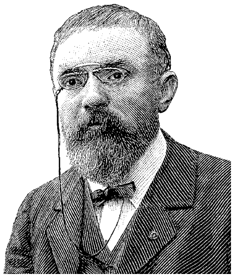

Volume 60, Number 2, 2024
ISSN 0246-0203

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES

Global contractivity for Langevin dynamics with distribution-dependent forces and uniform in time propagation of chaos . . . . *K. Schuh* 753–789

Fluctuations for mean field limits of interacting systems of spiking neurons . . . . *E. Löcherbach* 790–823

Empirical optimal transport between different measures adapts to lower complexity *S. Hundrieser, T. Staudt and A. Munk* 824–846

Central limit theorems for general transportation costs . . . . *E. del Barrio, A. González-Sanz and J.-M. Loubes* 847–873

Concentration of quasi-stationary distributions for one-dimensional diffusions with applications *Z. Shen, S. Wang and Y. Yi* 874–903

On the Itô–Alekseev–Gröbner formula for stochastic differential equations *A. Hudde, M. Hutzenthaler, A. Jentzen and S. Mazzonetto* 904–922

Matsumoto–Yor and Dufresne type theorems for a random walk on positive definite matrices *J. Arista, E. Bisi and N. O'Connell* 923–945

Linear spectral statistics of sequential sample covariance matrices . . . . *N. Dörnemann and H. Dette* 946–970

Robust subgaussian estimation with VC-dimension . . . . *J. Depersin* 971–989

Large deviations for random matrices in the orthogonal group and Stiefel manifold with applications to random projections of product distributions . . . . *Z. Kabluchko and J. Prochno* 990–1024

Random sorting networks: Edge limit . . . . *V. Gorin and J. Xu* 1025–1047

Log determinant of large correlation matrices under infinite fourth moment . . . . *J. Heiny and N. Parolya* 1048–1076

Exponential concentration for the number of roots of random trigonometric polynomials *H. H. Nguyen and O. Zeitouni* 1077–1089

On measures strongly log-concave on a subspace . . . . *P. Bizeul* 1090–1100

Metropolis–Hastings transition kernel couplings . . . . *J. O'Leary and G. Wang* 1101–1124

A central limit theorem for the variation of the sum of digits . . . . *Y. Hosten, É. Janvresse and T. de la Rue* 1125–1149

Number of visits in arbitrary sets for ϕ-mixing dynamics . . . . *S. Gallo, N. Haydn and S. Vaienti* 1150–1187

Active phase for activated random walks on the lattice in all dimensions . . . . *N. Forien and A. Gaudillière* 1188–1214

Hydrodynamics of the t-PNG model via a colored t-PNG model . . . . *H. Drillick and Y. Lin* 1215–1245

Non-stationary KPZ equation from ASEP with slow bonds . . . . *K. Yang* 1246–1294

Convergence of limit shapes for 2D near-critical first-passage percolation . . . . *C.-L. Yao* 1295–1333

Random walks in Dirichlet environments on $\mathbb{Z}$ with bounded jumps . . . . *D. J. Slonim* 1334–1355

Fractal properties of Aldous–Kendall random metric . . . . *G. Blanc* 1356–1386

The Seneta–Heyde scaling for supercritical super-Brownian motion . . . . *H. Hou, Y.-X. Ren and R. Song* 1387–1417

A branching process with deletions and mergers that matches the threshold for hypercube percolation *L. Eslava, S. Penington and F. Skerman* 1418–1457

Generalized range of slow random walks on trees . . . . *P. Andreoletti and A. Kagan* 1458–1509

60|2

ANNALES DE L'INSTITUT HENRI POINCARÉ PROBABILITÉS ET STATISTIQUES

Vol. 60, No. 2 (May, 2024) 753–1509

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES

Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques (ISSN 0246-0203), Volume 60, Number 2, May 2024. Published quarterly by Association des Publications de l'Institut Henri Poincaré.

POSTMASTER: Send address changes to Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, Dues and Subscriptions Office, PO Box 729, Middletown, Maryland 21769, USA.

Président et directeur de la publication : Sylvie Benzoni
Périodicité : 4 n^os^ / an

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 753–789
https://doi.org/10.1214/22-AIHP1337

Global contractivity for Langevin dynamics with distribution-dependent forces and uniform in time propagation of chaos

Katharina Schuh[a]

Universität Bonn, Institut für Angewandte Mathematik, Endenicher Allee 60, 53115 Bonn, Germany, [a]*katharina.schuh@uni-bonn.de*

Abstract. We study the long-time behaviour of both the classical second-order Langevin dynamics and the nonlinear second-order Langevin dynamics of McKean–Vlasov type. By a coupling approach, we establish global contraction in an L^1 Wasserstein distance with an explicit dimension-free rate for pairwise weak interactions. For external forces corresponding to a κ-strongly convex potential a contraction rate of order $\mathcal{O}(\sqrt{\kappa})$ is obtained in certain cases. But the result is not restricted to these forces. It rather includes multi-well potentials and non-gradient-type external forces as well as non-gradient-type repulsive and attractive interaction forces. The proof is based on a novel distance function which combines two contraction results for large and small distances and uses a coupling approach adjusted to the distance. By applying a componentwise adaptation of the coupling we provide uniform in time propagation of chaos bounds for the corresponding mean-field particle system.

Résumé. Nous étudions le comportement en temps long de la dynamique de second ordre de Langevin ainsi que sa version non linéaire de type McKean–Vlasov. Par une approche par couplage, nous établissons la contraction globale en distance de Wasserstein L_1 avec un taux explicite indépendant de la dimension dans le cas d'une faible interaction par paire. Lorsque la force de confinement correspond à un potentiel κ-fortement convexe, un taux de contraction de l'ordre de $\mathcal{O}(\sqrt{\kappa})$ est obtenu dans certains cas. Mais le résultat ne se limite pas à ce type de forces. En effet, il est possible de considérer également des confinements de type non gradient et multi-puits ainsi que des interactions non gradients atractives ou répulsives. Notre preuve repose sur une nouvelle fonction distance qui combine deux résultats de contraction pour les petites et grandes distances et utilise ainsi un couplage adapté. En utilisant une adaptation coordonnée par coordonnée du couplage nous obtenons la propagation du chaos uniforme en temps pour le système de particules à champ moyen associé.

MSC2020 subject classifications: Primary 60H10; 60J60; secondary 82C31

Keywords: Langevin dynamics; Coupling; Convergence to equilibrium; Wasserstein distance; Vlasov–Fokker–Planck equation; Propagation of chaos

References

[1] F. Achleitner, A. Arnold and D. Stürzer. Large-time behavior in non-symmetric Fokker–Planck equations. *Riv. Mat. Univ. Parma (N.S.)* **6** (1) (2015) 1–68. MR3468699

[2] D. Bakry, P. Cattiaux and A. Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.* **254** (3) (2008) 727–759. MR2381160 https://doi.org/10.1016/j.jfa.2007.11.002

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

[4] D. Benedetto, E. Caglioti, J. A. Carrillo and M. Pulvirenti. A non-Maxwellian steady distribution for one-dimensional granular media. *J. Stat. Phys.* **91** (5–6) (1998) 979–990. MR1637274 https://doi.org/10.1023/A:1023032000560

[5] F. Bolley, I. Gentil and A. Guillin. Convergence to equilibrium in Wasserstein distance for Fokker–Planck equations. *J. Funct. Anal.* **263** (8) (2012) 2430–2457. MR2964689 https://doi.org/10.1016/j.jfa.2012.07.007

[6] F. Bolley, I. Gentil and A. Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.* **208** (2) (2013) 429–445. MR3035983 https://doi.org/10.1007/s00205-012-0599-z

[7] F. Bolley, A. Guillin and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *ESAIM: M2AN* **44** (5) (2010) 867–884. MR2731396 https://doi.org/10.1051/m2an/2010045

[8] F. Bouchut and J. Dolbeault. On long time asymptotics of the Vlasov–Fokker–Planck equation and of the Vlasov–Poisson–Fokker–Planck system with Coulombic and Newtonian potentials. *Differential Integral Equations* **8** (3) (1995) 487–514. MR1306570

[9] Y. Cao, J. Lu and L. Wang On explicit L^2-convergence rate estimate for underdamped Langevin dynamics. arXiv preprint, 2019. Available at arXiv:1908.04746v4.
[10] J. A. Carrillo, R. J. McCann and C. Villani. Kinetic equilibration rates for granular media and related equations: Entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoam.* **19** (3) (2003) 971–1018. MR2053570 https://doi.org/10.4171/RMI/376
[11] J. A. Carrillo, R. J. McCann and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.* **179** (2) (2006) 217–263. MR2209130 https://doi.org/10.1007/s00205-005-0386-1
[12] P. Cattiaux, A. Guillin and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields* **140** (1–2) (2008) 19–40. MR2357669 https://doi.org/10.1007/s00440-007-0056-3
[13] L.-P. Chaintron and A. Diez. Propagation of chaos: A review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models* **15** (6) (2022) 895. MR4489768 https://doi.org/10.3934/krm.2022017
[14] L.-P. Chaintron and A. Diez. Propagation of chaos: A review of models, methods and applications. II. Applications. *Kinet. Relat. Models* **15** (6) (2022) 1017. MR4489769 https://doi.org/10.3934/krm.2022018
[15] X. Cheng, N. S. Chatterji, P. L. Bartlett and M. I. Jordan. Underdamped Langevin MCMC: A non-asymptotic analysis, 2017. Preprint. Available at arXiv:1707.03663v7.
[16] A. S. Dalalyan and L. Riou-Durand. On sampling from a log-concave density using kinetic Langevin diffusions. *Bernoulli* **26** (3) (2020) 1956–1988. MR4091098 https://doi.org/10.3150/19-BEJ1178
[17] J. Dolbeault, C. Mouhot and C. Schmeiser. Hypocoercivity for kinetic equations with linear relaxation terms. *C. R. Math. Acad. Sci. Paris* **347** (9–10) (2009) 511–516. MR2576899 https://doi.org/10.1016/j.crma.2009.02.025
[18] J. Dolbeault, C. Mouhot and C. Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Amer. Math. Soc.* **367** (6) (2015) 3807–3828. MR3324910 https://doi.org/10.1090/S0002-9947-2015-06012-7
[19] M. H. Duong and J. Tugaut. Stationary solutions of the Vlasov–Fokker–Planck equation: Existence, characterization and phase-transition. *Appl. Math. Lett.* **52** (2016) 38–45. MR3416384 https://doi.org/10.1016/j.aml.2015.08.003
[20] M. H. Duong and J. Tugaut. The Vlasov–Fokker–Planck equation in non-convex landscapes: Convergence to equilibrium. *Electron. Commun. Probab.* **23** (2018) 19. MR3779816 https://doi.org/10.1214/18-ECP116
[21] A. Durmus, A. Eberle, A. Guillin and K. Schuh. Sticky nonlinear SDEs and convergence of McKean–Vlasov equations without confinement, 2022. arXiv preprint. Available at arXiv:2201.07652.
[22] A. Durmus, A. Eberle, A. Guillin and R. Zimmer. An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.* **148** (12) (2020) 5387–5398. MR4163850 https://doi.org/10.1090/proc/14612
[23] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields* **166** (3–4) (2016) 851–886. MR3568041 https://doi.org/10.1007/s00440-015-0673-1
[24] A. Eberle, A. Guillin and R. Zimmer. Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47** (4) (2019) 1982–2010. MR3980913 https://doi.org/10.1214/18-AOP1299
[25] A. Guillin, P. Le Bris and P. Monmarché. Convergence rates for the Vlasov–Fokker–Planck equation and uniform in time propagation of chaos in non convex cases. *Electron. J. Probab.* **27** (2022). MR4489825 https://doi.org/10.1214/22-ejp853
[26] A. Guillin, W. Liu, L. Wu and C. Zhang. The kinetic Fokker–Planck equation with mean field interaction. *J. Math. Pures Appl.* **9** (150) (2021) 1–23. MR4248461 https://doi.org/10.1016/j.matpur.2021.04.001
[27] G. Arnaud and P. Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *J. Stat. Phys.* **185** (2) (2021) 15. MR4333408 https://doi.org/10.1007/s10955-021-02839-6
[28] B. Helffer and F. Nier. *Hypoelliptic Estimates and Spectral Theory for Fokker–Planck Operators and Witten Laplacians*. *Lecture Notes in Mathematics* **1862**. Springer-Verlag, Berlin, 2005. MR2130405 https://doi.org/10.1007/b104762
[29] F. Hérau. Short and long time behavior of the Fokker–Planck equation in a confining potential and applications. *J. Funct. Anal.* **244** (1) (2007) 95–118. MR2294477 https://doi.org/10.1016/j.jfa.2006.11.013
[30] F. Hérau and F. Nier. Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.* **171** (2) (2004) 151–218. MR2034753 https://doi.org/10.1007/s00205-003-0276-3
[31] K. Hu, Z. Ren, D. Šiška and Ł. Szpruch. Mean-field Langevin dynamics and energy landscape of neural networks. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** (4) (2021) 2043–2065. MR4328560 https://doi.org/10.1214/20-aihp1140
[32] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. III* 171–197. University of California Press, Berkeley-Los Angeles, Calif, 1956. MR0084985
[33] A. Kazeykina, Z. Ren, X. Tan and J. Yang Ergodicity of the underdamped mean-field Langevin dynamics. Preprint, 2020. Available at arXiv:2007.14660v2.
[34] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.* **95** (1) (2001) 109–132. MR1847094 https://doi.org/10.1016/S0304-4149(01)00095-3
[35] J. C. Mattingly, A. M. Stuart and D. J. Higham. Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.* **101** (2) (2002) 185–232. MR1931266 https://doi.org/10.1016/S0304-4149(02)00150-3
[36] H. P. McKean Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** (1966) 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907
[37] S. Mei, A. Montanari and P.-M. Nguyen. A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA* **115** (33) (2018) E7665–E7671. MR3845070 https://doi.org/10.1073/pnas.1806579115
[38] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations* 42–95. *Montecatini Terme*, *1995*. *Lecture Notes in Math.* **1627**. Springer, Berlin, 1996. MR1431299 https://doi.org/10.1007/BFb0093177
[39] P. Monmarché. Long-time behaviour and propagation of chaos for mean field kinetic particles. *Stochastic Process. Appl.* **127** (6) (2017) 1721–1737. MR3646428 https://doi.org/10.1016/j.spa.2016.10.003
[40] Y. Nesterov. *Lectures on Convex Optimization*, 2nd edition. *Springer Optimization and Its Applications.* **137**. Springer, Cham, 2018. MR3839649 https://doi.org/10.1007/978-3-319-91578-4
[41] G. A. Pavliotis. *Stochastic Processes and Applications: Diffusion Processes, the Fokker–Planck and Langevin Equations*. *Texts in Applied Mathematics* **60**. Springer, New York, 2014. MR3288096 https://doi.org/10.1007/978-1-4939-1323-7

[42] G. M. Rotskoff and E. Vanden-Eijnden. Trainability and accuracy of artificial neural networks: An interacting particle system approach. *Comm. Pure Appl. Math.* **75** (9) (2022) 1889–1935. MR4465905
[43] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989* 165–251. *Lecture Notes in Math.* **1464**. Springer, Berlin, 1991. MR1108185 https://doi.org/10.1007/BFb0085169
[44] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields* **8** (2) (2002) 163–198. Inhomogeneous random systems (Cergy-Pontoise, 2001). MR1924934
[45] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.* **202** (2009) 950. MR2562709 https://doi.org/10.1090/S0065-9266-09-00567-5
[46] L. Wu. Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. *Stochastic Process. Appl.* **91** (2) (2001) 205–238. MR1807683 https://doi.org/10.1016/S0304-4149(00)00061-2

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 790–823
https://doi.org/10.1214/22-AIHP1359

Fluctuations for mean field limits of interacting systems of spiking neurons

Eva Löcherbach[a]

SAMM, Statistique, Analyse et Modélisation Multidisciplinaire, Université Paris 1 Panthéon-Sorbonne, EA 4543 et FR FP2M 2036 CNRS, France,
[a]*eva.locherbach@univ-paris1.fr*

Abstract. We consider a system of N neurons, each spiking randomly with rate depending on its membrane potential. When a neuron spikes, its potential is reset to 0 and all other neurons receive an additional amount h/N of potential, where $h > 0$ is some fixed parameter. In between successive spikes, each neuron's potential follows a deterministic flow with drift b expressing both the attraction to an equilibrium potential and some leakage factors. While the propagation of chaos of the system, as $N \to \infty$, to a limit nonlinear jumping stochastic differential equation has already been established in a series of papers, see (*J. Stat. Phys.* **158** (2015) 866–902, *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 1844–1876, *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 343–378), the present paper is devoted to the associated central limit theorem. More precisely we study the measure valued process of fluctuations at scale $N^{-1/2}$ of the empirical measures of the membrane potentials, centered around the associated limit. We show that this fluctuation process, interpreted as càdlàg process taking values in a suitable weighted Sobolev space, converges in law to a limit process characterized by a system of stochastic differential equations driven by Gaussian white noise. We complete this picture by studying the fluctuations, at scale $N^{-1/2}$, of the membrane potential processes around their associated limit quantities, giving rise to a mesoscopic approximation of the membrane potentials that take into account the correlations within the finite system.

Résumé. Nous considérons un système de N neurones. Chaque neurone décharge un potentiel d'action à des instants aléatoires, à un taux qui dépend de son potentiel de membrane. Ce potentiel est alors remis à 0, et tous les autres neurones reçoivent une charge supplémentaire de h/N, où $h > 0$ est un paramètre fixé. Entre deux décharges successives, le potentiel de membrane de chaque neurone est attiré vers un potentiel d'équilibre et évolue selon un flot déterministe avec dérive b. Alors que la propriété de propagation du chaos du système, lorsque $N \to \infty$, vers la solution d'une équation différentielle stochastique non-linéaire à sauts a déjà été établie dans une série d'articles, voir (*J. Stat. Phys.* **158** (2015) 866–902, *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 1844–1876, *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 343–378), cet article est consacré à l'étude des fluctuations associées. Plus précisément, nous étudions le processus à valeurs mesures des fluctuations des mesures empiriques, centrées autour de leur limite, et renormalisées par $N^{-1/2}$. Nous montrons que ce processus, interprété comme processus càdlàg à valeurs dans un espace de Sobolev convenable, converge en loi vers un processus limite qui est caractérisé par un système d'équations différentielles stochastiques, dirigées par un bruit blanc gaussien. Nous complétons ce résultat par l'étude des fluctuations, à l'échelle $N^{-1/2}$, des processus de potentiel de membrane, autour de leurs limites respectives. Nous obtenons ainsi une approximation mésoscopique des processus potentiel de membrane qui tient compte des corrélations présentes dans le système fini.

MSC2020 subject classifications: 60G55; 60F05; 60G57; 92B20

Keywords: Convergence of fluctuations; Weighted Sobolev spaces; Systems of interacting neurons; Piecewise deterministic Markov processes; Mean field interactions

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev Spaces*, 2nd edition. *Pure and Applied Mathematics* **140**. Elsevier/Academic Press, Amsterdam, 2003. MR2424078
[2] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, New York, London, Sydney, Toronto, 1968. MR0233396
[3] J. Chevallier. Fluctuations for mean-field interacting age-dependent Hawkes processes. *Electron. J. Probab.* **22** (2017) 1–49. MR3646068 https://doi.org/10.1214/17-EJP63
[4] J. Chevallier, A. Duarte, E. Löcherbach and G. Ost. Mean field limits for nonlinear spatially extended Hawkes processes with exponential memory kernels. *Stochastic Process. Appl.* **129** (2019) 1–27. MR3906989 https://doi.org/10.1016/j.spa.2018.02.007
[5] Q. Cormier, E. Tanré and R. Veltz. Long time behavior of a mean-field model of interacting neurons. *Stochastic Process. Appl.* **130** (2020) 2553–2595. MR4080722 https://doi.org/10.1016/j.spa.2019.07.010

[6] Q. Cormier, E. Tanré and R. Veltz. Hopf bifurcation in a mean-field model of spiking neurons. *Electron. J. Probab.* **26** (2021) 1–40. MR4316639 https://doi.org/10.1214/21-ejp688

[7] A. De Masi, A. Galves, E. Löcherbach and E. Presutti. Hydrodynamic limit for interacting neurons. *J. Stat. Phys.* **158** (2015) 866–902. MR3311484 https://doi.org/10.1007/s10955-014-1145-1

[8] A. Duarte, A. A. Rodríguez and G. Ost. Hydrodynamic limit for spatially structured interacting neurons. *J. Stat. Phys.* **161** (2015) 1163–1202. MR3422922 https://doi.org/10.1007/s10955-015-1366-y

[9] X. Erny, E. Löcherbach and D. Loukianova. Conditional propagation of chaos for mean field systems of interacting neurons. *Electron. J. Probab.* **26** (2021) 1–25. MR4235471 https://doi.org/10.1214/21-EJP580

[10] X. Erny, E. Löcherbach and D. Loukianova. Strong error bounds for the convergence to its mean field limit for systems of interacting neurons in a diffusive scaling. *Ann. Appl. Probab.* To appear, 2023.

[11] R. Ferland, X. Fernique and G. Giroux. Compactness of the fluctuations associated with some generalized nonlinear Boltzmann equations. *Canad. J. Math.* **44** (1992) 1192–1205. MR1192413 https://doi.org/10.4153/CJM-1992-071-1

[12] B. Fernandez and S. Méléard. A Hilbertian approach for fluctuations on the McKean–Vlasov model. *Stochastic Process. Appl.* **71** (1997) 33–53. MR1480638 https://doi.org/10.1016/S0304-4149(97)00067-7

[13] N. Fournier and E. Löcherbach. On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.* **52** (2016) 1844–1876. MR3573298 https://doi.org/10.1214/15-AIHP701

[14] J. Jacod. Multivariate point processes: Predictable projection, Radon-Nikodym derivates, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **3** (1975) 235–253. MR0380978 https://doi.org/10.1007/BF00536010

[15] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*, 2nd edition. Springer-Verlag, Berlin, 2003. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

[16] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. Appl. Probab.* **18** (1986) 20–65. MR0827331 https://doi.org/10.2307/1427238

[17] E. Löcherbach and P. Monmarché. Metastability for systems of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (2022) 343–378. MR4374679 https://doi.org/10.1214/21-aihp1164

[18] P. Robert and J. Touboul. On the dynamics of random neuronal networks. *J. Stat. Phys.* **165** (2016) 545–584. MR3562424 https://doi.org/10.1007/s10955-016-1622-9

[19] V. Schmutz, E. Löcherbach and T. Schwalger. On a finite-size neuronal population equation. Preprint, 2022. Available at arXiv:2106.14721.

[20] T. Schwalger, M. Deger and W. Gerstner. Towards a theory of cortical columns: From spiking neurons to interacting neural populations of finite size. *PLoS Comput. Biol.* **13** (4) (2017) e1005507.

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

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 824–846
https://doi.org/10.1214/23-AIHP1369

Empirical optimal transport between different measures adapts to lower complexity

Shayan Hundrieser[1,2,a], Thomas Staudt[1,2,b] and Axel Munk[1,2,3,c]

[1]*Institute for Mathematical Stochastics, University of Göttingen, Göttingen, Germany,* [a]*s.hundrieser@math.uni-goettingen.de,* [b]*thomas.staudt@uni-goettingen.de,* [c]*munk@math.uni-goettingen.de*
[2]*Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), University of Göttingen, Göttingen, Germany*
[3]*Max Planck Institute for Multidisciplinary Sciences, University of Göttingen, Göttingen, Germany*

Abstract. The empirical optimal transport (OT) cost between two probability measures from random data is a fundamental quantity in transport based data analysis. In this work, we derive novel guarantees for its convergence rate when the involved measures are *different*, possibly supported on different spaces. Our central observation is that the statistical performance of the empirical OT cost is determined by the *less* complex measure, a phenomenon we refer to as *lower complexity adaptation* of empirical OT. For instance, under Lipschitz ground costs, we find that the expected error between the empirical OT cost based on n observations and the population quantity decreases with rate $n^{-1/d}$ if one of the two measures is concentrated on a d-dimensional manifold, while the other can be arbitrary. For semi-concave ground costs, we show that the upper bound for the rate improves to $n^{-2/d}$. Similarly, our theory establishes the general convergence rate $n^{-1/2}$ for semi-discrete OT. All of these results are valid in the two-sample case as well. Our findings therefore suggest that the curse of dimensionality only affects the estimation of the OT cost when *both* measures exhibit a high intrinsic dimension. Our proofs are based on the dual formulation of OT as a maximization over a suitable function class $\mathcal{F}_c$ and the observation that the c-transform of $\mathcal{F}_c$ under bounded costs has the same uniform metric entropy as $\mathcal{F}_c$ itself.

Résumé. Le coût empirique de transport optimal (OT) entre deux mesures de probabilité issues de données aléatoires est une quantité fondamentale pour l'analyse de données basée sur la théorie du transport optimal. Dans ce travail, nous dérivons de nouvelles garanties pour le taux de convergence de cette quantité quand les mesures en jeu sont *différentes* et potentiellement supportées sur des espaces différents. Notre observation centrale est que la performance statistique du coût de transport empirique est déterminée par la mesure dont la complexité est la plus faible, un phénomène que nous nommons l'*adaptivité à la complexité la plus faible* du coût de transport empirique. Par exemple, dans le cas d'une fonction de coût lipschitzienne, nous trouvons que l'espérance de l'erreur entre le coût de transport optimal empirique basé sur n observations et son équivalent dans la population décroît à un taux $n^{-1/d}$ si une des deux mesures est concentrée sur une variété de dimension d, l'autre mesure pouvant être arbitraire. Pour des fonctions de coût semi-concaves, nous montrons que la borne supérieure pour le taux est meilleure et d'ordre $n^{-2/d}$. De manière similaire, notre théorie montre que le taux $n^{-1/2}$ est atteint pour le transport semi-discret. Tous les résultats s'appliquent aussi au cas de deux échantillons. Nos résultats suggèrent que le fléau de la dimension n'affecte l'estimation du coût de transport optimal que quand *les deux* mesures ont une dimension intrinsèque élevée. Nos preuves se basent sur la formulation duale du problème de transport optimal, une maximisation sur une classe de fonctions $\mathcal{F}_c$, et l'observation que la c-transformée de $\mathcal{F}_c$, dans le cas de fonctions de coûts bornées, a la même entropie métrique uniforme que $\mathcal{F}_c$ elle-même.

