANSHENG LIU[2,**] and QI-MAN SHAO[3]

[1] *Center for Applied Mathematics, Tianjin University, Tianjin 300072, China.*
*E-mail: fanxiequan@hotmail.com*
[2] *Université de Bretagne-Sud, LMBA, UMR CNRS 6205, Campus de Tohannic, 56017 Vannes, France.*
*E-mail: *ion.grama@univ-ubs.fr; **quansheng.liu@univ-ubs.fr*
[3] *Department of Statistics, The Chinese University of Hong Kong, Shatin, NT, Hong Kong.*
*E-mail: qmshao@sta.cuhk.edu.hk*

Let $(X_i, \mathcal{F}_i)_{i\geq 1}$ be a sequence of martingale differences. Set $S_n = \sum_{i=1}^n X_i$ and $[S]_n = \sum_{i=1}^n X_i^2$. We prove a Cramér type moderate deviation expansion for $\mathbf{P}(S_n/\sqrt{[S]_n} \geq x)$ as $n \to +\infty$. Our results partly extend the earlier work of Jing, Shao and Wang (*Ann. Probab.* **31** (2003) 2167–2215) for independent random variables.

*Keywords:* Cramér's moderate deviations; martingales; self-normalized sequences

## References

[1] Bentkus, V., Bloznelis, M. and Götze, F. (1996). A Berry–Esséen bound for Student's statistic in the non-i.i.d. case. *J. Theoret. Probab.* **9** 765–796. MR1400598
[2] Bentkus, V. and Götze, F. (1996). The Berry–Esseen bound for Student's statistic. *Ann. Probab.* **24** 491–503. MR1387647
[3] Bercu, B., Delyon, B. and Rio, E. (2015). *Concentration Inequalities for Sums and Martingales. SpringerBriefs in Mathematics*. Cham: Springer. MR3363542
[4] Bercu, B. and Touati, A. (2008). Exponential inequalities for self-normalized martingales with applications. *Ann. Appl. Probab.* **18** 1848–1869. MR2462551
[5] Bolthausen, E. (1982). Exact convergence rates in some martingale central limit theorems. *Ann. Probab.* **10** 672–688. MR0659537
[6] Chen, S., Wang, Z., Xu, W. and Miao, Y. (2014). Exponential inequalities for self-normalized martingales. *J. Inequal. Appl.* **2014** 289. MR3317493
[7] Chung, K.-L. (1946). The approximate distribution of Student's statistic. *Ann. Math. Stat.* **17** 447–465. MR0018390
[8] Delyon, B. (2009). Exponential inequalities for sums of weakly dependent variables. *Electron. J. Probab.* **14** 752–779. MR2495559
[9] de la Peña, V.H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** 537–564. MR1681153
[10] de la Peña, V.H., Lai, T.L. and Shao, Q.-M. (2009). *Self-Normalized Processes: Limit Theory and Statistical Applications. Probability and Its Applications (New York)*. Berlin: Springer. MR2488094
[11] de la Peña, V.H. and Pang, G. (2009). Exponential inequalities for self-normalized processes with applications. *Electron. Commun. Probab.* **14** 372–381. MR2545288

1350-7265 

[12] Esscher, F. (1924). On a method of determining correlation from the ranks of the variates. *Scand*. *Actuar. J*. **1** 201–219.
[13] Fan, X., Grama, I., Liu, Q. and Shao, Q.-M. (2018). Supplement to "Self-normalized Cramér type moderate deviations for martingales." DOI:10.3150/18-BEJ1071SUPP.
[14] Giné, E., Götze, F. and Mason, D.M. (1997). When is the Student $t$-statistic asymptotically standard normal? *Ann*. *Probab*. **25** 1514–1531. MR1457629
[15] Grama, I. and Haeusler, E. (2000). Large deviations for martingales via Cramér's method. *Stochastic Process*. *Appl*. **85** 279–293. MR1731027
[16] Haeusler, E. (1988). On the rate of convergence in the central limit theorem for martingales with discrete and continuous time. *Ann*. *Probab*. **16** 275–299. MR0920271
[17] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*: *Probability and Mathematical Statistics*. New York: Academic Press [Harcourt Brace Jovanovich, Publishers]. MR0624435
[18] Hu, Z., Shao, Q.-M. and Wang, Q. (2009). Cramér type moderate deviations for the maximum of self-normalized sums. *Electron*. *J*. *Probab*. **14** 1181–1197. MR2511281
[19] Itô, K. and McKean, H.P. Jr. (1996). *Difussion Processes and Their Sample Paths*. Berlin: Springer.
[20] Jing, B., Liang, H. and Zhou, W. (2012). Self-normalized moderate deviations for independent random variables. *Sci*. *China Math*. **55** 2297–2315. MR2994121
[21] Jing, B.-Y., Shao, Q.-M. and Wang, Q. (2003). Self-normalized Cramér-type large deviations for independent random variables. *Ann*. *Probab*. **31** 2167–2215. MR2016616
[22] Linnik, Ju.V. (1961). On the probability of large deviations for the sums of independent variables. In *Proc*. 4*th Berkeley Sympos*. *Math*. *Statist*. *and Prob*., *Vol*. *II* 289–306. Berkeley, CA: Univ. California Press. MR0137142
[23] Liu, W., Shao, Q.-M. and Wang, Q. (2013). Self-normalized Cramér type moderate deviations for the maximum of sums. *Bernoulli* **19** 1006–1027. MR3079304
[24] Novak, S.Y. (2011). *Extreme Value Methods with Applications to Finance*. *Monographs on Statistics and Applied Probability* **122**. Boca Raton, FL: CRC Press. MR2933280
[25] Petrov, V.V. (1975). *Sums of Independent Random Variables*. New York: Springer. Translated from the Russian by A. A. Brown, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 82. MR0388499
[26] Shao, Q.-M. (1997). Self-normalized large deviations. *Ann*. *Probab*. **25** 285–328. MR1428510
[27] Shao, Q.-M. (1999). A Cramér type large deviation result for Student's $t$-statistic. *J*. *Theoret*. *Probab*. **12** 385–398. MR1684750
[28] Shao, Q.-M. and Wang, Q. (2013). Self-normalized limit theorems: A survey. *Probab*. *Surv*. **10** 69–93. MR3161676
[29] Shao, Q.-M. and Zhou, W.-X. (2016). Cramér type moderate deviation theorems for self-normalized processes. *Bernoulli* **22** 2029–2079. MR3498022
[30] Slavova, V.V. (1985). On the Berry–Esseen bound for Student's statistic. In *Stability Problems for Stochastic Models* (*Uzhgorod*, 1984). *Lecture Notes in Math*. **1155** 355–390. Berlin: Springer. MR0825335
[31] Talagrand, M. (1995). The missing factor in Hoeffding's inequalities. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **31** 689–702. MR1355613
[32] von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the $r$th absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann*. *Math*. *Stat*. **36** 299–303. MR0170407

*Bernoulli* **25**(4A), 2019, 2824–2853
https://doi.org/10.3150/18-BEJ1072

# A multivariate Berry–Esseen theorem with explicit constants

MARTIN RAIČ

*This paper is dedicated to the memory of Vidmantas Kastytis Bentkus (1949–2010).*

*University of Ljubljana, University of Primorska, and Institute of Mathematics, Physics and Mechanics, Slovenia. E-mail: martin.raic@fmf.uni-lj.si; url: valjhun.fmf.uni-lj.si/~raicm/*

We provide a Lyapunov type bound in the multivariate central limit theorem for sums of independent, but not necessarily identically distributed random vectors. The error in the normal approximation is estimated for certain classes of sets, which include the class of measurable convex sets. The error bound is stated with explicit constants. The result is proved by means of Stein's method. In addition, we improve the constant in the bound of the Gaussian perimeter of convex sets.

*Keywords:* Berry–Esseen theorem; explicit constants; Lyapunov bound; multivariate central limit theorem; Stein's method

## References

[1] Abramowitz, M. and Stegun, I.A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series* **55**. Washington, DC: U.S. Government Printing Office. MR0167642
[2] Ball, K. (1993). The reverse isoperimetric problem for Gaussian measure. *Discrete Comput. Geom.* **10** 411–420. MR1243336
[3] Banach, S. (1938). Über homogene Polynome in $L^2$. *Studia Math.* **7** 36–44.
[4] Barbour, A.D. and Čekanavičius, V. (2002). Total variation asymptotics for sums of independent integer random variables. *Ann. Probab.* **30** 509–545. MR1905850
[5] Bentkus, V. (2003). On the dependence of the Berry–Esseen bound on dimension. *J. Statist. Plann. Inference* **113** 385–402. MR1965117
[6] Bentkus, V. (2005). A Lyapunov type bound in $\mathbf{R}^d$. *Theory Probab. Appl.* **49** 311–323. MR2144310
[7] Bentkus, V.Yu. (1986). Dependence of the Berry–Esseen estimate on the dimension. *Lith. Math. J.* **26** 110–114. MR0862741
[8] Bhattacharya, R. and Holmes, S. (2010). An exposition in Götze's estimation of the rate of convergence in the multivariate central limit theorem. Available at arXiv:1003.4251v1.
[9] Bochnak, J. and Siciak, J. (1971). Polynomials and multilinear mappings in topological vector spaces. *Studia Math.* **39** 59–76. MR0313810
[10] Chen, L.H.Y. and Fang, X. (2015). Multivariate normal approximation by Stein's method: The concentration inequality approach. Available at arXiv:1111.4073v2.
[11] Chen, L.H.Y. and Shao, Q.-M. (2001). A non-uniform Berry–Esseen bound via Stein's method. *Probab. Theory Related Fields* **120** 236–254. MR1841329
[12] Chernozhukov, V., Chetverikov, D. and Kato, K. (2013). Supplement to "Gaussian approximations and multiplier bootstrap for maxima of sums of high-dimensional random vectors". doi:10.1214/13-AOS1161SUPP. MR3161448

1350-7265 

[13] Esseen, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math*. **77** 1–125. MR0014626

[14] Federer, H. (1969). *Geometric Measure Theory*. *Die Grundlehren der Mathematischen Wissenschaften*, *Band* 153. New York: Springer. MR0257325

[15] Gaunt, R.E., Pickett, A.M. and Reinert, G. (2017). Chi-square approximation by Stein's method with application to Pearson's statistic. *Ann*. *Appl*. *Probab*. **27** 720–756. MR3655852

[16] Götze, F. (1991). On the rate of convergence in the multivariate CLT. *Ann*. *Probab*. **19** 724–739. MR1106283

[17] Livshyts, G. (2014). Maximal surface area of a convex set in $\mathbb{R}^n$ with respect to log concave rotation invariant measures. In *Geometric Aspects of Functional Analysis*. *Lecture Notes in Math*. **2116** 355–383. Cham: Springer. MR3364697

[18] Nagaev, S.V. (1976). An estimate of the remainder term in the multidimensional central limit theorem. In *Proceedings of the Third Japan–USSR Symposium on Probability Theory* (*Tashkent*, 1975). *Lecture Notes in Math*. **550** 419–438. Berlin: Springer. MR0443043

[19] Nazarov, F. (2003). On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a Gaussian measure. In *Geometric Aspects of Functional Analysis*. *Lecture Notes in Math*. **1807** 169–187. Berlin: Springer. MR2083397

[20] Raič, M. (2018). Supplement to "A multivariate Berry–Esseen theorem with explicit constants." DOI:10.3150/18-BEJ1072SUPP.

[21] Röllin, A. (2013). Stein's method in high dimensions with applications. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **49** 529–549. MR3088380

[22] Sazonov, V.V. (1972). On a bound for the rate of convergence in the multidimensional central limit theorem. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif*., 1970/1971), *Vol. II*: *Probability Theory* 563–581. Berkeley, CA: Univ. California Press. MR0400351

[23] Sazonov, V.V. (1981). *Normal Approximation – Some Recent Advances*. *Lecture Notes in Math*. **879**. Berlin: Springer. MR0643968

[24] Shevtsova, I.G. (2010). Refinement of estimates for the rate of convergence in Lyapunov's theorem. *Dokl*. *Math*. **82** 862–864. MR2790498

[25] Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Syst*. *Tech*. *J*. **41** 463–501. MR0133183

[26] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif*., 1970/1971), *Vol. II*: *Probability Theory* 583–602. Berkeley, CA: Univ. California Press. MR0402873

[27] Tihomirov, A.N. (1980). Convergence rate in the central limit theorem for weakly dependent random variables. *Teor. Veroyatn*. *Primen*. **25** 800–818. MR0595140

*Bernoulli* **25**(4A), 2019, 2854–2882
https://doi.org/10.3150/18-BEJ1073

# High-dimensional Bayesian inference via the unadjusted Langevin algorithm

ALAIN DURMUS[1] and ÉRIC MOULINES[2]

[1]*CMLA, ENS Cachan, CNRS, Université Paris-Saclay, 94235 Cachan, France. E-mail: alain.durmus@cmla.ens-cachan.fr*
[2]*Centre de Mathématiques Appliquées, UMR 7641, Ecole Polytechnique, route de Saclay, 91128 Palaiseau cedex, France. E-mail: eric.moulines@polytechnique.edu*

We consider in this paper the problem of sampling a high-dimensional probability distribution $\pi$ having a density w.r.t. the Lebesgue measure on $\mathbb{R}^d$, known up to a normalization constant $x \mapsto \pi(x) = \mathrm{e}^{-U(x)}/\int_{\mathbb{R}^d} \mathrm{e}^{-U(y)}\,\mathrm{d}y$. Such problem naturally occurs for example in Bayesian inference and machine learning. Under the assumption that $U$ is continuously differentiable, $\nabla U$ is globally Lipschitz and $U$ is strongly convex, we obtain non-asymptotic bounds for the convergence to stationarity in Wasserstein distance of order 2 and total variation distance of the sampling method based on the Euler discretization of the Langevin stochastic differential equation, for both constant and decreasing step sizes. The dependence on the dimension of the state space of these bounds is explicit. The convergence of an appropriately weighted empirical measure is also investigated and bounds for the mean square error and exponential deviation inequality are reported for functions which are measurable and bounded. An illustration to Bayesian inference for binary regression is presented to support our claims.

*Keywords:* Langevin diffusion; Markov chain Monte Carlo; Metropolis adjusted Langevin algorithm; rate of convergence; total variation distance

## References

[1] Albert, J.H. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. MR1224394
[2] Borodin, A.N. and Salminen, P. (2002). *Handbook of Brownian Motion—Facts and Formulae*, 2nd ed. *Probability and Its Applications*. Basel: Birkhäuser. MR1912205
[3] Bubeck, S., Eldan, R. and Lehec, J. (2015). Finite-time analysis of projected Langevin Monte Carlo. In *Proceedings of the* 28*th International Conference on Neural Information Processing Systems*, *NIPS*'15 1243–1251. Cambridge, MA, USA: MIT Press.
[4] Bubley, R., Dyer, M. and Jerrum, M. (1998). An elementary analysis of a procedure for sampling points in a convex body. *Random Structures Algorithms* **12** 213–235. MR1635248
[5] Chen, M.F. and Li, S.F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.* **17** 151–177. MR0972776
[6] Choi, H.M. and Hobert, J.P. (2013). The Polya-gamma Gibbs sampler for Bayesian logistic regression is uniformly ergodic. *Electron. J. Stat.* **7** 2054–2064. MR3091616
[7] Chopin, N. and Ridgway, J. (2017). Leave Pima Indians alone: Binary regression as a benchmark for Bayesian computation. *Statist. Sci.* **32** 64–87. MR3634307
[8] Dalalyan, A.S. Further and stronger analogy between sampling and optimization: Langevin Monte Carlo and gradient descent. In *Proceedings of the* 30*th Annual Conference on Learning Theory*.