MSC2020 subject classifications: Primary 62R07; 62G20; 62G30; 49Q22; secondary 62E20; 62F35; 60B10

Keywords: Wasserstein distance; Convergence rate; Curse of dimensionality; Lower complexity adaptation; Metric entropy; Semi-discrete; Manifolds

References

[1] M. Ajtai, J. Komlós and G. Tusnády. On optimal matchings. *Combinatorica* **4** (1984) 259–264. MR0779885 https://doi.org/10.1007/BF02579135
[2] J. Altschuler, J. Niles-Weed and P. Rigollet. Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. In *Advances in Neural Information Processing Systems*, I. Guyon, U. von Luxburg et al(Eds) **30**. Curran Associates, Red Hook, 2017.
[3] M. Arjovsky, S. Chintala and L. Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning 214–223. Proceedings of Machine Learning Research*, 2017.

[4] F. Aurenhammer, F. Hoffmann and B. Aronov. Minkowski-type theorems and least-squares clustering. *Algorithmica* **20** (1998) 61–76. MR1483422 https://doi.org/10.1007/PL00009187
[5] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization. Athena Scientific Series in Optimization and Neural Computation*. Athena Scientific, 1997.
[6] P. J. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** (1981) 1196–1217. MR0630103 https://doi.org/10.1214/aos/1176345637
[7] E. Boissard and T. Le Gouic. On the mean speed of convergence of empirical and occupation measures in Wasserstein distance. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** (2014) 539–563. MR3189084 https://doi.org/10.1214/12-AIHP517
[8] N. Bonneel, M. van de Panne, S. Paris and W. Heidrich. Displacement interpolation using Lagrangian mass transport. *ACM Trans. Graph.* **2011** (2011) 30.
[9] E. M. Bronshtein. ε-entropy of convex sets and functions. *Sib. Math. J.* **17** (1976) 393–398. MR0415155
[10] V. Chernozhukov, A. Galichon, M. Hallin and M. Henry. Monge–Kantorovich depth, quantiles, ranks and signs. *Ann. Statist.* **45** (2017) 223–256. MR3611491 https://doi.org/10.1214/16-AOS1450
[11] L. Chizat, P. Roussillon, F. Léger, F.-X. Vialard and G. Peyré. Faster Wasserstein distance estimation with the Sinkhorn divergence. In *Advances in Neural Information Processing Systems* 2257–2269. H. Larochelle, M. Ranzato et al(Eds) **33**. Curran Associates, Red Hook, 2020.
[12] N. Deb, P. Ghosal and B. Sen. Rates of estimation of optimal transport maps using plug-in estimators via barycentric projections. In *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer et al(Eds) **34**. Curran Associates, Red Hook, 2021.
[13] N. Deb and B. Sen. Multivariate rank-based distribution-free nonparametric testing using measure transportation. *J. Amer. Statist. Assoc.* **118** (2023) 192–207. MR4571116 https://doi.org/10.1080/01621459.2021.1923508
[14] E. del Barrio and J. A. Cuesta-Albertos. Tests of goodness of fit based on the L_2-Wasserstein distance. *Ann. Statist.* **27** (1999) 1230–1239. MR1740113 https://doi.org/10.1214/aos/1017938923
[15] E. del Barrio, A. González-Sanz and J.-M. Loubes Central limit theorems for semidiscrete Wasserstein distances, 2022. Preprint. Available at arXiv:2202.06380.
[16] S. Dereich, M. Scheutzow and R. Schottstedt. Constructive quantization: Approximation by empirical measures. *Ann. Inst. Henri Poincaré Probab. Stat.* **49** (2013) 1183–1203. MR3127919 https://doi.org/10.1214/12-AIHP489
[17] V. Divol. Measure estimation on manifolds: An optimal transport approach. *Probab. Theory Related Fields* **183** (2022) 581–647. MR4421180 https://doi.org/10.1007/s00440-022-01118-z
[18] V. Dobrić and J. E. Yukich. Asymptotics for transportation cost in high dimensions. *J. Theor. Probab.* **8** (1995) 97–118. MR1308672 https://doi.org/10.1007/BF02213456
[19] F. Dragomirescu and C. Ivan. The smallest convex extensions of a convex function. *Optimization* **24** (1992) 193–206. MR1247630 https://doi.org/10.1080/02331939208843789
[20] R. M. Dudley. The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Stat.* **40** (1969) 40–50. MR0236977 https://doi.org/10.1214/aoms/1177697802
[21] P. Dvurechensky, A. Gasnikov and A. Kroshnin. Computational optimal transport: Complexity by accelerated gradient descent is better than by Sinkhorn's algorithm. In *Proceedings of the 35th International Conference on Machine Learning* 1367–1376. J. Dy and A. Krause (Eds) *Proceedings of Machine Learning Research* **80**. 2018.
[22] S. N. Evans and F. A. Matsen. The phylogenetic Kantorovich–Rubinstein metric for environmental sequence samples. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** (2012) 569–592. MR2925374 https://doi.org/10.1111/j.1467-9868.2011.01018.x
[23] A. Forrow, J. Hütter, M. Nitzan and P. Rigollet. Statistical optimal transport via factored couplings. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics* 2454–2465. K. Chaudhuri and M. Sugiyama (Eds) *Proceedings of Machine Learning Research* **89**. PMLR, 2019.
[24] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** (2015) 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7
[25] A. Galichon. *Optimal Transport Methods in Economics*. Princeton University Press, Princeton, 2018. MR3586373 https://doi.org/10.1515/9781400883592
[26] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.* **177** (1996) 113–161. MR1440931 https://doi.org/10.1007/BF02392620
[27] D. Geiß, R. Klein, R. Penninger and G. Rote. Optimally solving a transportation problem using Voronoi diagrams. *Comput. Geom.* **46** (2013) 1009–1016. MR3061462 https://doi.org/10.1016/j.comgeo.2013.05.005
[28] A. Guntuboyina and B. Sen. Covering numbers for convex functions. *IEEE Trans. Inf. Theory* **59** (2013) 1957–1965. MR3043776 https://doi.org/10.1109/TIT.2012.2235172
[29] M. Hallin, E. del Barrio, J. Cuesta-Albertos and C. Matrán. Distribution and quantile functions, ranks and signs in dimension d: A measure transportation approach. *Ann. Statist.* **49** (2021) 1139–1165. MR4255122 https://doi.org/10.1214/20-aos1996
[30] M. Hallin, D. Hlubinka and Š. Hudecová. Efficient fully distribution-free center-outward rank tests for multiple-output regression and MANOVA. *J. Amer. Statist. Assoc.* (2022). In press, preprint available online.
[31] M. Hallin and G. Mordant Center-Outward Multiple-Output Lorenz Curves and Gini Indices a measure transportation approach, 2022. Preprint. Available at arXiv:2211.10822.
[32] M. Hallin, G. Mordant and J. Segers. Multivariate goodness-of-fit tests based on Wasserstein distance. *Electron. J. Stat.* **15** (2021) 1328–1371. MR4255302 https://doi.org/10.1214/21-ejs1816
[33] V. Hartmann and D. Schuhmacher. Semi-discrete optimal transport: A solution procedure for the unsquared Euclidean distance case. *Math. Methods Oper. Res.* **92** (2020) 133–163. MR4152920 https://doi.org/10.1007/s00186-020-00703-z
[34] F. Heinemann, A. Munk and Y. Zemel. Randomized Wasserstein barycenter computation: Resampling with statistical guarantees. *SIAM J. Math. Data Sci.* **4** (2022) 229–259. MR4386483 https://doi.org/10.1137/20M1385263
[35] S. Hundrieser, M. Klatt, T. Staudt and A. Munk A unifying approach to distributional limits for empirical optimal transport, 2022. Preprint. Available at arXiv:2202.12790.
[36] S. Hundrieser, G. Mordant, C. A. Weitkamp and A. Munk Empirical optimal transport under estimated costs: Distributional limits and statistical applications, 2023. Preprint. Available at arXiv:2301.01287.

[37] L. Kantorovich. On the translocation of masses. *Manage. Sci.* **5** (1958) 1–4. MR0096552 https://doi.org/10.1287/mnsc.5.1.1
[38] L. V. Kantorovich. On the translocation of masses. *Dokl. Akad. Nauk SSSR* **37** (1942) 7–8. MR0009619
[39] A. N. Kolmogorov and V. M. Tikhomirov. ε-Entropy and ε-capacity of sets in functional spaces. In *Twelve Papers on Algebra and Real Functions* 277–364. S. N. Cernikov, N. V. Cernikova, A. N. Kolmogorov, A. I. Mal'cev and B. I. Plotkin (Eds) *American Mathematical Society Translations–Series* **2**. Am. Math. Soc., Providence, 1961. MR0124720
[40] M. Ledoux. On optimal matching of Gaussian samples. *J. Math. Sci.* **238** (2019) 495–522. MR3723584 https://doi.org/10.1007/s10958-019-04253-6
[41] J. M. Lee. *Introduction to Smooth Manifolds*. *Graduate Texts in Mathematics* **218**. Springer, Berlin, 2013. MR2954043
[42] T. Liang On the minimax optimality of estimating the Wasserstein metric, 2019. Preprint. Available at arXiv:1908.10324.
[43] D. G. Luenberger. *Linear and Nonlinear Programming*, 2nd edition. Springer, New York, 2003. MR2012832
[44] C. L. Mallows. A note on asymptotic joint normality. *Ann. Math. Stat.* **43** (1972) 508–515. MR0298812 https://doi.org/10.1214/aoms/1177692631
[45] T. Manole, S. Balakrishnan, J. Niles-Weed and L. Wasserman Plugin estimation of smooth optimal transport maps, 2021. Preprint. Available at arXiv:2107.12364.
[46] T. Manole and J. Niles-Weed Sharp convergence rates for empirical optimal transport with smooth costs, 2021. Preprint. Available at arXiv:2106.13181v2.
[47] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*. *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. MR1333890 https://doi.org/10.1017/CBO9780511623813
[48] E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.* **40** (1934) 837–842. MR1562984 https://doi.org/10.1090/S0002-9904-1934-05978-0
[49] Q. Mérigot. A multiscale approach to optimal transport. In *Computer Graphics Forum* 1583–1592, **30**. Wiley, New York, 2011.
[50] G. Monge. Mémoire sur la théorie des déblais et des remblais. In *Histoire de l'Académie Royale des Sciences de Paris* 666–704, 1781.
[51] G. Mordant and J. Segers. Measuring dependence between random vectors via optimal transport. *J. Multivariate Anal.* **189** (2022), 104912. MR4349898 https://doi.org/10.1016/j.jmva.2021.104912
[52] A. Munk and C. Czado. Nonparametric validation of similar distributions and assessment of goodness of fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** (1998) 223–241. MR1625620 https://doi.org/10.1111/1467-9868.00121
[53] B. Muzellec, A. Vacher, F. Bach, F.-X. Vialard and A. Rudi Near-optimal estimation of smooth transport maps with kernel sums-of-squares, 2021. Preprint. Available at arXiv:2112.01907.
[54] T. G. Nies, T. Staudt and A. Munk Transport dependency: Optimal transport based dependency measures, 2021. Preprint. Available at arXiv:2105.02073.
[55] J. Niles-Weed and Q. Berthet. Minimax estimation of smooth densities in Wasserstein distance. *Ann. Statist.* **50** (2022) 1519–1540. MR4441130 https://doi.org/10.1214/21-aos2161
[56] J. Niles-Weed and P. Rigollet. Estimation of Wasserstein distances in the spiked transport model. *Bernoulli* **28** (2022) 2663–2688. MR4474558 https://doi.org/10.3150/21-bej1433
[57] V. M. Panaretos and Y. Zemel. Statistical aspects of Wasserstein distances. *Annu. Rev. Stat. Appl.* **6** (2019) 405–431. MR3939527 https://doi.org/10.1146/annurev-statistics-030718-104938
[58] G. Peyré and M. Cuturi. Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** (2019) 355–607.
[59] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume I: Theory. Probability and Its Applications*. Springer, Berlin, 1998. MR1619171
[60] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume II: Applications. Probability and Its Applications*. Springer, Berlin, 1998. MR1619171
[61] F. Santambrogio. *Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling*. *Progress in Nonlinear Differential Equations and Their Applications.* Springer, Berlin, 2015. MR3409718 https://doi.org/10.1007/978-3-319-20828-2
[62] G. Schiebinger, J. Shu, M. Tabaka, B. Cleary, V. Subramanian, A. Solomon, J. Gould, S. Liu, S. Lin, P. Berube, L. Lee, J. Chen, J. Brumbaugh, P. Rigollet, K. Hochedlinger, R. Jaenisch, A. Regev and E. S. Lander. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* **176** (2019) 928–943.e22. https://doi.org/10.1016/j.cell.2019.01.006
[63] G. R. Shorack and J. A. Wellner. *Empirical Processes with Applications to Statistics*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York, 1986. MR0838963
[64] S. Singh and B. Póczos Minimax distribution estimation in Wasserstein distance, 2018. Preprint. Available at arXiv:1802.08855.
[65] M. Sommerfeld and A. Munk. Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** (2018) 219–238. MR3744719 https://doi.org/10.1111/rssb.12236
[66] M. Sommerfeld, J. Schrieber and Y. Zemel. Optimal transport: Fast probabilistic approximation with exact solvers. *J. Mach. Learn. Res.* **20** (2019) 1–23. MR3990459
[67] B. K. Sriperumbudur, K. Fukumizu, A. Gretton and B. Schölkopf. On the empirical estimation of integral probability metrics. *Electron. J. Stat.* **6** (2012) 1550–1599. MR2988458 https://doi.org/10.1214/12-EJS722
[68] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series* **30**. Princeton University Press, Princeton, 1971. MR0290095
[69] M. Talagrand. Matching theorems and empirical discrepancy computations using majorizing measures. *J. Amer. Math. Soc.* **7** (1994) 455–537. MR1227476 https://doi.org/10.2307/2152764
[70] M. Talagrand. *Upper and Lower Bounds for Stochastic Processes Modern Methods and Classical Problems*. *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge/A Series of Modern Surveys in Mathematics*. Springer, Berlin, 2014. MR3184689 https://doi.org/10.1007/978-3-642-54075-2
[71] A. Talwalkar, S. Kumar and H. Rowley. Large-scale manifold learning. In *2008 IEEE Conference on Computer Vision and Pattern Recognition 1–8*. IEEE, Los Alamitos, 2008.
[72] C. Tameling, S. Stoldt, T. Stephan and J. Naas. Colocalization for super-resolution microscopy via optimal transport. *Nat. Comput. Sci.* **1** (2021) 199–211.
[73] A. Vacher, B. Muzellec and A. Rudi. A dimension-free computational upper-bound for smooth optimal transport estimation. In *Proceedings of Thirty Fourth Conference on Learning Theory (Mikhail Belkin and Samory Kpotufe* 4143–4173. *Proceedings of Machine Learning Research* **134**. 2021.

[74] C. Villani. *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Am. Math. Soc., Providence, 2003. MR1964483 https://doi.org/10.1090/gsm/058

[75] C. Villani. *Optimal Transport: Old and New*. *A Series of Comprehensive Studies in Mathematics* **338**. Springer, Berlin, 2008.

[76] U. von Luxburg and O. Bousquet. Distance-based classification with Lipschitz functions. *J. Mach. Learn. Res.* **5** (2004) 669–695. MR2247996

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

[78] S. Wang, T. T. Cai and H. Li. Optimal estimation of Wasserstein distance on a tree with an application to microbiome studies. *J. Amer. Statist. Assoc.* **116** (2021) 1237–1253. MR4309269 https://doi.org/10.1080/01621459.2019.1699422

[79] J. Weed and F. Bach. Sharp asymptotic and finite-sample rates of convergence of empirical measures in Wasserstein distance. *Bernoulli* **25** (2019) 2620–2648. MR4003560 https://doi.org/10.3150/18-BEJ1065

[80] C. A. Weitkamp, K. Proksch, C. Tameling and A. Munk. Distribution of distances based object matching: Asymptotic inference. *J. Amer. Statist. Assoc.* (2022). In press, preprint available online.

[81] H. Whitney. Analytic extensions of differentiable functions defined in closed sets. *Trans. Amer. Math. Soc.* **36** (1934) 63–89. MR1501735 https://doi.org/10.2307/1989708

[82] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen and M. S. Rosen. Image reconstruction by domain-transform manifold learning. *Nature* **555** (2018) 487–492.

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 847–873
https://doi.org/10.1214/22-AIHP1356

Central limit theorems for general transportation costs

Eustasio del Barrio[1,a], Alberto González-Sanz[2,b] and Jean-Michel Loubes[2,c]

[1] *IMUVA, Universidad de Valladolid, Spain,* [a]*tasio@eio.uva.es*
[2] *IMT, Université de Toulouse, France,* [b]*alberto.gonzalez.sanz@uva.es,* [c]*loubes@math.univ-toulouse.fr*

Abstract. We consider the problem of optimal transportation with general cost between an empirical measure and a general target probability on $\mathbb{R}^d$, with $d \geq 1$. We provide results on asymptotic stability of optimal transport potentials under minimal regularity assumptions on the costs or the underlying probability. This stability is combined with a refined linearization technique based on the sequential compactness of the closed unit ball in $L^2(P)$ for the weak topology and the strong convergence of Cesàro means along subsequences. As a result we obtain a CLT for the transportation cost under sharp smoothness and moment assumptions, giving a positive answer to a conjecture in (*Ann. Probab.* **47** (2019) 926–951) for the quadratic costs.

Résumé. Nous considérons le problème du transport optimal avec coût général entre une mesure empirique et une probabilité cible générale sur $\mathbb{R}^d$, avec $d \geq 1$. Nous fournissons des résultats sur la stabilité asymptotique des potentiels de transport optimal sous des hypothèses minimales de régularité sur les coûts ou les probabilités sous-jacentes. Cette stabilité est combinée avec une technique de linéarisation affinée basée sur la compacité séquentielle de la boule fermée unité dans $L^2(P)$ pour la topologie faible et la convergence forte des moyennes de Cesàro le long de sous-suites. En conséquence, nous obtenons un Théorème Limite en distribution pour le coût de transport sous des hypothèses minimales, donnant une réponse positive à une conjecture dans (*Ann. Probab.* **47** (2019) 926–951) pour les coûts quadratiques.

MSC2020 subject classifications: 60F05; 62E20

Keywords: Optimal transport; Banach–Saks property; CLT; Efron–Stein's inequality; Cesàro means

References

[1] G. Acosta and R. G. Durán. An optimal Poincaré inequality in L^1 for convex domains: Theory and algorithms. *Proc. Amer. Math. Soc.* **132** (2004) 195–202. MR2021262 https://doi.org/10.1090/S0002-9939-03-07004-7
[2] M. Ajtai, J. Komlós and G. Tusnády. On optimal matchings. *Combinatorica* **4** (1984) 259–264. MR0779885 https://doi.org/10.1007/BF02579135
[3] L. Ambrosio, F. Stra and D. Trevisan. A PDE approach to a 2-dimensional matching problem. *Probab. Theory Related Fields* **173** (2019) 433–477. MR3916111 https://doi.org/10.1007/s00440-018-0837-x
[4] F. Bachoc, F. Gamboa, J. M. Loubes and N. Venet. A Gaussian process regression model for distribution inputs. *IEEE Trans. Inf. Theory* **64** (10) (2017) 6620–6637. MR3860751 https://doi.org/10.1109/TIT.2017.2762322
[5] P. Berthet, J. C. Fort and T. Klein. A central limit theorem for Wasserstein type distances between two distinct univariate distributions. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** (2) (2017) 954–982. MR4076772 https://doi.org/10.1214/19-AIHP990
[6] P. Billingsley. *Convergence of Probability Measures*. *Wiley Series in Probability and Statistics*. Wiley, New York, United States, 1999. MR1700749 https://doi.org/10.1002/9780470316962
[7] E. Black, S. Yeom and M. Fredrikson. FlipTest: Fairness testing via optimal transport. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* 111–121, 2020.
[8] S. Boucheron, G. Lugosi and P. Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford, 2013. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001
[9] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York, United States, 2011. MR2759829
[10] N. Courty, R. Flamary and M. Ducoffe. Learning Wasserstein Embeddings. *International Conference on Learning Representations (ICLR)* (2018).
[11] J. A. Cuesta-Albertos, C. Matrán and A. Tuero-Díaz. Optimal transportation plans and convergence in distribution. *J. Multivariate Anal.* **60** (1997) 72–83. MR1441460 https://doi.org/10.1006/jmva.1996.1627
[12] M. Cuturi and G. Peyré. Special issue on optimal transport in data sciences. *Inf. Inference* **8** (4) (2019) 655–656. MR4045124 https://doi.org/10.1093/imaiai/iaz032
[13] M. Cuturi and G. Peyré. Computational optimal transport: With applications to data science. *Found. Trends Mach. Learn.* **11** (5–6) (2019) 355–607.
[14] E. del Barrio, E. Giné and C. Matrán. Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** (1999) 1009–1071. MR1698999 https://doi.org/10.1214/aop/1022677394

[15] E. del Barrio, E. Giné and F. Utzet. Asymptotics for L^2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* **11** (2005) 131–189. MR2121458 https://doi.org/10.3150/bj/1110228245
[16] E. del Barrio, A. González Sanz and J. M. Loubes. Central limit theorems for semidiscrete wasserstein distances, 2022. Available at arXiv:2202.06380.
[17] E. del Barrio, A. González Sanz, J. M. Loubes and J. Niles-Weed. An improved central limit theorem and fast convergence rates for entropic transportation costs, 2022. Available at arXiv:2204.09105.
[18] E. del Barrio, P. Gordaliza and J. M. Loubes. A central limit theorem for Lp transportation cost on the real line with application to fairness assessment in machine learning. *Inf. Inference* **8** (4) (2019) 817–849. MR4045479 https://doi.org/10.1093/imaiai/iaz016
[19] E. del Barrio and J. M. Loubes. Central limit theorems for empirical transportation cost in general dimension. *Ann. Probab.* **47** (2019) 926–951. MR3916938 https://doi.org/10.1214/18-AOP1275
[20] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** (2015) 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7
[21] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.* **177** (2) (1996) 113–161. MR1440931 https://doi.org/10.1007/BF02392620
[22] J. González-Delgado, A. González-Sanz, J. Cort'es and P. Neuvial. Two-sample goodness-of-fit tests on the flat torus based on Wasserstein distance and their relevance to structural biology, 2021. Available at arXiv:2108.00165.
[23] P. Gordaliza, E. del Barrio, F. Gamboa and J. M. Loubes. Obtaining fairness using optimal transport theory. In *International Conference on Machine Learning* 2357–2365, 2019.
[24] S. Hundrieser, M. Klatt, T. Staudt and A. Munk. A unifying approach to distributional limits for empirical optimal transport, 2022. Available at arXiv:2202.12790.
[25] S. Hundrieser, T. Staudt and A. Munk. Empirical optimal transport between different measures adapts to lower complexity, 2022. Available at arXiv:2202.10434.
[26] M. Ledoux. On optimal matching of Gaussian samples. *J. Math. Sci.* **238** (2019) 495–522. MR3723584
[27] M. Ledoux and M. Talagrand. *Probability in Banach Spaces*. Springer, Berlin, Heidelberg, 1991. MR1102015 https://doi.org/10.1007/978-3-642-20212-4
[28] T. Manole, S. Balakrishnan, J. Niles-Weed and L. A. Wasserman. Plugin estimation of smooth optimal transport maps, 2021. Available at arXiv:2107.12364.
[29] T. Manole and J. Niles-Weed. Sharp convergence rates for empirical optimal transport with smooth costs, 2021. Available at arXiv:2106.13181.
[30] G. Mena and J. Niles-Weed. Statistical bounds for entropic optimal transport: Sample complexity and the central limit theorem. *Adv. Neural Inf. Process. Syst.* **32** (2019) 4541–4551.
[31] L. Risser, A. González Sanz, Q. Vincenot and J. M. Loubes. Tackling algorithmic bias in neural-network classifiers using Wasserstein-2 regularization. *J. Math. Imaging Vision* **64** (2022) 672–689. MR4458660 https://doi.org/10.1007/s10851-022-01090-2
[32] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.* **17** (3) (1966) 497–510. MR0193549
[33] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, New Jersey, United States, 1970. MR0274683
[34] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. Springer, Berlin, 2009.
[35] L. Rüschendorf. Optimal solutions of multivariate coupling problems. *Appl. Math.* **23** (3) (1995) 325–338. MR1360058 https://doi.org/10.4064/am-23-3-325-338
[36] L. Rüschendorf. On c-optimal random variables. *Statist. Probab. Lett.* **27** (3) (1996) 267–270. MR1395577 https://doi.org/10.1016/0167-7152(95)00078-X
[37] F. Santambrogio. *Optimal Transport for Applied Mathematicians*. Birkhauser, NY, 2015. MR3409718 https://doi.org/10.1007/978-3-319-20828-2
[38] G. Schiebinger et al. Optimal-transport analysis of single-cell gene expression identifies developmental trajectories in reprogramming. *Cell* **176** (4) (2019) 928–943.
[39] C. Smith and M. Knott. On Hoeffding-frechet bounds and cyclic monotone relations. *J. Multivariate Anal.* **4** (1992) 328–334. MR1150616 https://doi.org/10.1016/0047-259X(92)90029-F
[40] M. Sommerfeld and A. Munk. Inference for empirical Wasserstein distances on finite spaces. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80** (1) (2018) 219–238. MR3744719 https://doi.org/10.1111/rssb.12236
[41] T. Staudt, S. Hundrieser and A. Munk. On the uniqueness of kantorovich potentials, 2022. Available at arXiv:2201.08316.
[42] M. Talagrand. Matching random samples in many dimensions. *Ann. Appl. Probab.* **2** (1992) 846–856. MR1189420
[43] M. Talagrand. The transportation cost from the uniform measure to the empirical measure in dimension ≥ 3. *Ann. Probab.* **22** (1994) 919–959. MR1288137
[44] M. Talagrand. Scaling and non-standard matching theorems. *C. R., Math.* **356** (2018) 692–695. MR3806901 https://doi.org/10.1016/j.crma.2018.04.018
[45] M. Talagrand and J. E. Yukich. The integrability of the square exponential transportation cost. *Ann. Appl. Probab.* **3** (1993) 1100–1111. MR1241036
[46] C. Tameling, M. Sommerfeld and A. Munk. Empirical optimal transport on countable metric spaces: Distributional limits and statistical applications. *Ann. Appl. Probab.* **29** (5) (2019) 2744–2781. MR4019874 https://doi.org/10.1214/19-AAP1463
[47] C. Villani. *Optimal Transport: Old and New*. Springer, Berlin, Heidelberg, Germany, 2008.

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 874–903
https://doi.org/10.1214/23-AIHP1362

Concentration of quasi-stationary distributions for one-dimensional diffusions with applications

Zhongwei Shen[1,a], Shirou Wang[2,1,b] and Yingfei Yi[1,2,c]

[1]*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada,* [a]*zhongwei@ualberta.ca,* [b]*shirou@jlu.edu.cn,* [c]*yingfei@ualberta.ca*
[2]*School of Mathematics, Jilin University, Changchun 130012, PRC*

Abstract. We consider small noise perturbations to an ordinary differential equation (ODE) that have a uniform absorbing state and exhibit transient dynamics in the sense that interesting dynamical behaviors governed by transient states display over finite time intervals and the eventual dynamics is simply controlled by the absorbing state. To capture the transient states, we study the noise-vanishing concentration of the so-called quasi-stationary distributions (QSDs) that describe the dynamics before reaching the absorbing state. By establishing concentration estimates based on constructed uniform-in-noises Lyapunov functions, we show that QSDs tend to concentrate on the global attractor of the ODE as noises vanish, and that each limiting measure of QSDs, if exists, must be an invariant measure of the ODE. Overcoming difficulties caused by the degeneracy and singularity of noises at the absorbing state, we further show the tightness of the family of QSDs under additional assumptions motivated by applications, that not only validates a priori information on the concentration of QSDs, but also asserts the reasonability of using QSDs in the mathematical modeling of transient states. Our approaches to the concentration and tightness of QSDs are purely analytic without probabilistic heuristics. Applications to diffusion approximations of chemical reactions and birth-and-death processes of logistic type are also discussed. Rigorously studying the transient dynamics and characterizing the transient states, our study is of both theoretical and practical significance.

Résumé. Nous considérons de petites perturbations par le bruit d'une équation différentielle ordinaire (EDO) qui ont un état d'absorption uniforme et présentent une dynamique transiente dans le sens où des comportements dynamiques intéressants régis par des états transients apparaissent sur des intervalles de temps finis et la dynamique finale est simplement contrôlée par l'état d'absorption. Pour capturer les états transients, nous étudions la concentration en fonction du bruit des distributions dites quasi-stationnaires (QSD) qui décrivent la dynamique avant d'atteindre l'état absorbant. En établissant des estimations de concentration basées sur des fonctions de Lyapunov construites uniformément par rapport au bruit, nous montrons que les QSD ont tendance à se concentrer sur l'attracteur global de l'EDO lorsque le bruit disparaît, et que chaque mesure limite des QSD, si elle existe, doit être une mesure invariante de l'EDO. Surmontant les difficultés causées par la dégénérescence et la singularité des bruits à l'état absorbant, nous montrons en outre la tension de la famille des QSD sous des hypothèses supplémentaires motivées par des applications, ce qui non seulement vérifie les informations a priori sur la concentration des QSD, mais aussi confirme le bien-fondé de l'utilisation des QSD dans la modélisation mathématique des états transients. Nos approches de la concentration et de la tension des QSD sont purement analytiques, sans heuristique probabiliste. Les applications aux approximations de diffusion des réactions chimiques et processus de naissance et de mort de type logistique sont également discutées. En étudiant rigoureusement la dynamique transiente et en caractérisant les états transients, notre étude présente un intérêt à la fois théorique et pratique.

MSC2020 subject classifications: Primary 60J60; 60J70; 34F05; secondary 92D25; 60H10

Keywords: One-dimensional diffusion process; Quasi-stationary distribution; Concentration; Tightness; Transient dynamics; Transient state; Keizer's paradox; Birth-and-death process; Diffusion approximation

References

[1] D. F. Anderson and T. G. Kurtz. *Stochastic Analysis of Biochemical Systems*. *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems* **1**. Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015. MR3363610 https://doi.org/10.1007/978-3-319-16895-1

[2] P. Cattiaux, P. Collet, A. Lambert, S. Martinez, S. Meleard and J. San Martin. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.* **37** (5) (2009) 1926–1969. MR2561437 https://doi.org/10.1214/09-AOP451

[3] P. Cattiaux and S. Méléard. Competitive or weak cooperative stochastic Lotka–Volterra systems conditioned on non-extinction. *J. Math. Biol.* **60** (6) (2010) 797–829. MR2606515 https://doi.org/10.1007/s00285-009-0285-4
[4] J. A. Cavender. Quasi-stationary distributions of birth-and-death processes. *Adv. in Appl. Probab.* **10** (3) (1978) 570–586. MR0501388 https://doi.org/10.2307/1426635
[5] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probab. Theory Related Fields* **164** (1–2) (2016) 243–283. MR3449390 https://doi.org/10.1007/s00440-014-0611-7
[6] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.* **14** (1) (2017) 177–199. MR3622466
[7] N. Champagnat and D. Villemonais. Uniform convergence of conditional distributions for absorbed one-dimensional diffusions. *Adv. in Appl. Probab.* **50** (1) (2018) 178–203. MR3781982 https://doi.org/10.1017/apr.2018.9
[8] N. Champagnat and D. Villemonais. Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *Stochastic Process. Appl.* **135** (2021) 51–74. MR4222402 https://doi.org/10.1016/j.spa.2020.12.005
[9] J.-R. Chazottes, P. Collet and S. Méléard. Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes. *Probab. Theory Related Fields* **164** (1–2) (2016) 285–332. MR3449391 https://doi.org/10.1007/s00440-014-0612-6
[10] P. Childs and J. P. Keener. Slow manifold reduction of a stochastic chemical reaction: Exploring Keizer's paradox. *Discrete Contin. Dyn. Syst. Ser. B* **17** (6) (2012) 1775–1794. MR2924439 https://doi.org/10.3934/dcdsb.2012.17.1775
[11] P. Collet, S. Martínez and J. San Martín. Asymptotic laws for one-dimensional diffusions conditioned to nonabsorption. *Ann. Probab.* **23** (3) (1995) 1300–1314. MR1349173
[12] P. Collet, S. Martínez and J. San Martín. *Quasi-Stationary Distributions. Markov Chains, Diffusions and Dynamical Systems*. *Probability and Its Applications (New York)*. Springer, Heidelberg, 2013. MR2986807 https://doi.org/10.1007/978-3-642-33131-2
[13] M. V. Day. Recent progress on the small parameter exit problem. *Stochastics* **20** (2) (1987) 121–150. MR0877726 https://doi.org/10.1080/17442508708833440
[14] M. V. Day. Mathematical approaches to the problem of noise-induced exit. In *Stochastic Analysis, Control, Optimization and Applications* 269–287. *Systems Control Found. Appl.*, Birkhäuser Boston, Boston, MA, 1999.
[15] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, 2nd edition. *Applications of Mathematics (New York)* **38**. Springer-Verlag, New York, 1998. MR1619036 https://doi.org/10.1007/978-1-4612-5320-4
[16] A. Devinatz, R. Ellis and A. Friedman. The asymptotic behavior of the first real eigenvalue of second order elliptic operators with a small parameter in the highest derivatives. II. *Indiana Univ. Math. J.* **23** (1973/74) 991–1011. MR0344709 https://doi.org/10.1512/iumj.1974.23.23081
[17] G. Di Gesú, T. Leliévre, D. Le Peutrec and B. Nectoux. Sharp asymptotics of the first exit point density. *Ann. PDE* **5** (1) (2019) 5. MR3975562 https://doi.org/10.1007/s40818-019-0059-2
[18] G. Di Gesú, T. Leliévre, D. Le Peutrec and B. Nectoux. The exit from a metastable state: Concentration of the exit point distribution on the low energy saddle points, part 1. *J. Math. Pures Appl. (9)* **138** (2020) 242–306. MR4098769 https://doi.org/10.1016/j.matpur.2019.06.003
[19] S. N. Ethier and T. G. Kurtz. *Markov Processes. Characterization and Convergence*. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1986. MR0838085 https://doi.org/10.1002/9780470316658
[20] M. Faure and S. J. Schreiber. Quasi-stationary distributions for randomly perturbed dynamical systems. *Ann. Appl. Probab.* **24** (2) (2014) 553–598. MR3178491 https://doi.org/10.1214/13-AAP923
[21] W. Feller. The parabolic differential equations and the associated semi-groups of transformations. *Ann. of Math. (2)* **55** (1952) 468–519. MR0047886 https://doi.org/10.2307/1969644
[22] W. Feller. Diffusion processes in one dimension. *Trans. Amer. Math. Soc.* **77** (1954) 1–31. MR0063607 https://doi.org/10.2307/1990677
[23] P. A. Ferrari, H. Kesten, S. Martinez and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.* **23** (2) (1995) 501–521. MR1334159
[24] R. Fierro, S. Martínez and J. San Martín. Limiting conditional and conditional invariant distributions for the Poisson process with negative drift. *J. Appl. Probab.* **36** (4) (1999) 1194–1209. MR1742160 https://doi.org/10.1239/jap/1032374765
[25] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*, 3rd edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **260**. Springer, Heidelberg, 2012. Translated from the 1979 Russian original by Joseph Szücs. MR2953753 https://doi.org/10.1007/978-3-642-25847-3
[26] A. Friedman. The asymptotic behavior of the first real eigenvalue of a second order elliptic operator with a small parameter in the highest derivatives. *Indiana Univ. Math. J.* **22** (1972/73) 1005–1015. MR0320551 https://doi.org/10.1512/iumj.1973.22.22084
[27] M. Fukushima. *Dirichlet Forms and Markov Processes*. *North-Holland Mathematical Library* **23**. North-Holland Publishing Co., Amsterdam–New York; Kodansha, Ltd., Tokyo, 1980. MR0569058
[28] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. *Classics in Mathematics*. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition. MR1814364
[29] A. Hastings. Transients: The key to long-term ecological understanding? *Trends Ecol. Evol.* **19** (1) (2004) 39–45.
[30] A. Hastings, K. C. Abbott, K. Cuddington, T. Francis, G. Gellner, Y. C. Lai et al. Transient phenomena in ecology. *Science* **361** (6406) (2018), eaat6412.
[31] A. Hening and M. Kolb. Quasistationary distributions for one-dimensional diffusions with singular boundary points. *Stochastic Process. Appl.* **129** (5) (2019) 1659–1696. MR3944780 https://doi.org/10.1016/j.spa.2018.05.012
[32] A. Hening, W. Qi, Z. Shen and Y. Yi. Quasi-stationary distributions of multi-dimensional diffusion processes. Available at https://sites.ualberta.ca/~zhongwei/manuscript-Hening-Qi-Shen-Yi-QSD.pdf.
[33] G. Högnäs. On the quasi-stationary distribution of a stochastic Ricker model. *Stochastic Process. Appl.* **70** (2) (1997) 243–263. MR1475665 https://doi.org/10.1016/S0304-4149(97)00064-1
[34] W. Huang, M. Ji, Z. Liu and Y. Yi. Integral identity and measure estimates for stationary Fokker–Planck equations. *Ann. Probab.* **43** (4) (2015) 1712–1730. MR3353813 https://doi.org/10.1214/14-AOP917
[35] W. Huang, M. Ji, Z. Liu and Y. Yi. Concentration and limit behaviors of stationary measures. *Phys. D* **369** (2018) 1–17. MR3771195 https://doi.org/10.1016/j.physd.2017.12.009
[36] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*, 2nd edition. *North-Holland Mathematical Library* **24**. North-Holland Publishing Co., Amsterdam; Kodansha, Ltd., Tokyo, 1989. MR0637061

[37] F. Jacobs and S. J. Schreiber. Random perturbations of dynamical systems with absorbing states. *SIAM J. Appl. Dyn. Syst.* **5** (2) (2006) 293–312. MR2237149 https://doi.org/10.1137/050626417
[38] J. Keizer. Master equations, Langevin equations, and the effect of diffusion on concentration fluctuations. *J. Chem. Phys.* **67** (4) (1977) 1473–1476.
[39] J. Keizer. *Statistical Thermodynamics of Nonequilibrium Processes*. Springer-Verlag, New York, 1987. MR1082346
[40] R. Khasminskii. *Stochastic Stability of Differential Equations. With Contributions by G. N. Milstein and M. B. Nevelson Completely Revised and Enlarged* 2nd edition. *Stochastic Modelling and Applied Probability* **66**. Springer, Heidelberg, 2012. MR2894052 https://doi.org/10.1007/978-3-642-23280-0
[41] F. C. Klebaner, J. Lazar and O. Zeitouni. On the quasi-stationary distribution for some randomly perturbed transformations of an interval. *Ann. Appl. Probab.* **8** (1) (1998) 300–315. MR1620378 https://doi.org/10.1214/aoap/1027961045
[42] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.* **40** (1) (2012) 162–212. MR2917771 https://doi.org/10.1214/10-AOP623
[43] T. G. Kurtz. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Probab.* **8** (1971) 344–356. MR0287609 https://doi.org/10.1017/s002190020003535x
[44] T. G. Kurtz. Limit theorems and diffusion approximations for density dependent Markov chains. *Math. Program. Stud.* **5** (1976) 67–78. MR0445626 https://doi.org/10.1007/bfb0120765
[45] T. Leliévre, D. Le Peutrec and B. Nectoux. The exit from a metastable state: concentration of the exit point distribution on the low energy saddle points, part 2. Stoch PDE. *Anal Comp* (2021). MR4385411 https://doi.org/10.1007/s40072-021-00202-0
[46] T. Leliévre and F. Nier. Low temperature asymptotics for quasistationary distributions in a bounded domain. *Anal. PDE* **8** (3) (2015) 561–628. MR3353826 https://doi.org/10.2140/apde.2015.8.561
[47] X. Liao, L. Wang and P. Yu. *Stability of Dynamical Systems*. *Monograph Series on Nonlinear Science and Complexity* **5**. Elsevier B. V., Amsterdam, 2007. MR2535211
[48] J. Littin C. Uniqueness of quasistationary distributions and discrete spectra when ∞ is an entrance boundary and 0 is singular. *J. Appl. Probab.* **49** (3) (2012) 719–730. MR3012095 https://doi.org/10.1239/jap/1346955329
[49] M. Lladser and J. San Martín. Domain of attraction of the quasi-stationary distributions for the Ornstein–Uhlenbeck process. *J. Appl. Probab.* **37** (2) (2000) 511–520. MR1781008 https://doi.org/10.1017/s0021900200015692
[50] P. Mandl. Spectral theory of semi-groups connected with diffusion processes and its application. *Czechoslovak Math. J.* **11** (86) (1961) 558–569. MR0137143
[51] S. Martínez and J. San Martín. Quasi-stationary distributions for a Brownian motion with drift and associated limit laws. *J. Appl. Probab.* **31** (4) (1994) 911–920. MR1303922 https://doi.org/10.1017/s0021900200099447
[52] S. Martínez and J. San Martín. Classification of killed one-dimensional diffusions. *Ann. Probab.* **32** (2004) 530–552. MR2040791 https://doi.org/10.1214/aop/1078415844
[53] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probab. Surv.* **9** (2012) 340–410. MR2994898 https://doi.org/10.1214/11-PS191
[54] S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*, 2nd edition. Cambridge University Press, Cambridge, 2009. With a prologue by Peter W. Glynn. MR2509253 https://doi.org/10.1017/CBO9780511626630
[55] Y. Miura. Ultracontractivity for Markov semigroups and quasi-stationary distributions. *Stoch. Anal. Appl.* **32** (4) (2014) 591–601. MR3219695 https://doi.org/10.1080/07362994.2014.905865
[56] A. Morozov et al. Long transients in ecology: Theory and applications. *Phys. Life Rev.* (2019). https://doi.org/10.1016/j.plrev.2019.09.004
[57] H. Qian. Nonlinear stochastic dynamics of mesoscopic homogeneous biochemical reaction systems – an analytical theory. *Nonlinearity* **24** (6) (2011) R19–R49. MR2793894 https://doi.org/10.1088/0951-7715/24/6/R01
[58] K. Ramanan and O. Zeitouni. The quasi-stationary distribution for small random perturbations of certain one-dimensional maps. *Stochastic Process. Appl.* **84** (1) (1999) 25–51. MR1720096 https://doi.org/10.1016/S0304-4149(99)00044-7
[59] S. K. Scott, B. Peng, A. S. Tomlin and K. Showalter. Transient chaos in a closed chemical system. *J. Chem. Phys.* **94** (1991) 1134.
[60] D. Steinsaltz and S. N. Evans. Quasistationary distributions for one-dimensional diffusions with killing. *Trans. Amer. Math. Soc.* **359** (3) (2007) 1285–1324. MR2262851 https://doi.org/10.1090/S0002-9947-06-03980-8
[61] M. Vellela and H. Qian. A quasistationary analysis of a stochastic chemical reaction: Keizer's paradox. *Bull. Math. Biol.* **69** (5) (2007) 1727–1746. MR2326546 https://doi.org/10.1007/s11538-006-9188-3
[62] M. Vidyasagar. *Nonlinear Systems Analysis*. *Classics in Applied Mathematics* **42**. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the second (1993) edition. MR1946479 https://doi.org/10.1137/1.9780898719185
[63] J. Wang, P. G. Soerensen and F. Hynne. Transient period doublings, torus oscillations, and chaos in a closed chemical system. *J. Phys. Chem.* **98** (3) (1994) 725–727.
[64] K. Yamato. A unifying approach to non-minimal quasi-stationary distributions for one-dimensional diffusions. *J. Appl. Probab.* **59** (4) (2022) 1106–1128. MR4507684 https://doi.org/10.1017/jpr.2022.2
[65] H. Zhang and G. He. Existence and construction of quasi-stationary distributions for one-dimensional diffusions. *J. Math. Anal. Appl.* **434** (1) (2016) 171–181. MR3404554 https://doi.org/10.1016/j.jmaa.2015.09.010

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 904–922
https://doi.org/10.1214/21-AIHP1199

On the Itô–Alekseev–Gröbner formula for stochastic differential equations

Anselm Hudde[1,a], Martin Hutzenthaler[2,b], Arnulf Jentzen[3,4,c,d] and Sara Mazzonetto[5,e]

[1] *Faculty of Mathematics and Technology, University of Applied Sciences Koblenz, Remagen, Germany,* [a]*hudde@rheinahrcampus.de*
[2] *Faculty of Mathematics, University of Duisburg-Essen, Essen, Germany,* [b]*martin.hutzenthaler@uni-due.de*
[3] *School of Data Science and Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Shenzhen, China,* [c]*ajentzen@cuhk.edu.cn*
[4] *Applied Mathematics: Institute for Analysis and Numerics, Faculty of Mathematics and Computer Science, University of Münster, Münster, Germany,* [d]*ajentzen@uni-muenster.de*
[5] *Université de Lorraine, Institut Élie Cartan de Lorraine, Nancy, France,* [e]*sara.mazzonetto@univ-lorraine.fr*

Abstract. In this article we establish a new formula for the difference of a test function of the solution of a stochastic differential equation and of the test function of an Itô process. The introduced formula essentially generalizes both the classical Alekseev–Gröbner formula from the literature on deterministic differential equations as well as the classical Itô formula from stochastic analysis. The discovered formula, which we suggest to refer to as Itô–Alekseev–Gröbner formula, is a powerful tool for deriving strong approximation rates for perturbations and approximations of stochastic ordinary and partial differential equations.

Résumé. Dans cet article nous présentons une nouvelle formule qui exprime la différence entre une fonction test appliquée à une solution d'une équation différentielle stochastique et la même fonction test appliquée à un processus d'Itô. Cette formule généralise à la fois la formule classique d'Alekseev–Gröbner pour les équations différentielles déterministes et la formule d'Itô de l'analyse stochastique. Ainsi, nous suggérons de l'appeler formule d'Itô–Alekseev–Gröbner. Il s'agit d'un outil puissant pour dériver les taux d'approximation forte pour des perturbations et approximations d'équations différentielles stochastiques et d'équations à dérivées partielles stochastiques.

MSC2020 subject classifications: 60H10

Keywords: Itô formula; Alekseev–Gröbner formula; Nonlinear variation-of-constants formula; Nonlinear integration-by-parts formula; Perturbation of stochastic differential equations; Strong convergence rate; Non-globally monotone coefficients; Small-noise analysis

References

[1] V. Alekseev. An estimate for the perturbations of the solution of ordinary differential equations (Russian). *Vestn. Mosk. Univ., Ser. I, Math. Meh.* **2** (1961). MR0133536
[2] E. Alòs and D. Nualart. An extension of Itô's formula for anticipating processes. *J. Theor. Probab.* **11** (2) (1998) 493–514. MR1622583 https://doi.org/10.1023/A:1022692024364
[3] M. Arnaudon and P. Del Moral. A variational approach to nonlinear and interacting diffusions. *Stoch. Anal. Appl.* **37** (5) (2019) 717–748. MR3991057 https://doi.org/10.1080/07362994.2019.1609985
[4] S. Becker, B. Gess, A. Jentzen et al. Strong convergence rates for explicit space-time discrete numerical approximations of stochastic Allen-Cahn equations. *Stoch. Partial Differ. Equ. Anal. Comput.* **11** (2023) 211–268. MR4563700 https://doi.org/10.1007/s40072-021-00226-6
[5] S. Becker, B. Gess, A. Jentzen and P. E. Kloeden. Lower and upper bounds for strong approximation errors for numerical approximations of stochastic heat equations. *BIT Numer. Math.* **60** (2020) 1057–1073. MR4179715 https://doi.org/10.1007/s10543-020-00807-2
[6] S. Becker, B. Gess, A. Jentzen and P. E. Kloeden. Strong convergence rates for explicit space-time discrete numerical approximations of stochastic Allen–Cahn equations. *Stoch. Partial Differ. Equ. Anal. Comput.* **11** (1) (2023) 211–268. MR4563700 https://doi.org/10.1007/s40072-021-00226-6
[7] S. Becker and A. Jentzen. Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg–Landau equations. *Stochastic Process. Appl.* **129** (1) (2019) 28–69. MR3906990 https://doi.org/10.1016/j.spa.2018.02.008
[8] C.-E. Bréhier, J. Cui and J. Hong. Strong convergence rates of semidiscrete splitting approximations for the stochastic Allen–Cahn equation. *IMA J. Numer. Anal.* **39** (4) (2019) 2096–2134. MR4019051 https://doi.org/10.1093/imanum/dry052