[9] 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
[10] Durmus, A. and Moulines, É. (2017). Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *Ann. Appl. Probab.* **27** 1551–1587. MR3678479
[11] Durmus, A. and Moulines, É. (2019). Supplement to "High-dimensional Bayesian inference via the unadjusted Langevin algorithm." DOI:10.3150/18-BEJ1073SUPP.
[12] Eberle, A. Quantitative contraction rates for Markov chains on continuous state spaces. In preparation.
[13] Eberle, A. (2016). Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** 851–886. MR3568041
[14] Eberle, A., Guillin, A. and Zimmer, R. (2018). Quantitative Harris type theorems for diffusions and McKean–Vlasov processes. *Trans. Amer. Math. Soc.* To appear.
[15] Ermak, D.L. (1975). A computer simulation of charged particles in solution. I. Technique and equilibrium properties. *J. Chem. Phys.* **62** 4189–4196.
[16] Faes, C., Ormerod, J.T. and Wand, M.P. (2011). Variational Bayesian inference for parametric and nonparametric regression with missing data. *J. Amer. Statist. Assoc.* **106** 959–971. MR2894756
[17] Frühwirth-Schnatter, S. and Frühwirth, R. (2010). Data augmentation and MCMC for binary and multinomial logic models. In *Statistical Modelling and Regression Structures* 111–132. Heidelberg: Physica-Verlag/Springer. MR2664631
[18] Gramacy, R.B. and Polson, N.G. (2012). Simulation-based regularized logistic regression. *Bayesian Anal.* **7** 567–589. MR2981628
[19] Grenander, U. (1996). *Elements of Pattern Theory. Johns Hopkins Studies in the Mathematical Sciences*. Baltimore, MD: Johns Hopkins Univ. Press. MR1405164
[20] Grenander, U. and Miller, M.I. (1994). Representations of knowledge in complex systems. *J. Roy. Statist. Soc. Ser. B* **56** 549–603. With discussion and a reply by the authors. MR1293234
[21] Hanson, T.E., Branscum, A.J. and Johnson, W.O. (2014). Informative $g$-priors for logistic regression. *Bayesian Anal.* **9** 597–611. MR3256057
[22] Holmes, C.C. and Held, L. (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Anal.* **1** 145–168. MR2227368
[23] Joulin, A. and Ollivier, Y. (2010). Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.* **38** 2418–2442. MR2683634
[24] Karatzas, I. and Shreve, S.E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. New York: Springer. MR1121940
[25] Klartag, B. (2007). A central limit theorem for convex sets. *Invent. Math.* **168** 91–131. MR2285748
[26] Lamberton, D. and Pagès, G. (2002). Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8** 367–405. MR1913112
[27] Lamberton, D. and Pagès, G. (2003). Recursive computation of the invariant distribution of a diffusion: The case of a weakly mean reverting drift. *Stoch. Dyn.* **3** 435–451. MR2030742
[28] Lemaire, V. (2005). Estimation de la mesure invariante d'un processus de diffusion. Ph.D. thesis, Université Paris-Est.
[29] Lindvall, T. and Rogers, L.C.G. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872. MR0841588
[30] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266
[31] Neal, R.M. (1993). Bayesian learning via stochastic dynamics. In *Advances in Neural Information Processing Systems* 5, [*NIPS Conference*] 475–482. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

[32] Nesterov, Y. (2004). *Introductory Lectures on Convex Optimization*: *A Basic Course*. *Applied Optimization* **87**. Boston, MA: Kluwer Academic. MR2142598

[33] Parisi, G. (1981). Correlation functions and computer simulations. *Nuclear Phys*. *B* **180** 378–384. MR0615176

[34] Polson, N.G., Scott, J.G. and Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *J. Amer. Statist*. *Assoc*. **108** 1339–1349. MR3174712

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

[36] Rossky, P.J., Doll, J.D. and Friedman, H.L. (1978). Brownian dynamics as smart Monte Carlo simulation. *J. Chem*. *Phys*. **69** 4628–4633.

[37] Sabanés Bové, D. and Held, L. (2011). Hyper-$g$ priors for generalized linear models. *Bayesian Anal*. **6** 387–410. MR2843537

[38] Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch*. *Anal*. *Appl*. **8** 483–509. MR1091544

[39] Villani, C. (2009). *Optimal Transport*: *Old and New*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Berlin: Springer. MR2459454

[40] Welling, M. and Teh, Y.W. (2011). Bayesian learning via stochastic gradient Langevin dynamics. In *Proceedings of the* 28*th International Conference on Machine Learning* (*ICML*-11) 681–688.

[41] Windle, J., Polson, N.G. and Scott, J.G. (2013). Bayeslogit: Bayesian logistic regression. R package version 0.2. Available at http://cran.r-project.org/web/packages/BayesLogit/index.html.

*Bernoulli* **25**(4A), 2019, 2883–2919
https://doi.org/10.3150/18-BEJ1074

# A supermartingale approach to Gaussian process based sequential design of experiments

JULIEN BECT[1,*], FRANÇOIS BACHOC[2] and DAVID GINSBOURGER[3,4]

[1]*Laboratoire des Signaux et Systèmes (L2S), CentraleSupélec, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France. E-mail: julien.bect@centralesupelec.fr*

[2]*Toulouse Mathematics Institute, University Paul Sabatier, France. E-mail: francois.bachoc@math.univ-toulouse.fr*

[3]*Uncertainty Quantification and Optimal Design group, Idiap Research Institute, Martigny, Switzerland. E-mail: ginsbourger@idiap.ch*

[4]*Institute of Mathematical Statistics and Actuarial Science, Department of Mathematics and Statistics, University of Bern, Switzerland. E-mail: ginsbourger@stat.unibe.ch*

Gaussian process (GP) models have become a well-established framework for the adaptive design of costly experiments, and notably of computer experiments. GP-based sequential designs have been found practically efficient for various objectives, such as global optimization (estimating the global maximum or maximizer(s) of a function), reliability analysis (estimating a probability of failure) or the estimation of level sets and excursion sets. In this paper, we study the consistency of an important class of sequential designs, known as stepwise uncertainty reduction (SUR) strategies. Our approach relies on the key observation that the sequence of residual uncertainty measures, in SUR strategies, is generally a supermartingale with respect to the filtration generated by the observations. This observation enables us to establish generic consistency results for a broad class of SUR strategies. The consistency of several popular sequential design strategies is then obtained by means of this general result. Notably, we establish the consistency of two SUR strategies proposed by Bect, Ginsbourger, Li, Picheny and Vazquez (*Stat. Comput.* **22** (2012) 773–793) – to the best of our knowledge, these are the first proofs of consistency for GP-based sequential design algorithms dedicated to the estimation of excursion sets and their measure. We also establish a new, more general proof of consistency for the expected improvement algorithm for global optimization which, unlike previous results in the literature, applies to any GP with continuous sample paths.

*Keywords:* active learning; convergence; sequential design of experiments; stepwise uncertainty reduction; supermartingale; uncertainty functional

## References

[1] Adler, R.J. (1981). *The Geometry of Random Fields*. *Wiley Series in Probability and Mathematical Statistics*. Chichester: Wiley. MR0611857

[2] Azaïs, J.-M. and Wschebor, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. Hoboken, NJ: Wiley. MR2478201

[3] Azzimonti, D., Ginsbourger, D., Chevalier, C., Bect, J. and Richet, Y. (2018). Adaptive design of experiments for conservative estimation of excursion sets. ArXiv preprint, arXiv:1611.07256v3.

1350-7265 

[4] Bayarri, M.J., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.-H. and Tu, J. (2007). A framework for validation of computer models. *Technometrics* **49** 138–154. MR2380530
[5] Bect, J., Li, L. and Vazquez, E. (2017). Bayesian subset simulation. *SIAM/ASA J. Uncertain. Quantificat.* **5** 762–786. MR3683699
[6] Bect, J., Ginsbourger, D., Li, L., Picheny, V. and Vazquez, E. (2012). Sequential design of computer experiments for the estimation of a probability of failure. *Stat. Comput.* **22** 773–793. MR2909621
[7] 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
[8] Binois, M. (2015). Uncertainty quantification on Pareto fronts and high-dimensional strategies in Bayesian optimization, with applications in multi-objective automotive design. Ph.D. thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne.
[9] Bogachev, V.I. (1998). *Gaussian Measures*. *Mathematical Surveys and Monographs* **62**. Providence, RI: Amer. Math. Soc.. MR1642391
[10] Brezis, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. *Universitext*. New York: Springer. MR2759829
[11] Bull, A.D. (2011). Convergence rates of efficient global optimization algorithms. *J. Mach. Learn. Res.* **12** 2879–2904. MR2854351
[12] Chevalier, C. (2013). Fast uncertainty reduction strategies relying on Gaussian process models. Ph.D. thesis, University of Bern.
[13] Chevalier, C., Ginsbourger, D., Bect, J., Vazquez, E., Picheny, V. and Richet, Y. (2014). Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics* **56** 455–465. MR3290615
[14] Cohn, D.A., Ghahramani, Z. and Jordan, M.I. (1996). Active learning with statistical models. *J. Artificial Intelligence Res.* **4** 129–145.
[15] Cover, T.M. and Thomas, J.A. (1991). *Elements of Information Theory*. *Wiley Series in Telecommunications*. New York: Wiley. A Wiley-Interscience Publication. MR1122806
[16] Cressie, N.A.C. (1993). *Statistics for Spatial Data*. *Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. Revised reprint of the 1991 edition, A Wiley-Interscience Publication. MR1239641
[17] DeGroot, M.H. (1962). Uncertainty, information, and sequential experiments. *Ann. Math. Stat.* **33** 404–419. MR0139242
[18] DeGroot, M.H. (1984). Changes in utility as information. *Theory and Decision* **17** 287–303. MR0781339
[19] DeGroot, M.H. (1986). Concepts of information based on utility. In *Recent Developments in the Foundations of Utility and Risk Theory* (L. Daboni, A. Montesano and M. Lines, eds.) 265–275. Netherlands, Dordrecht: Springer.
[20] Diaconis, P. (1988). Bayesian numerical analysis. In *Statistical Decision Theory and Related Topics, IV, Vol.* 1 (*West Lafayette, Ind.*, 1986) 163–175. New York: Springer. MR0927099
[21] Emmerich, M., Giannakoglou, K. and Naujoks, B. (2006). Single-and multiobjective optimization assisted by Gaussian random field metamodels. *IEEE Trans. Evol. Comput.* **10**.
[22] Feliot, P., Bect, J. and Vazquez, E. (2017). A Bayesian approach to constrained single- and multi-objective optimization. *J. Global Optim.* **67** 97–133. MR3596831
[23] Forrester, A., Sobester, A. and Keane, A. (2008). *Engineering Design Via Surrogate Modelling: A Practical Guide*. Wiley.
[24] Frazier, P. (2009). *Knowledge-Gradient Methods for Statistical Learning*. Ann Arbor, MI: ProQuest LLC. Thesis (Ph.D.)–Princeton University. MR2713050

[25] Frazier, P.I., Powell, W.B. and Dayanik, S. (2008). A knowledge-gradient policy for sequential information collection. *SIAM J. Control Optim.* **47** 2410–2439. MR2448467
[26] Frazier, P., Powell, W. and Dayanik, S. (2009). The knowledge-gradient policy for correlated normal beliefs. *INFORMS J. Comput.* **21** 599–613. MR2588343
[27] Geman, D. and Jedynak, B. (1996). An active testing model for tracking roads in satellite images. *IEEE Trans. Pattern Anal. Mach. Intell.* **18** 1–14.
[28] Ginsbourger, D., Roustant, O. and Durrande, N. (2016). On degeneracy and invariances of random fields paths with applications in Gaussian process modelling. *J. Statist. Plann. Inference* **170** 117–128. MR3427864
[29] Ginsbourger, D., Baccou, J., Chevalier, C., Perales, F., Garland, N. and Monerie, Y. (2014). Bayesian adaptive reconstruction of profile optima and optimizers. *SIAM/ASA J. Uncertain. Quantificat.* **2** 490–510. MR3283918
[30] Gramacy, R.B., Gray, G.A., Le Digabel, S., Lee, H.K.H., Ranjan, P., Wells, G. and Wild, S.M. (2016). Modeling an augmented Lagrangian for blackbox constrained optimization. *Technometrics* **58** 1–11. MR3463148
[31] Grünewälder, S., Audibert, J.-Y., Opper, M. and Shawe-Taylor, J. (2010). Regret bounds for Gaussian process bandit problems. In *International Conference on Artificial Intelligence and Statistics*.
[32] Hainy, M., Müller, W.G. and Wynn, H.P. (2014). Learning functions and approximate Bayesian computation design: ABCD. *Entropy* **16** 4353–4374. MR3255991
[33] Hennig, P., Osborne, M.A. and Girolami, M. (2015). Probabilistic numerics and uncertainty in computations. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **471** 20150142, 17. MR3378744
[34] Ibragimov, I.d.A. and Rozanov, Y.A. (1978). *Gaussian Random Processes. Applications of Mathematics* **9**. New York–Berlin: Springer. Translated from the Russian by A.B. Aries. MR0543837
[35] Johnson, R. (1960). An information theory approach to diagnosis. In *Proceedings of the* 6*th Symposium on Reliability and Quality Control* 102–109.
[36] Jones, D.R., Schonlau, M. and Welch, W.J. (1998). Efficient global optimization of expensive blackbox functions. *J. Global Optim.* **13** 455–492. Workshop on Global Optimization (Trier, 1997). MR1673460
[37] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169
[38] King-Smith, P.E., Grigsby, S.S., Vingrys, A.J., Benes, S.C. and Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Res.* **34** 885–912.
[39] Koehler, J.R., Puhalskii, A.A. and Simon, B. (1998). Estimating functions evaluated by simulation: A Bayesian/analytic approach. *Ann. Appl. Probab.* **8** 1184–1215. MR1661156
[40] Ledoux, M. and Talagrand, M. (2011). *Probability in Banach Spaces*: *Isoperimetry and Processes*. *Classics in Mathematics*. Berlin: Springer. Reprint of the 1991 edition. MR2814399
[41] Lukić, M.N. and Beder, J.H. (2001). Stochastic processes with sample paths in reproducing kernel Hilbert spaces. *Trans. Amer. Math. Soc.* **353** 3945–3969. MR1837215
[42] MacKay, D.J.C. (1992). Information-based objective functions for active data selection. *Neural Comput.* **4** 590–604.
[43] Mockus, J.B., Tiesis, V. and Žilinskas, A. (1978). The application of Bayesian methods for seeking the extremum. In *Towards Global Optimization*, *Volume* 2 (L.C.W. Dixon and G.P. Szegö, eds.) 117–129. New York: North Holland.
[44] Molchanov, I. (2005). *Theory of Random Sets*. *Probability and Its Applications* (*New York*). London: Springer London, Ltd.. MR2132405
[45] O'Geran, J.H., Wynn, H.P. and Zhiglyavsky, A.A. (1993). Mastermind as a test-bed for search algorithms. *Chance* **6** 31–37.

[46] O'Hagan, A. (1991). Bayes–Hermite quadrature. *J. Statist. Plann. Inference* **29** 245–260. MR1144171
[47] Perlman, M.D. (1974). Jensen's inequality for a convex vector-valued function on an infinite-dimensional space. *J. Multivariate Anal.* **4** 52–65. MR0362421
[48] Picheny, V. (2014). A stepwise uncertainty reduction approach to constrained global optimization. In *Proceedings of the* 17*th International Conference on Artificial Intelligence and Statistics* (*AISTATS*).
[49] Picheny, V., Ginsbourger, D., Roustant, O., Haftka, R. and Kim, N.-H. (2010). Adaptive designs of experiments for accurate approximation of target regions. *J. Mech. Des.* **132**.
[50] Ranjan, P., Bingham, D. and Michailidis, G. (2008). Sequential experiment design for contour estimation from complex computer codes. *Technometrics* **50** 527–541. MR2655651
[51] Ritter, K. (2000). *Average-Case Analysis of Numerical Problems*. *Lecture Notes in Math.* **1733**. Berlin: Springer. MR1763973
[52] Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989). Design and analysis of computer experiments. *Statist. Sci.* **4** 409–435. With comments and a rejoinder by the authors. MR1041765
[53] Santner, T.J., Williams, B.J. and Notz, W.I. (2003). *The Design and Analysis of Computer Experiments*. *Springer Series in Statistics*. New York: Springer. MR2160708
[54] Schönfeld, P. (1973). A note on the measurability of the pseudo-inverse. *J. Econometrics* **1** 313–314.
[55] Scott, W., Frazier, P. and Powell, W. (2011). The correlated knowledge gradient for simulation optimization of continuous parameters using Gaussian process regression. *SIAM J. Optim.* **21** 996–1026. MR2837561
[56] Shahriari, B., Swersky, K., Wang, Z., Adams, R. and de Freitas, N. (2016). Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **104** 148–175.
[57] Srinivas, N., Krause, A., Kakade, S.M. and Seeger, M.W. (2012). Information-theoretic regret bounds for Gaussian process optimization in the bandit setting. *IEEE Trans. Inform. Theory* **58** 3250–3265. MR2952544
[58] Stein, M.L. (1999). *Interpolation of Spatial Data*: *Some Theory for Kriging*. *Springer Series in Statistics*. New York: Springer. MR1697409
[59] Stroock, D.W. (2011). *Probability Theory*: *An Analytic View*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2760872
[60] 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. Translated from the Russian and with a preface by Wojbor A. Woyczynski. MR1435288
[61] 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
[62] Vazquez, E. and Bect, J. (2009). A sequential Bayesian algorithm to estimate a probability of failure. In 15*th IFAC Symposium on System Identification* (*SYSID* 2009).
[63] Vazquez, E. and Bect, J. (2010). Convergence properties of the expected improvement algorithm with fixed mean and covariance functions. *J. Statist. Plann. Inference* **140** 3088–3095. MR2659839
[64] Vazquez, E. and Bect, J. (2010). Pointwise consistency of the kriging predictor with known mean and covariance functions. In *MODa* 9*–Advances in Model-Oriented Design and Analysis* 221–228. Springer.
[65] Vazquez, E. and Bect, J. (2012). Sequential search based on kriging: Convergence analysis of some algorithms. In *Bulletin of the ISI* 58*th World Statistics Congress of the International Statistical Institute*, 2011, *The Hague*, *The Netherlands* International Statistical Institute.
[66] Villemonteix, J., Vazquez, E. and Walter, E. (2009). An informational approach to the global optimization of expensive-to-evaluate functions. *J. Global Optim.* **44** 509–534. MR2525048
[67] Wang, H., Lin, G. and Li, J. (2016). Gaussian process surrogates for failure detection: A Bayesian experimental design approach. *J. Comput. Phys.* **313** 247–259. MR3481014

[68] Williams, B.J., Santner, T.J. and Notz, W.I. (2000). Sequential design of computer experiments to minimize integrated response functions. *Statist*. *Sinica* **10** 1133–1152. MR1804554

[69] Yarotsky, D. (2013). Examples of inconsistency in optimization by expected improvement. *J*. *Global Optim*. **56** 1773–1790. MR3078332

[70] Yarotsky, D. (2013). Univariate interpolation by exponential functions and Gaussian RBFs for generic sets of nodes. *J*. *Approx*. *Theory* **166** 163–175. MR3003954

[71] Zuluaga, M., Krause, A., Sergent, G. and Püschel, M. (2013). Active learning for level set estimation. In *International Joint Conference on Artificial Intelligence* (*IJCAI*).

*Bernoulli* **25**(4A), 2019, 2920–2948
https://doi.org/10.3150/18-BEJ1075

# On rate of convergence in non-central limit theorems

VO ANH[1], NIKOLAI LEONENKO[2], ANDRIY OLENKO[3] and VOLODYMYR VASKOVYCH[3]

[1]*School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD, 4001, Australia. E-mail: v.anh@qut.edu.au*
[2]*School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK. E-mail: LeonenkoN@cardiff.ac.uk*
[3]*Department of Mathematics and Statistics, La Trobe University, Melbourne, VIC, 3086, Australia. E-mail: a.olenko@latrobe.edu.au; vaskovych.v@students.latrobe.edu.au*

The main result of this paper is the rate of convergence to Hermite-type distributions in non-central limit theorems. To the best of our knowledge, this is the first result in the literature on rates of convergence of functionals of random fields to Hermite-type distributions with ranks greater than 2. The results were obtained under rather general assumptions on the spectral densities of random fields. These assumptions are even weaker than in the known convergence results for the case of Rosenblatt distributions. Additionally, Lévy concentration functions for Hermite-type distributions were investigated.

*Keywords:* Hermite-type distribution; long-range dependence; non-central limit theorems; random field; rate of convergence

## References

[1] Anh, V., Leonenko, N. and Olenko, A. (2015). On the rate of convergence to Rosenblatt-type distribution. *J. Math. Anal. Appl.* **425** 111–132. MR3299653
[2] Arcones, M.A. (1994). Limit theorems for nonlinear functionals of a stationary Gaussian sequence of vectors. *Ann. Probab.* **22** 2242–2274. MR1331224
[3] Arras, B., Azmoodeh, E., Poly, G. and Swan, Y. (2016). Stein's method on the second Wiener chaos: 2-Wasserstein distance. Available at arXiv:1601.03301.
[4] Bai, S. and Taqqu, M.S. (2017). The impact of the diagonals of polynomial forms on limit theorems with long memory. *Bernoulli* **23** 710–742. MR3556790
[5] 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
[6] Brandolini, L., Hofmann, S. and Iosevich, A. (2003). Sharp rate of average decay of the Fourier transform of a bounded set. *Geom. Funct. Anal.* **13** 671–680. MR2006553
[7] Breton, J.-C. (2011). On the rate of convergence in non-central asymptotics of the Hermite variations of fractional Brownian sheet. *Probab. Math. Statist.* **31** 301–311. MR2853680
[8] Breton, J.-C. and Nourdin, I. (2008). Error bounds on the non-normal approximation of Hermite power variations of fractional Brownian motion. *Electron. Commun. Probab.* **13** 482–493. MR2447835
[9] Davydov, Y.A. (1990). Distributions of multiple Wiener–Itô integrals. *Theory Probab. Appl.* **35** 51–62. MR1050053
[10] Davydov, Y.A. and Manukyan, R.R. (1995). A local limit theorem for multiple stochastic Wiener–Itô integrals. *Theory Probab. Appl.* **40** 423–430. MR1346478

1350-7265 

[11] Davydov, Y.A. and Martynova, G.V. (1989). Limit behavior of distributions of multiple stochastic integrals. In *Statistics and Control of Random Processes* (*Russian*) (*Preila*, 1987) 55–57. Moscow: "Nauka". MR1079335
[12] 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
[13] Giraitis, L. and Surgailis, D. (1985). CLT and other limit theorems for functionals of Gaussian processes. *Z. Wahrsch. Verw. Gebiete* **70** 191–212. MR0799146
[14] Gubner, J.A. (2005). Theorems and fallacies in the theory of long-range-dependent processes. *IEEE Trans. Inform. Theory* **51** 1234–1239. MR2237996
[15] Houdré, C. and Pérez-Abreu, V. (1994). *Chaos Expansions, Multiple Wiener–Itô Integrals and Their Applications. Probab. Stochastics Ser.* Boca Raton, FL: CRC.
[16] Hu, Y., Lu, F. and Nualart, D. (2014). Convergence of densities of some functionals of Gaussian processes. *J. Funct. Anal.* **266** 814–875. MR3132731
[17] Ivanov, A.V. and Leonenko, N.N. (1989). *Statistical Analysis of Random Fields. Mathematics and Its Applications* (*Soviet Series*) **28**. Dordrecht: Kluwer Academic. MR1009786
[18] Kim, Y.T. and Park, H.S. (2015). Kolmogorov distance for the central limit theorems of the Wiener chaos expansion and applications. *J. Korean Statist. Soc.* **44** 565–576. MR3421893
[19] Leonenko, N. and Olenko, A. (2013). Tauberian and Abelian theorems for long-range dependent random fields. *Methodol. Comput. Appl. Probab.* **15** 715–742. MR3117624
[20] Leonenko, N. and Olenko, A. (2014). Sojourn measures of Student and Fisher–Snedecor random fields. *Bernoulli* **20** 1454–1483. MR3217450
[21] Leonenko, N.N. (1988). On the accuracy of the normal approximation of functionals of strongly correlated Gaussian random fields. *Math. Notes* **43** 161–171.
[22] Leonenko, N.N. and Anh, V.V. (2001). Rate of convergence to the Rosenblatt distribution for additive functionals of stochastic processes with long-range dependence. *J. Appl. Math. Stoch. Anal.* **14** 27–46. MR1825099
[23] Leonenko, N.N., Ruiz-Medina, M.D. and Taqqu, M.S. (2017). Rosenblatt distribution subordinated to Gaussian random fields with long-range dependence. *Stoch. Anal. Appl.* **35** 144–177. MR3581700
[24] Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere. London Mathematical Society Lecture Note Series* **389**. Cambridge: Cambridge Univ. Press. MR2840154
[25] Neufcourt, L. and Viens, F.G. (2016). A third-moment theorem and precise asymptotics for variations of stationary Gaussian sequences. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** 239–264. MR3479476
[26] Nourdin, I. (2018). Malliavin-Stein approach, website, accessed 10 September 2018. Available at https://sites.google.com/site/malliavinstein/home.
[27] Nourdin, I., Nualart, D. and Poly, G. (2013). Absolute continuity and convergence of densities for random vectors on Wiener chaos. *Electron. J. Probab.* **18** 22. MR3035750
[28] Nourdin, I. and Peccati, G. (2009). Stein's method on Wiener chaos. *Probab. Theory Related Fields* **145** 75–118. MR2520122
[29] Nourdin, I. and Peccati, G. (2009). Stein's method and exact Berry–Esseen asymptotics for functionals of Gaussian fields. *Ann. Probab.* **37** 2231–2261. MR2573557
[30] Nourdin, I., Peccati, G. and Podolskij, M. (2011). Quantitative Breuer–Major theorems. *Stochastic Process. Appl.* **121** 793–812. MR2770907
[31] Nourdin, I. and Poly, G. (2013). Convergence in total variation on Wiener chaos. *Stochastic Process. Appl.* **123** 651–674. MR3003367
[32] Olenko, A.Y. (2005). A Tauberian theorem for fields with the OR spectrum. I. *Theory Probab. Math. Statist.* **73** 120–133. MR2213848
[33] Olver, F.W.J., Lozier, D.W., Boisvert, R.F. and Clark, C.W. (2010). *NIST Handbook of Mathematical Functions*. Cambridge: Cambridge Univ. Press. MR2723248

[34] Peccati, G. and Taqqu, M.S. (2011). *Wiener Chaos*: *Moments*, *Cumulants and Diagrams*. *Bocconi & Springer Series* **1**. Milan: Springer; Milan: Bocconi Univ. Press. MR2791919

[35] Petrov, V.V. (1995). *Limit Theorems of Probability Theory*. *Oxford Studies in Probability* **4**. New York: The Clarendon Press, Oxford University Press. MR1353441

[36] Pham, V.-H. (2013). On the rate of convergence for central limit theorems of sojourn times of Gaussian fields. *Stochastic Process*. *Appl*. **123** 2158–2174. MR3038501

[37] Shigekawa, I. (1980). Derivatives of Wiener functionals and absolute continuity of induced measures. *J*. *Math*. *Kyoto Univ*. **20** 263–289. MR0582167

[38] Taqqu, M.S. (1974/75). Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z*. *Wahrsch*. *Verw*. *Gebiete* **31** 287–302. MR0400329

[39] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z*. *Wahrsch*. *Verw*. *Gebiete* **50** 53–83. MR0550123

[40] Veillette, M.S. and Taqqu, M.S. (2013). Properties and numerical evaluation of the Rosenblatt distribution. *Bernoulli* **19** 982–1005. MR3079303

*Bernoulli* **25**(4A), 2019, 2949–2981
https://doi.org/10.3150/18-BEJ1076

# On logarithmically optimal exact simulation of max-stable and related random fields on a compact set

ZHIPENG LIU[1,*], JOSE H. BLANCHET[2], A.B. DIEKER[1,**] and THOMAS MIKOSCH[3]

[1]*Industrial Engineering and Operations Research Department, School of Engineering and Applied Sciences, Columbia University, 500 West 120th Street, New York, NY 10027, U.S.A.*
*E-mail:* [*]*zl2337@columbia.edu;* [**]*ton.dieker@ieor.columbia.edu*
[2]*Department of Management Science and Engineering, School of Engineering, Stanford University, Huang Engineering Center, 475 Via Ortega Avenue, Stanford, CA 94305-6015, U.S.A.*
*E-mail: jose.blanchet@stanford.edu*
[3]*Department of Mathematics, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. E-mail: mikosch@math.ku.dk*

We consider the random field

$$M(t) = \sup_{n\geq 1}\{-\log A_n + X_n(t)\}, \qquad t \in T,$$

for a set $T \subset \mathbb{R}^m$, where $(X_n)$ is an i.i.d. sequence of centered Gaussian random fields on $T$ and $0 < A_1 < A_2 < \cdots$ are the arrivals of a general renewal process on $(0, \infty)$, independent of $(X_n)$. In particular, a large class of max-stable random fields with Gumbel marginals have such a representation. Assume that one needs $c(d) = c(\{t_1, \dots, t_d\})$ function evaluations to sample $X_n$ at $d$ locations $t_1, \dots, t_d \in T$. We provide an algorithm which samples $M(t_1), \dots, M(t_d)$ with complexity $O(c(d)^{1+o(1)})$ as measured in the $L_p$ norm sense for any $p \geq 1$. Moreover, if $X_n$ has an a.s. converging series representation, then $M$ can be a.s. approximated with error $\delta$ uniformly over $T$ and with complexity $O(1/(\delta \log(1/\delta))^{1/\alpha})$, where $\alpha$ relates to the Hölder continuity exponent of the process $X_n$ (so, if $X_n$ is Brownian motion, $\alpha = 1/2$).