[9] C.-E. Bréhier and L. Goudenège. Analysis of some splitting schemes for the stochastic Allen–Cahn equation. *Discrete Contin. Dyn. Syst. Ser. B* **24** (8) (2019) 4169. MR3986273 https://doi.org/10.3934/dcdsb.2019077
[10] S. Cerrai. *Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach*. *Lecture Notes in Mathematics* **1762**. Springer-Verlag, Berlin, 2001. MR1840644 https://doi.org/10.1007/b80743
[11] S. G. Cox, M. Hutzenthaler and A. Jentzen. Local Lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations, 1–90. Preprint, 2013. Available at arXiv:1309.5595.
[12] K. Dareiotis, C. Kumar and S. Sabanis. On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations. *SIAM J. Numer. Anal.* **54** (3) (2016) 1840–1872. MR3513865 https://doi.org/10.1137/151004872
[13] A. M. Davie and J. G. Gaines. Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comp.* **70** (233) (2001) 121–134 (electronic). MR1803132 https://doi.org/10.1090/S0025-5718-00-01224-2
[14] K. Debrabant and A. Rößler. Continuous weak approximation for stochastic differential equations. *J. Comput. Appl. Math.* **214** (1) (2008) 259–273. MR2391687 https://doi.org/10.1016/j.cam.2007.02.040
[15] P. del Moral and S. S. Singh. A forward–backward stochastic analysis of diffusion flows. Preprint, 2019. Available at arXiv:1906.09145v3.
[16] P. K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. Springer, Cham, 2014. MR3289027 https://doi.org/10.1007/978-3-319-08332-2
[17] C. Graham and D. Talay. *Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations of Stochastic Simulation*, **68**. Springer Science & Business Media, Heidelberg, 2013. MR3097957 https://doi.org/10.1007/978-3-642-39363-1
[18] W. Gröbner. *Die Lie-Reihen und Ihre Anwendungen*. VEB Deutscher Verlag der Wiss, Berlin, 1960. MR0116102
[19] I. Gyöngy and A. Millet. Rate of convergence of implicit approximations for stochastic evolution equations. In *Stochastic Differential Equations: Theory and Applications* 281–310. *Interdiscip. Math. Sci.* **2**. World Sci. Publ., Hackensack, NJ, 2007. MR2393581 https://doi.org/10.1142/9789812770639_0011
[20] E. Hairer, S. P. Nørsett and G. Wanner. *Solving Ordinary Differential Equations I*, 2nd edition. Springer-Verlag, Berlin, 1993. MR1227985
[21] M. Hairer, M. Hutzenthaler and A. Jentzen. Loss of regularity for Kolmogorov equations. *Ann. Probab.* **43** (2) (2015) 468–527. MR3305998 https://doi.org/10.1214/13-AOP838
[22] A. Hudde, M. Hutzenthaler, A. Jentzen and S. Mazzonetto. On the Itô–Alekseev–Gröbner formula for stochastic differential equations, 1–29. Preprint, 2018. Available at arXiv:1812.09857v1.
[23] A. Hudde, M. Hutzenthaler and S. Mazzonetto. Existence of spatially differentiable solutions of stochastic differential equations with non-globally monotone coefficient functions. Preprint, 2019. Available at arXiv:1903.09707.
[24] M. Hutzenthaler and A. Jentzen. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with nonglobally monotone coefficients. *Ann. Probab.* **48** (1) (2020) 53–93. MR4079431 https://doi.org/10.1214/19-AOP1345
[25] M. Hutzenthaler, A. Jentzen and P. E. Kloeden. Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22** (4) (2012) 1611–1641. MR2985171 https://doi.org/10.1214/11-AAP803
[26] A. Iserles. *A First Course in the Numerical Analysis of Differential Equations*, 2nd edition. *Cambridge Texts in Applied Mathematics*. Cambridge University Press, Cambridge, 2009. MR2478556
[27] A. Iserles and G. Söderlind. Global bounds on numerical error for ordinary differential equations. *J. Complexity* **9** (1) (1993) 97–112. MR1213489 https://doi.org/10.1006/jcom.1993.1007
[28] A. Jentzen and P. Pušnik. Strong convergence rates for an explicit numerical approximation method for stochastic evolution equations with non-globally Lipschitz continuous nonlinearities. *IMA J. Numer. Anal.* **40** (2) (2020) 1005–1050. MR4092277 https://doi.org/10.1093/imanum/drz009
[29] T. Kato. *Perturbation Theory for Linear Operators*, **132**. Springer Science & Business Media, Berlin, 1980. MR0678094
[30] C. Kelly and G. J. Lord. Adaptive time-stepping strategies for nonlinear stochastic systems. *IMA J. Numer. Anal.* **38** (3) (2018) 1523–1549. MR3829168 https://doi.org/10.1093/imanum/drx036
[31] A. Klenke. *Probability Theory: A Comprehensive Course*. *Universitext*. Springer-Verlag London Ltd., London, 2008. Translated from the 2006 German original. MR2372119 https://doi.org/10.1007/978-1-84800-048-3
[32] M. Kovács, S. Larsson and F. Lindgren. On the backward Euler approximation of the stochastic Allen–Cahn equation. *J. Appl. Probab.* **52** (2) (2015) 323–338. MR3372078 https://doi.org/10.1239/jap/1437658601
[33] R. Kruse. *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*. *Lecture Notes in Mathematics* **2093**. Springer, Cham, 2014. MR3154916 https://doi.org/10.1007/978-3-319-02231-4
[34] C. Leonhard and A. Rößler. Enhancing the order of the Milstein scheme for stochastic partial differential equations with commutative noise. *SIAM J. Numer. Anal.* **56** (4) (2018) 2585–2622. MR3842926 https://doi.org/10.1137/16M1094087
[35] X.-M. Li. Strong p-completeness of stochastic differential equations and the existence of smooth flows on noncompact manifolds. *Probab. Theory Related Fields* **100** (4) (1994) 485–511. MR1305784 https://doi.org/10.1007/BF01268991
[36] X.-M. Li and M. Scheutzow. Lack of strong completeness for stochastic flows. *Ann. Probab.* **39** (4) (2011) 1407–1421. MR2857244 https://doi.org/10.1214/10-AOP585
[37] I. Lie and S. P. Nørsett. Superconvergence for multistep collocation. *Math. Comp.* **52** (185) (1989) 65–79. MR0971403 https://doi.org/10.2307/2008653
[38] W. Liu and M. Röckner. *Stochastic Partial Differential Equations: An Introduction*. Springer, Cham, 2015. MR3410409 https://doi.org/10.1007/978-3-319-22354-4
[39] Z. Liu and Z. Qiao. Strong approximation of monotone stochastic partial differential equations driven by white noise. *IMA J. Numer. Anal.* **40** (2) (2020) 1074–1093. MR4092279 https://doi.org/10.1093/imanum/dry088
[40] Z. Liu and Z. Qiao. Strong approximation of monotone stochastic partial differential equations driven by multiplicative noise. *Stoch. Partial Differ. Equ. Anal. Comput.* **9** (2021) 559–602. MR4297233 https://doi.org/10.1007/s40072-020-00179-2
[41] A. K. Majee and A. Prohl. Optimal strong rates of convergence for a space-time discretization of the stochastic Allen–Cahn equation with multiplicative noise. *Comput. Methods Appl. Math.* **18** (2) (2018) 297–311. MR3776047 https://doi.org/10.1515/cmam-2017-0023
[42] T. Müller-Gronbach and K. Ritter. Lower bounds and nonuniform time discretization for approximation of stochastic heat equations. *Found. Comput. Math.* **7** (2) (2007) 135–181. MR2324415 https://doi.org/10.1007/s10208-005-0166-6
[43] T. Müller-Gronbach, K. Ritter and T. Wagner. Optimal pointwise approximation of a linear stochastic heat equation with additive space-time white noise. In *Monte Carlo and Quasi-Monte Carlo Methods 2006* 577–589. Springer, Berlin, 2008. MR2479247 https://doi.org/10.1007/978-3-540-74496-2_34

[44] T. Müller-Gronbach, K. Ritter and T. Wagner. Optimal pointwise approximation of infinite-dimensional Ornstein–Uhlenbeck processes. *Stoch. Dyn.* **8** (3) (2008) 519–541. MR2444516 https://doi.org/10.1142/S0219493708002433

[45] H.-L. Ngo and D. Taguchi. Strong rate of convergence for the Euler–Maruyama approximation of stochastic differential equations with irregular coefficients. *Math. Comp.* **85** (300) (2016) 1793–1819. MR3471108 https://doi.org/10.1090/mcom3042

[46] S. P. Nørsett and G. Wanner. The real-pole sandwich for rational approximations and oscillation equations. *BIT Numer. Math.* **19** (1) (1979) 79–94. MR0530118 https://doi.org/10.1007/BF01931224

[47] D. Nualart. *The Malliavin Calculus and Related Topics*, **1995**. Springer, New York, 2006. MR1344217 https://doi.org/10.1007/978-1-4757-2437-0

[48] D. Nualart and É. Pardoux. Stochastic calculus with anticipating integrands. *Probab. Theory Related Fields* **78** (4) (1988) 535–581. MR0950346 https://doi.org/10.1007/BF00353876

[49] E. Pardoux and P. Protter. A two-sided stochastic integral and its calculus. *Probab. Theory Related Fields* **76** (1) (1987) 15–49. MR0899443 https://doi.org/10.1007/BF00390274

[50] P. E. Protter. *Stochastic Integration and Differential Equations*, 2nd edition. *Stochastic Modelling and Applied Probability* **21**. Springer-Verlag, Berlin, 2005. Version 2.1, Corrected third printing. MR2273672 https://doi.org/10.1007/978-3-662-10061-5

[51] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 2nd edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Springer-Verlag, Berlin, 1994. MR1303781

[52] S. Sabanis. Euler approximations with varying coefficients: The case of superlinearly growing diffusion coefficients. *Ann. Appl. Probab.* **26** (4) (2016) 2083–2105. MR3543890 https://doi.org/10.1214/15-AAP1140

[53] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8** (4) (1990) 483–509. MR1091544 https://doi.org/10.1080/07362999008809220

[54] A. Tambue and J. M. T. Ngnotchouye. Weak convergence for a stochastic exponential integrator and finite element discretization of stochastic partial differential equation with multiplicative & additive noise. *Appl. Numer. Math.* **108** (2016) 57–86. MR3528302 https://doi.org/10.1016/j.apnum.2016.04.013

[55] X. Wang. An efficient explicit full-discrete scheme for strong approximation of stochastic Allen–Cahn equation. *Stochastic Process. Appl.* **130** (10) (2020) 6271–6299. MR4140034 https://doi.org/10.1016/j.spa.2020.05.011

[56] X. Zhang. Stochastic flows and Bismut formulas for stochastic Hamiltonian systems. *Stochastic Process. Appl.* **120** (10) (2010) 1929–1949. MR2673982 https://doi.org/10.1016/j.spa.2010.05.015

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 923–945
https://doi.org/10.1214/22-AIHP1338

Matsumoto–Yor and Dufresne type theorems for a random walk on positive definite matrices

Jonas Arista[1,a], Elia Bisi[2,b] and Neil O'Connell[3,c]

[1] *Universität Bielefeld, Fakultät für Mathematik, Universitätsstraße 25, 33615 Bielefeld, Germany,* [a]*jarista@math.uni-bielefeld.de*
[2] *Technische Universität Wien, Institut für Stochastik und Wirtschaftsmathematik, E 105-07, Wiedner Hauptstraße 8-10, 1040 Wien, Austria,* [b]*elia.bisi@tuwien.ac.at*
[3] *School of Mathematics and Statistics, University College Dublin, Dublin 4, Ireland,* [c]*neil.oconnell@ucd.ie*

Abstract. We establish analogues of the geometric Pitman $2M - X$ theorem of Matsumoto and Yor and of the classical Dufresne identity, for a multiplicative random walk on positive definite matrices with Beta type II distributed increments. The Dufresne type identity provides another example of a stochastic matrix recursion, as considered by Chamayou and Letac (J. Theoret. Probab. 12, 1999), that admits an explicit solution.

Résumé. Nous établissons des analogues de la version géométrique du théorème 2M-X de Pitman, démontrée par Matsumoto et Yor, et de l'identité de Dufresne classique, pour une marche aléatoire multiplicative sur l'ensemble des matrices définies positives d'incréments distribués selon une loi Bêta II. L'identité de Dufresne proposée fournit un autre exemple de récursion matricielle stochastique, comme l'ont considéré Chamayou et Letac (J. Theoret. Probab. 12, 1999), qui admet une solution explicite.

MSC2020 subject classifications: Primary 60K35; 82B23; 60B20; secondary 60G10; 22E30; 62H10

Keywords: Matrix Dufresne identity; Matrix Matsumoto–Yor theorem; Intertwining relations; Stochastic matrix recursions and equations; Matrix variate distributions; Wishart and Beta distributions; Lyapunov exponents

References

[1] J. Arista, E. Bisi and N. O'Connell. Matrix Whittaker processes, 2022. Available at arXiv:2203.14868.
[2] J.-P. Bouchaud, A. Comtet, A. Georges and P. Le Doussal. Classical diffusion of a particle in a one-dimensional random force field. *Ann. Phys.* **201** (2) (1990) 284–341. MR1062911 https://doi.org/10.1016/0003-4916(90)90043-N
[3] P. Bougerol and J. Lacroix. *Products of Random Matrices with Applications to Schrödinger Operators*. *Progress in Probability* **8**. Birkhaüser, Boston, MA, 1985. MR0886674 https://doi.org/10.1007/978-1-4684-9172-2
[4] M. Casalis and G. Letac. The Lukacs–Olkin–Rubin characterization of Wishart distributions on symmetric cones. *Ann. Statist.* **24** (2) (1996) 763–786. MR1394987 https://doi.org/10.1214/aos/1032894464
[5] J. Chamayou and G. Letac. Explicit stationary distributions for compositions of random functions and products of random matrices. *J. Theoret. Probab.* **4** (1991) 3–36. MR1088391 https://doi.org/10.1007/BF01046992
[6] J. Chamayou and G. Letac. Additive properties of the Dufresne laws and their multivariate extension. *J. Theoret. Probab.* **12** (1999) 1045–1066. MR1729469 https://doi.org/10.1023/A:1021649305082
[7] I. Corwin, N. O'Connell, T. Seppäläinen and N. Zygouras. Tropical combinatorics and Whittaker functions. *Duke Math. J.* **163** (3) (2014) 513–563. MR3165422 https://doi.org/10.1215/00127094-2410289
[8] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.* **41** (1) (1999) 45–76. MR1669737 https://doi.org/10.1137/S0036144598338446
[9] D. Dufresne. The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuar. J.* **1990** (1) (1990) 39–79. MR1129194 https://doi.org/10.1080/03461238.1990.10413872
[10] D. Dufresne. On the stochastic equation $\mathcal{L}(X) = \mathcal{L}[B(X + C)]$ and a property of gamma distributions. *Bernoulli* **2** (3) (1996) 287–291. MR1416868 https://doi.org/10.2307/3318525
[11] E. B. Dynkin. Non-negative eigenfunctions of the Laplace–Beltrami operator and Brownian motion in certain symmetric spaces. *Dokl. Akad. Nauk SSSR* **141** (2) (1961) 288–291. MR0132607
[12] A. Furman. Random walks on groups and random transformations. In *Handbook of Dynamical Systems, Vol. 1A* 931–1014. B. Hasselblatt and A. Katok (Eds). Elsevier, Amsterdam, 2002. MR1928529 https://doi.org/10.1016/S1874-575X(02)80014-5
[13] T. Gautié, J.-P. Bouchaud and P. Le Doussal. Matrix Kesten recursion, inverse-Wishart ensemble and fermions in a Morse potential. *J. Phys. A: Math. Theor.* **54** (25) (2021) 255201. MR4271318 https://doi.org/10.1088/1751-8121/abfc7f

[14] C. M. Goldie. Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1** (1) (1991) 126–166. MR1097468
[15] A. Grabsch and C. Texier. Wigner–Smith matrix, exponential functional of the matrix Brownian motion and matrix Dufresne identity. *J. Phys. A: Math. Theor.* **53** (42) (2020) 425003. MR4177049 https://doi.org/10.1088/1751-8121/aba215
[16] A. K. Gupta and D. K. Nagar. *Matrix Variate Distributions*, 1st edition. Chapman & Hall/CRC, London, 1999. MR1738933
[17] R. A. Horn and C. R. Johnson. *Matrix Analysis*, 2nd edition. Cambridge University Press, New York, 2013. MR2978290
[18] F. P. Kelly. Markovian functions of a Markov chain. *Sankhyā Ser. A* **44** (3) (1982) 372–379. MR0705461
[19] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. Springer, New York, NY, 1976. MR0410929
[20] H. Kesten. Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** (1973) 207–248. MR0440724 https://doi.org/10.1007/BF02392040
[21] T. G. Kurtz. Martingale problems for conditional distributions of Markov processes. *Electron. J. Probab.* **3** (9) (1998) 1–29. MR1637085 https://doi.org/10.1214/EJP.v3-31
[22] G. Letac. A contraction principle for certain Markov chains and its applications. In *Random Matrices and Their Applications* 263–273. *Contemp. Math.* **50**, 1986. MR0841098 https://doi.org/10.1090/conm/050/841098
[23] E. Lukacs. A characterization of the gamma distribution. *Ann. Math. Stat.* **26** (2) (1955) 319–324. MR0069408 https://doi.org/10.1214/aoms/1177728549
[24] A. M. Mathai. *Jacobians of Matrix Transformations and Functions of Matrix Argument*. World Scientific, Singapore, 1997. MR1607602 https://doi.org/10.1142/3438
[25] H. Matsumoto and M. Yor. An analogue of Pitman's $2M - X$ theorem for exponential Wiener functionals: Part I: A time-inversion approach. *Nagoya Math. J.* **159** (2000) 125–166. MR1783567 https://doi.org/10.1017/S0027763000007455
[26] C. M. Newman. The distribution of Lyapunov exponents: Exact results for random matrices. *Comm. Math. Phys.* **103** (1986) 121–126. MR0826860
[27] J. R. Norris, L. C. G. Rogers and D. Williams. Brownian motions of ellipsoids. *Trans. Amer. Math. Soc.* **294** (2) (1986) 757–765. MR0825735 https://doi.org/10.2307/2000214
[28] N. O'Connell. A path-transformation for random walks and the Robinson–Schensted correspondence. *Trans. Amer. Math. Soc.* **355** (9) (2003) 3669–3697. MR1990168 https://doi.org/10.1090/S0002-9947-03-03226-4
[29] N. O'Connell. Conditioned random walks and the RSK correspondence. *J. Phys. A: Math. Gen.* **36** (12) (2003) 3049–3066. MR1986407 https://doi.org/10.1088/0305-4470/36/12/312
[30] N. O'Connell. Directed polymers and the quantum Toda lattice. *Ann. Probab.* **40** (2) (2012) 437–458. MR2952082 https://doi.org/10.1214/10-AOP632
[31] N. O'Connell. Interacting diffusions on positive definite matrices. *Probab. Theory Related Fields* **180** (2021) 679–726. MR4288330 https://doi.org/10.1007/s00440-021-01039-3
[32] N. O'Connell and M. Yor. A representation for non-colliding random walks. *Electron. Commun. Probab.* **7** (2002) 1–12. MR1887169 https://doi.org/10.1214/ECP.v7-1042
[33] I. Olkin and H. Rubin. A characterization of the Wishart distribution. *Ann. Math. Stat.* **33** (4) (1962) 1272–1280. MR0141186 https://doi.org/10.1214/aoms/1177704359
[34] J. W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. Appl. Probab.* **7** (3) (1975) 511–526. MR0375485 https://doi.org/10.2307/1426125
[35] B. Rider and B. Valkó. Matrix Dufresne identities. *Int. Math. Res. Not.* **2016** (1) (2016) 174–218. MR3514061 https://doi.org/10.1093/imrn/rnv127
[36] L. C. G. Rogers and J. W. Pitman. Markov functions. *Ann. Probab.* **33** (4) (1981) 573–582. MR0624684
[37] T. Seppäläinen. Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* **40** (1) (2012) 19–73. MR2917766 https://doi.org/10.1214/10-AOP617
[38] F. Solomon. Random walks in a random environment. *Ann. Probab.* **3** (1) (1975) 1–31. MR0362503 https://doi.org/10.1214/aop/1176996444
[39] A. Terras. *Harmonic Analysis on Symmetric Spaces – Higher Rank Spaces, Positive Definite Matrix Space and Generalizations*, 2nd edition. Springer, New York, NY, 2016. MR3496932 https://doi.org/10.1007/978-1-4939-3408-9
[40] W. Vervaat. On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Probab.* **11** (4) (1979) 750–783. MR0544194 https://doi.org/10.2307/1426858

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 946–970
https://doi.org/10.1214/22-AIHP1339

Linear spectral statistics of sequential sample covariance matrices

Nina Dörnemann[a] and Holger Dette[b]

Faculty of Mathematics, Ruhr University Bochum, Bochum, Germany, [a]*nina.doernemann@rub.de,* [b]*holger.dette@rub.de*

Abstract. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ denote independent p-dimensional vectors with independent complex or real valued entries such that $\mathbb{E}[\mathbf{x}_i] = \mathbf{0}$, $\operatorname{Var}(\mathbf{x}_i) = \mathbf{I}_p$, $i = 1, \dots, n$, let $\mathbf{T}_n$ be a $p \times p$ Hermitian nonnegative definite matrix and f be a given function. We prove that an appropriately standardized version of the stochastic process $(\operatorname{tr}(f(\mathbf{B}_{n,t})))_{t\in[t_0,1]}$ corresponding to a linear spectral statistic of the sequential empirical covariance estimator

$$(\mathbf{B}_{n,t})_{t\in[t_0,1]} = \left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \mathbf{T}_n^{1/2}\mathbf{x}_i\mathbf{x}_i^\star\mathbf{T}_n^{1/2}\right)_{t\in[t_0,1]}$$

converges weakly to a non-standard Gaussian process for $n, p \to \infty$. As an application, we use these results to develop a novel approach for monitoring the sphericity assumption in a high-dimensional framework, even if the dimension of the underlying data is larger than the sample size.

Résumé. Soient $\mathbf{x}_1, \dots, \mathbf{x}_n$ des vecteurs p-dimensionnels indépendants aux composantes complexes ou réelles indépendantes tels que $\mathbb{E}[\mathbf{x}_i] = \mathbf{0}$, $\operatorname{Var}(\mathbf{x}_i) = \mathbf{I}_p$, $i = 1, \dots, n$. Soit $\mathbf{T}_n$ une matrice hermitienne positive définie d'ordre p et soit f une fonction donnée. Nous démontrons qu'une certaine version normalisée du processus $(\operatorname{tr}(f(\mathbf{B}_{n,t})))_{t\in[t_0,1]}$ que l'on peut voir comme une statistique spectrale linéaire du processus de covariance empirique

$$(\mathbf{B}_{n,t})_{t\in[t_0,1]} = \left(\frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} \mathbf{T}_n^{1/2}\mathbf{x}_i\mathbf{x}_i^\star\mathbf{T}_n^{1/2}\right)_{t\in[t_0,1]}$$

converge faiblement vers un processus gaussien non standard lorsque $n, p \to \infty$. Afin d'illustrer l'utilité de ce résultat, nous construisons un nouveau test de sphéricité dans le cadre de données en grande dimension, qui reste valide dans un régime où la dimension des données est supérieure à la taille de l'échantillon.

MSC2020 subject classifications: Primary 15A18; 60F17; secondary 62H15

Keywords: Linear spectral statistic; Sequential sample covariance matrix; Sequential process; Stieltjes transform; Monitoring spherictiy

References

[1] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*, 2nd edition. *Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. John Wiley & Sons, Inc., New York, 1984. MR0771294
[2] A. Aue, S. Hörmann, L. Horváth and M. Reimherr. Break detection in the covariance structure of multivariate time series models. *Ann. Statist.* **37** (2009) 4046–4087. MR2572452 https://doi.org/10.1214/09-AOS707
[3] Z. Bai, D. Jiang, J. F. Yao and S. Zheng. Corrections to LRT on large-dimensional covariance matrix by RMT. *Ann. Statist.* **37** (2009) 3822–3840. MR2572444 https://doi.org/10.1214/09-AOS694
[4] Z. Bai and J. W. Silverstein. CLT for linear spectral statistics of large-dimensional sample covariance matrices. *Ann. Probab.* **32** (2004) 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845
[5] Z. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices* **20**. Springer, 2010. MR2567175 https://doi.org/10.1007/978-1-4419-0661-8
[6] M. Banna, J. Najim and J. Yao. A CLT for linear spectral statistics of large random information-plus-noise matrices. *Stochastic Process. Appl.* **130** (2020) 2250–2281. MR4074699 https://doi.org/10.1016/j.spa.2019.06.017
[7] Z. Bao, L.-C. Lin, G. Pan and W. Zhou. Spectral statistics of large dimensional Spearman's rank correlation matrix and its application. *Ann. Statist.* **43** (2015) 2588–2623. MR3405605 https://doi.org/10.1214/15-AOS1353

[8] Z. Bao, G. Pan and W. Zhou. The logarithmic law of random determinant. *Bernoulli* **21** (2015) 1600–1628. MR3352055 https://doi.org/10.3150/14-BEJ615
[9] M. Birke and H. Dette. A note on testing the covariance matrix for large dimension. *Statist. Probab. Lett.* **74** (2005) 281–289. MR2189467 https://doi.org/10.1016/j.spl.2005.04.051
[10] T. Bodnar, H. Dette and N. Parolya. Testing for independence of large dimensional vectors. *Ann. Statist.* **47** (2019) 2977–3008. MR3988779 https://doi.org/10.1214/18-AOS1771
[11] A. Borodin CLT for spectra of submatrices of Wigner random matrices, 2010. arXiv preprint. Available at arXiv:1010.0898. MR3221945 https://doi.org/10.17323/1609-4514-2014-14-1-29-38
[12] T. T. Cai, T. Liang and H. H. Zhou. Law of log determinant of sample covariance matrix and optimal estimation of differential entropy for high-dimensional Gaussian distributions. *J. Multivariate Anal.* **137** (2015) 161–172. MR3332804 https://doi.org/10.1016/j.jmva.2015.02.003
[13] S. X. Chen, L. X. Zhang and P. S. Zhong. Testing high dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105** (2010) 810–819. MR2724863 https://doi.org/10.1198/jasa.2010.tm09560
[14] A. D'Aristotile. An invariance principle for triangular arrays. *J. Theoret. Probab.* **13** (2000) 327–341. MR1777537 https://doi.org/10.1023/A:1007801726073
[15] A. D'Aristotile, P. Diaconis and C. M. Newman. Brownian motion and the classical groups. *IMS Lecture Notes Monogr. Ser.* **41** (2003) 97–116. MR1999417 https://doi.org/10.1214/lnms/1215091660
[16] H. Dette and N. Dörnemann. Likelihood ratio tests for many groups in high dimensions. *J. Multivariate Anal.* **178** (2020), 104605. MR4079037 https://doi.org/10.1016/j.jmva.2020.104605
[17] H. Dette and J. Gösmann. A likelihood ratio approach to sequential change point detection for a general class of parameters. *J. Amer. Statist. Assoc.* **115** (2020) 1361–1377. MR4143471 https://doi.org/10.1080/01621459.2019.1630562
[18] H. Dette and D. Tomecki. Determinants of block Hankel matrices for random matrix-valued measures. *Stochastic Process. Appl.* **129** (2019) 5200–5235. MR4025703 https://doi.org/10.1016/j.spa.2019.02.010
[19] N. Dörnemann and H. Dette. Supplement to "Linear spectral statistics of sequential sample covariance matrices" (2024). https://doi.org/10.1214/22-AIHP1339SUPP
[20] I. Dumitriu and E. Paquette. Spectra of overlapping Wishart matrices and the Gaussian free field. *Random Matrices Theory Appl.* **7** (2018), 1850003. MR3786884 https://doi.org/10.1142/S201032631850003X
[21] J. Fan and R. Li. Statistical challenges with high dimensionality: Feature selection in knowledge discovery. In *Proceedings of the International Congress of Mathematicians* 595–622, M. Sanz-Solé, J. Soria, J. L. Varona and J. Verdera (Eds). European Mathematical Society, Madrid, Spain, 2006. MR2275698
[22] T. J. Fisher, X. Sun and C. M. Gallagher. A new test for sphericity of the covariance matrix for high dimensional data. *J. Multivariate Anal.* **101** (2010) 2554–2570. MR2719881 https://doi.org/10.1016/j.jmva.2010.07.004
[23] A. K. Gupta and J. Xu. On some tests of the covariance matrix under general conditions. *Ann. Inst. Statist. Math.* **58** (2006) 101–114. MR2281208 https://doi.org/10.1007/s10463-005-0010-z
[24] T. Jiang and F. Yang. Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Statist.* **41** (2013) 2029–2074. MR3127857 https://doi.org/10.1214/13-AOS1134
[25] B. Jin, C. Wang, Z. D. Bai, K. K. Nair and M. Harding. Limiting spectral distribution of a symmetrized auto-cross covariance matrix. *Ann. Appl. Probab.* **24** (2014) 1199–1225. MR3199984 https://doi.org/10.1214/13-AAP945
[26] S. John. Some optimal multivariate tests. *Biometrika* **58** (1971) 123–127. MR0275568 https://doi.org/10.1093/biomet/58.1.123
[27] I. Johnstone. High dimensional statistical inference and random matrices, 2006. MR2334195 https://doi.org/10.4171/022-1/13
[28] D. Jonsson. Some limit theorems for the eigenvalues of a sample covariance matrix. *J. Multivariate Anal.* **12** (1982) 1–38. MR0650926 https://doi.org/10.1016/0047-259X(82)90080-X
[29] O. Ledoit and M. Wolf. Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *Ann. Statist.* **30** (2002) 1081–1102. MR1926169 https://doi.org/10.1214/aos/1031689018
[30] Z. Li, Q. Wang and R. Li. Central limit theorem for linear spectral statistics of large dimensional Kendall's rank correlation matrices and its applications, 2019. MR3828819 https://doi.org/10.1016/j.laa.2018.05.031
[31] A. Lytova and L. Pastur. Central limit theorem for linear eigenvalue statistics of the Wigner and sample covariance random matrices. *Metrika* **69** (2009) 153–172. MR2481919 https://doi.org/10.1007/s00184-008-0212-5
[32] J. W. Mauchly. Significance test for sphericity of a normal N-variate distribution. *Ann. Math. Stat.* **11** (1940) 204–209. MR0002084 https://doi.org/10.1214/aoms/1177731915
[33] R. J. Muirhead. *Aspects of Multivariate Statistical Theory* **197**. John Wiley & Sons, 2009. MR0652932
[34] J. Nagel. A functional CLT for partial traces of random matrices. *J. Theoret. Probab.* **98** (2020). MR4259455 https://doi.org/10.1007/s10959-019-00982-1
[35] J. Najim and J. Yao. Gaussian fluctuations for linear spectral statistics of large random covariance matrices. *Ann. Appl. Probab.* **26** (2016) 1837–1887. MR3513608 https://doi.org/10.1214/15-AAP1135
[36] H. H. Nguyen and V. Vu. Random matrices: Law of the determinant. *Ann. Probab.* **42** (2014) 146–167. MR3161483 https://doi.org/10.1214/12-AOP791
[37] G. Pan. Comparison between two types of large sample covariance matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **50** (2014) 655–677. MR3189088 https://doi.org/10.1214/12-AIHP506
[38] G. Pan and W. Zhou. Central limit theorem for signal-to-interference ratio of reduced rank linear receiver. *Ann. Appl. Probab.* **18** (2008) 1232–1270. MR2418244 https://doi.org/10.1214/07-AAP477
[39] N. Parolya, J. Heiny and D. Kurowicka Logarithmic law of large random correlation matrix, 2021. arXiv preprint. Available at arXiv:2103.13900.
[40] A. W. Van Der Vaart and J. A. Wellner. *Weak Convergence*. Springer, 1996.
[41] Q. Wang and J. Yao. On the sphericity test with large-dimensional observations. *Electron. J. Stat.* **7** (2013) 2164–2192. MR3104916 https://doi.org/10.1214/13-EJS842
[42] X. Wang, X. Han and G. Pan. The logarithmic law of sample covariance matrices near singularity. *Bernoulli* **24** (2018) 80–114. MR3706751 https://doi.org/10.3150/16-BEJ867
[43] J. Yao, S. Zheng and Z. Bai. *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, 2015. MR3468554 https://doi.org/10.1017/CBO9781107588080