*Keywords:* Brown–Resnick process; exact simulation; Gaussian field; max-stable random fields; record-breaking

## References

[1] Adler, R.J. and Taylor, J.E. (2007). *Random Fields and Geometry*. *Springer Monographs in Mathematics*. New York: Springer. MR2319516
[2] Asmussen, S. (2003). *Applied Probability and Queues*: *Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. New York: Springer. MR1978607
[3] Asmussen, S., Glynn, P. and Pitman, J. (1995). Discretization error in simulation of one-dimensional reflecting Brownian motion. *Ann*. *Appl*. *Probab*. **5** 875–896. MR1384357
[4] Asmussen, S. and Glynn, P.W. (2007). *Stochastic Simulation*: *Algorithms and Analysis*. *Stochastic Modelling and Applied Probability* **57**. New York: Springer. MR2331321

1350-7265 

[5] Ayache, A. and Taqqu, M.S. (2003). Rate optimality of wavelet series approximations of fractional Brownian motion. *J. Fourier Anal. Appl.* **9** 451–471. MR2027888

[6] Blanchet, J. and Chen, X. (2015). Steady-state simulation of reflected Brownian motion and related stochastic networks. *Ann. Appl. Probab.* **25** 3209–3250. MR3404635

[7] Blanchet, J., Chen, X. and Dong, J. (2017). $\varepsilon$-strong simulation for multidimensional stochastic differential equations via rough path analysis. *Ann. Appl. Probab.* **27** 275–336. MR3619789

[8] Blanchet, J. and Wallwater, A. (2015). Exact sampling of stationary and time-reversed queues. *ACM Trans. Model. Comput. Simul.* **25** Art. 26. MR3432614

[9] Blanchet, J.H. and Sigman, K. (2011). On exact sampling of stochastic perpetuities. *J. Appl. Probab.* **48A** 165–182. MR2865624

[10] Brown, B.M. and Resnick, S.I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.* **14** 732–739. MR0517438

[11] de Haan, L. (1984). A spectral representation for max-stable processes. *Ann. Probab.* **12** 1194–1204. MR0757776

[12] de Haan, L. and Zhou, C. (2008). On extreme value analysis of a spatial process. *REVSTAT* **6** 71–81. MR2386300

[13] Dieker, A.B. and Mikosch, T. (2015). Exact simulation of Brown–Resnick random fields at a finite number of locations. *Extremes* **18** 301–314. MR3351818

[14] Dombry, C., Engelke, S. and Oesting, M. (2016). Exact simulation of max-stable processes. *Biometrika* **103** 303–317. Available at arXiv:1506.04430. MR3509888

[15] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613

[16] Gut, A. (2009). *Stopped Random Walks: Limit Theorems and Applications*, 2nd ed. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2489436

[17] Kabluchko, Z., Schlather, M. and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37** 2042–2065. MR2561440

[18] Kenealy, B. (2013, August 11). New York's MTA buys $200 million cat bond to avoid storm surge losses. *Bus. Insur.* Available at https://www.businessinsurance.com.

[19] Oesting, M., Schlather, M. and Zhou, C. (2018). Exact and fast simulation of max-stable processes on a compact set using the normalized spectral representation. *Bernoulli* **24** 1497–1530. MR3706800

[20] Pollock, M., Johansen, A.M. and Roberts, G.O. (2016). On the exact and $\varepsilon$-strong simulation of (jump) diffusions. *Bernoulli* **22** 794–856. MR3449801

[21] Schilling, R.L. and Partzsch, L. (2012). *Brownian Motion: An Introduction to Stochastic Processes*. Berlin: de Gruyter. With a chapter on simulation by Björn Böttcher. MR2962168

[22] Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes* **5** 33–44. MR1947786

[23] Smith, R.L. (1990). Max-stable processes and spatial extremes. Unpublished manuscript.

[24] Steele, J.M. (2001). *Stochastic Calculus and Financial Applications. Applications of Mathematics (New York)* **45**. New York: Springer. MR1783083

[25] Thibaud, E., Aalto, J., Cooley, D.S., Davison, A.C. and Heikkinen, J. (2016). Bayesian inference for the Brown–Resnick process, with an application to extreme low temperatures. *Ann. Appl. Stat.* **10** 2303–2324. Available at arXiv:1506.07836. MR3592058

*Bernoulli* **25**(4A), 2019, 2982–3015
https://doi.org/10.3150/18-BEJ1077

# Uniform rates of the Glivenko–Cantelli convergence and their use in approximating Bayesian inferences

EMANUELE DOLERA* and EUGENIO REGAZZINI**

*Dipartimento di Matematica, via Ferrata 5, 27100 Pavia (PV), Italy. E-mail:* *emanuele.dolera@unipv.it; **eugenio.regazzini@unipv.it

This paper deals with suitable quantifications in approximating a probability measure by an "empirical" random probability measure $\hat{\mathfrak{p}}_n$, depending on the first $n$ terms of a sequence $\{\tilde{\xi}_i\}_{i\geq 1}$ of random elements. Section 2 studies the range of oscillation near zero of the Wasserstein distance $\mathrm{d}^{(p)}_{[\mathbb{S}]}$ between $\mathfrak{p}_0$ and $\hat{\mathfrak{p}}_n$, assuming the $\tilde{\xi}_i$'s i.i.d. from $\mathfrak{p}_0$. In Theorem 2.1 $\mathfrak{p}_0$ can be fixed in the space of all probability measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ and $\hat{\mathfrak{p}}_n$ coincides with the empirical measure $\tilde{\mathfrak{e}}_n := \frac{1}{n}\sum_{i=1}^n \delta_{\tilde{\xi}_i}$. In Theorem 2.2 (Theorem 2.3, respectively), $\mathfrak{p}_0$ is a $d$-dimensional Gaussian distribution (an element of a distinguished statistical exponential family, respectively) and $\hat{\mathfrak{p}}_n$ is another $d$-dimensional Gaussian distribution with estimated mean and covariance matrix (another element of the same family with an estimated parameter, respectively). These new results improve on allied recent works by providing also uniform bounds with respect to $n$, meaning the finiteness of the $p$-moment of $\sup_{n\geq 1} b_n \mathrm{d}^{(p)}_{[\mathbb{S}]}(\mathfrak{p}_0, \hat{\mathfrak{p}}_n)$ is proved for some diverging sequence $b_n$ of positive numbers. In Section 3, assuming the $\tilde{\xi}_i$'s exchangeable, one studies the range of oscillation near zero of the Wasserstein distance between the conditional distribution – also called posterior – of the directing measure of the sequence, given $\tilde{\xi}_1, \ldots, \tilde{\xi}_n$, and the point mass at $\hat{\mathfrak{p}}_n$. Similarly, a bound for the approximation of predictive distributions is given. Finally, Theorems from 3.3 to 3.5 reconsider Theorems from 2.1 to 2.3, respectively, according to a Bayesian perspective.

*Keywords:* dominated ergodic theorem; empirical measure; exchangeability; Glivenko–Cantelli theorem; law of the iterated logarithm; posterior distribution; predictive distribution; Wasserstein distance

## References

[1] Ajtai, M., Komlós, J. and Tusnády, G. (1984). On optimal matchings. *Combinatorica* **4** 259–264. MR0779885

[2] Aldous, D.J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII* – 1983. *Lecture Notes in Math.* **1117** 1–198. Berlin: Springer. MR0883646

[3] 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

[4] Ambrosio, L., Stra, F. and Trevisan, D. (2018). A PDE approach to a 2-dimensional matching problem. To appear on *Probab. Theory Related Fields*. DOI:10.1007/s00440-018-0837-x.

[5] Anderson, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*, 2nd ed. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0771294

[6] Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. *Wiley Series in Probability and Mathematical Statistics*. Chichester: Wiley. MR0489333

[7] Berti, P., Crimaldi, I., Pratelli, L. and Rigo, P. (2009). Rate of convergence of predictive distributions for dependent data. *Bernoulli* **15** 1351–1367. MR2597596

[8] Berti, P., Pratelli, L. and Rigo, P. (2012). Limit theorems for empirical processes based on dependent data. *Electron*. *J*. *Probab*. **17** no. 9, 18. MR2878788

[9] Berti, P., Pratelli, L. and Rigo, P. (2018). Asymptotic predictive inference with exchangeable data. *Braz*. *J*. *Probab*. *Stat*. **32** 815–833. MR3845031

[10] Blackwell, D. and Dubins, L. (1962). Merging of opinions with increasing information. *Ann*. *Math*. *Stat*. **33** 882–886. MR0149577

[11] Bobkov, S. and Ledoux, M. (2016). One-dimensional empirical measures, order statistics, and Kantorovich transport distances. To appear on *Mem*. *Amer*. *Math*. *Soc*.

[12] Boissard, E. (2011). Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance. *Electron*. *J*. *Probab*. **16** 2296–2333. MR2861675

[13] Boissard, E. and Le Gouic, T. (2014). On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **50** 539–563. MR3189084

[14] Bolley, F., Guillin, A. and Villani, C. (2007). Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab*. *Theory Related Fields* **137** 541–593. MR2280433

[15] Borovkov, A.A. and Mogul'skiĭ, A.A. (2011). Chebyshev-type exponential inequalities for sums of random vectors and for trajectories of random walks. *Teor*. *Veroyatn*. *Primen*. **56** 3–29. MR2848414

[16] Brown, L.D. (1986). *Fundamentals of Statistical Exponential Families with Applications in Statistical Decision Theory*. *Institute of Mathematical Statistics Lecture Notes – Monograph Series* **9**. Hayward, CA: IMS. MR0882001

[17] Cantelli, F.P. (1933). Sulla determinazione empirica delle leggi di probabilità. *G*. *Ist*. *Ital*. *Attuari* **4** 421–424.

[18] Caracciolo, S., Lucibello, G. and Parisi, G. (2014). Scaling hypothesis for the Euclidean bipartite matching problem. *Phys*. *Rev*. *E* **90** 012118.

[19] Chow, Y.S. and Teicher, H. (1997). *Probability Theory*: *Independence*, *Interchangeability*, *Martingales*, 3rd ed. *Springer Texts in Statistics*. New York: Springer. MR1476912

[20] Cifarelli, D.M., Dolera, E. and Regazzini, E. (2016). Frequentistic approximations to Bayesian prevision of exchangeable random elements. *Internat*. *J*. *Approx*. *Reason*. **78** 138–152. MR3543878

[21] Cifarelli, D.M., Dolera, E. and Regazzini, E. (2017). Note on "Frequentistic approximations to Bayesian prevision of exchangeable random elements" [Int. J. Approx. Reason. 78 (2016) 138–152] [MR3543878]. *Internat*. *J*. *Approx*. *Reason*. **86** 26–27. MR3653971

[22] Cox, D.R. (1975). Partial likelihood. *Biometrika* **62** 269–276. MR0400509

[23] de Finetti, B. (1930). Funzione caratteristica di un fenomeno aleatorio. *Atti R*. *Accad*. *Naz*. *Lincei*, *Mem*. **4** 86–133.

[24] de Finetti, B. (1933). Sull'approssimazione empirica di una legge di probabilità. *G*. *Ist*. *Ital*. *Attuari* **4** 415–420.

[25] de Finetti, B. (1933). La legge dei grandi numeri nel caso dei numeri aleatori equivalenti. *Atti R*. *Accad*. *Naz*. *Lincei*, *Rend*. **18** 203–207.

[26] de Finetti, B. (1937). La prévision: Ses lois logiques, ses sources subjectives. *Ann*. *Inst*. *Henri Poincaré* **7** 1–68. MR1508036

[27] Dereich, S., Scheutzow, M. and Schottstedt, R. (2013). Constructive quantization: Approximation by empirical measures. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **49** 1183–1203. MR3127919

[28] Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. *Applications of Mathematics* (*New York*) **31**. New York: Springer. MR1383093

[29] Diaconis, P. and Freedman, D. (1986). On the consistency of Bayes estimates. *Ann*. *Statist*. **14** 1–67. MR0829555

[30] Dowson, D.C. and Landau, B.V. (1982). The Fréchet distance between multivariate normal distributions. *J*. *Multivariate Anal*. **12** 450–455. MR0666017

[31] Dudley, R.M. (1968). The speed of mean Glivenko–Cantelli convergence. *Ann*. *Math*. *Stat*. **40** 40–50. MR0236977

[32] Dudley, R.M. (1999). *Uniform Central Limit Theorems*. *Cambridge Studies in Advanced Mathematics* **63**. Cambridge: Cambridge Univ. Press. MR1720712

[33] Dudley, R.M. (2002). *Real Analysis and Probability*. *Cambridge Studies in Advanced Mathematics* **74**. Cambridge: Cambridge Univ. Press. MR1932358

[34] Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann*. *Statist*. **7** 1–26. MR0515681

[35] Efron, B. (2010). *Large-Scale Inference*: *Empirical Bayes Methods for Estimation*, *Testing*, *and Prediction*. *Institute of Mathematical Statistics* (*IMS*) *Monographs* **1**. Cambridge: Cambridge Univ. Press. MR2724758

[36] Einmahl, U. (2016). Law of the iterated logarithm type results for random vectors with infinite second moments. *Math*. *Appl*. (*Warsaw*) **44** 167–181. MR3557096

[37] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab*. *Theory Related Fields* **162** 707–738. MR3383341

[38] Glivenko, V.I. (1933). Sulla determinazione empirica delle leggi di probabilità. *G*. *Ist*. *Ital*. *Attuari* **4** 92–99.

[39] Gozlan, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *Ann*. *Probab*. **37** 2480–2498. MR2573565

[40] Gozlan, N. and Léonard, C. (2007). A large deviation approach to some transportation cost inequalities. *Probab*. *Theory Related Fields* **139** 235–283. MR2322697

[41] Gozlan, N. and Léonard, C. (2010). Transport inequalities. A survey. *Markov Process*. *Related Fields* **16** 635–736. MR2895086

[42] Horn, R.A. and Johnson, C.R. (2013). *Matrix Analysis*, 2nd ed. Cambridge: Cambridge Univ. Press. MR2978290

[43] Horowitz, J. and Karandikar, R.L. (1994). Mean rates of convergence of empirical measures in the Wasserstein metric. *J*. *Comput*. *Appl*. *Math*. **55** 261–273. MR1329874

[44] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169

[45] Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. *G*. *Ist*. *Ital*. *Attuari* **4** 83–91.

[46] Maritz, J.S. and Lwin, T. (1989). *Empirical Bayes Methods*, 2nd ed. *Monographs on Statistics and Applied Probability* **35**. London: CRC Press. MR1019835

[47] Massart, P. (1988). About the Prohorov [Prokhorov] distance between the uniform distribution over the unit cube in $\mathbf{R}^d$ and its empirical measure. *Probab*. *Theory Related Fields* **79** 431–450. MR0959517

[48] Morris, C.N. (1983). Parametric empirical Bayes inference: Theory and applications. *J*. *Amer*. *Statist*. *Assoc*. **78** 47–65. MR0696849

[49] Olkin, I. and Pukelsheim, F. (1982). The distance between two random vectors with given dispersion matrices. *Linear Algebra Appl*. **48** 257–263. MR0683223

[50] Petrone, S., Rizzelli, S., Rousseau, J. and Scricciolo, C. (2014). Empirical Bayes methods in classical and Bayesian inference. *Metron* **72** 201–215. MR3233149

[51] Petrone, S., Rousseau, J. and Scricciolo, C. (2014). Bayes and empirical Bayes: Do they merge? *Biometrika* **101** 285–302. MR3215348

[52] Robbins, H. (1956). An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol*. *I* 157–163. Berkeley and Los Angeles: Univ. California Press. MR0084919

[53] Rockafellar, R.T. (1970). *Convex Analysis*. *Princeton Mathematical Series* **28**. Princeton, NJ: Princeton Univ. Press. MR0274683

[54] Rousseau, J. (2016). On the frequentist properties of Bayesian nonparametric methods. *Annu*. *Rev*. *Sci*. *Stat*. **3** 211–231.

[55] Severini, T.A. (2000). *Likelihood Methods in Statistics*. *Oxford Statistical Science Series* **22**. Oxford: Oxford Univ. Press. MR1854870

[56] Shor, P.W. and Yukich, J.E. (1991). Minimax grid matching and empirical measures. *Ann*. *Probab*. **19** 1338–1348. MR1112419

[57] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. New York: Wiley. MR0838963

[58] Siegmund, D. (1969). On moments of the maximum of normed partial sums. *Ann*. *Math*. *Stat*. **40** 527–531. MR0239695

[59] Talagrand, M. (1994). The transportation cost from the uniform measure to the empirical measure in dimension $\geq 3$. *Ann*. *Probab*. **22** 919–959. MR1288137

[60] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geom*. *Funct*. *Anal*. **6** 587–600. MR1392331

[61] Teicher, H. (1971). Completion of a dominated ergodic theorem. *Ann*. *Math*. *Stat*. **42** 2156–2158. MR0300327

[62] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes*: *With Applications to Statistics*. *Springer Series in Statistics*. New York: Springer. MR1385671

[63] Vapnik, V.N. and Červonenkis, A.J. (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab*. *Appl*. **16** 264–280.

[64] Villani, C. (2009). *Optimal Transport*: *Old and New*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **338**. Berlin: Springer. MR2459454

[65] Weed, J. and Bach, F. (2017). Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. Preprint. Available at arXiv:1707.00087.