[44] S. Zheng. Central limit theorems for linear spectral statistics of large dimensional F-matrices. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (2012) 444–476. MR2954263 https://doi.org/10.1214/11-AIHP414

[45] S. Zheng, Z. Bai and J. Yao. Substitution principle for CLT of linear spectral statistics of high-dimensional sample covariance matrices with applications to hypothesis testing. *Ann. Statist.* **43** (2015) 546–591. MR3316190 https://doi.org/10.1214/14-AOS1292

[46] S. Zheng, Z. Bai and J. Yao. CLT for eigenvalue statistics of large-dimensional general Fisher matrices with applications. *Bernoulli* **23** (2017) 1130–1178. MR3606762 https://doi.org/10.3150/15-BEJ772

[47] T. Zou, S. Zheng, Z. Bai, J. Yao and H. Zhu. CLT for linear spectral statistics of large dimensional sample covariance matrices with dependent data. *Statist. Papers* **63** (2022) 605–664. MR4394872 https://doi.org/10.1007/s00362-021-01250-3

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 971–989
https://doi.org/10.1214/22-AIHP1347

Robust subgaussian estimation with VC-dimension

Jules Depersin[a]

University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, [a]*j.d.p.depersin@uva.nl*

Abstract. Median-of-means (MOM) based procedures provide non-asymptotic and strong deviation bounds even when data are heavy tailed and/or corrupted. This work proposes a new general and systematic way to bound the excess risk for MOM estimators. The core technique is the use of the VC-dimension (instead of Rademacher complexity) to measure the statistical complexity. In particular, this allows one to give the first robust estimators for sparse estimation which achieves the so-called subgaussian rate, only assuming a finite second moment for the uncorrupted data.

By comparison, previous works using Rademacher complexities required a number of finite moments that grows logarithmically with the dimension. With this technique, we derive new robust subgaussian bounds for mean estimation in any norm.

Résumé. Les procédures basées sur la médiane des moyennes (MOM) fournissent des bornes non asymptotiques et fortes, même lorsque les données ont des queues de distribution lourdes et/ou sont corrompues. Ce travail propose une nouvelle méthode générale et systématique pour limiter le risque des estimateurs MOM. La technique de base est l'utilisation de la dimension VC (au lieu de la complexité de Rademacher) pour mesurer la complexité statistique. Cela permet en particulier de trouver des estimateurs robustes pour l'estimation sparse qui atteignent le taux dit sous-gaussien en supposant seulement un second moment fini pour les données non corrompues.

En comparaison, les travaux précédents utilisant les complexités de Rademacher nécessitaient un nombre de moments finis de l'ordre du logarithme de la dimension, donc dépendant de la dimension. Grâce à cette technique, nous proposons de nouvelles bornes sous-gaussiennes robustes pour l'estimation de la moyenne dans n'importe quelle norme.

MSC2020 subject classifications: 62F35

Keywords: Robustness; Heavy tailed data; Sparse estimation

References

[1] M. E. Ahsen and M. Vidyasagar. An approach to one-bit compressed sensing based on probably approximately correct learning theory. *J. Mach. Learn. Res.* **20** (11) (2019) 1–23. MR3911418
[2] N. Alon, Y. Matias and M. Szegedy. The space complexity of approximating the frequency moments. *J. Comput. System Sci.* **58** (1) (1999) 137–147. MR1688610 https://doi.org/10.1006/jcss.1997.1545
[3] J.-Y. Audibert and O. Catoni. Robust linear least squares regression. *Ann. Statist.* **39** (5) (2011) 2766–2794. MR2906886 https://doi.org/10.1214/11-AOS918
[4] S. Balakrishnan, S. S. Du, J. Li and A. Singh. Computationally efficient robust sparse estimation in high dimensions. In *Proceedings of the 2017 Conference on Learning Theory, Proceedings of Machine Learning Research, PMLR* 169–212. *07–10 Jul 2017*, 2017.
[5] L. Birgé. Stabilité et instabilité du risque minimax pour des variables indépendantes équidistribuées. *Ann. Inst. Henri Poincaré Probab. Stat.* **20** (3) (1984) 201–223. MR0762855
[6] S. Boucheron, G. Lugosi and P. Massart. *Concentration Inequalities A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford, 2013. With a foreword by Michel Ledoux. MR3185193 https://doi.org/10.1093/acprof:oso/9780199535255.001.0001
[7] S. Bubeck, N. Cesa-Bianchi and G. Lugosi. Bandits with heavy tail. *IEEE Trans. Inf. Theory* **59** (11) (2013) 7711–7717. MR3124669 https://doi.org/10.1109/TIT.2013.2277869
[8] O. Catoni. Challenging the empirical mean and empirical variance: A deviation study. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (4) (2012) 1148–1185. MR3052407 https://doi.org/10.1214/11-AIHP454
[9] M. Chen, C. Gao and Z. Ren. Robust covariance and scatter matrix estimation under Huber's contamination model. *Ann. Statist.* **46** (5) (2018) 1932–1960. MR3845006 https://doi.org/10.1214/17-AOS1607
[10] Y. Cheng, I. Diakonikolas and R. Ge. High-dimensional robust mean estimation in nearly-linear time. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms* 2755–2771. SIAM, Philadelphia, PA, 2019. MR3909640 https://doi.org/10.1137/1.9781611975482.171
[11] Y. Cheng, I. Diakonikolas, R. Ge and D. P. Woodruff. Faster algorithms for high-dimensional robust covariance estimation. In *Proceedings of the Thirty-Second Conference on Learning Theory, Proceedings of Machine Learning Research, PMLR* 727–757. *25–28 Jun 2019*, 2019.

[12] Y. Cherapanamjeri, N. Flammarion and P. L. Bartlett. Fast mean estimation with sub-Gaussian rates. In *Proceedings of the Thirty-Second Conference on Learning Theory, Proceedings of Machine Learning Research, PMLR* 786–806. *25–28 Jun 2019*, 2019.
[13] Y. Cherapanamjeri, S. B. Hopkins, T. Kathuria, P. Raghavendra and N. Tripuraneni. Algorithms for heavy-tailed statistics: Regression, covariance estimation, and beyond, 2019. MR4141785
[14] G. Chinot, G. Lecué and M. Lerasle. Robust statistical learning with Lipschitz and convex loss functions. *Probab. Theory Related Fields* **176** (3) (2020) 897–940. MR4087486 https://doi.org/10.1007/s00440-019-00931-3
[15] J. Depersin and G. Lecué. Robust sub-Gaussian estimation of a mean vector in nearly linear time. *Ann. Statist.* **50** (1) (2022) 511–536. MR4382026 https://doi.org/10.1214/21-aos2118
[16] L. Devroye, M. Lerasle, G. Lugosi and R. I. Oliveira. Sub-Gaussian mean estimators. *Ann. Statist.* **44** (6) (2016) 2695–2725. MR3576558 https://doi.org/10.1214/16-AOS1440
[17] I. Diakonikolas, G. Kamath, D. M. Kane, J. Li, A. Moitra and A. Stewart. Robust estimators in high dimensions without the computational intractability. In *57th Annual IEEE Symposium on Foundations of Computer Science – FOCS 2016* 655–664. IEEE Computer Soc., Los Alamitos, CA, 2016. MR3631028 https://doi.org/10.1109/FOCS.2016.85
[18] I. Diakonikolas and D. M. Kane. Recent advances in algorithmic high-dimensional robust statistics, 2019.
[19] I. Diakonikolas, W. Kong and A. Stewart. Efficient algorithms and lower bounds for robust linear regression. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA'19* 2745–2754. Society for Industrial and Applied Mathematics, USA, 2019. MR3909639 https://doi.org/10.1137/1.9781611975482.170
[20] D. L. Donoho. Compressed sensing. *IEEE Trans. Inf. Theory* **52** (4) (2006) 1289–1306. MR2241189 https://doi.org/10.1109/TIT.2006.871582
[21] R. M. Dudley. Central limit theorems for empirical measures. *Ann. Probab.* **6** (6) (1978) 899–929. MR0512411
[22] C. Gao. Robust regression via multivariate regression depth. *Bernoulli* **26** (2) (2020) 1139–1170. MR4058363 https://doi.org/10.3150/19-BEJ1144
[23] F. R. Hampel. A general qualitative definition of robustness. *Ann. Math. Stat.* **42** (1971) 1887–1896. MR0301858 https://doi.org/10.1214/aoms/1177693054
[24] F. R. Hampel. Robust estimation: A condensed partial survey. *Z. Wahrsch. Verw. Gebiete* **27** (1973) 87–104. MR0359096 https://doi.org/10.1007/BF00536619
[25] S. B. Hopkins. Mean estimation with sub-Gaussian rates in polynomial time. *Ann. Statist.* **48** (2) (2020) 1193–1213. MR4102693 https://doi.org/10.1214/19-AOS1843
[26] D. Hsu and S. Sabato. Loss minimization and parameter estimation with heavy tails. *J. Mach. Learn. Res.* **17** (18) (2016) 1–40. MR3491112
[27] P. J. Huber. Robust estimation of a location parameter. *Ann. Math. Stat.* **35** (1964) 73–101. MR0161415 https://doi.org/10.1214/aoms/1177703732
[28] P. J. Huber and E. M. Ronchetti. *Robust Statistics*, 2nd edition. *Wiley Series in Probability and Statistics*. John Wiley & Sons, Inc., Hoboken, NJ, 2009. MR2488795 https://doi.org/10.1002/9780470434697
[29] M. R. Jerrum, L. G. Valiant and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoret. Comput. Sci.* **43** (2–3) (1986) 169–188. MR0855970 https://doi.org/10.1016/0304-3975(86)90174-X
[30] G. Lecué and M. Lerasle. Learning from MOM's principles: Le Cam's approach. *Stochastic Process. Appl.* **129** (11) (2019) 4385–4410. MR4013866 https://doi.org/10.1016/j.spa.2018.11.024
[31] G. Lecué and M. Lerasle. Robust machine learning by median-of-means: Theory and practice. *Ann. Statist.* **48** (2) (2020) 906–931. MR4102681 https://doi.org/10.1214/19-AOS1828
[32] G. Lecué and S. Mendelson. Regularization and the small-ball method I: Sparse recovery. *Ann. Statist.* **46** (2) (2018) 611–641. MR3782379 https://doi.org/10.1214/17-AOS1562
[33] M. Ledoux. *The Concentration of Measure Phenomenon*. *Mathematical Surveys and Monographs* **89**. American Mathematical Society, Providence, RI, 2001. MR1849347 https://doi.org/10.1090/surv/089
[34] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: Isoperimetry and Processes*. *Classics in Mathematics*. Springer-Verlag, Berlin, 2011. Reprint of the 1991 edition. MR2814399
[35] M. Lerasle. *Lecture Notes: Selected Topics on Robust Statistical Learning Theory*, 2019.
[36] M. Lerasle and R. Oliveira. Robust empirical mean estimators. Technical report, IMPA and CNRS, 2011.
[37] M. Lerasle, Z. Szabo Thimotée Matthieu and G. Lecué. Monk – outliers-robust mean embedding estimation by median-of-means. Technical report, CNRS, University of Paris 11, Ecole Polytechnique and CREST, 2017.
[38] J. Li. Robust sparse estimation tasks in high dimensions, 2017.
[39] G. Lugosi and S. Mendelson. Mean estimation and regression under heavy-tailed distributions: A survey. *Found. Comput. Math.* **19** (5) (2019) 1145–1190. MR4017683 https://doi.org/10.1007/s10208-019-09427-x
[40] G. Lugosi and S. Mendelson. Near-optimal mean estimators with respect to general norms. *Probab. Theory Related Fields* **175** (3) (2019) 957–973. MR4026610 https://doi.org/10.1007/s00440-019-00906-4
[41] G. Lugosi and S. Mendelson. Regularization, sparse recovery, and median-of-means tournaments. *Bernoulli* **25** (3) (2019) 2075–2106. MR3961241 https://doi.org/10.3150/18-BEJ1046
[42] G. Lugosi and S. Mendelson. Robust multivariate mean estimation: The optimality of trimmed mean. *Ann. Statist.* **49** (1) (2021) 393–410. MR4206683 https://doi.org/10.1214/20-AOS1961
[43] G. Lugosi, S. Mendelson et al. Sub-Gaussian estimators of the mean of a random vector. *Ann. Statist.* **47** (2) (2019) 783–794. MR3909950 https://doi.org/10.1214/17-AOS1639
[44] S. Mendelson. Extending the small-ball method, 2017.
[45] S. Mendelson and N. Zhivotovskiy. Robust covariance estimation under L_4–L_2 norm equivalence. *Ann. Statist.* **48** (3) (2020) 1648–1664. MR4124338 https://doi.org/10.1214/19-AOS1862
[46] S. Minsker. Geometric median and robust estimation in Banach spaces. *Bernoulli* **21** (4) (2015) 2308–2335. MR3378468 https://doi.org/10.3150/14-BEJ645
[47] S. Minsker. Distributed statistical estimation and rates of convergence in normal approximation. *Electron. J. Stat.* **13** (2) (2019) 5213–5252. MR4043072 https://doi.org/10.1214/19-EJS1647
[48] A. S. Nemirovsky and D. B. Yudin. *Problem Complexity and Method Efficiency in Optimization. A Wiley-Interscience Publication*. *Wiley-Interscience Series in Discrete Mathematics*. John Wiley & Sons, Inc., New York, 1983. Translated from the Russian and with a preface by Dawson, E. R. MR0702836

[49] A. Prasad. Sivaraman Balakrishnan, and Pradeep Ravikumar. A unified approach to robust mean estimation, 2019.
[50] A. Prasad, S. Balakrishnan and P. Ravikumar. A robust univariate mean estimator is all you need. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research, PMLR* 4034–4044. *26–28 Aug 2020*, 2020.
[51] N. Sauer. On the density of families of sets. *J. Combin. Theory Ser. A* **13** (1) (1972) 145–147. MR0307902 https://doi.org/10.1016/0097-3165(72)90019-2
[52] H. Tsukuma. Proper Bayes minimax estimators of the normal mean matrix with common unknown variances. *J. Statist. Plann. Inference* **140** (9) (2010) 2596–2606. MR2644081 https://doi.org/10.1016/j.jspi.2010.03.031
[53] J. W. Tukey. The future of data analysis. *Ann. Math. Stat.* **33** (1962) 1–67. MR0133937 https://doi.org/10.1214/aoms/1177704711
[54] A. van der Vaart and J. A. Wellner. A note on bounds for VC dimensions. In *Collections* 103–107, **5**. Institute of Mathematical Statistics, Beachwood, Ohio, USA, 2009. MR2797943 https://doi.org/10.1214/09-IMSCOLL508
[55] R. van Handel. *Probability in High Dimension*, 2016.
[56] R. Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2018. MR3837109 https://doi.org/10.1017/9781108231596
[57] M. Vidyasagar. *A Theory of Learning and Generalization: With Applications to Neural Networks and Control Systems*. Springer-Verlag, Berlin, Heidelberg, 1997. MR1482231
[58] Y. Vorobeychik and M. Kantarcioglu. Adversarial machine learning. In *Synthesis Lectures on Artificial Intelligence and Machine Learning* 1–169, **12**, 2018.
[59] H. E. Warren. Lower bounds for approximation by nonlinear manifolds. *Trans. Amer. Math. Soc.* **133** (1) (1968) 167–178. MR0226281 https://doi.org/10.2307/1994937
[60] X. Wei and S. Minsker. Estimation of the covariance structure of heavy-tailed distributions. In *NIPS*, 2017.
[61] L. Wolf, H. Jhuang and T. Hazan. Modeling appearances with low-rank SVM. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* 1–6, 2007.
[62] Z. Zhang, Y. Xu, J. Yang, X. Li and D. Zhang. A survey of sparse representation: Algorithms and applications. *IEEE Access* **3** (2015) 490–530.
[63] S. Zinodiny, S. Rezaei and S. Nadarajah. Bayes minimax estimation of the mean matrix of matrix-variate normal distribution under balanced loss function. *Statist. Probab. Lett.* **125** (2017) 110–120. MR3626075 https://doi.org/10.1016/j.spl.2017.02.003
[64] S. Zinodiny, S. Rezaei and S. Nadarajah. Minimax estimation of the mean matrix of the matrix variate normal distribution under the divergence loss function. *Statistica* **77** (4) (2018) 369–384. MR3626075 https://doi.org/10.1016/j.spl.2017.02.003

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 990–1024
https://doi.org/10.1214/22-AIHP1340

Large deviations for random matrices in the orthogonal group and Stiefel manifold with applications to random projections of product distributions

Zakhar Kabluchko[1,a] and Joscha Prochno[2,b]

[1]*Faculty of Mathematics, University of Münster, Orléans-Ring 10, 48149 Münster, Germany,* [a]*zakhar.kabluchko@uni-muenster.de*
[2]*Faculty of Computer Science and Mathematics, University of Passau, Dr.-Hans-Kapfinger-Stra(ß)e 30, 94032 Passau, Germany,* [b]*joscha.prochno@uni-passau.de*

Abstract. We prove large deviation principles (LDPs) for random matrices in the orthogonal group and Stiefel manifold, determining both the speed and good convex rate functions that are explicitly given in terms of certain log-determinants of trace-class operators and are finite on the set of Hilbert-Schmidt operators M satisfying $\|MM^*\| < 1$. As an application of those LDPs, we determine the precise large deviation behavior of k-dimensional random projections of high-dimensional product distributions using an appropriate interpretation in terms of point processes, also characterizing the space of all possible deviations. The case of uniform distributions on ℓ_p-balls, $1 \le p \le \infty$, is then considered and reduced to appropriate product measures. Those applications generalize considerably the recent work (*Studia Mathematica* **264** (2022) 103–119).

Résumé. Nous prouvons des principes de grandes déviations (LDPs) pour les matrices aléatoires uniformes sur le groupe orthogonal et les variétés de Stiefel, en déterminant à la fois la vitesse et les bonnes fonctions de taux convexes qui sont explicitement données en termes de certains log-déterminants d'opérateurs à trace, et sont finies sur l'ensemble des opérateurs de Hilbert-Schmidt M satisfaisant $\|MM^*\| < 1$. Comme application de ces LDPs, nous déterminons le comportement précis des grandes déviations des projections aléatoires de dimension k des lois de produit de grande dimension en utilisant une interprétation appropriée en termes de processus ponctuels, caractérisant également l'espace de toutes les déviations possibles. Le cas des lois uniformes sur les boules ℓ_p, $1 \le p \le \infty$, est ensuite considéré et réduit à des mesures produit appropriées. Ces applications généralisent considérablement les travaux récents (*Studia Mathematica* **264** (2022) 103–119).

MSC2020 subject classifications: Primary 52A23; 60F10; 60B20; secondary 52A22; 46B06

Keywords: Large deviation principle; Matrix variate distributions; Orthogonal group; Product measure; Projective limit; Random matrix; Random projection; Stiefel manifold

References

[1] R. Adamczak, R. Latała, A. E. Litvak, A. Pajor and N. Tomczak-Jaegermann. Geometry of log-concave ensembles of random matrices and approximate reconstruction. *C. R. Math. Acad. Sci. Paris* **349** (13–14) (2011) 783–786. MR2825941 https://doi.org/10.1016/j.crma.2011.06.025
[2] G. Akemann, J. Baik and P. Di Francesco (Eds) *The Oxford Handbook of Random Matrix Theory*. Oxford University Press, Oxford, 2015. MR3410165
[3] G. Akinwande and M. Reitzner. Multivariate central limit theorems for random simplicial complexes. *Adv. in Appl. Math.* **121** (2020), 102076. MR4126722 https://doi.org/10.1016/j.aam.2020.102076
[4] D. Alonso-Gutiérrez and J. Prochno. Thin-shell concentration for random vectors in Orlicz balls via moderate deviations and Gibbs measures. *J. Funct. Anal.* **282** (1) (2022), 109291. MR4330166 https://doi.org/10.1016/j.jfa.2021.109291
[5] D. Alonso-Gutiérrez, J. Prochno and C. Thäle. Large deviations for high-dimensional random projections of ℓ_p^n-balls. *Adv. in Appl. Math.* **99** (2018) 1–35. MR3806754 https://doi.org/10.1016/j.aam.2018.04.003
[6] D. Alonso-Gutiérrez, J. Prochno and C. Thäle. Gaussian fluctuations for high-dimensional random projections of ℓ_p^n-balls. *Bernoulli* **25** (4A) (2019) 3139–3174. MR4003577 https://doi.org/10.3150/18-BEJ1084
[7] D. Alonso-Gutiérrez, J. Prochno and C. Thäle. Large deviations, moderate deviations, and the KLS conjecture. *J. Funct. Anal.* **280** (1) (2021), 108779. MR4156132 https://doi.org/10.1016/j.jfa.2020.108779
[8] G. W. Anderson, A. Guionnet and O. Zeitouni. *An Introduction to Random Matrices*. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge University Press, Cambridge, 2010. MR2760897