[66] Yosida, K. and Kakutani, S. (1939). Birkhoff's ergodic theorem and the maximal ergodic theorem. *Proc*. *Imp*. *Acad*. (*Tokyo*) **15** 165–168. MR0000355

[67] Yukich, J.E. (1989). Optimal matching and empirical measures. *Proc*. *Amer*. *Math*. *Soc*. **107** 1051–1059. MR1000171

[68] Yurinskiĭ, V.V. (1976). Exponential inequalities for sums of random vectors. *J*. *Multivariate Anal*. **6** 473–499. MR0428401

*Bernoulli* **25**(4A), 2019, 3016–3040
https://doi.org/10.3150/18-BEJ1078

# Localized Gaussian width of $M$-convex hulls with applications to Lasso and convex aggregation

PIERRE C. BELLEC[1]

[1]*Department of Statistics and Biostatistics, Rutgers University, 501 Hill Center, 110 Frelinghuysen Road, Piscataway, NJ 08854, USA. E-mail: pcb71@stat.rutgers.edu*

Upper and lower bounds are derived for the Gaussian mean width of a convex hull of $M$ points intersected with a Euclidean ball of a given radius. The upper bound holds for any collection of extreme points bounded in Euclidean norm. The upper bound and the lower bound match up to a multiplicative constant whenever the extreme points satisfy a one sided Restricted Isometry Property.

An appealing aspect of the upper bound is that no assumption on the covariance structure of the extreme points is needed. This aspect is especially useful to study regression problems with anisotropic design distributions. We provide applications of this bound to the Lasso estimator in fixed-design regression, the Empirical Risk Minimizer in the anisotropic persistence problem, and the convex aggregation problem in density estimation.

*Keywords:* anisotropic design; convex aggregation; convex hull; Gaussian mean width; Lasso; localized Gaussian width

## References

[1] Audibert, J.-Y. and Tsybakov, A.B. (2007). Fast learning rates for plug-in classifiers. *Ann. Statist.* **35** 608–633. MR2336861

[2] Bartlett, P.L., Bousquet, O. and Mendelson, S. (2005). Local Rademacher complexities. *Ann. Statist.* **33** 1497–1537. MR2166554

[3] Bartlett, P.L. and Mendelson, S. (2006). Empirical minimization. *Probab. Theory Related Fields* **135** 311–334. MR2240689

[4] Bartlett, P.L., Mendelson, S. and Neeman, J. (2012). $\ell_1$-Regularized linear regression: Persistence and oracle inequalities. *Probab. Theory Related Fields* **154** 193–224. MR2981422

[5] Bellec, P. and Tsybakov, A. (2017). Bounds on the prediction error of penalized least squares estimators with convex penalty. In *Modern Problems of Stochastic Analysis and Statistics*. *Springer Proc. Math. Stat.* **208** (V. Panov, ed.) 315–333. Cham: Springer. MR3747672

[6] Bellec, P.C. (2017). Optimal exponential bounds for aggregation of density estimators. *Bernoulli* **23** 219–248. MR3556772

[7] Bellec, P.C. (2017). Optimistic lower bounds for convex regularized least-squares. ArXiv preprint. Available at arXiv:1703.01332.

[8] Bellec, P.C. (2018). Optimal bounds for aggregation of affine estimators. *Ann. Statist.* **46** 30–59. MR3766945

[9] Bellec, P.C. (2018). Sharp oracle inequalities for least squares estimators in shape restricted regression. *Ann. Statist.* **46** 745–780. MR3782383

[10] Bellec, P.C. (2018). The noise barrier and the large signal bias of the lasso and other convex estimators. ArXiv preprint. Available at arXiv:1804.01230.
[11] Bellec, P.C., Dalalyan, A.S., Grappin, E. and Paris, Q. (2018). On the prediction loss of the lasso in the partially labeled setting. *Electron*. *J*. *Stat*. **12** 3443–3472. MR3864589
[12] Bellec, P.C., Lecué, G. and Tsybakov, A.B. (2018). Slope meets Lasso: Improved oracle bounds and optimality. *Ann*. *Statist*. **46** 3603–3642. MR3852663
[13] Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of lasso and Dantzig selector. *Ann*. *Statist*. **37** 1705–1732. MR2533469
[14] Boucheron, S., Lugosi, G. and Massart, P. (2013). *Concentration Inequalities*: *A Nonasymptotic Theory of Independence*. Oxford: Oxford Univ. Press. With a foreword by Michel Ledoux. MR3185193
[15] Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data*: *Methods*, *Theory and Applications*. *Springer Series in Statistics*. Heidelberg: Springer. MR2807761
[16] Candès, E.J. and Plan, Y. (2009). Near-ideal model selection by $\ell_1$ minimization. *Ann*. *Statist*. **37** 2145–2177. MR2543688
[17] Chatterjee, S. (2014). A new perspective on least squares under convex constraint. *Ann*. *Statist*. **42** 2340–2381. MR3269982
[18] Dai, D., Rigollet, P. and Zhang, T. (2012). Deviation optimal learning using greedy $Q$-aggregation. *Ann*. *Statist*. **40** 1878–1905. MR3015047
[19] Dalalyan, A.S., Hebiri, M. and Lederer, J. (2017). On the prediction performance of the Lasso. *Bernoulli* **23** 552–581. MR3556784
[20] Dalalyan, A.S. and Tsybakov, A.B. (2007). Aggregation by exponential weighting and sharp oracle inequalities. In *Learning Theory*. *Lecture Notes in Computer Science* **4539** 97–111. Berlin: Springer. MR2397581
[21] den Hollander, F. (2012). Probability theory: The coupling method.
[22] Giraud, C. (2014). *Introduction to High-Dimensional Statistics*. *Chapman and Hall/CRC Monographs on Statistics and Applied Probability* **138**. Boca Raton, FL: CRC Press.
[23] Gordon, Y., Litvak, A.E., Mendelson, S. and Pajor, A. (2007). Gaussian averages of interpolated bodies and applications to approximate reconstruction. *J*. *Approx*. *Theory* **149** 59–73. MR2371614
[24] Greenshtein, E. and Ritov, Y. (2004). Persistence in high-dimensional linear predictor selection and the virtue of overparametrization. *Bernoulli* **10** 971–988. MR2108039
[25] Hastie, T., Tibshirani, R. and Wainwright, M. (2015). *Statistical Learning with Sparsity*: *The Lasso and Generalizations*. *Monographs on Statistics and Applied Probability* **143**. Boca Raton, FL: CRC Press. MR3616141
[26] Juditsky, A., Rigollet, P. and Tsybakov, A.B. (2008). Learning by mirror averaging. *Ann*. *Statist*. **36** 2183–2206. MR2458184
[27] Knight, K. and Fu, W. (2000). Asymptotics for Lasso-type estimators. *Ann*. *Statist*. **28** 1356–1378. MR1805787
[28] Koltchinskii, V. (2006). Local Rademacher complexities and oracle inequalities in risk minimization. *Ann*. *Statist*. **34** 2593–2656. MR2329442
[29] Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*. *Lecture Notes in Math*. **2033**. Heidelberg: Springer. Lectures from the 38th Probability Summer School held in Saint-Flour, 2008, École d'Été de Probabilités de Saint-Flour. [Saint-Flour Probability Summer School]. MR2829871
[30] Koltchinskii, V., Lounici, K. and Tsybakov, A.B. (2011). Nuclear-norm penalization and optimal rates for noisy low-rank matrix completion. *Ann*. *Statist*. **39** 2302–2329. MR2906869
[31] Lecué, G. (2006). Lower bounds and aggregation in density estimation. *J*. *Mach*. *Learn*. *Res*. **7** 971–981. MR2274393

[32] Lecué, G. (2013). Empirical risk minimization is optimal for the convex aggregation problem. *Bernoulli* **19** 2153–2166. MR3160549

[33] Lecué, G. and Mendelson, S. (2009). Aggregation via empirical risk minimization. *Probab. Theory Related Fields* **145** 591–613. MR2529440

[34] Lecué, G. and Mendelson, S. (2013). Learning subGaussian classes: Upper and minimax bounds. ArXiv preprint. Available at arXiv:1305.4825.

[35] Lecué, G. and Rigollet, P. (2014). Optimal learning with $Q$-aggregation. *Ann. Statist.* **42** 211–224. MR3178462

[36] Leung, G. and Barron, A.R. (2006). Information theory and mixing least-squares regressions. *IEEE Trans. Inform. Theory* **52** 3396–3410. MR2242356

[37] Lounici, K. (2007). Generalized mirror averaging and $D$-convex aggregation. *Math. Methods Statist.* **16** 246–259. MR2356820

[38] Lounici, K., Pontil, M., van de Geer, S. and Tsybakov, A.B. (2011). Oracle inequalities and optimal inference under group sparsity. *Ann. Statist.* **39** 2164–2204. MR2893865

[39] Mendelson, S. (2015). Learning without concentration. *J. ACM* **62** Article ID 21. MR3367000

[40] Nemirovski, A. (2000). Topics in non-parametric statistics. In *Lectures on Probability Theory and Statistics* (*Saint-Flour*, 1998). *Lecture Notes in Math.* **1738** 85–277. Berlin: Springer. MR1775640

[41] Pisier, G. (1981). Remarques sur un résultat non publié de B. Maurey. In *Seminar on Functional Analysis*, 1980–1981, *Exp. No.* 5 1–12. Palaiseau: École Polytech. MR0659306

[42] Plan, Y. and Vershynin, R. (2016). The generalized Lasso with non-linear observations. *IEEE Trans. Inform. Theory* **62** 1528–1537. MR3472264

[43] Plan, Y., Vershynin, R. and Yudovina, E. (2017). High-dimensional estimation with geometric constraints. *Inf. Inference* **6** 1–40. MR3636866

[44] Rigollet, P. and Tsybakov, A. (2011). Exponential screening and optimal rates of sparse estimation. *Ann. Statist.* **39** 731–771. MR2816337

[45] Rigollet, P. and Tsybakov, A.B. (2012). Sparse estimation by exponential weighting. *Statist. Sci.* **27** 558–575. MR3025134

[46] Rigollet, P. and Tsybakov, A.B. (2007). Linear and convex aggregation of density estimators. *Math. Methods Statist.* **16** 260–280. MR2356821

[47] Rudelson, M. and Zhou, S. (2013). Reconstruction from anisotropic random measurements. *IEEE Trans. Inform. Theory* **59** 3434–3447. MR3061256

[48] Sun, T. and Zhang, C.-H. (2012). Scaled sparse linear regression. *Biometrika* **99** 879–898. MR2999166

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

[50] Tsybakov, A.B. (2003). Optimal rates of aggregation. In *Learning Theory and Kernel Machines* 303–313. Springer.

[51] Tsybakov, A.B. (2009). *Introduction to Nonparametric Estimation*. *Springer Series in Statistics*. New York: Springer. Revised and extended from the 2004 French original. Translated by Vladimir Zaiats. MR2724359

[52] Tsybakov, A.B. (2014). Aggregation and minimax optimality in high-dimensional estimation. In *Proceedings of the International Congress of Mathematicians—Seoul* 2014. *Vol. IV* 225–246. Seoul: Kyung Moon Sa. MR3727610

[53] Vershynin, R. (2012). Introduction to the non-asymptotic analysis of random matrices. In *Compressed Sensing* 210–268. Cambridge: Cambridge Univ. Press. MR2963170

[54] Yang, Y. (2000). Mixing strategies for density estimation. *Ann. Statist.* **28** 75–87. MR1762904

[55] Zhang, C.-H. and Huang, J. (2008). The sparsity and bias of the LASSO selection in high-dimensional linear regression. *Ann. Statist.* **36** 1567–1594. MR2435448

*Bernoulli* **25**(4A), 2019, 3041–3068
https://doi.org/10.3150/18-BEJ1079

# Signal detection via Phi-divergences for general mixtures

MARC DITZHAUS

*Mathematical Institute, Heinrich-Heine University Düsseldorf, Universitätststraße 1, 40225 Düsseldorf, Germany. E-mail: marc.ditzhaus@hhu.de*

The family of goodness-of-fit tests based on Φ-divergences is known to be optimal for detecting signals hidden in high-dimensional noise data when the heterogeneous normal mixture model is underlying. This test family includes Tukey's popular higher criticism test and the famous Berk–Jones test. In this paper we address the open question whether the tests' optimality is still present beyond the prime normal mixture model. On the one hand, we transfer the known optimality of the higher criticism test for different models, for example, for the heteroscedastic normal, general Gaussian and exponential-$\chi^2$-mixture models, to the whole test family. On the other hand, we discuss the optimality for new model classes based on exponential families including the scale exponential, the scale Fréchet and the location Gumbel models. For all these examples we apply a general machinery which might be used to show the tests' optimality for further models/model classes in future.

*Keywords:* Berk and Jones test; detection boundary; Φ-divergences; sparse and dense signal detection; Tukey's higher criticism

## References

[1] Ali, S.M. and Silvey, S.D. (1966). A general class of coefficients of divergence of one distribution from another. *J. Roy. Statist. Soc. Ser. B* **28** 131–142. MR0196777

[2] Arias-Castro, E., Candès, E.J. and Plan, Y. (2011). Global testing under sparse alternatives: ANOVA, multiple comparisons and the higher criticism. *Ann. Statist.* **39** 2533–2556. MR2906877

[3] Arias-Castro, E. and Wang, M. (2015). The sparse Poisson means model. *Electron. J. Stat.* **9** 2170–2201. MR3406276

[4] Arias-Castro, E. and Wang, M. (2017). Distribution-free tests for sparse heterogeneous mixtures. *TEST* **26** 71–94. MR3613606

[5] Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Chichester: Wiley. MR0489333

[6] Berk, R.H. and Jones, D.H. (1979). Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Z. Wahrsch. Verw. Gebiete* **47** 47–59. MR0521531

[7] 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

[8] Cai, T.T., Jeng, X.J. and Jin, J. (2011). Optimal detection of heterogeneous and heteroscedastic mixtures. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 629–662. MR2867452

[9] Cai, T.T. and Wu, Y. (2014). Optimal detection of sparse mixtures against a given null distribution. *IEEE Trans. Inform. Theory* **60** 2217–2232. MR3181520

[10] Cayon, L., Jin, J. and Treaster, A. (2004). Higher criticism statistic: Detecting and identifying non-Gaussianity in the WMAP first year data. *Mon. Not. R. Astron. Soc.* **362** 826–832.

[11] Chang, L.-C. (1955). On the ratio of an empirical distribution function to the theoretical distribution function. *Acta Math. Sinica* **5** 347–368. MR0076213

[12] Cressie, N. and Read, T.R.C. (1984). Multinomial goodness-of-fit tests. *J. Roy. Statist. Soc. Ser. B* **46** 440–464. MR0790631

[13] Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Sci. Math. Hungar.* **2** 299–318. MR0219345

[14] Dai, H., Charnigo, R., Srivastava, T., Talebizadeh, Z. and Qing, S. (2012). Integrating P-values for genetic and genomic data analysis. *J. Biom. Biostat.* 3–7.

[15] Ditzhaus, M. (2017). The power of tests for signal detection under high-dimensional data. Ph.D. thesis, Heinrich-Heine-Univ. Duesseldorf. Available at https://docserv.uni-duesseldorf.de/servlets/DocumentServlet?id=42808.

[16] Ditzhaus, M. and Janssen, A. (2017). Detectability of nonparametric signals: Higher criticism versus likelihood ratio. Available at 1709.07264v2.