[9] M. Anttila, K. M. Ball and I. Perissinaki. The central limit problem for convex bodies. *Trans. Amer. Math. Soc.* **355** (12) (2003) 4723–4735. MR1997580 https://doi.org/10.1090/S0002-9947-03-03085-X
[10] S. Artstein, K. M. Ball, F. Barthe and A. Naor. On the rate of convergence in the entropic central limit theorem. *Probab. Theory Related Fields* **129** (3) (2004) 381–390. MR2128238 https://doi.org/10.1007/s00440-003-0329-4
[11] F. Barthe, F. Gamboa, L. Lozada-Chang and A. Rouault. Generalized Dirichlet distributions on the ball and moments. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** (2010) 319–340. MR2737561
[12] F. Barthe and P. Wolff. Volume properties of high-dimensional Orlicz balls. 2021. Available at arXiv:2106.01675.
[13] P. Billingsley. *Convergence of Probability Measures*, 2nd edition. *Wiley Series in Probability and Statistics: Probability and Statistics*. John Wiley & Sons, Inc., New York, 1999. MR1700749 https://doi.org/10.1002/9780470316962
[14] D. Chafaï, O. Guédon, G. Lecué and A. Pajor. *Interactions Between Compressed Sensing Random Matrices and High Dimensional Geometry*. *Panoramas et Synthèses [Panoramas and Syntheses]* **37**. Société Mathématique de France, Paris, 2012. MR3113826
[15] B. Dadoun, M. Fradelizi, O. Guédon and P.-A. Zitt Asymptotics of the inertia moments and the variance conjecture in Schatten balls. Preprint, 2021. MR4500732 https://doi.org/10.1016/j.jfa.2022.109741
[16] D. A. Dawson and J. Gärtner. Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** (4) (1987) 247–308. MR0885876 https://doi.org/10.1080/17442508708833446
[17] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability*, 2nd edition. Springer-Verlag, Berlin Heidelberg, 2010. MR2571413 https://doi.org/10.1007/978-3-642-03311-7
[18] F. den Hollander. *Large Deviations*. *Fields Institute Monographs* **14**. American Mathematical Society, Providence, RI, 2000. MR1739680 https://doi.org/10.1007/s00440-009-0235-5
[19] P. W. Diaconis, M. L. Eaton and S. L. Lauritzen. Finite de Finetti theorems in linear models and multivariate analysis. *Scand. J. Stat.* **19** (4) (1992) 289–315. MR1211786
[20] J. M. Dickey. Matrixvariate generalizations of the multivariate t distribution and the inverted multivariate t distribution. *Ann. Math. Stat.* **38** (1967) 511–518. MR0208752 https://doi.org/10.1214/aoms/1177698967
[21] H. Dym. *Linear Algebra in Action*, 2nd edition. *Graduate Studies in Mathematics* **78**. American Mathematical Society, Providence, RI, 2013. MR3154813 https://doi.org/10.1090/gsm/078
[22] M. L. Eaton. Group invariance applications in statistics. *NSF-CBMS Regional Conference Series in Probability and Statistics. Institute of Mathematical Statistics* **1** (1989). MR1089423
[23] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing. Applied and Numerical Harmonic Analysis*. Birkhäuser/Springer, New York, 2013. MR3100033 https://doi.org/10.1007/978-0-8176-4948-7
[24] N. Gantert, S. S. Kim and K. Ramanan. Cramér's theorem is atypical. In *Advances in the Mathematical Sciences* 253–270. *Assoc. Women Math. Ser.* **6**. Springer, Cham, 2016. MR3654500 https://doi.org/10.1007/978-3-319-34139-2_11
[25] N. Gantert, S. S. Kim and K. Ramanan. Large deviations for random projections of ℓ^p balls. *Ann. Probab.* **45** (6B) (2017) 4419–4476. MR3737915 https://doi.org/10.1214/16-AOP1169
[26] I. Gohberg, S. Goldberg and N. Krupnik. *Traces and Determinants of Linear Operators*. *Operator Theory: Advances and Applications.* **116**. Birkhäuser Verlag, Basel, 2000. MR1744872 https://doi.org/10.1007/978-3-0348-8401-3
[27] I. C. Gohberg and M. G. Kreĭn. *Introduction to the Theory of Linear Nonselfadjoint Operators*. *Translations of Mathematical Monographs* **18**. American Mathematical Society, Providence, R.I., 1969. MR0246142
[28] A. Guionnet. Large deviations and stochastic calculus for large random matrices. *Probab. Surv.* **1** (2004) 72–172. MR2095566 https://doi.org/10.1214/154957804100000033
[29] A. K. Gupta and D. K. Nagar. *Matrix Variate Distributions*. *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics* **104**. Chapman & Hall/CRC, Boca Raton, FL, 2000. MR1738933
[30] A. Hinrichs, D. Krieg, E. Novak, J. Prochno and M. Ullrich. Random sections of ellipsoids and the power of random information. *Trans. Amer. Math. Soc.* **374** (12) (2021) 8691–8713. MR4337926 https://doi.org/10.1090/tran/8502
[31] A. Hinrichs, J. Prochno and M. Sonnleitner. Random sections of ℓ_p-ellipsoids, optimal recovery and Gelfand numbers of diagonal operators, 2021. Available at arXiv:2109.14504.
[32] A. Hinrichs, J. Prochno and M. Ullrich. The curse of dimensionality for numerical integration on general domains. *J. Complexity* **50** (2019) 25–42. MR3907362 https://doi.org/10.1016/j.jco.2018.08.003
[33] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1994. MR1288752
[34] T. Jiang. Maxima of entries of Haar distributed matrices. *Probab. Theory Related Fields* **131** (1) (2005) 121–144. MR2105046 https://doi.org/10.1007/s00440-004-0376-5
[35] T. Jiang. How many entries of a typical orthogonal matrix can be approximated by independent normals? *Ann. Probab.* **34** (4) (2006) 1497–1529. MR2257653 https://doi.org/10.1214/009117906000000205
[36] T. Jiang. The entries of Haar-invariant matrices from the classical compact groups. *J. Theoret. Probab.* **23** (4) (2010) 1227–1243. MR2735744 https://doi.org/10.1007/s10959-009-0241-7
[37] S. G. G. Johnston, Z. Kabluchko and J. Prochno. Projections of the uniform distribution on the cube – a large deviation perspective. *Studia Math.* **264** (2022) 103–119. MR4440685 https://doi.org/10.4064/sm210413-16-9
[38] S. G. G. Johnston and J. Prochno. A Maxwell principle for generalized Orlicz balls. *Ann. Inst. H. Poincaré Probab. Statist.* To appear (2022).
[39] Z. Kabluchko and J. Prochno. The maximum entropy principle and volumetric properties of Orlicz balls. *J. Math. Anal. Appl.* **495** (1) (2021), 124687. MR4172842 https://doi.org/10.1016/j.jmaa.2020.124687
[40] Z. Kabluchko, J. Prochno and C. Thäle. High-dimensional limit theorems for random vectors in ℓ_p^n-balls. *Commun. Contemp. Math.* **21** (1) (2019), 1750092. MR3904638 https://doi.org/10.1142/S0219199717500924
[41] Z. Kabluchko, J. Prochno and C. Thäle. Sanov-type large deviations in Schatten classes. *Ann. Inst. Henri Poincaré Probab. Stat.* **56** (2) (2020) 928–953. MR4076771 https://doi.org/10.1214/19-AIHP989
[42] Z. Kabluchko, J. Prochno and C. Thäle. High-dimensional limit theorems for random vectors in ℓ_p^n-balls. II. *Commun. Contemp. Math.* **23** (3) (2021), 1950073. MR4216415 https://doi.org/10.1142/S0219199719500731
[43] Z. Kabluchko, J. Prochno and C. Thäle. A new look at random projections of the cube and general product measures. *Bernoulli* **27** (3) (2021) 2117–2138. MR4260512 https://doi.org/10.3150/20-bej1303

[44] A. M. Kagan, Yu. V. Linnik and C. R. Rao. *Characterization Problems in Mathematical Statistics*. *Wiley Series in Probability and Mathematical Statistics*. John Wiley & Sons, New York–London–Sydney, 1973. MR0346969
[45] O. Kallenberg. *Foundations of Modern Probability. Probability and Its Applications*, 2nd edition. Springer-Verlag, New York, 2002. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[46] T. Kaufmann. Sharp asymptotics for q-norms of random vectors in high-dimensional ℓ_p^n-balls. *Mod. Stoch. Theory Appl.* **8** (2) (2021) 239–274. MR4279877 https://doi.org/10.15559/21-vmsta182
[47] J. Kerstan, K. Matthes and J. Mecke. *Infinitely Divisible Point Processes*. *Wiley Series in Probability and Mathematical Statistics*. Wiley, New York, 1978. MR0517931
[48] C. G. Khatri. On the mutual independence of certain statistics. *Ann. Math. Stat.* **30** (1959) 1258–1262. MR0110135 https://doi.org/10.1214/aoms/1177706112
[49] C. G. Khatri. A note on Mitra's paper "A density free approach to the matrix variate beta distribution". *Sankhyā Ser. A* **32** (3) (1970) 311–318. MR0293757
[50] S. S. Kim, Y.-T. Liao and K. Ramanan. An asymptotic thin shell condition and large deviations for random multidimensional projections. *Adv. in Appl. Math.* **134** (2022), 102306. MR4352552 https://doi.org/10.1016/j.aam.2021.102306
[51] S. S. Kim and K. Ramanan. A conditional limit theorem for high-dimensional ℓ^p-spheres. *J. Appl. Probab.* **55** (4) (2018) 1060–1077. MR3899928 https://doi.org/10.1017/jpr.2018.71
[52] S. S. Kim and K. Ramanan. Large deviation principles induced by the Stiefel manifold, and random multi-dimensional projections, 2021. Available at arXiv:2105.04685.
[53] B. Klartag. A central limit theorem for convex sets. *Invent. Math.* **168** (1) (2007) 91–131. MR2285748 https://doi.org/10.1007/s00222-006-0028-8
[54] A. Klenke. *Probability Theory: A Comprehensive Course*, 2nd edition. *Universitext.* Springer, London, 2014. MR3112259 https://doi.org/10.1007/978-1-4471-5361-0
[55] H. König. *Eigenvalue Distribution of Compact Operators*. *Operator Theory: Advances and Applications.* **16**. Birkhäuser Verlag, Basel, 1986. MR0889455 https://doi.org/10.1007/978-3-0348-6278-3
[56] D. Krieg and M. Sonnleitner. Random points are optimal for the approximation of Sobolev functions, 2020. Available at arXiv:2009.11275.
[57] Y.-T. Liao and K. Ramanan. Geometric sharp large deviations for random projections of ℓ_p^n spheres, 2020. Available at arXiv:2001.04053. MR4352552 https://doi.org/10.1016/j.aam.2021.102306
[58] Ju. V. Linnik. Linear forms and statistical criteria. II. In *Selected Transl. Math. Statist. and Prob.* 41–90, **3**. Amer. Math. Soc., Providence, R.I., 1962. MR0154361
[59] Ju. V. Linnik. Linear forms and statistical criteria. I. In *Selected Transl. Math. Statist. and Prob.* 1–40, **3**. Amer. Math. Soc., Providence, R.I., 1963. MR0154360
[60] Yu. V. Linnik. Linear forms and statistical criteria. I, II. *Ukr. Mat. Zh.* **5** (1953) 207–243, 247–290. MR0060767
[61] E. Lukacs. Some extensions of a theorem of Marcinkiewicz. *Pacific J. Math.* **8** (1958) 487–501. MR0101543
[62] E. Lukacs. *Characteristic Functions*. Charles Griffin and Company, London, 1970. 350 p. MR0346874
[63] J. Marcinkiewicz. Sur une propriété de la loi de Gauß. *Math. Z.* **44** (1) (1939) 612–618. MR1545791 https://doi.org/10.1007/BF01210677
[64] E. S. Meckes and M. W. Meckes. The central limit problem for random vectors with symmetries. *J. Theoret. Probab.* **20** (4) (2007) 697–720. MR2359052 https://doi.org/10.1007/s10959-007-0119-5
[65] M. Moeller and T. Ullrich. L_2-norm sampling discretization and recovery of functions from RKHS with finite trace. *Sampl. Theory Signal Process. Data Anal.* **19** (2021) 13. MR4354442 https://doi.org/10.1007/s43670-021-00013-3
[66] G. Paouris, P. Pivovarov and J. Zinn. A central limit theorem for projections of the cube. *Probab. Theory Related Fields* **159** (3–4) (2014) 701–719. MR3230006 https://doi.org/10.1007/s00440-013-0518-8
[67] A. Pietsch. *Operator Ideals*. *North-Holland Mathematical Library* **20**. North-Holland Publishing Co., Amsterdam-New York, 1980. MR0582655
[68] A. Pietsch. *Eigenvalues and s-Numbers*. *Cambridge Studies in Advanced Mathematics* **13**. Cambridge University Press, Cambridge, 1987. MR0890520
[69] Y. V. Prohorov. Convergence of random processes and limit theorems in probability theory. *Teor. Veroyatn. Primen.* **1** (1956) 177–238. MR0084896
[70] M. Reitzner. Central limit theorems for random polytopes. *Probab. Theory Related Fields* **133** (4) (2005) 483–507. MR2197111 https://doi.org/10.1007/s00440-005-0441-8
[71] S. I. Resnick. *Extreme Values, Regular Variation and Point Processes*. *Springer Series in Operations Research and Financial Engineering.* Springer, New York, 2008. MR2364939
[72] G. Schechtman and M. Schmuckenschläger. Another remark on the volume of the intersection of two L_p^n balls. In *Geometric Aspects of Functional Analysis (1989–90)* 174–178. *Lecture Notes in Math.* **1469**. Springer, Berlin, 1991. MR1122622 https://doi.org/10.1007/BFb0089224
[73] G. Schechtman and J. Zinn. On the volume of the intersection of two L_p^n balls. *Proc. Amer. Math. Soc.* **110** (1) (1990) 217–224. MR1015684 https://doi.org/10.2307/2048262
[74] M. Schmuckenschläger. CLT and the volume of intersections of l_p^n-balls. *Geom. Dedicata* **85** (1–3) (2001) 189–195. MR1845607 https://doi.org/10.1023/A:1010353121014
[75] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A** (1/2) (1928) 32–52. MR0050223 https://doi.org/10.1093/biomet/39.1-2.1
[76] A. A. Zinger. On samples with identically distributed linear statistics. *Theory Probab. Appl.* **20** (3) (1977) 655–660. MR0381077
[77] A. A. Zinger. On the characterization of the normal law by identically distributed linear statistics. *Sankhyā Ser. A* **39** (3) (1977) 232–242. MR0518913

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1025–1047
https://doi.org/10.1214/22-AIHP1357

Random sorting networks: Edge limit

Vadim Gorin[1,a] and Jiaming Xu[2,b]

[1]*Departments of Statistics and Mathematics, University of California, Berkeley, USA,* [a]*vadicgor@gmail.com*
[2]*Department of Mathematics, University of Wisconsin – Madison, USA,* [b]*jxu385@wisc.edu*

Abstract. A sorting network is a shortest path from $1\ 2\ \dots\ n$ to $n\ \dots\ 2\ 1$ in the Cayley graph of the symmetric group $\mathfrak{S}_n$ spanned by adjacent transpositions. The paper computes the edge local limit of the uniformly random sorting networks as $n \to \infty$. We find the asymptotic distribution of the first occurrence of a given swap $(k, k+1)$ and identify it with the law of the smallest positive eigenvalue of a $2k \times 2k$ aGUE (an aGUE matrix has purely imaginary Gaussian entries that are independently distributed subject to skew-symmetry). Next, we give two different formal definitions of a spacing – the time distance between the occurrence of a given swap $(k, k+1)$ in a uniformly random sorting network. Two definitions lead to two different expressions for the asymptotic laws expressed in terms of derivatives of Fredholm determinants.

Résumé. Un réseau de tri est un chemin le plus court de $1\ 2\ \dots\ n$ à $n\ \dots\ 2\ 1$ dans le graphe de Cayley du groupe symétrique $\mathfrak{S}_n$, engendré par des transpositions des éléments adjacents. Dans cet article nous calculons la limite locale au bord des réseaux de tri choisi uniformément quand $n \to \infty$. Nous trouvons la distribution asymptotique de la première occurrence d'une transposition donnée $(k, k+1)$ et l'identifions avec la loi de la plus petite valeur propre positive d'un $2k \times 2k$ aGUE (une matrice aGUE a des entrées gaussiennes purement imaginaires qui sont distribuées indépendamment sous condition d'antisymétrie). Ensuite, nous considerons des espacements entre deux occurrences consecutives d'un échange donné (k, k + 1) pour un réseau de tri aléatoire choisi uniformément. Nous prenons deux formalisations pour un choix aléatoire d'un tel espacement. En passant à limite, ces deux définitions conduisent à deux expressions différentes pour des lois asymptotiques exprimées en termes de dérivées des déterminants de Fredholm.

MSC2020 subject classifications: 60B20; 60B15

Keywords: Sorting network; Spacing; Antisymmetric Gaussian Unitary Ensemble

References

[1] O. Angel, D. Dauvergne, A. E. Holroyd and B. Virag. The local limit of random sorting networks. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (1) (2019) 412–440. Available at arXiv:1702.08368.
[2] O. Angel, V. Gorin and A. E. Holroyd. A pattern theorem for random sorting networks. *Electron. J. Probab.* **17** (2012) 99. Available at arXiv:1110.0160.
[3] O. Angel, A. Holroyd, D. Romik and B. Virág. Random sorting networks. *Adv. Math.* **215** (2) (2007) 839–864. Available at arXiv:0609.538.
[4] A. Borodin. *Determinantal Point Processes*. *Oxford Handbook of Random Matrix Theory*. Oxford University Press, London, 2011. Available at arXiv:0911.1153.
[5] D. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Vol. I. Elementary Theory and Methods*. Springer–Verlag, New York, 2003.
[6] D. Dauvergne. The Archimedean limit of random sorting networks. *J. Amer. Math. Soc.* **35** (2022) 1215–1267. Available at arXiv:1802.08934.
[7] D. Dauvergne and B. Virag. Circular support in random sorting networks. *Trans. Amer. Math. Soc.* **373** (2020) 1529–1553. Available at arXiv:1802.08933.
[8] P. Edelman and C. Greene. Balanced tableaux. *Adv. Math.* **63** (1) (1987) 42–99.
[9] P. J. Forrester and E. Nordenstam. The anti-symmetric GUE minor process. *Mosc. Math. J.* **9** (4) (2009) 749–774. Available at arXiv:0804.3293.
[10] J. S. Frame, G. B. Robinson and R. M. Thrall. The hook graphs of the symmetric groups. *Canad. J. Math.* **6** (1954) 316–324.
[11] V. Gorin and M. Rahman. Random sorting networks: Local statistics via random matrix laws. *Probab. Theory Related Fields* **175** (1) (2019) 45–96.
[12] N. Johnson, S. Kotz and N. Balakrishnan. *Continuous Univariate Distributions*, **2**, 2nd edition. Wiley, New York, 1995.
[13] M. Kotowski. Limits of random permuton processes and large deviations for the interchange process. Ph.D. Thesis, University of Toronto, 2016.
[14] A. Lenard. Correlation functions and the uniqueness of the state in classical statistical mechanics. *Comm. Math. Phys.* **30** (1973) 35–44.
[15] M. L. Mehta. *Random Matrices*, 3rd edition. *Pure and Applied Mathematics (Amsterdam)* **142**. Elsevier/Academic Press, Amsterdam, 2004.
[16] L. Petrov. Asymptotics of random lozenge tilings via Gelfand–Tsetlin schemes. *Probab. Theory Related Fields* **160** (2014) 429–487.

[17] M. Rahman, B. Virag and M. Vizer. Geometry of permutation limits. *Combinatorica* **39** (2019) 933–960. Available at arXiv:1609.03891.
[18] A. Rozinov. Statistics of Random Sorting Networks. Ph.D. Thesis, Courant Institute, NYU, 2016.
[19] R. P. Stanley. On the number of reduced decompositions of elements of Coxeter groups. *European J. Combin.* **5** (4) (1984) 359–372.
[20] G. Szego. *Orthogonal Polynomials*. American Mathematical Society, Providence, 1939.

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1048–1076
https://doi.org/10.1214/23-AIHP1368

Log determinant of large correlation matrices under infinite fourth moment

Johannes Heiny[1,a] and Nestor Parolya[2,b]

[1]*Department of Mathematics, Stockholm University, Stockholm, Sweden,* [a]*johannes.heiny@math.su.se*
[2]*Delft Institute of Applied Mathematics, Delft Universtity of Technology, Delft, the Netherlands,* [b]*n.parolya@tudelft.nl*

Abstract. In this paper, we show the central limit theorem for the logarithmic determinant of the sample correlation matrix $\mathbf{R}$ constructed from the $(p \times n)$-dimensional data matrix $\mathbf{X}$ containing independent and identically distributed random entries with mean zero, variance one and infinite fourth moments. Precisely, we show that for $p/n \to \gamma \in (0,1)$ as $n, p \to \infty$ the *logarithmic law*

$$\frac{\log\det\mathbf{R} - (p - n + \frac{1}{2})\log(1 - p/n) + p - p/n}{\sqrt{-2\log(1 - p/n) - 2p/n}} \xrightarrow{d} N(0,1)$$

is still valid if the entries of the data matrix $\mathbf{X}$ follow a symmetric distribution with a regularly varying tail of index $\alpha \in (3,4)$. The latter assumptions seem to be crucial, which is justified by the simulations: if the entries of $\mathbf{X}$ have the infinite absolute third moment and/or their distribution is not symmetric, the logarithmic law is not valid anymore. The derived results highlight that the logarithmic determinant of the sample correlation matrix is a very stable and flexible statistic for heavy-tailed big data and open a novel way of analysis of high-dimensional random matrices with self-normalized entries.

Résumé. Dans cet article, nous démontrons le théorème de la limite centrale pour le déterminant logarithmique d'une matrice de corrélation $\mathbf{R}$ construite à partir d'une matrice de données $\mathbf{X}$ de taille $(p \times n)$ contenant des entrées indépendantes d'espérance 0, variance 1 et quatrième moment infini. Plus précisément, nous démontrons que dans le régime $p/n \to \gamma \in (0,1)$ quand $n, p \to \infty$ la *loi logarithmique*

$$\frac{\log\det\mathbf{R} - (p - n + \frac{1}{2})\log(1 - p/n) + p - p/n}{\sqrt{-2\log(1 - p/n) - 2p/n}} \xrightarrow{d} N(0,1)$$

est toujours valable si les entrées de la matrice de données $\mathbf{X}$ suivent une distribution symétrique avec une queue à variation régulière d'indice $\alpha \in (3,4)$. Ces dernières conditions semblent être cruciales, ce qui est justifié par les simulations : si les entrées de $\mathbf{X}$ n'ont pas de troisième moment et/ou si leur distribution n'est pas symétrique, la loi logarithmique n'est plus valable. Les résultats obtenus mettent en évidence que le déterminant logarithmique d'une matrice de corrélation est une statistique très stable et flexible pour les données massives à queue lourde et ouvrent une nouvelle voie pour analyser les grandes matrices aléatoires avec entrées auto-normalisées.

MSC2020 subject classifications: Primary 60B20; secondary 60F05; 60G10; 60G57; 60G70

Keywords: Sample correlation matrix; Logarithmic determinant; Random matrix theory; Heavy tails; Infinite fourth moment

References

[1] H. Albrecher and J. L. Teugels. Asymptotic analysis of a measure of variation. *Theory Probab. Math. Statist.* **74** (2007) 1–10. MR2336773 https://doi.org/10.1090/S0094-9000-07-00692-8
[2] S. Anatolyev and P. Yaskov. Asymptotics of diagonal elements of projection matrices under many instruments/regressors. *Econometric Theory* **33** (3) (2017) 717–738. MR3637973 https://doi.org/10.1017/S0266466616000165
[3] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, New Jersey, 2003. MR0091588
[4] Z. Bai and J. W. Silverstein. *Spectral Analysis of Large Dimensional Random Matrices*, 2nd edition. *Springer Series in Statistics*, xvi+551. Springer, New York, 2010. MR2567175 https://doi.org/10.1007/978-1-4419-0661-8
[5] Z. Bai and W. Zhou. Large sample covariance matrices without independence structures in columns. *Statist. Sinica* **18** (2) (2008) 425–442. MR2411613
[6] Z. D. Bai and J. W. Silverstein. CLT for linear spectral statistics of large dimensional sample covariance matrices. *Ann. Probab.* **32** (2004) 553–605. MR2040792 https://doi.org/10.1214/aop/1078415845

[7] Z. Bao, G. Pan and W. Zhou. Tracy-Widom law for the extreme eigenvalues of sample correlation matrices. *Electron. J. Probab.* **17** (2012), 32 pp. MR2988403 https://doi.org/10.1214/EJP.v17-1962

[8] Z. Bao, G. Pan and W. Zhou. The logarithmic law of random determinant. *Bernoulli* **21** (3) (2015) 1600–1628. MR3352055 https://doi.org/10.3150/14-BEJ615

[9] N. H. Bingham, C. M. Goldie and J. L. Teugels. *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**, xx+491. Cambridge University Press, Cambridge, 1987. MR898871 https://doi.org/10.1017/CBO9780511721434

[10] T. Bodnar, H. Dette and N. Parolya. Testing for independence of large dimensional vectors. *Ann. Statist.* **47** (5) (2019) 2977–3008. MR3988779 https://doi.org/10.1214/18-AOS1771

[11] A. Dembo. On random determinants. *Quart. Appl. Math.* **47** (2) (1989) 185–195.

[12] H. Dette and N. Dörnemann. Likelihood ratio tests for many groups in high dimensions. *J. Multivariate Anal.* **178** (2020), 104605, 16. MR4079037 https://doi.org/10.1016/j.jmva.2020.104605

[13] N. Dörnemann. Likelihood ratio tests under model misspecification in high dimensions. *J. Multivariate Anal.* **193** (2023), Paper No. 105122. MR4504582 https://doi.org/10.1016/j.jmva.2022.105122

[14] N. El Karoui. Tracy-Widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.* **35** (2) (2007) 663–714. MR2308592 https://doi.org/10.1214/009117906000000917

[15] N. El Karoui. Concentration of measure and spectra of random matrices: Applications to correlation matrices, elliptical distributions and beyond. *Ann. Appl. Probab.* **19** (6) (2009) 2362–2405. MR2588248 https://doi.org/10.1214/08-AAP548

[16] L. Erdös and H.-T. Yau. *A Dynamical Approach to Random Matrix Theory*. *Courant Lecture Notes in Mathematics* **28**, ix+226. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2017. Available at http://www.math.harvard.edu/~htyau/RM-Aug-2016.pdf. MR3699468

[17] A. Fuchs, A. Joffe and J. Teugels. Expectation of the ratio of the sum of squares to the square of the sum: Exact and asymptotic results. *Teor. Veroyatn. Primen.* **46** (2) (2001) 297–310. MR1968687 https://doi.org/10.1137/S0040585X97978919

[18] J. Gao, X. Han, G. Pan and Y. Yang. High dimensional correlation matrices: The central limit theorem and its applications. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** (3) (2017) 677–693. https://doi.org/10.1111/rssb.12189

[19] V. Girko. Central limit theorem for random determinants. In *Theory of Probability & Its Applications* 846–846, **23**, 1979. SIAM. MR0550240

[20] V. L. Girko. A refinement of the central limit theorem for random determinants. *Theory Probab. Appl.* **42** (1) (1998) 121–129. MR1453330 https://doi.org/10.1137/S0040585X97975939

[21] N. R. Goodman. The distribution of the determinant of a complex Wishart distributed matrix. *Ann. Math. Stat.* **34** (1) (1963) 178–180. MR0145619 https://doi.org/10.1214/aoms/1177704251

[22] J. Grote, Z. Kabluchko and C. Thäle. Limit theorems for random simplices in high dimensions. *ALEA Lat. Am. J. Probab. Math. Stat.* **16** (1) (2019) 141–177. MR3903027 https://doi.org/10.30757/alea.v16-06

[23] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application*. *Probability and Mathematical Statistics*, xii+308. Academic Press, New York–London, 1980. MR624435

[24] J. Heiny. Large sample correlation matrices: A comparison theorem and its applications. *Electron. J. Probab.* **27** (2022), Paper No. 94, 20. MR4456777 https://doi.org/10.1214/22-ejp817

[25] J. Heiny, S. Johnston and J. Prochno. Thin-shell theory for rotationally invariant random simplices. *Electron. J. Probab.* **27** (2022) 1–41. MR4362288 https://doi.org/10.1214/21-ejp734

[26] J. Heiny and T. Mikosch. Almost sure convergence of the largest and smallest eigenvalues of high-dimensional sample correlation matrices. *Stochastic Process. Appl.* **128** (8) (2018) 2779–2815. Available at https://www.sciencedirect.com/science/article/pii/S0304414917302533. MR3811704 https://doi.org/10.1016/j.spa.2017.10.002

[27] J. Heiny and J. Yao. Limiting distributions for eigenvalues of sample correlation matrices from heavy-tailed populations. *Ann. Statist.* **50** (6) (2022) 3249–3280. MR4524496 https://doi.org/10.1214/22-aos2226

[28] T. Jiang. The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā* **66** (1) (2004) 35–48. Available at http://www.jstor.org/stable/25053330.