[17] Donoho, D. and Jin, J. (2004). Higher criticism for detecting sparse heterogeneous mixtures. *Ann. Statist.* **32** 962–994. MR2065195

[18] Donoho, D. and Jin, J. (2009). Feature selection by higher criticism thresholding achieves the optimal phase diagram. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **367** 4449–4470. MR2546396

[19] Donoho, D. and Jin, J. (2015). Higher criticism for large-scale inference, especially for rare and weak effects. *Statist. Sci.* **30** 1–25. MR3317751

[20] Eicker, F. (1979). The asymptotic distribution of the suprema of the standardized empirical processes. *Ann. Statist.* **7** 116–138. MR0515688

[21] Feller, W. (1966). *An Introduction to Probability Theory and Its Applications. Vol. II*. New York: Wiley. MR0210154

[22] Goldstein, D.B. (2009). Common genetic variation and human traits. *N. Engl. J. Med.* **360** 1696–1698.

[23] Gontscharuk, V., Landwehr, S. and Finner, H. (2015). The intermediates take it all: Asymptotics of higher criticism statistics and a powerful alternative based on equal local levels. *Biom. J.* **57** 159–180. MR3298224

[24] Gontscharuk, V., Landwehr, S. and Finner, H. (2016). Goodness of fit tests in terms of local levels with special emphasis on higher criticism tests. *Bernoulli* **22** 1331–1363. MR3474818

[25] Hall, P., Pittelkow, Y. and Ghosh, M. (2008). Theoretical measures of relative performance of classifiers for high dimensional data with small sample sizes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 159–173. MR2412636

[26] Ingster, Y.I., Tsybakov, A.B. and Verzelen, N. (2010). Detection boundary in sparse regression. *Electron. J. Stat.* **4** 1476–1526. MR2747131

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

[28] Iyengar, S.K. and Elston, R.C. (2007). The genetic basis of complex traits: Rare variants or "common gene, common disease"? *Methods Mol. Biol.* **376** 71–84.

[29] Jaeschke, D. (1979). The asymptotic distribution of the supremum of the standardized empirical distribution function on subintervals. *Ann. Statist.* **7** 108–115. MR0515687

[30] Jager, L. and Wellner, J. (2004). A new goodness of fit test: The reversed Berk–Jones statistic. Technical report, Dept. Statistics, Univ. Washington, DC.

[31] Jager, L. and Wellner, J.A. (2007). Goodness-of-fit tests via phi-divergences. *Ann. Statist.* **35** 2018–2053. MR2363962

[32] Jin, J. (2009). Impossibility of successful classification when useful features are rare and weak. *Proc. Natl. Acad. Sci. USA* **106** 8859–8864. MR2520682

[33] Jin, J., Starck, J.-L., Donoho, D.L., Aghanim, N. and Forni, O. (2005). Cosmological non-Gaussian signature detection: Comparing performance of different statistical tests. *EURASIP J. Appl. Signal Process.* **15** 2470–2485. MR2210857

[34] Khmaladze, E. and Shinjikashvili, E. (2001). Calculation of noncrossing probabilities for Poisson processes and its corollaries. *Adv. in Appl. Probab.* **33** 702–716. MR1860097

[35] Khmaladze, E.V. (1998). Goodness of fit tests for "chimeric" alternatives. *Stat. Neerl.* **52** 90–111. MR1615550

[36] Kulldorff, M., Heffernan, R., Hartman, J., Assunção, R. and Mostashari, F. (2005). A space–time permutation scan statistic for disease outbreak detection. *PLoS Med.* **2** e59.

[37] Mukherjee, R., Pillai, N.S. and Lin, X. (2015). Hypothesis testing for high-dimensional sparse binary regression. *Ann. Statist.* **43** 352–381. MR3311863

[38] Neill, D. and Lingwall, J. (2007). A nonparametric scan statistic for multivariate disease surveillance. *Advances in Disease Surveillance* **4** 106–116.

[39] Saligrama, V. and Zhao, M. (2012). Local anomaly detection. *JMLR W&CP* **22** 969–983.

[40] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. New York: Wiley. MR0838963

[41] Strasser, H. (1985). *Mathematical Theory of Statistics*. *De Gruyter Studies in Mathematics* **7**. Berlin: de Gruyter. MR0812467

[42] Tukey, J.W. (1976). *T*13 *N*: *The Higher Criticism*. *Coures Notes*. *Stat* **411**. Princeton: Princeton University Press.

[43] Tukey, J.W. (1989). Higher Criticism for individual significances in several tables or parts of tables. Internal working paper, Princeton Univ.

[44] Tukey, J.W. (1994). *The Collected Works of John W. Tukey. Vol. VIII*. London: Chapman and Hall. MR1263027

[45] Wellner, J.A. (1978). Limit theorems for the ratio of the empirical distribution function to the true distribution function. *Z. Wahrsch. Verw. Gebiete* **45** 73–88. MR0651392

*Bernoulli* **25**(4A), 2019, 3069–3089
https://doi.org/10.3150/18-BEJ1080

# Second order Lyapunov exponents for parabolic and hyperbolic Anderson models

RALUCA M. BALAN[1] and JIAN SONG[2]

[1]*Department of Mathematics and Statistics, University of Ottawa, 585 King Edward Avenue, Ottawa, ON, K1N 6N5, Canada. E-mail:* rbalan@uottawa.ca
[2]*School of Mathematics, Shandong University, Jinan, Shandong 250100, P.R. China. E-mail:* txjsong@hotmail.com

In this article, we consider the hyperbolic and parabolic Anderson models in arbitrary space dimension $d$, with constant initial condition, driven by a Gaussian noise which is white in time. We consider two spatial covariance structures: (i) the Fourier transform of the spectral measure of the noise is a non-negative locally-integrable function; (ii) $d = 1$ and the noise is a fractional Brownian motion in space with index $1/4 < H < 1/2$. In both cases, we show that there is striking similarity between the Laplace transforms of the second moment of the solutions to these two models. Building on this connection and the recent powerful results of [*Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 1305–1340] for the parabolic model, we compute the second order (upper) Lyapunov exponent for the hyperbolic model. In case (i), when the spatial covariance of the noise is given by the Riesz kernel, we present a unified method for calculating the second order Lyapunov exponents for the two models.

*Keywords:* hyperbolic Anderson model; Lyapunov exponent; parabolic Anderson model; spatially homogeneous Gaussian noise

## References

[1] Balan, R.M. and Chen, L. (2018). Parabolic Anderson model with space–time homogeneous Gaussian noise and rough initial condition. *J. Theoret. Probab.* **31** 2216–2265. MR3866613
[2] Balan, R.M., Jolis, M. and Quer-Sardanyons, L. (2017). Intermittency for the hyperbolic Anderson model with rough noise in space. *Stochastic Process. Appl.* **127** 2316–2338. MR3652415
[3] Balan, R.M. and Song, J. (2017). Hyperbolic Anderson model with space–time homogeneous Gaussian noise. *ALEA Lat. Am. J. Probab. Math. Stat.* **14** 799–849. MR3731795
[4] Bass, R., Chen, X. and Rosen, J. (2009). Large deviations for Riesz potentials of additive processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 626–666. MR2548497
[5] Chen, X. (2017). Moment asymptotics for parabolic Anderson equation with fractional time–space noise: In Skorokhod regime. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 819–841. MR3634276
[6] Chen, X., Hu, Y., Song, J. and Song, X. (2018). Temporal asymptotics for fractional parabolic Anderson model. *Electron. J. Probab.* **23** Paper No. 14, 39. MR3771751
[7] Chen, X., Hu, Y., Song, J. and Xing, F. (2015). Exponential asymptotics for time-space Hamiltonians. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** 1529–1561. MR3414457
[8] Conus, D. and Dalang, R.C. (2008). The non-linear stochastic wave equation in high dimensions. *Electron. J. Probab.* **13** 629–670. MR2399293
[9] Dalang, R.C. (2001). Corrections to: "Extending the martingale measure stochastic integral with applications to spatially homogeneous s.p.d.e.'s" [Electron J. Probab. **4** (1999), no. 6, 29 pp. (electronic); MR1684157 (2000b:60132)]. *Electron. J. Probab.* **6** no. 6, 5. MR1825714

[10] Dalang, R.C. and Mueller, C. (2009). Intermittency properties in a hyperbolic Anderson problem. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 1150–1164. MR2572169

[11] Dawson, D.A. and Salehi, H. (1980). Spatially homogeneous random evolutions. *J. Multivariate Anal.* **10** 141–180. MR0575923

[12] Folland, G.B. (1995). *Introduction to Partial Differential Equations*, 2nd ed. Princeton, NJ: Princeton Univ. Press. MR1357411

[13] Foondun, M. and Khoshnevisan, D. (2013). On the stochastic heat equation with spatially-colored random forcing. *Trans. Amer. Math. Soc.* **365** 409–458. Erratum in *Trans. AMS* **66**, 561-562. MR2984063

[14] Foondun, M., Khoshnevisan, D. and Nualart, E. (2011). A local-time correspondence for stochastic partial differential equations. *Trans. Amer. Math. Soc.* **363** 2481–2515. MR2763724

[15] Gel'fand, I.M. and Shilov, G.E. (1964). *Generalized Functions. Vol. I: Properties and Operations. Translated by Eugene Saletan*. New York: Academic Press. MR0166596

[16] Huang, J., Lê, K. and Nualart, D. (2017). Large time asymptotics for the parabolic Anderson model driven by spatially correlated noise. *Ann. Inst. Henri Poincaré Probab. Stat.* **53** 1305–1340. MR3689969

[17] Khoshnevisan, D. (2014). *Analysis of Stochastic Partial Differential Equations. CBMS Regional Conference Series in Mathematics* **119**. Providence, RI: Amer. Math. Soc.. MR3222416

[18] Podlubny, I. (1999). *Fractional Differential Equations. Mathematics in Science and Engineering* **198**. San Diego, CA: Academic Press. MR1658022

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

[20] Song, J. (2017). On a class of stochastic partial differential equations. *Stochastic Process. Appl.* **127** 37–79. MR3575535

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

*Bernoulli* **25**(4A), 2019, 3090–3108
https://doi.org/10.3150/18-BEJ1082

# Φ-entropy inequalities and asymmetric covariance estimates for convex measures

VAN HOANG NGUYEN

*Institute of Research and Development, Duy Tan University, Da Nang, Vietnam.*
*E-mail: vanhoang0610@yahoo.com*

In this paper, we use the semi-group method and an adaptation of the $L^2$-method of Hörmander to establish some Φ-entropy inequalities and asymmetric covariance estimates for the strictly convex measures in $\mathbb{R}^n$. These inequalities extends the ones for the strictly log-concave measures to more general setting of convex measures. The Φ-entropy inequalities are turned out to be sharp in the special case of Cauchy measures. Finally, we show that the similar inequalities for log-concave measures can be obtained from our results in the limiting case.

*Keywords:* Φ-entropy inequalities; asymmetric covariance estimates; Beckner type inequalities; Brascamp–Lieb type inequalities; convex measures; $L^2$-method of Hörmander; Poincaré type inequalities; semi-group

## References

[1] Arnaudon, M., Bonnefont, M. and Joulin, A. (2018). Intertwinings and generalized Brascamp–Lieb inequalities. *Rev. Mat. Iberoam.* **34** 1021–1054. MR3850277

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

[3] 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

[4] Bakry, D., Gentil, I. and Scheffer, G. (2018). Sharp Beckner-type inequalities for Cauchy and spherical distributions. Preprint. Available atarXiv:1804.03374.

[5] Beckner, W. (1989). A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.* **105** 397–400. MR0954373

[6] Blanchet, A., Bonforte, M., Dolbeault, J., Grillo, G. and Vázquez, J.-L. (2007). Hardy–Poincaré inequalities and applications to nonlinear diffusions. *C. R. Math. Acad. Sci. Paris* **344** 431–436. MR2320246

[7] Blanchet, A., Bonforte, M., Dolbeault, J., Grillo, G. and Vázquez, J.L. (2009). Asymptotics of the fast diffusion equation via entropy estimates. *Arch. Ration. Mech. Anal.* **191** 347–385. MR2481073

[8] Bobkov, S.G. and Ledoux, M. (2009). Weighted Poincaré-type inequalities for Cauchy and other convex measures. *Ann. Probab.* **37** 403–427. MR2510011

[9] Bolley, F. and Gentil, I. (2010). Phi-entropy inequalities for diffusion semigroups. *J. Math. Pures Appl.* (9) **93** 449–473. MR2609029

[10] Bonforte, M., Dolbeault, J., Grillo, G. and Vázquez, J.L. (2010). Sharp rates of decay of solutions to the nonlinear fast diffusion equation via functional inequalities. *Proc. Natl. Acad. Sci. USA* **107** 16459–16464. MR2726546

1350-7265 

[11] Bonnefont, M., Joulin, A. and Ma, Y. (2016). A note on spectral gap and weighted Poincaré inequalities for some one-dimensional diffusions. *ESAIM Probab. Stat.* **20** 18–29. MR3519678
[12] Brascamp, H.J. and Lieb, E.H. (1976). On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Funct. Anal.* **22** 366–389. MR0450480
[13] Carlen, E.A., Cordero-Erausquin, D. and Lieb, E.H. (2013). Asymmetric covariance estimates of Brascamp–Lieb type and related inequalities for log-concave measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** 1–12. MR3060145
[14] Cattiaux, P., Guillin, A. and Wu, L.-M. (2011). Some remarks on weighted logarithmic Sobolev inequality. *Indiana Univ. Math. J.* **60** 1885–1904. MR3008255
[15] Chafaï, D. (2004). Entropies, convexity, and functional inequalities: On $\Phi$-entropies and $\Phi$-Sobolev inequalities. *J. Math. Kyoto Univ.* **44** 325–363. MR2081075
[16] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. MR0420249
[17] Hörmander, L. (1965). $L^2$ estimates and existence theorems for the $\bar{\partial}$ operator. *Acta Math.* **113** 89–152. MR0179443
[18] Menz, G. and Otto, F. (2013). Uniform logarithmic Sobolev inequalities for conservative spin systems with super-quadratic single-site potential. *Ann. Probab.* **41** 2182–2224. MR3098070
[19] Nguyen, V.H. (2014). Dimensional variance inequalities of Brascamp–Lieb type and a local approach to dimensional Prékopa's theorem. *J. Funct. Anal.* **266** 931–955. MR3132733
[20] Scheffer, G. (2003). Local Poincaré inequalities in non-negative curvature and finite dimension. *J. Funct. Anal.* **198** 197–228. MR1962358

*Bernoulli* **25**(4A), 2019, 3109–3138
https://doi.org/10.3150/18-BEJ1083

# On the geometric ergodicity of Hamiltonian Monte Carlo

SAMUEL LIVINGSTONE[1,*], MICHAEL BETANCOURT[2,**], SIMON BYRNE[1,†] and MARK GIROLAMI[2,3,4,‡]

[1]*Department of Statistical Science, University College, Gower Street, London WC1E 6BT, United Kingdom. E-mail:* [*]*samuel.livingstone@ucl.ac.uk;* [†]*simonbyrne@gmail.com*
[2]*Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom. E-mail:* [**]*betanalpha@gmail.com;* [‡]*m.girolami@warwick.ac.uk*
[3]*Department of Mathematics, South Kensington Campus, Imperial College London, London SW7 2AZ, United Kingdom. E-mail: m.girolami@imperial.ac.uk*
[4]*The Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, United Kingdom*

We establish general conditions under which Markov chains produced by the Hamiltonian Monte Carlo method will and will not be geometrically ergodic. We consider implementations with both position-independent and position-dependent integration times. In the former case, we find that the conditions for geometric ergodicity are essentially a gradient of the log-density which asymptotically points towards the centre of the space and grows no faster than linearly. In an idealised scenario in which the integration time is allowed to change in different regions of the space, we show that geometric ergodicity can be recovered for a much broader class of tail behaviours, leading to some guidelines for the choice of this free parameter in practice.

*Keywords:* geometric ergodicity; Hamiltonian dynamics; Hamiltonian Monte Carlo; hybrid Monte Carlo; Markov chain Monte Carlo; Markov chains; stochastic simulation

## References

[1] Alder, B.J. and Wainwright, T.E. (1959). Studies in molecular dynamics. I. General method. *J. Chem. Phys.* **31** 459–466. MR0115313

[2] Andrieu, C., De Freitas, N., Doucet, A. and Jordan, M.I. (2003). An introduction to MCMC for machine learning. *Mach. Learn.* **50** 5–43.

[3] Beskos, A., Pillai, N., Roberts, G., Sanz-Serna, J.-M. and Stuart, A. (2013). Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli* **19** 1501–1534. MR3129023

[4] Beskos, A., Pinski, F.J., Sanz-Serna, J.M. and Stuart, A.M. (2011). Hybrid Monte Carlo on Hilbert spaces. *Stochastic Process. Appl.* **121** 2201–2230. MR2822774

[5] Betancourt, M. (2013). A general metric for Riemannian manifold Hamiltonian Monte Carlo. In *Geometric Science of Information. Lecture Notes in Computer Science* **8085** 327–334. Heidelberg: Springer. MR3126061

[6] Betancourt, M. (2016). Identifying the Optimal Integration Time in Hamiltonian Monte Carlo. Preprint. Available at arXiv:1601.00225.

[7] Betancourt, M., Byrne, S., Livingstone, S. and Girolami, M. (2017). The geometric foundations of Hamiltonian Monte Carlo. *Bernoulli* **23** 2257–2298. MR3648031

[8] Bou-Rabee, N. and Hairer, M. (2013). Nonasymptotic mixing of the MALA algorithm. *IMA J. Numer. Anal.* **33** 80–110. MR3020951

[9] Bou-Rabee, N. and Sanz-Serna, J.M. (2017). Randomized Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **27** 2159–2194. Available at arXiv:1511.09382. MR3693523

[10] Bou-Rabee, N. and Sanz-Serna, J.M. (2018). Geometric integrators and the Hamiltonian Monte Carlo method. *Acta Numer.* **27** 113–206. MR3826507

[11] Byrne, S. and Girolami, M. (2013). Geodesic Monte Carlo on embedded manifolds. *Scand. J. Stat.* **40** 825–845. MR3145120

[12] Campos, C.M. and Sanz-Serna, J.M. (2015). Extra chance generalized hybrid Monte Carlo. *J. Comput. Phys.* **281** 365–374. MR3281978

[13] Cancès, E., Legoll, F. and Stoltz, G. (2007). Theoretical and numerical comparison of some sampling methods for molecular dynamics. *ESAIM Math. Model. Numer. Anal.* **41** 351–389. MR2339633

[14] Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P. and Riddell, A. (2016). Stan: A probabilistic programming language. *J. Stat. Softw.*.

[15] Diaconis, P. (2013). Some things we've learned (about Markov chain Monte Carlo). *Bernoulli* **19** 1294–1305. MR3102552

[16] Diaconis, P. and Freedman, D. (1999). Iterated random functions. *SIAM Rev.* **41** 45–76. MR1669737

[17] Diaconis, P., Seiler, C. and Holmes, S. (2014). Connections and extensions: A discussion of the paper by Girolami and Byrne [MR3145120]. *Scand. J. Stat.* **41** 3–7. MR3181123

[18] Duane, S., Kennedy, A.D., Pendleton, B.J. and Roweth, D. (1987). Hybrid Monte Carlo. *Phys. Lett. B* **195** 216–222.

[19] Durmus, A. and Moulines, É. (2015). Quantitative bounds of convergence for geometrically ergodic Markov chain in the Wasserstein distance with application to the Metropolis adjusted Langevin algorithm. *Stat. Comput.* **25** 5–19. MR3304898

[20] Durmus, A., Moulines, E. and Saksman, E. (2017). On the convergence of Hamiltonian Monte Carlo. Preprint. Available at arXiv:1705.00166.

[21] Eberle, A. (2014). Error bounds for Metropolis–Hastings algorithms applied to perturbations of Gaussian measures in high dimensions. *Ann. Appl. Probab.* **24** 337–377. MR3161650

[22] Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (2014). *Bayesian Data Analysis* **2**. Taylor & Francis.

[23] 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. With discussion and a reply by the authors. MR2814492

[24] Goldstein, H. (1965). *Classical Mechanics*. Pearson Education India.

[25] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** 97–109. MR3363437

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

[27] Horowitz, A.M. (1991). A generalized guided Monte Carlo algorithm. *Phys. Lett. B* **268** 247–252.

[28] Jarner, S.F. and Hansen, E. (2000). Geometric ergodicity of Metropolis algorithms. *Stochastic Process. Appl.* **85** 341–361. MR1731030

[29] Jarner, S.F. and Tweedie, R.L. (2003). Necessary conditions for geometric and polynomial ergodicity of random-walk-type Markov chains. *Bernoulli* **9** 559–578. MR1996270

[30] Jones, G.L. and Hobert, J.P. (2001). Honest exploration of intractable probability distributions via Markov chain Monte Carlo. *Statist. Sci.* **16** 312–334. MR1888447

[31] Lee, J.M. (2012). *Introduction to Smooth Manifolds*. Springer.

[32] Leimkuhler, B. and Reich, S. (2004). *Simulating Hamiltonian Dynamics*. *Cambridge Monographs on Applied and Computational Mathematics* **14**. Cambridge: Cambridge Univ. Press. MR2132573

[33] Livingstone, S., Betancourt, M., Byrne, S. and Girolami, M. (2018). Supplement to "On the geometric ergodicity of Hamiltonian Monte Carlo." DOI:10.3150/18-BEJ1083SUPP.
[34] Livingstone, S. and Girolami, M. (2014). Information-geometric Markov chain Monte Carlo methods using diffusions. *Entropy* **16** 3074–3102. MR3234224
[35] Mattingly, J.C., Stuart, A.M. and Higham, D.J. (2002). Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** 185–232. MR1931266
[36] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21** 1087–1092.
[37] Meyn, S.P. and Tweedie, R.L. (2012). *Markov Chains and Stochastic Stability*. Springer Science & Business Media.
[38] Neal, R.M. (2011). 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
[39] Ollivier, Y. (2009). Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **256** 810–864. MR2484937
[40] 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
[41] Roberts, G.O. and Rosenthal, J.S. (1997). Geometric ergodicity and hybrid Markov chains. *Electron. Commun. Probab.* **2** 13–25. MR1448322
[42] Roberts, G.O. and Rosenthal, J.S. (2004). General state space Markov chains and MCMC algorithms. *Probab. Surv.* **1** 20–71. MR2095565
[43] Roberts, G.O. and Tweedie, R.L. (1996). Geometric convergence and central limit theorems for multidimensional Hastings and Metropolis algorithms. *Biometrika* **83** 95–110. MR1399158
[44] Roberts, G.O. and Tweedie, R.L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* **2** 341–363. MR1440273
[45] Seiler, C., Rubinstein-Salzedo, S. and Holmes, S. (2014). Positive curvature and Hamiltonian Monte Carlo. In *Advances in Neural Information Processing Systems* 586–594.
[46] Stuart, A.M. (2010). Inverse problems: A Bayesian perspective. *Acta Numer.* **19** 451–559. MR2652785
[47] Tierney, L. (1994). Markov chains for exploring posterior distributions. *Ann. Statist.* **22** 1701–1762. With discussion and a rejoinder by the author. MR1329166
[48] Tierney, L. (1998). A note on Metropolis–Hastings kernels for general state spaces. *Ann. Appl. Probab.* **8** 1–9. MR1620401

*Bernoulli* **25**(4A), 2019, 3139–3174
https://doi.org/10.3150/18-BEJ1084

# Gaussian fluctuations for high-dimensional random projections of $\ell_p^n$-balls

DAVID ALONSO-GUTIÉRREZ[1], JOSCHA PROCHNO[2] and CHRISTOPH THÄLE[3]

[1]*Departamento de Matemáticas, Universidad de Zaragoza, Spain. E-mail: alonsod@unizar.es*
[2]*Institute of Mathematics & Scientific Computing, University of Graz, Austria. E-mail: joscha.prochno@uni-graz.at*
[3]*Faculty of Mathematics, Ruhr University Bochum, Germany. E-mail: christoph.thaele@rub.de*

In this paper, we study high-dimensional random projections of $\ell_p^n$-balls. More precisely, for any $n \in \mathbb{N}$ let $E_n$ be a random subspace of dimension $k_n \in \{1, \dots, n\}$ and $X_n$ be a random point in the unit ball of $\ell_p^n$. Our work provides a description of the Gaussian fluctuations of the Euclidean norm $\|P_{E_n} X_n\|_2$ of random orthogonal projections of $X_n$ onto $E_n$. In particular, under the condition that $k_n \to \infty$ it is shown that these random variables satisfy a central limit theorem, as the space dimension $n$ tends to infinity. Moreover, if $k_n \to \infty$ fast enough, we provide a Berry–Esseen bound on the rate of convergence in the central limit theorem. At the end, we provide a discussion of the large deviations counterpart to our central limit theorem.

*Keywords:* $\ell_p^n$-ball; Berry–Esseen bound; central limit theorem; cone measure; large deviations; random projection; uniform distribution

## References

[1] Alonso-Gutiérrez, D., Prochno, J. and Thäle, C. (2018). Large deviations for high-dimensional random projections of $\ell_p^n$-balls. *Adv. in Appl. Math.* **99** 1–35. MR3806754
[2] Barthe, F., Guédon, O., Mendelson, S. and Naor, A. (2005). A probabilistic approach to the geometry of the $l_p^n$-ball. *Ann. Probab.* **33** 480–513. MR2123199
[3] Bass, R.F. (2011). *Stochastic Processes*. *Cambridge Series in Statistical and Probabilistic Mathematics* **33**. Cambridge: Cambridge Univ. Press. MR2856623
[4] Basu, S.K. (1974). Density versions of the univariate central limit theorem. *Ann. Probab.* **2** 270–276. MR0356176
[5] 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
[6] den Hollander, F. (2000). *Large Deviations*. *Fields Institute Monographs* **14**. Providence, RI: Amer. Math. Soc. MR1739680
[7] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*. *Vol. II*. 2nd ed. New York: Wiley. MR0270403
[8] Gantert, N., Kim, S.S. and Ramanan, K. (2017). Large deviations for random projections of $\ell^p$ balls. *Ann. Probab.* **45** 4419–4476. MR3737915
[9] Goldstein, L. and Shao, Q.-M. (2009). Berry–Esseen bounds for projections of coordinate symmetric random vectors. *Electron. Commun. Probab.* **14** 474–485. MR2559097

1350-7265 

[10] Grote, J., Kabluchko, Z. and Thäle, C. (2017). Limit theorems for random simplices in high dimensions. ArXiv e-prints.

[11] Kabluchko, Z., Litvak, A.E. and Zaporozhets, D. (2015). Mean width of regular polytopes and expected maxima of correlated Gaussian variables. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov.* (*POMI*) **442** 75–96. MR3506845

[12] Kabluchko, Z., Prochno, J. and Thäle, C. (2017). High-dimensional limit theorems for random vectors in $\ell_p^n$-balls. *Commun. Contemp. Math.* To appear.

[13] Kallenberg, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1876169

[14] Klartag, B. (2007). A central limit theorem for convex sets. *Invent. Math.* **168** 91–131. MR2285748

[15] Klartag, B. (2007). Power-law estimates for the central limit theorem for convex sets. *J. Funct. Anal.* **245** 284–310. MR2311626

[16] Klartag, B. (2009). A Berry–Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Related Fields* **145** 1–33. MR2520120

[17] Meckes, M.W. (2009). Gaussian marginals of convex bodies with symmetries. *Beitr. Algebra Geom.* **50** 101–118. MR2499783

[18] Naor, A. (2007). The surface measure and cone measure on the sphere of $l_p^n$. *Trans. Amer. Math. Soc.* **359** 1045–1079. MR2262841

[19] 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

[20] Schechtman, G. and Zinn, J. (1990). On the volume of the intersection of two $L_p^n$ balls. *Proc. Amer. Math. Soc.* **110** 217–224. MR1015684

[21] Simon, B. (2015). *Basic Complex Analysis. A Comprehensive Course in Analysis, Part* 2*A*. Providence, RI: Amer. Math. Soc. MR3443339

*Bernoulli* **25**(4A), 2019, 3175–3200
https://doi.org/10.3150/18-BEJ1085

# On Lasso refitting strategies

EVGENII CHZHEN[1,*], MOHAMED HEBIRI[1,**] and JOSEPH SALMON[2]

[1]*LAMA, Université Paris-Est, 5 Boulevard Descartes, 77420, Champs-sur-Marne, France.*
*E-mail:* [*]*evgenii.chzhen@univ-paris-est.fr;* [**]*mohamed.hebiri@u-pem.fr*
[2]*IMAG, Université de Montpellier, Place Eugène Bataillon, 34095 Montpellier Cedex 5, France.*
*E-mail: joseph.salmon@umontpellier.fr*

A well-known drawback of $\ell_1$-penalized estimators is the systematic shrinkage of the large coefficients towards zero. A simple remedy is to treat Lasso as a model-selection procedure and to perform a second refitting step on the selected support. In this work, we formalize the notion of refitting and provide oracle bounds for arbitrary refitting procedures of the Lasso solution. One of the most widely used refitting techniques which is based on Least-Squares may bring a problem of interpretability, since the signs of the refitted estimator might be flipped with respect to the original estimator. This problem arises from the fact that the Least-Squares refitting considers only the support of the Lasso solution, avoiding any information about signs or amplitudes. To this end, we define a sign consistent refitting as an arbitrary refitting procedure, preserving the signs of the first step Lasso solution and provide Oracle inequalities for such estimators. Finally, we consider special refitting strategies: Bregman Lasso and Boosted Lasso. Bregman Lasso has a fruitful property to converge to the Sign-Least-Squares refitting (Least-Squares with sign constraints), which provides with greater interpretability. We additionally study the Bregman Lasso refitting in the case of orthogonal design, providing with simple intuition behind the proposed method. Boosted Lasso, in contrast, considers information about magnitudes of the first Lasso step and allows to develop better oracle rates for prediction. Finally, we conduct an extensive numerical study to show advantages of one approach over others in different synthetic and semi-real scenarios.

*Keywords:* Bregman; Lasso; linear regression; refitting

## References

[1] Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19** 521–547. MR3037163

[2] Bickel, P.J., Ritov, Y. and Tsybakov, A.B. (2009). Simultaneous analysis of Lasso and Dantzig selector. *Ann. Statist.* **37** 1705–1732. MR2533469

[3] Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge: Cambridge Univ. Press. MR2061575

[4] Brinkmann, E.-M., Burger, M., Rasch, J. and Sutour, C. (2017). Bias reduction in variational regularization. *J. Math. Imaging Vision* **59** 534–566. MR3712429

[5] Bühlmann, P. and Mandozzi, J. (2014). High-dimensional variable screening and bias in subsequent inference, with an empirical comparison. *Comput. Statist.* **29** 407–430. MR3261821

[6] Bühlmann, P. and Yu, B. (2003). Boosting with the $L_2$ loss: Regression and classification. *J. Amer. Statist. Assoc.* **98** 324–339. MR1995709

[7] Chzhen, E., Hebiri, M. and Salmon, J. (2019). Supplement to "On Lasso refitting strategies." DOI:10.3150/18-BEJ1085SUPP.