[29] T. Jiang and F. Yang. Central limit theorems for classical likelihood ratio tests for high-dimensional normal distributions. *Ann. Statist.* **41** (4) (2013) 2029–2074. https://doi.org/10.1214/13-AOS1134

[30] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)* **72** (114) (1967) 507–536. MR0208649

[31] M. Mohammadi. On the bounds for diagonal and off-diagonal elements of the hat matrix in the linear regression model. *REVSTAT* **14** (2016) 75–87. MR3474087

[32] D. Morales-Jimenez, I. M. Johnstone, M. R. McKay and J. Yang. Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. *Statist. Sinica* **31** (2) (2021) 571–601. MR4286186 https://doi.org/10.5705/ss.20

[33] H. H. Nguyen and V. Vu. Random matrices: Law of the determinant. *Ann. Probab.* **42** (1) (2014) 146–167. https://doi.org/10.1214/12-AOP791

[34] J. Nielsen. The distribution of volume reductions induced by isotropic random projections. *Adv. in Appl. Probab.* **31** (4) (1999) 985–994. MR1747452 https://doi.org/10.1239/aap/1029955254

[35] N. Parolya, J. Heiny and D. Kurowicka. Logarithmic law of large random correlation matrix. *arXiv preprint arXiv:2103.13900* (2021).

[36] J. W. Silverstein and S.-I. Choi. Analysis of the limiting spectral distribution of large dimensional random matrices. *J. Multivariate Anal.* **54** (1995) 295–309. MR1345541 https://doi.org/10.1006/jmva.1995.1058

[37] T. Tao and V. Vu. A central limit theorem for the determinant of a Wigner matrix. *Adv. Math.* **231** (1) (2012) 74–101. Available at https://www.sciencedirect.com/science/article/pii/S0001870812001806. MR2935384 https://doi.org/10.1016/j.aim.2012.05.006

[38] Q. Wang and J. Yao. On the sphericity test with large-dimensional observations. *Electron. J. Stat.* **7** (2013) 2164–2192. MR3104916 https://doi.org/10.1214/13-EJS842

[39] X. Wang, X. Han and G. Pan. The logarithmic law of sample covariance matrices near singularity. *Bernoulli* **24** (1) (2018) 80–114. MR3706751 https://doi.org/10.3150/16-BEJ867

[40] D. P. Wiens. On moments of quadratic forms in non-spherically distributed variables. *Statistics* **23** (3) (1992) 265–270. MR1237804 https://doi.org/10.1080/02331889208802374

[41] X. Yang, X. Zheng and J. Chen. Testing high-dimensional covariance matrices under the elliptical distribution and beyond. *J. Econometrics* **221** (2) (2021) 409–423. Available at https://www.sciencedirect.com/science/article/pii/S0304407620302384. MR4215033 https://doi.org/10.1016/j.jeconom.2020.05.017
[42] J. Yao, Z. Bai and S. Zheng. *Large Sample Covariance Matrices and High-Dimensional Data Analysis (No. 39)*. Cambridge University Press, New York, 2015. MR3468554 https://doi.org/10.1017/CBO9781107588080

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1077–1089
https://doi.org/10.1214/23-AIHP1366

Exponential concentration for the number of roots of random trigonometric polynomials

Hoi H. Nguyen[1,a] and Ofer Zeitouni[2,b]

[1]*Department of Mathematics, The Ohio State University, 231 W 18th Ave, Columbus, OH 43210, USA,* [a]*nguyen.1261@osu.edu*
[2]*Faculty of Mathematics, Weizmann Institute and Courant Institute, NYU, Rehovot 76100, Israel and NY 10012, USA,* [b]*ofer.zeitouni@weizmann.ac.il*

Abstract. We show that the number of real roots of random trigonometric polynomials with i.i.d. coefficients, which are either bounded or satisfy the logarithmic Sobolev inequality, satisfies an exponential concentration of measure.

Résumé. Nous montrons que le nombre des racines réelles de polynômes trigonométriques aléatoires avec des coefficients i.i.d., qui sont soit bornés soit satisfont l'inégalité de Sobolev logarithmique, vérifie une concentration exponentielle de mesure.

MSC2020 subject classifications: 60F10; 30C15

Keywords: Random polynomials; Concentration; Universality

References

[1] J. Angst and G. Poly. Variations on Salem–Zygmund results for random trigonometric polynomials: Application to almost sure nodal asymptotics. *Electron. J. Probab.* **26** (2021) 1–36. MR4350983 https://doi.org/10.1214/21-ejp716
[2] J. M. Azaïs and J. León. CLT for crossings of random trigonometric polynomials. *Electron. J. Probab.* **18** (68) (2013) 1–17. MR3084654 https://doi.org/10.1214/EJP.v18-2403
[3] R. Basu, A. Dembo, N. Feldheim and O. Zeitouni. Exponential concentration for zeroes of stationary Gaussian processes. *Int. Math. Res. Not.* **23** (2020) 9769–9796. MR4182810 https://doi.org/10.1093/imrn/rny277
[4] V. Bally, L. Caramellino and G. Poly. Non universality for the variance of the number of real roots of random trigonometric polynomials. *Probab. Theory Related Fields* **174** (2019) 887–927. MR3980307 https://doi.org/10.1007/s00440-018-0869-2
[5] P. Borwein and T. Erdélyi. *Polynomials and Polynomial Inequalities*. *Graduate Texts in Mathematics*. Springer, Berlin–New York, 1995. MR1367960 https://doi.org/10.1007/978-1-4612-0793-1
[6] Y. Do, H. Nguyen and V. Vu. Real roots of random polynomials: Expectation and repulsion. *Proc. Lond. Math. Soc.* **111** (6) (2015) 1231–1260. MR3447793 https://doi.org/10.1112/plms/pdv055
[7] Y. Do, H. Nguyen and O. Nguyen. Random trigonometric polynomials: Universality and non-universality of the variance of the number of real roots. *Ann. Inst. Henri Poincaré Probab. Stat.* **58** (3) (2022) 1460–1504. MR4452640 https://doi.org/10.1214/21-aihp1206
[8] A. Edelman and E. Kostlan. How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc. (N.S.)* **32** (1995) 1–37. Erratum: Bull. Amer. Math. Soc. (N.S.) 33 (1996), 325.. MR1290398 https://doi.org/10.1090/S0273-0979-1995-00571-9
[9] O. N. Feldheim and A. Sen. Double roots of random polynomials with integer coefficients. *Electron. J. Probab.* **22** (2017) paper no. 10 MR3613703 https://doi.org/10.1214/17-EJP24
[10] D. Gayet and J.-Y. Welschinger. Exponential rarefaction of real curves with many components. *Publ. Math. IHES* **113** (2011) 69–93. MR2805598 https://doi.org/10.1007/s10240-011-0033-3
[11] A. Granville and I. Wigman. The distribution of the zeros of random trigonometric polynomials. *Amer. J. Math.* **133** (2) (2011) 295–357. MR2797349 https://doi.org/10.1353/ajm.2011.0015
[12] H. Iwaniec and E. Kowalski. *Analytic Number Theorem*. *Colloqium Publications* **53**. AMS, Providence, RI, 2004. MR2061214 https://doi.org/10.1090/coll/053
[13] M. Ledoux. *The Concentration of Measure Phenomenon*. *Mathematical Surveys and Monographs* **89**. AMS, Providence, RI, 2001. MR1849347 https://doi.org/10.1090/surv/089
[14] D. S. Lubinsky, A. Mate and P. Nevai. Quadrature sums involving p-th powers of polynomials. *SIAM J. Math. Anal.* **18** (1987) 531–544. MR0876290 https://doi.org/10.1137/0518041
[15] F. Nazarov and M. Sodin. On the number of nodal domains of random spherical harmonics. *Amer. J. Math.* **131** (2009) 1337–1357. MR2555843 https://doi.org/10.1353/ajm.0.0070
[16] H. Nguyen, O. Nguyen and V. Vu. On the number of real roots of random polynomials. *Commun. Contemp. Math.* **18** (4) (2016) 1550052. MR3493213 https://doi.org/10.1142/S0219199715500522

[17] O. Nguyen and V. Vu. Roots of random functions: A framework for local universality. *Amer. J. Math.* **144** (01) (2022) 1–74. MR4367414 https://doi.org/10.1353/ajm.2022.0000

[18] H. H. Nguyen and O. Zeitouni. Exponential concentration for the number of roots of random trigonometric polynomials. Available at arXiv:1912.12051v1.

[19] R. Peled, A. Sen and O. Zeitouni. Double roots of random Littlewood polynomials. *Israel J. Math.* **213** (2016) 55–77. MR3509468 https://doi.org/10.1007/s11856-016-1328-3

[20] C. Qualls. On the number of zeros of a stationary Gaussian random trigonometric polynomial. *J. Lond. Math. Soc. (2)* **2** (1970) 216–220. MR0258110 https://doi.org/10.1112/jlms/s2-2.2.216

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1090–1100
https://doi.org/10.1214/23-AIHP1363

On measures strongly log-concave on a subspace

Pierre Bizeul[a]

Institut de Mathématiques de Jussieu-Paris Rive Gauche, Sorbonne Université, Paris, France, [a]*pbizeul@imj-prg.fr*

Abstract. In this work we study the concentration properties of log-concave measures which potential is curved only on a subspace of directions. Proofs use an adapted version of the stochastic localization process.

Résumé. Dans cet article, nous étudions les propriétés de concentration des mesures log-concaves dont le potentiel est courbé sur un sous-espace de directions. L'étude se fait via une version adaptée de la localisation stochastique.

MSC2020 subject classifications: Primary 60D05; 60H30; secondary 52A23

Keywords: KLS conjecture; Spectral gap; Isoperimetric inequality; Log-concave measure

References

[1] C. Borell. Convex measures on locally convex spaces. *Ark. Mat.* **12** (1974) 239–252. MR0388475 https://doi.org/10.1007/BF02384761
[2] Y. Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geom. Funct. Anal.* **31** (2021) 34–61. MR4244847 https://doi.org/10.1007/s00039-021-00558-4
[3] R. Eldan. Distribution of Mass in Convex Bodies. Tel Aviv University, 2012.
[4] R. Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geom. Funct. Anal.* **23** (2013) 532–569. MR3053755 https://doi.org/10.1007/s00039-013-0214-y
[5] D. A. Freedman. On tail probabilities for martingales. *Ann. Probab.* (1975) 100–118. MR0380971 https://doi.org/10.1214/aop/1176996452
[6] M. Gromov and V. D. Milman. A topological application of the isoperimetric inequality. *Amer. J. Math.* **105** (1983) 843–854. MR0708367 https://doi.org/10.2307/2374298
[7] R. Kannan, L. Lovász and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete Comput. Geom.* **13** (1995) 541–559. MR1318794 https://doi.org/10.1007/BF02574061
[8] B. Klartag. Eldan's stochastic localization and tubular neighborhoods of complex-analytic sets. *J. Geom. Anal.* **28** (2018) 2008–2027. MR3833784 https://doi.org/10.1007/s12220-017-9894-0
[9] B. Klartag and J. Lehec. Bourgain's slicing problem and KLS isoperimetry up to polylog, 2022. ArXiv preprint. Available at arXiv:2203.15551. MR4498841 https://doi.org/10.1007/s00039-022-00612-9
[10] M. Ledoux. *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs*. American Mathematical Society, Providence, 2001. MR1849347 https://doi.org/10.1090/surv/089
[11] Y. T. Lee and S. S. Srinivas. Eldan's stochastic localization and the KLS hyperplane conjecture: An improved lower bound for expansion. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)* 998–1007, 2017. MR3734299 https://doi.org/10.1109/FOCS.2017.96
[12] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.* **177** (2009) 1–43. MR2507637 https://doi.org/10.1007/s00222-009-0175-9
[13] B. Øksendal. Stochastic differential equations. In *Stochastic Differential Equations* 65–84. Springer, Berlin, 2003. MR2001996 https://doi.org/10.1007/978-3-642-14394-6
[14] G. Paouris. Concentration of mass on convex bodies. *Geom. Funct. Anal.* **16** (2006) 1021–1049. MR2276533 https://doi.org/10.1007/s00039-006-0584-5

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1101–1124
https://doi.org/10.1214/22-AIHP1360

Metropolis–Hastings transition kernel couplings

John O'Leary[1,a] and Guanyang Wang[2,b]

[1]*Tudor Investment Corporation, Boston, USA,* [a]*john.oleary@tudor.com*
[2]*Department of Statistics, Rutgers University, Piscataway, USA,* [b]*guanyang.wang@rutgers.edu*

Abstract. Couplings play a central role in the analysis of Markov chain convergence and in the construction of novel Markov chain Monte Carlo estimators, diagnostics, and variance reduction techniques. The set of possible couplings is often intractable, frustrating the search for tight bounds and efficient estimators. To address this challenge for algorithms in the Metropolis–Hastings (MH) family, we establish a simple characterization of the set of MH transition kernel couplings. We then extend this result to describe the set of maximal couplings of the MH kernel, resolving an open question of O'Leary, Wang and Jacob (In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* (2021) 1225–1233 PMLR). Our results represent an advance in understanding the MH transition kernel and a step forward for coupling this popular class of algorithms.

Résumé. Les techniques de couplage jouent un rôle central dans l'analyse de la convergence des chaînes de Markov et dans la construction de nouveaux estimateurs à partir de Chaînes de Markov par Monte Carlo, ainsi que de diagnostiques et techniques de réduction de la variance. Souvent, l'ensemble des couplages possibles n'est pas calculable et la recherche de bornes précises et d'estimateurs efficaces semble hors d'atteinte. Pour aborder un tel défi pour les algorithmes de la famille Metropolis–Hastings (MH), nous établissons une caractérisation simple de l'ensemble des couplages des noyaux de transition MH. Nous étendons ensuite ce résultat en décrivant l'ensemble des couplages maximaux des noyaux MH, résolvant une question ouverte de O'Leary, Wang and Jacob (In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics* (2021) 1225–1233 PMLR). Nos résultats représentent un progrès dans la compréhension des noyaux de transition MH et des couplages pour cette classe d'algorithmes populaire.

MSC2020 subject classifications: 60J05; 60J22; 65C05

Keywords: Metropolis–Hastings algorithm; Couplings; Markov chain Monte Carlo

References

[1] D. Aldous. Random walks on finite groups and rapidly mixing Markov chains. In *Séminaire de Probabilités XVII 1981/82* 243–297. Springer, Berlin, 1983. MR0770418 https://doi.org/10.1007/BFb0068322
[2] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. University of California Press, Berkeley, 1995.
[3] C. D. Aliprantis and O. Burkinshaw. *Principles of Real Analysis*. Gulf Pub., Houston, 1998. MR1669668
[4] C. Andrieu, A. Doucet and R. Holenstein. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** (3) (2010) 269–342. MR2758115 https://doi.org/10.1111/j.1467-9868.2009.00736.x
[5] C. Andrieu, A. Lee and S. Livingstone. A general perspective on the Metropolis-Hastings kernel, 2020. Preprint. Available at arXiv:2012.14881.
[6] Y. F. Atchadé and F. Perron. On the geometric ergodicity of Metropolis-Hastings algorithms. *Statistics* **41** (1) (2007) 77–84. MR2303970 https://doi.org/10.1080/10485250601033214
[7] K. B. Athreya and P. Ney. A new approach to the limit theory of recurrent Markov chains. *Trans. Amer. Math. Soc.* **245** (1978) 493–501. MR0511425 https://doi.org/10.2307/1998882
[8] S. Banerjee and W. S. Kendall. Rigidity for Markovian maximal couplings of elliptic diffusions. *Probab. Theory Related Fields* **168** (1–2) (2017) 55–112. MR3651049 https://doi.org/10.1007/s00440-016-0706-4
[9] A. A. Barker. Monte Carlo calculations of the radial distribution functions for a proton–electron plasma. *Aust. J. Phys.* **18** (2) (1965) 119–134.
[10] G. Birkhoff. Three observations on linear algebra. *Univ. Nac. Tacuman, Rev. Ser. A* **5** (1946) 147–151. MR0020547
[11] N. Biswas, P. E. Jacob and P. Vanetti. Estimating convergence of Markov chains with L-lag couplings. In *Advances in Neural Information Processing Systems* 7391–7401, 2019.
[12] B. Böttcher Markovian maximal coupling of Markov processes. Preprint, 2017. Available at arXiv:1710.09654.
[13] N. Bou-Rabee, A. Eberle and R. Zimmer. Coupling and convergence for Hamiltonian Monte Carlo. *Ann. Appl. Probab.* **30** (3) (2020) 1209–1250. MR4133372 https://doi.org/10.1214/19-AAP1528
[14] S. Boyd, P. Diaconis, P. Parrilo and L. Xiao. Fastest mixing Markov chain on graphs with symmetries. *SIAM J. Optim.* **20** (2) (2009) 792–819. MR2515797 https://doi.org/10.1137/070689413

[15] S. Boyd, P. Diaconis, J. Sun and L. Xiao. Fastest mixing Markov chain on a path. *Amer. Math. Monthly* **113** (1) (2006) 70–74. MR2202924 https://doi.org/10.2307/27641840
[16] S. Boyd, P. Diaconis and L. Xiao. Fastest mixing Markov chain on a graph. *SIAM Rev.* **46** (4) (2004) 667–689. MR2124681 https://doi.org/10.1137/S0036144503423264
[17] K. Burdzy and W. S. Kendall. Efficient Markovian couplings: Examples and counterexamples. *Ann. Appl. Probab.* (2000) 362–409. MR1768241 https://doi.org/10.1214/aoap/1019487348
[18] S. Connor and S. Jacka. Optimal co-adapted coupling for the symmetric random walk on the hypercube. *J. Appl. Probab.* **45** (3) (2008) 703–713. MR2455179 https://doi.org/10.1239/jap/1222441824
[19] W. Doeblin. Exposé de la théorie des chaînes simples constantes de Markov à un nombre fini d'états. *Mathématique de l'Union Interbalkanique* **2** (77–105) (1938) 78–80.
[20] R. Douc, E. Moulines, P. Priouret and P. Soulier. *Markov Chains*. Springer, Berlin, 2018. MR3889011 https://doi.org/10.1007/978-3-319-97704-1
[21] J. H. Dshalalow. *Foundations of Abstract Analysis*. Springer, Berlin, 2012. MR3012328 https://doi.org/10.1007/978-1-4614-5962-0
[22] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth. Hybrid Monte Carlo. *Phys. Lett. B* **195** (2) (1987) 216–222. MR3960671 https://doi.org/10.1016/0370-2693(87)91197-x
[23] R. Dudley. *Real Analysis and Probability*. Cambridge University Press, New York, 2002. MR1932358 https://doi.org/10.1017/CBO9780511755347
[24] P. A. Ernst, W. S. Kendall, G. O. Roberts and J. S. Rosenthal. MEXIT: Maximal un-coupling times for stochastic processes. *Stochastic Process. Appl.* **129** (2) (2019) 355–380. MR3907003 https://doi.org/10.1016/j.spa.2018.03.001
[25] J. A. Fill. An interruptible algorithm for perfect sampling via Markov chains. In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing* 688–695, 1997. MR1753391
[26] J. M. Flegal and R. Herbei. Exact sampling for intractable probability distributions via a Bernoulli factory. *Electron. J. Stat.* **6** (2012) 10–37. MR2879671 https://doi.org/10.1214/11-EJS663
[27] A. E. Gelfand and A. F. Smith. Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85** (410) (1990) 398–409. MR1141740
[28] A. Gelman, S. Brooks, G. Jones and X. Meng. *Handbook of Markov Chain Monte Carlo: Methods and Applications. Chapman & Hall/CRC Handbooks of Modern Statistical Methods*. CRC Press, Boca Raton, 2010. MR2742422 https://doi.org/10.1201/b10905
[29] P. W. Glynn and C.-H. Rhee. Exact estimation for Markov chain equilibrium expectations. *J. Appl. Probab.* **51** (A) (2014) 377–389. MR3317370 https://doi.org/10.1239/jap/1417528487
[30] S. Goldstein. Maximal coupling. *Probab. Theory Related Fields* **46** (2) (1979) 193–204. MR0516740 https://doi.org/10.1007/BF00533259
[31] J. B. Goodman and K. K. Lin. Coupling control variates for Markov chain Monte Carlo. *J. Comput. Phys.* **228** (19) (2009) 7127–7136. MR2568586 https://doi.org/10.1016/j.jcp.2009.03.043
[32] D. Griffeath. A maximal coupling for Markov chains. *Probab. Theory Related Fields* **31** (2) (1975) 95–106. MR0370771 https://doi.org/10.1007/BF00539434
[33] P. Gustafson. A guided walk Metropolis algorithm. *Stat. Comput.* **8** (4) (1998) 357–364.
[34] T. E. Harris. On chains of infinite order. *Pacific J. Math.* **5** (Suppl. 1) (1955) 707–724. MR0075482
[35] T. P. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. In *46th Annual IEEE Symposium on Foundations of Computer Science* 511–520. IEEE, Pittsburgh, PA, USA, 2005. MR2326236 https://doi.org/10.1214/105051607000000104
[36] T. P. Hayes and E. Vigoda. A non-Markovian coupling for randomly sampling colorings. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings* 618–627. IEEE, Cambridge, MA, USA, 2003.
[37] J. Heng, J. Houssineau and A. Jasra. On unbiased score estimation for partially observed diffusions, 2021. Preprint. Available at arXiv:2105.04912.
[38] J. Heng and P. E. Jacob. Unbiased Hamiltonian Monte Carlo with couplings. *Biometrika* **106** (2) (2019) 287–302. MR3949304 https://doi.org/10.1093/biomet/asy074
[39] J. Heng, A. Jasra, K. J. Law and A. Tarakanov. On unbiased estimation for discretized models, 2021. Preprint. Available at arXiv:2102.12230.
[40] E. P. Hsu and K.-T. Sturm. Maximal coupling of Euclidean Brownian motions. *Commun. Math. Stat.* **1** (1) (2013) 93–104. MR3197874 https://doi.org/10.1007/s40304-013-0007-5
[41] P. E. Jacob, J. O'Leary and Y. F. Atchadé. Unbiased Markov chain Monte Carlo methods with couplings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** (3) (2020) 543–600. MR4112777 https://doi.org/10.1111/rssb.12336
[42] V. E. Johnson. Studying convergence of Markov chain Monte Carlo algorithms using coupled sample paths. *J. Amer. Statist. Assoc.* **91** (433) (1996) 154–166. MR1394069 https://doi.org/10.2307/2291391
[43] V. E. Johnson. A coupling-regeneration scheme for diagnosing convergence in Markov chain Monte Carlo algorithms. *J. Amer. Statist. Assoc.* **93** (441) (1998) 238–248. MR1614640 https://doi.org/10.2307/2669620
[44] G. L. Jones and J. P. Hobert. Honest exploration of intractable probability distributions via Markov chain. *Monte Carlo. Statistical Science* (2001) 312–334. MR1888447 https://doi.org/10.1214/ss/1015346317
[45] M. Kartashov and V. Golomozyi. Maximal coupling procedure and stability of discrete Markov chains. II. *Theory Probab. Math. Statist.* **87** (2013) 65–78. MR3241447 https://doi.org/10.1090/S0094-9000-2014-00905-9
[46] W. S. Kendall. Coupling, local times, immersions. *Bernoulli* **21** (2) (2015) 1014–1046. MR3338655 https://doi.org/10.3150/14-BEJ596
[47] W. S. Kendall. *Lectures on Probabilistic Coupling*, 2017. MR3338655 https://doi.org/10.3150/14-BEJ596
[48] K. Kuwada. Characterization of maximal Markovian couplings for diffusion processes. *Electron. J. Probab.* **14** (2009) 633–662. MR2486817 https://doi.org/10.1214/EJP.v14-634
[49] D. A. Levin, Y. Peres and E. L. Wilmer. *Markov Chains and Mixing Times*, **107**. Am. Statist. Assoc., Alexandria, 2017. MR3726904 https://doi.org/10.1090/mbk/107
[50] T. Lindvall. *Lectures on the Coupling Method*. Dover, New York, 1992. MR1180522
[51] J. S. Liu, F. Liang and W. H. Wong. The multiple-try method and local optimization in Metropolis sampling. *J. Amer. Statist. Assoc.* **95** (449) (2000) 121–134. MR1803145 https://doi.org/10.2307/2669532
[52] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21** (6) (1953) 1087–1092.
[53] S. P. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Springer, Berlin, 2012. MR2509253 https://doi.org/10.1017/CBO9780511626630

[54] L. Middleton, G. Deligiannidis, A. Doucet and P. E. Jacob. Unbiased smoothing using particle independent Metropolis-Hastings. In *Proceedings of the 22th International Conference on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research. PMLR* 16–18, 2019.