1350-7265 

[8] Dalalyan, A.S., Hebiri, M. and Lederer, J. (2017). On the prediction performance of the Lasso. *Bernoulli* **23** 552–581. MR3556784

[9] Deledalle, C.-A., Papadakis, N. and Salmon, J. (2015). On debiasing restoration algorithms: Applications to total-variation and nonlocal-means. In *Scale Space and Variational Methods in Computer Vision*. *Lecture Notes in Computer Science* **9087** 129–141. Cham: Springer. MR3394925

[10] Deledalle, C.-A., Papadakis, N., Salmon, J. and Vaiter, S. (2017). CLEAR: Covariant LEAst-square refitting with applications to image restoration. *SIAM J. Imaging Sci.* **10** 243–284. MR3615463

[11] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. MR1946581

[12] Gao, H.-Y. and Bruce, A.G. (1997). WaveShrink with firm shrinkage. *Statist. Sinica* **7** 855–874. MR1488646

[13] Gasso, G., Rakotomamonjy, A. and Canu, S. (2009). Recovering sparse signals with a certain family of nonconvex penalties and DC programming. *IEEE Trans. Signal Process.* **57** 4686–4698. MR2722328

[14] Giraud, C. (2014). *Introduction to High-Dimensional Statistics*. *Chapman and Hall/CRC Monographs on Statistics and Applied Probability*. Boca Raton: Chapman and Hall/CRC.

[15] Koltchinskii, V. (2011). *Oracle Inequalities in Empirical Risk Minimization and Sparse Recovery Problems*. *Lecture Notes in Math.* **2033**. Heidelberg: Springer. MR2829871

[16] Lederer, J. (2013). Trust, but verify: Benefits and pitfalls of least-squares refitting in high dimensions. Preprint. Available at arXiv:1306.0113.

[17] Meinshausen, N. (2007). Relaxed Lasso. *Comput. Statist. Data Anal.* **52** 374–393. MR2409990

[18] Osher, S., Burger, M., Goldfarb, D., Xu, J. and Yin, W. (2005). An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.* **4** 460–489. MR2162864

[19] Osher, S., Ruan, F., Xiong, J., Yao, Y. and Yin, W. (2016). Sparse recovery via differential inclusions. *Appl. Comput. Harmon. Anal.* **41** 436–469. MR3534446

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

[21] Tibshirani, R.J. (2013). The lasso problem and uniqueness. *Electron. J. Stat.* **7** 1456–1490. MR3066375

[22] Tukey, J.W. (1977). *Exploratory Data Analysis*. Reading: Addison-Wesley.

[23] Xu, J. and Osher, S. (2007). Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising. *IEEE Trans. Image Process.* **16** 534–544. MR2462742

[24] Yin, W., Osher, S., Goldfarb, D. and Darbon, J. (2008). Bregman iterative algorithms for $l_1$-minimization with applications to compressed sensing. *SIAM J. Imaging Sci.* **1** 143–168. MR2475828

[25] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.* **38** 894–942. MR2604701

*Bernoulli* **25**(4A), 2019, 3201
https://doi.org/10.3150/19-BEJ1136
Main article: https://doi.org/10.3150/14-BEJ689

# Corrigendum: Analysis of the forward search using some new results for martingales and empirical processes

VANESSA BERENGUER-RICO[1,*], SØREN JOHANSEN[2] and BENT NIELSEN[1,**]

[1]*Department of Economics, University of Oxford, Manor Road, Oxford, OX1 3UQ, United Kingdom. E-mail:* [*]*vanessa.berenguer-rico@economics.ox.ac.uk;* [**]*bent.nielsen@nuffield.ox.ac.uk*
[2]*Department of Economics, University of Copenhagen and CREATES, Department of Economics and Business, Aarhus University, DK-8000 Aarhus C, Denmark. E-mail: soren.johansen@econ.ku.dk*

## References

[1] Berenguer-Rico, V., Johansen, S. and Nielsen, B. (2019). The analysis of marked and weighted empirical processes of estimated residuals. Discussion Paper 19-05, Department of Economics, University of Copenhagen.

[2] Johansen, S. and Nielsen, B. (2016). Analysis of the forward search using some new results for martingales and empirical processes. *Bernoulli* **22** 1131–1183. MR3449811

# Bernoulli Forthcoming Papers

LEI, T.
Scaling limit of random forests with prescribed degree sequences

ICHIBA, T. and SARANTSEV, A.
Stationary distributions and convergence for Walsh diffusions

LIMIC, V.
The eternal multiplicative coalescent encoding via excursions of Lévy-type processes

BUHL, S., DAVIS, R.A., KLÜPPELBERG, C. and STEINKOHL, C.
Semiparametric estimation for isotropic max-stable space-time processes

GÖTZE, F., NAUMOV, A., SPOKOINY, V. and ULYANOV, V.
Large ball probabilities, Gaussian comparison and anti-concentration

FANG, X., PENG, S., SHAO, Q.-M. and SONG, Y.
Limit theorems with rate of convergence under sublinear expectations

REISS, M. and WAHL, M.
Functional estimation and hypothesis testing in nonparametric boundary models

WEED, J. and BACH, F.
Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance

MARGUET, A.
Uniform sampling in a structured branching population

ABRAHAM, K.
Nonparametric Bayesian posterior contraction rates for scalar diffusions with high-frequency data

HUESMANN, M. and TREVISAN, D.
A Benamou–Brenier formulation of martingale optimal transport

BHATTACHARJEE, C. and GOLDSTEIN, L.
Dickman approximation in simulation, summations and perpetuities

FAN, X., GRAMA, I., LIU, Q. and SHAO, Q.-M.
Self-normalized Cramér type moderate deviations for martingales

RAIČ, M.
A multivariate Berry–Esseen theorem with explicit constants

DURMUS, A. and MOULINES, É.
High-dimensional Bayesian inference via the unadjusted Langevin algorithm

BECT, J., BACHOC, F. and GINSBOURGER, D.
A supermartingale approach to Gaussian process based sequential design of experiments

ANH, V., LEONENKO, N., OLENKO, A. and VASKOVYCH, V.
On rate of convergence in non-central limit theorems

*Continues*

# Bernoulli Forthcoming Papers—*Continued*

LIU, Z., BLANCHET, J.H., DIEKER, A.B. and MIKOSCH, T.
On logarithmically optimal exact simulation of max-stable and related random fields on a compact set

DOLERA, E. and REGAZZINI, E.
Uniform rates of the Glivenko–Cantelli convergence and their use in approximating Bayesian inferences

BELLEC, P.C.
Localized Gaussian width of $M$-convex hulls with applications to Lasso and convex aggregation

DITZHAUS, M.
Signal detection via Phi-divergences for general mixtures

BALAN, R.M. and SONG, J.
Second order Lyapunov exponents for parabolic and hyperbolic Anderson models

NGUYEN, V.H.
$\Phi$-entropy inequalities and asymmetric covariance estimates for convex measures

LIVINGSTONE, S., BETANCOURT, M., BYRNE, S. and GIROLAMI, M.
On the geometric ergodicity of Hamiltonian Monte Carlo

ALONSO-GUTIÉRREZ, D., PROCHNO, J. and THÄLE, C.
Gaussian fluctuations for high-dimensional random projections of $\ell_p^n$-balls

CHZHEN, E., HEBIRI, M. and SALMON, J.
On Lasso refitting strategies

BERENGUER-RICO, V., JOHANSEN, S. and NIELSEN, B.
Corrigendum: Analysis of the forward search using some new results for martingales and empirical processes

THAI, M.-N.
Birth and death process in mean field type interaction

HAUSER, R.A. and MATZINGER, H.
Microscopic path structure of optimally aligned random sequences

MERLEVÈDE, F., PELIGRAD, M. and UTEV, S.
Functional CLT for martingale-like nonstationary dependent structures

VARVENNE, M.
Rate of convergence to equilibrium for discrete-time stochastic dynamics with memory

BALABDAOUI, F., DUROT, C. and JANKOWSKI, H.
Least squares estimation in the monotone single index model

HONDA, T., ING, C.-K. and WU, W.-Y.
Adaptively weighted group Lasso for semiparametric quantile regression models

*Continues*

# Bernoulli Forthcoming Papers—*Continued*

ALETTI, G., CRIMALDI, I. and GHIGLIETTI, A.
Networks of reinforced stochastic processes: Asymptotics for the empirical means

BUTLER, R.W. and WOOD, A.T.A.
Limiting saddlepoint relative errors in large deviation regions under purely Tauberian conditions

JIANG, W. and ZHANG, C.-H.
Rate of divergence of the nonparametric likelihood ratio test for Gaussian mixtures

BLANCHARD, G. and ZADOROZHNYI, O.
Concentration of weakly dependent Banach-valued sums and applications to statistical learning methods

GREENSHTEIN, E., MANTZURA, A. and RITOV, Y.
Nonparametric empirical Bayes improvement of shrinkage estimators with applications to time series

MALLEIN, B. and RAMASSAMY, S.
Two-sided infinite-bin models and analyticity for Barak–Erdős graphs

CÉNAC, P., DE LOYNES, B., OFFRET, Y. and ROUSSELLE, A.
Recurrence of multidimensional persistent random walks. Fourier and series criteria

PILAVAKIS, D., PAPARODITIS, E. and SAPATINAS, T.
Moving block and tapered block bootstrap for functional time series with an application to the $K$-sample mean problem

FAN, J.Y., HAMZA, K., JAGERS, P. and KLEBANER, F.
Convergence of the age structure of general schemes of population processes

BERTAIL, P. and CLÉMENÇON, S.
Bernstein-type exponential inequalities in survey sampling: Conditional Poisson sampling schemes

BARTHELMÉ, S., AMBLARD, P.-O. and TREMBLAY, N.
Asymptotic equivalence of fixed-size and varying-size determinantal point processes

HEINY, J. and MIKOSCH, T.
The eigenstructure of the sample covariance matrices of high-dimensional stochastic volatility models with heavy tails

PITMAN, J. and YAKUBOVICH, Y.
Gaps and interleaving of point processes in sampling from a residual allocation model

LIN, S.
Harmonic measure for biased random walk in a supercritical Galton–Watson tree

FRANCESCHI, S. and RASCHEL, K.
Integral expression for the stationary distribution of reflected Brownian motion in a wedge

*Continues*

# Bernoulli Forthcoming Papers—*Continued*

GOUÉRÉ, J.-B. and THÉRET, M.
Equivalence of some subcritical properties in continuum percolation

RAVNER, L., BOXMA, O. and MANDJES, M.
Estimating the input of a Lévy-driven queue by Poisson sampling of the workload process

COLLING, B. and VAN KEILEGOM, I.
Estimation of fully nonparametric transformation models

SHIOZAWA, Y. and WANG, J.
Long-time heat kernel estimates and upper rate functions of Brownian motion type for symmetric jump processes

CHAKRABORTY, S. and KHARE, K.
Consistent estimation of the spectrum of trace class Data Augmentation algorithms

KOKOSZKA, P., STOEV, S. and XIONG, Q.
Principal components analysis of regularly varying functions

KLOPP, O., LU, Y., TSYBAKOV, A.B. and ZHOU, H.H.
Structured matrix estimation and completion

BERTAIL, P. and PORTIER, F.
Rademacher complexity for Markov chains: Applications to kernel smoothing and Metropolis–Hastings

BURDZY, K., KOŁODZIEJEK, B. and TADIĆ, T.
Inverse exponential decay: Stochastic fixed point equation and ARMA models

SAUMARD, A.
Weighted Poincaré inequalities, concentration inequalities and tail bounds related to Stein kernels in dimension one

GÖTZE, F. and SAMBALE, H.
Second order concentration via logarithmic Sobolev inequalities

CASTILLO, I. and SZABÓ, B.
Spike and slab empirical Bayes sparse credible sets

GEISS, C., LABART, C. and LUOTO, A.
Random walk approximation of BSDEs with Hölder continuous terminal condition

HIRSCH, C. and MÖNCH, C.
Distances and large deviations in the spatial preferential attachment model

BELITSER, E. and NURUSHEV, N.
Needles and straw in a haystack: Robust confidence for possibly sparse sequences

BONNET, G. and CHENAVIER, N.
The maximal degree in a Poisson–Delaunay graph

MAZUR, S., OTRYAKHIN, D. and PODOLSKIJ, M.
Estimation of the linear fractional stable motion

*Continues*

# Bernoulli Forthcoming Papers—*Continued*

BOGERD, K., CASTRO, R.M. and VAN DER HOFSTAD, R.
Cliques in rank-1 random graphs: The role of inhomogeneity

GAYNANOVA, I.
Prediction and estimation consistency of sparse multi-class penalized optimal scoring

ZWINGMANN, T. and HOLZMANN, H.
Weak convergence of quantile and expectile processes under general assumptions

GIORDANO, L.M., JOLIS, M. and QUER-SARDANYONS, L.
SPDEs with fractional noise in space: Continuity in law with respect to the Hurst index

DING, X.
High dimensional deformed rectangular matrices with applications in matrix denoising

SHI, C. and TANG, B.
Construction results for strong orthogonal arrays of strength three

MA, Z. and LI, X.
Subspace perspective on canonical correlation analysis: Dimension reduction and minimax rates

FOKIANOS, K., STØVE, B., TJØSTHEIM, D. and DOUKHAN, P.
Multivariate count autoregression

SHEN, Y. and WANG, Y.
Operator-scaling Gaussian random fields via aggregation

DÖRING, L. and WEISSMANN, P.
Stable processes conditioned to hit an interval continuously from the outside

MUKHERJEE, S.
Degeneracy in sparse ERGMs with functions of degrees as sufficient statistics

DAOUIA, A., GIRARD, S. and STUPFLER, G.
Tail expectile process and risk assessment

CONTI, P.L., MARELLA, D., MECATTI, F. and ANDREIS, F.
A unified principled framework for resampling based on pseudo-populations: Asymptotic theory

MARIUCCI, E., RAY, K. and SZABÓ, B.
A Bayesian nonparametric approach to log-concave density estimation

BEUTNER, E., BORDES, L. and DOYEN, L.
Consistent semiparametric estimators for recurrent event times models with application to virtual age models

PRIVAULT, N. and SERAFIN, G.
Normal approximation for sums of weighted $U$-statistics – application to Kolmogorov bounds in random subgraph counting

*Continues*

# Bernoulli Forthcoming Papers—*Continued*

YANO, K. and KATO, K.
On frequentist coverage errors of Bayesian credible sets in moderately high dimensions

ALETTI, G., CRIMALDI, I. and GHIGLIETTI, A.
Interacting reinforced stochastic processes: Statistical inference based on the weighted empirical means

CHENG, D., CAMMAROTA, V., FANTAYE, Y., MARRINUCCI, D. and SCHWARTZMAN, A.
Multiple testing of local maxima for detection of peaks on the (celestial) sphere

BERHARD, H. and DAS, B.
Heavy-tailed random walks, buffered queues and hidden large deviations

DAVIS, R.A., NIELSEN, M.S. and ROHDE, V.
Stochastic differential equations with a fractionally filtered delay: a semimartingale model for long-range dependent processes

HO, N., NGUYEN, X. and RITOV, Y.
Robust estimation of mixing measures in finite mixture models

NOURDIN, I., PECCATI, G. and SEURET, S.
Sojourn time dimensions of fractional Brownian motion

GAO, C.
Robust regression via mutivariate regression depth

SCHNURR, A.
The fourth characteristic of a semimartingale

LIU, Y. and PAGES, G.
Characterization of probability distribution convergence in Wasserstein distance by $L^p$-quantization error function

KLUSOWSKI, J. M. and WU, Y.
Estimating the number of connected components in a graph via subgraph sampling

LIANG, M., SCHILLING, R.L. and WANG, J.
A unified approach to coupling SDEs driven by Levy noise and some applications

MINSKER, S. and WEI, X.
Robust modifications of U-statistics and applications to covariance estimation problems

ISSOGLIO, E. and RUSSO, F.
A Feynman–Kac result via Markov BSDEs with generalised drivers

LEI, J.
Convergence and concentration of empirical measures under Wasserstein distance in unbounded functional spaces

MARCHAND, D.C. and MANOLESCU, I.
Influence of the seed in affine preferential attachment trees