[55] L. Middleton, G. Deligiannidis, A. Doucet and P. E. Jacob. Unbiased Markov chain Monte Carlo for intractable target distributions. *Electron. J. Stat.* **14** (2) (2020) 2842–2891. MR4132645 https://doi.org/10.1214/20-EJS1727

[56] W. K. H. Monte. Carlo sampling methods using Markov chains and their applications. *Biometrika* **57** (1) (1970) 97–109. MR3363437 https://doi.org/10.1093/biomet/57.1.97

[57] R. Neal and R. Pinto. Improving Markov chain Monte Carlo estimators by coupling to an approximating chain. Technical report, Department of Statistics, University of Toronto, 2001.

[58] R. M. Neal. Bayesian learning via stochastic dynamics. In *Advances in Neural Information Processing Systems* 475–482, 1993.

[59] R. M. Neal. Circularly-coupled Markov chain sampling. Technical report, Department of Statistics, University of Toronto, 1999.

[60] R. M. Neal. MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo* 2, **2**, 2011. MR2858447

[61] E. Nummelin. Uniform and ratio limit theorems for Markov renewal and semi-regenerative processes on a general state space. *Ann. Inst. Henri Poincaré Probab. Stat.* **14** (2) (1978) 119–143. MR0507729

[62] J. O'Leary, 2021. Couplings of the Random-Walk Metropolis algorithm. Preprint. Available at arXiv:2102.01790.

[63] J. O'Leary, G. Wang and P. E. Jacob. Maximal couplings of the Metropolis-Hastings algorithm. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics* 1225–1233. *Proceedings of Machine Learning Research* **130**. PMLR, Virtual Event, 2021.

[64] P. H. Peskun. Optimum Monte-Carlo sampling using Markov chains. *Biometrika* **60** (3) (1973) 607–612. MR0362823 https://doi.org/10.1093/biomet/60.3.607

[65] N. S. Pillai and A. Smith. Mixing times for a constrained Ising process on the two-dimensional torus at low density. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (2019) 1649–1678.

[66] D. Piponi, M. Hoffman and P. Sountsov. Hamiltonian Monte Carlo swindles. In *International Conference on Artificial Intelligence and Statistics* 3774–3783. PMLR, Palermo, Sicily, Italy, 2020.

[67] J. Pitman. On coupling of Markov chains. *Probab. Theory Related Fields* **35** (4) (1976) 315–322. MR0415775 https://doi.org/10.1007/BF00532957

[68] J. G. Propp and D. B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9** (1–2) (1996) 223–252. MR1611693 https://doi.org/10.1002/(SICI)1098-2418(199608/09)9:1/2<223::AID-RSA14>3.3.CO;2-R

[69] Q. Qin and J. P. Hobert. Wasserstein-based methods for convergence complexity analysis of MCMC with applications. *Ann. Appl. Probab.* (2021). MR4386523 https://doi.org/10.1214/21-aap1673

[70] A. Reutter and V. E. Johnson. General strategies for assessing convergence of MCMC algorithms using coupled sample paths. Technical report, Institute of Statistics and Decision Sciences, Duke University, 1995.

[71] G. O. Roberts and J. S. Rosenthal. Optimal scaling of discrete approximations to Langevin diffusions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** (1) (1998) 255–268. MR1625691 https://doi.org/10.1111/1467-9868.00123

[72] G. O. Roberts and R. L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli* **2** (4) (1996) 341–363. MR1440273 https://doi.org/10.2307/3318418

[73] J. S. Rosenthal. Minorization conditions and convergence rates for Markov chain Monte Carlo. *J. Amer. Statist. Assoc.* **90** (430) (1995) 558–566. MR1340509

[74] J. S. Rosenthal. Analysis of the Gibbs sampler for a model related to James-Stein estimators. *Stat. Comput.* **6** (3) (1996) 269–275.

[75] J. S. Rosenthal. Quantitative convergence rates of Markov chains: A simple account. *Electron. Commun. Probab.* **7** (2002) 123–128. MR1917546 https://doi.org/10.1214/ECP.v7-1054

[76] L. Tierney. A note on Metropolis-Hastings kernels for general state spaces. *Annals of Applied Probability* (1998) 1–9. MR1620401 https://doi.org/10.1214/aoap/1027961031

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1125–1149
https://doi.org/10.1214/22-AIHP1346

A central limit theorem for the variation of the sum of digits

Yohan Hosten[1,a], Élise Janvresse[1,b] and Thierry de la Rue[2,c],

[1]*LAMFA, Université de Picardie Jules Verne, CNRS UMR 7352, 33, rue Saint-Leu, 80000, Amiens, France,* [a]*yohan.hosten@math.cnrs.fr,* [b]*elise.janvresse@u-picardie.fr*
[2]*LMRS, Université de Rouen Normandie, CNRS UMR 6085, Avenue de l'Université, 76801, Saint Étienne du Rouvray, France,* [c]*Thierry.de-la-Rue@univ-rouen.fr*

Abstract. We prove a Central Limit Theorem for probability measures defined via the variation of the sum-of-digits function, in base $b \geq 2$. For $r \geq 0$ and $d \in \mathbb{Z}$, we consider $\mu^{(r)}(d)$ as the density of integers $n \in \mathbb{N}$ for which the sum of digits increases by d when we add r to n. We give a probabilistic interpretation of $\mu^{(r)}$ on the probability space given by the group of b-adic integers equipped with the normalized Haar measure. We split the base-b expansion of the integer r into so-called "blocks", and we consider the asymptotic behaviour of $\mu^{(r)}$ as the number of blocks goes to infinity. We show that, up to renormalization, $\mu^{(r)}$ converges to the standard normal law as the number of blocks of r grows to infinity. We provide an estimate of the speed of convergence. The proof relies, in particular, on a ϕ-mixing process defined on the b-adic integers.

Résumé. On prouve un Théorème Central Limite pour des mesures de probabilités définies grâce à la variation de la somme des chiffres en base $b \geq 2$. Pour $r \geq 0$ et $d \in \mathbb{Z}$, on considère $\mu^{(r)}(d)$, la densité des entiers $n \in \mathbb{N}$ pour lesquels la somme des chiffres augmente de d quand on ajoute r à n. On donne une interprétation probabiliste de $\mu^{(r)}$ sur l'espace de probabilités donné par le groupe des entiers b-adiques muni de la mesure de Haar renormalisée. On décompose l'écriture en base b d'un entier r en ce que l'on appelle des "blocs", et nous considérons le comportement asymptotique de $\mu^{(r)}$ quand le nombre de blocs tend vers l'infini. On montre qu'à renormalisation près, $\mu^{(r)}$ converge vers une loi normale centrée réduite quand le nombre de blocs de r tend vers l'infini. Nous fournissons une estimation de la vitesse de convergence. La preuve repose, entre autres, sur un processus ϕ-mélangeant défini sur les entiers b-adiques.

MSC2020 subject classifications: 11A63; 37A44; 60F05

Keywords: Sum of digits; Central Limit Theorem; b-adic odometer; ϕ-mixing

References

[1] J. Bésineau. Indépendance statistique d'ensembles liés à la fonction "somme des chiffres". In *Séminaire de Théorie des Nombres, 1970–1971 (Univ. Bordeaux I, Talence), Exp. No. 21*, 20 pp., Lab. Théorie des Nombres, 1971. MR0392881
[2] R. C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** (2005) 107–144. Update of, and a supplement to, the 1986 original. MR2178042 https://doi.org/10.1214/154957805100000104
[3] T. Downarowicz. Survey of odometers and Toeplitz flows. In *Algebraic and Topological Dynamics* 7–37. *Contemp. Math.* **385**. Amer. Math. Soc., Providence, RI, 2005. MR2180227 https://doi.org/10.1090/conm/385/07188
[4] J. Emme and P. Hubert. Central limit theorem for probability measures defined by sum-of-digits function in base 2. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **19** (2) (2019) 757–780. MR3964413
[5] J. Emme and A. Prikhod'ko. On the asymptotic behavior of density of sets defined by sum-of-digits function in base 2. *Integers* **17** (2017) Paper No. A58, 28. MR3731560
[6] N. P. Fogg. *Substitutions in Dynamics Arithmetics and Combinatorics*. *Lecture Notes in Mathematics* **1794**. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. MR1970385 https://doi.org/10.1007/b13861
[7] J. F. Morgenbesser and L. Spiegelhofer. A reverse order property of correlation measures of the sum-of-digits function. *Integers* **12** (2012) Paper No. A47, 5. MR3083420
[8] L. Spiegelhofer and M. Wallner. The digits of n+t, 2021. Available at arXiv:2005.07167.
[9] J. K. Sunklodas. Some estimates of the normal approximation for ϕ-mixing random variables. *Lith. Math. J.* **51** (2) (2011) 260–273. MR2805743 https://doi.org/10.1007/s10986-011-9124-6

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1150–1187
https://doi.org/10.1214/22-AIHP1350

Number of visits in arbitrary sets for ϕ-mixing dynamics

Sandro Gallo[1,a], Nicolai Haydn[2,b] and Sandro Vaienti[3,c]

[1]*Departamento de Estatística, Universidade Federal de São Carlos, CEP: 13565-905, São Paulo, Brasil,* [a]*sandro.gallo@ufscar.br*
[2]*Mathematics Department, USC, Los Angeles, 90089-2532,* [b]*nhaydn@math.usc.edu*
[3]*Aix Marseille Université, Université de Toulon, CNRS, CPT, 13009 Marseille, France,* [c]*vaienti@cpt.univ-mrs.fr*

Abstract. It is well-known that, for sufficiently mixing dynamical systems, the number of visits to balls and cylinders of vanishing measure is approximately Poisson compound distributed in the Kac scaling. Here we extend this kind of results when the target set is an arbitrary set with vanishing measure in the case of ϕ-mixing systems. The error of approximation in total variation is derived using Stein–Chen method. An important part of the paper is dedicated to examples to illustrate the assumptions, as well as applications to temporal synchronisation of g-measures.

Résumé. Il est bien connu que, pour les systèmes dynamiques suffisamment mélangeant, la loi du nombre de visites dans les boules et les cylindres de mesure tendant vers zéro, est proche d'une loi de Poisson composée à l'échelle de Kac. Ici, nous étendons ce type de résultats lorsque l'ensemble cible est un ensemble arbitraire de mesure qui tend vers zéro, dans le cas des systèmes ϕ-mélangeants. L'erreur d'approximation en variation totale est obtenue à l'aide de la méthode de Stein–Chen. Une partie importante de l'article est consacrée à des exemples pour illustrer les hypothèses, ainsi qu'à des applications à la synchronisation temporelle de g-mesures.

MSC2020 subject classifications: Primary 37B20; 37D35; secondary 37A25; 37A50; 60G70

Keywords: Poincaré recurrence; Synchronization of dynamical systems; g-measures; Mixing processes; Compound poisson distribution

References

[1] M. Abadi. Exponential approximation for hitting times in mixing processes. *Math. Phys. Electron. J.* **7** (2001) Paper 19. MR1871384
[2] M. Abadi, V. Amorim, J.-R. Chazottes and S. Gallo. Return-time-spectrum for equilibrium states with potentials of summable variation. *Ergodic Theory Dynam. Systems* **43** (8) (2023) 2489–2515. MR4611739 https://doi.org/10.1017/etds.2022.40
[3] M. Abadi, L. Cardeño and S. Gallo. Potential well spectrum and hitting time in renewal processes. *J. Stat. Phys.* **159** (5) (2015) 1087–1106. MR3345411 https://doi.org/10.1007/s10955-015-1216-y
[4] M. Abadi, A. C. M. Freitas and J. M. Freitas. Clustering indices and decay of correlations in non-Markovian models. *Nonlinearity* **32** (12) (2019) 4853. MR4030595 https://doi.org/10.1088/1361-6544/ab37b8
[5] M. Abadi, A. C. M. Freitas and J. M. Freitas. Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution. *J. Lond. Math. Soc.* **102** (2) (2020) 670–694. MR4171430 https://doi.org/10.1112/jlms.12332
[6] S. Asmussen. *Applied Probability and Queues Stochastic Modelling and Applied Probability* 2nd edition. *Applications of Mathematics (New York)* **51**. Springer-Verlag, New York, 2003. MR1978607
[7] W. Bahsoun, H. Hu and S. Vaienti. Pseudo-orbits, stationary measures and metastability. *Dyn. Syst.* **29** (3) (2014) 322–336. MR3227776 https://doi.org/10.1080/14689367.2014.890172
[8] A. D. Barbour, L. H. Chen and W.-L. Loh. Compound poisson approximation for nonnegative random variables via stein's method. *The Annals of Probability* (1992) 1843–1866. MR1188044
[9] R. C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probab. Surv.* **2** (2) (2005) 107–144. MR2178042 https://doi.org/10.1214/154957805100000104
[10] R. C. Bradley. Introduction to strong mixing conditions, 2007. Kendrick press. MR2325295
[11] X. Bressaud, R. Fernández and A. Galves. Decay of correlations for non-Hölderian dynamics. A coupling approach. *Electron. J. Probab.* **4** (3) (1999) 19. MR1675304 https://doi.org/10.1214/EJP.v4-40
[12] M. Carney, M. Holland and M. Nicol. Extremes and extremal indices for level set observables on hyperbolic systems. *Nonlinearity* **34** (2) (2021) 1136. MR4228013 https://doi.org/10.1088/1361-6544/abd85f
[13] J.-R. Chazottes, Z. Coelho and P. Collet. Poisson processes for subsystems of finite type in symbolic dynamics. *Stoch. Dyn.* **9** (03) (2009) 393–422. MR2566908 https://doi.org/10.1142/S0219493709002713
[14] J.-R. Chazottes and P. Collet. Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems. *Ergodic Theory Dynam. Systems* **33** (1) (2013) 49–80. MR3009103 https://doi.org/10.1017/S0143385711000897

[15] L. H. Y. Chen and A. Barbour. *Stein's Method and Applications*, **5**. World Scientific, Singapore, 2005. MR2201882 https://doi.org/10.1142/9789812567673
[16] E. Cinlar. Introduction to stochastic processes. *Courier Corporation.* (2013). MR0380912
[17] Z. Coelho and P. Collet. Asymptotic limit law for the close approach of two trajectories in expanding maps of the circle. *Probab. Theory Related Fields* **99** (2) (1994) 237–250. MR1278884 https://doi.org/10.1007/BF01199024
[18] D. Faranda, H. Ghoudi, P. Guiraud and S. Vaienti. Extreme value theory for synchronization of coupled map lattices. *Nonlinearity* **37** (7) (2018) 3326. MR3816758 https://doi.org/10.1088/1361-6544/aabc8e
[19] R. Fernández. Gibbsianness and non-gibbsianness in lattice random fields. *Les Houches* **LXXXIII** (2005) 731–799. MR2581896 https://doi.org/10.1016/S0924-8099(06)80052-1
[20] R. F. Ferreira, S. Gallo and F. Paccaut. Non-regular g-measures and variable length memory chains. *Nonlinearity* **33** (11) (2020) 6026. MR4164670 https://doi.org/10.1088/1361-6544/aba0c5
[21] A. C. M. Freitas, J. M. Freitas and M. Magalhães. Convergence of marked point processes of excesses for dynamical systems. *J. Eur. Math. Soc.* **20** (9) (2018) 2131–2179. MR3836843 https://doi.org/10.4171/JEMS/808
[22] A. C. M. Freitas, J. M. Freitas, M. Magalhães and S. Vaienti. Point processes of non stationary sequences generated by sequential and random dynamical systems. *J. Stat. Phys.* **181** (4) (2020) 1365–1409. MR4163505 https://doi.org/10.1007/s10955-020-02630-z
[23] A. C. M. Freitas, J. M. Freitas and M. Todd. The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. *Comm. Math. Phys.* **321** (2) (2013) 483–527. MR3063917 https://doi.org/10.1007/s00220-013-1695-0
[24] J. M. Freitas, N. Haydn and M. Nicol. Convergence of rare event point processes to the Poisson process for planar billiards. *Nonlinearity* **27** (7) (2014) 1669. MR3232197 https://doi.org/10.1088/0951-7715/27/7/1669
[25] H. Furstenberg. *Stationary Processes and Prediction Theory*, **44**. Princeton University Press, Princeton, 1960. MR0140151
[26] A. Galves and E. Löcherbach. Infinite systems of interacting chains with memory of variable length—a stochastic model for biological neural nets. *J. Stat. Phys.* **151** (5) (2013) 896–921. MR3055382 https://doi.org/10.1007/s10955-013-0733-9
[27] N. Haydn and Y. Psiloyenis. Return times distribution for Markov towers with decay of correlations. *Nonlinearity* **27** (6) (2014) 1323. MR3215837 https://doi.org/10.1088/0951-7715/27/6/1323
[28] N. Haydn and S. Vaienti. The distribution of return times near periodic orbits. *Probab. Theory Related Fields* **144** (2009) 517–542. MR2496441 https://doi.org/10.1007/s00440-008-0153-y
[29] N. Haydn and S. Vaienti. Limiting entry and return times distribution for arbitrary null sets. *Comm. Math. Phys.* **378** (1) (2020) 149–184. MR4124984 https://doi.org/10.1007/s00220-020-03795-0
[30] N. Haydn and K. Wasilewska. Limiting distribution and error terms for the number of visits to balls in non-uniformly hyperbolic dynamical systems. *Discrete Contin. Dyn. Syst.* **36** (5) (2016) 2585–2611. MR3485409 https://doi.org/10.3934/dcds.2016.36.2585
[31] N. Haydn and F. Yang. A derivation of the poisson law for returns of smooth maps with certain geometrical properties. *Contemporary Mathematics Proceedings in memoriam Chernov.* (2017). MR3716090 https://doi.org/10.1090/conm/698/14031
[32] N. T. Haydn. Entry and return times distribution. *Dyn. Syst.* **28** (3) (2013) 333–353. MR3170620 https://doi.org/10.1080/14689367.2013.822459
[33] M. Hirata. Poisson law for axiom a diffeomorphisms. *Ergodic Theory Dynam. Systems* **13** (3) (1993) 533–556. MR1245828 https://doi.org/10.1017/S0143385700007513
[34] J. Janssen and R. Manca. *Applied Semi-Markov Processes*. Springer Science & Business Media, Berlin, 2006. MR2175472
[35] M. Keane. Strongly mixing g-measures. *Invent. Math.* **16** (1972) 309–324. MR0310193 https://doi.org/10.1007/BF01425715
[36] G. Keller. Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dyn. Syst.* **27** (1) (2012) 11–27. MR2903242 https://doi.org/10.1080/14689367.2011.653329
[37] G. Keller and C. Liverani. Rare events, escape rates and quasistationarity: Some exact formulae. *J. Stat. Phys.* **135** (3) (2009) 519–534. MR2535206 https://doi.org/10.1007/s10955-009-9747-8
[38] Y. Kifer and A. Rapaport. Poisson and compound Poisson approximations in conventional and nonconventional setups. *Probab. Theory Related Fields* **160** (3–4) (2014) 797–831. MR3278921 https://doi.org/10.1007/s00440-013-0541-9
[39] Y. Kifer and F. Yang. Geometric law for numbers of returns until a hazard under ϕ-mixing, 2018. ArXiv preprint. Available at arXiv:1812.09927. MR4344031 https://doi.org/10.1007/s11856-021-2182-5
[40] F. Ledrappier. Principe variationnel et systèmes dynamiques symboliques. *Z. Wahrsch. Verw. Gebiete* **30** (1974) 185–202. MR0404584 https://doi.org/10.1007/BF00533471
[41] V. Lucarini, D. Faranda, J. M. M. de Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti et al. *Extremes and Recurrence in Dynamical Systems*. Wiley, New York, 2016. MR3558780 https://doi.org/10.1002/9781118632321
[42] M. R. Palmer, W. Parry and P. Walters. Large sets of endomorphisms and of g-measures. In *The Structure of Attractors in Dynamical Systems* 191–210. Springer, Berlin, 1978. MR0518560
[43] F. Pène and B. Saussol. Back to balls in billiards. *Comm. Math. Phys.* **293** (3) (2010) 837–866. MR2566164 https://doi.org/10.1007/s00220-009-0911-4
[44] F. Pène and B. Saussol. Poisson law for some non-uniformly hyperbolic dynamical systems with polynomial rate of mixing. *Ergodic Theory Dynam. Systems* **36** (8) (2016) 2602–2626. MR3570026 https://doi.org/10.1017/etds.2015.28
[45] F. Pène and B. Saussol. Spatio-temporal Poisson processes for visits to small sets. *Israel J. Math.* **240** (2) (2020) 625–665. MR4193145 https://doi.org/10.1007/s11856-020-2074-0
[46] B. Pitskel. Poisson limit law for Markov chains. *Ergodic Theory Dynam. Systems* **11** (3) (1991) 501–513. MR1125886 https://doi.org/10.1017/S0143385700006301
[47] M. Roos et al. Stein's method for compound Poisson approximation: The local approach. *Ann. Appl. Probab.* **4** (4) (1994) 1177–1187. MR1304780
[48] P. C. Shields. *The Ergodic Theory of Discrete Sample Paths*. *Graduate Studies in Mathematics* **13**. American Mathematical Society, Providence, RI, 1996. MR1400225 https://doi.org/10.1090/gsm/013
[49] C. Stein Approximate computation of expectations, 1986. IMS. MR0882007
[50] E. Verbitskiy. On factors of g-measures. *Indag. Math. (N.S.)* **22** (3–4) (2011) 315–329. MR2853610 https://doi.org/10.1016/j.indag.2011.09.001
[51] E. Verbitskiy. Hidden gibbs models: Theory and applications. Unpublished notes, 2015.
[52] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.* **214** (1975) 375–387. MR0412389 https://doi.org/10.2307/1997113
[53] F. Yang. Rare event process and entry times distribution for arbitrary null sets on compact manifolds. *Ann. Inst. Henri Poincaré Probab. Stat.* **57** (2) (2021) 1103–1135. MR4260497 https://doi.org/10.1214/20-aihp1109

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1188–1214
https://doi.org/10.1214/22-AIHP1341

Active phase for activated random walks on the lattice in all dimensions

Nicolas Forien[1,2,a] and Alexandre Gaudillière[1,b]

[1]Aix Marseille Univ., CNRS, I2M, Marseille, France, [a]nicolas.forien@uniroma1.it, [b]alexandre.gaudilliere@math.cnrs.fr
[2]Dipartimento di Matematica, Sapienza Università di Roma, Roma, Italy

Abstract. We show that the critical density of the Activated Random Walk model on $\mathbb{Z}^d$ is strictly less than one when the sleep rate λ is small enough, and tends to 0 when $\lambda \to 0$, in any dimension $d \geq 1$. As far as we know, the result is new for $d = 2$.

We prove this by showing that, for high enough density and small enough sleep rate, the stabilization time of the model on the d-dimensional torus is exponentially large. To do so, we fix the the set of sites where the particles eventually fall asleep, which reduces the problem to a simpler model with density one. Taking advantage of the Abelian property of the model, we show that the stabilization time stochastically dominates the escape time of a one-dimensional random walk with a negative drift. We then check that this slow phase for the finite volume dynamics implies the existence of an active phase on the infinite lattice.

Résumé. Nous démontrons que la densité critique du modèle des Marches Aléatoires Activées sur $\mathbb{Z}^d$ est strictement inférieure à 1 quand le taux d'endormissement λ est suffisamment petit, et tend vers 0 quand $\lambda \to 0$, en toute dimension $d \geq 1$. À notre connaissance, le résultat est nouveau pour $d = 2$.

Nous obtenons ce résultat en prouvant que, pour une densité suffisamment élevée et un taux d'endormissement suffisamment petit, le temps de stabilisation du modèle sur le tore en dimension d est exponentiellement grand. Pour cela, nous fixons l'ensemble des sites sur lesquels les particules s'endorment, ce qui réduit le problème à un modèle plus simple avec densité 1. En utilisant la propriété d'Abélianité du modèle, nous montrons que le temps de stabilisation domine stochastiquement le temps d'atteinte de 0 pour une marche aléatoire en dimension 1 avec une dérive négative. Nous vérifions ensuite que cette phase de stabilisation lente pour la dynamique en volume fini implique l'existence d'une phase active sur le réseau infini.

MSC2020 subject classifications: Primary 60K35; secondary 82B26

Keywords: Activated random walks; Phase transition; Self-organized criticality

References

[1] O. Alves, F. Machado and S. Popov. Phase transition for the frog model. *Electron. J. Probab.* **7** (2002) 1–21. MR1943889 https://doi.org/10.1214/EJP.v7-115
[2] O. S. Alves, F. P. Machado and S. Y. Popov. The shape theorem for the frog model. *Ann. Appl. Probab.* **12** (2) (2002) 533–546. MR1910638 https://doi.org/10.1214/aoap/1026915614
[3] G. Amir and O. Gurel-Gurevich. On fixation of activated random walks. *Electron. Commun. Probab.* **15** (2010) 119–123. MR2643591 https://doi.org/10.1214/ECP.v15-1536
[4] A. Asselah, L. T. Rolla and B. Schapira. Diffusive bounds for the critical density of activated random walks. arXiv preprint. Available at arXiv:1907.12694 (2019). MR4394304 https://doi.org/10.30757/alea.v19-17
[5] P. Bak, C. Tang and K. Wiesenfeld. Self-organized criticality: An explanation of the 1/f noise. *Phys. Rev. Lett.* **59** (4) (1987) 381. MR0949160 https://doi.org/10.1103/PhysRevA.38.364
[6] R. Basu, S. Ganguly and C. Hoffman. Non-fixation for conservative stochastic dynamics on the line. *Comm. Math. Phys.* **358** (3) (2018) 1151–1185. MR3778354 https://doi.org/10.1007/s00220-017-3059-7
[7] R. Basu, S. Ganguly, C. Hoffman and J. Richey. Activated random walk on a cycle. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (3) (2019) 1258–1277. MR4010935 https://doi.org/10.1214/18-aihp918
[8] M. Cabezas, L. T. Rolla and V. Sidoravicius. Non-equilibrium phase transitions: Activated random walks at criticality. *J. Stat. Phys.* **155** (6) (2014) 1112–1125. MR3207731 https://doi.org/10.1007/s10955-013-0909-3
[9] D. Dhar. Theoretical studies of self-organized criticality. *Phys. A* **369** (1) (2006) 29–70. MR2246566 https://doi.org/10.1016/j.physa.2006.04.004
[10] P. Diaconis and W. Fulton. A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Semin. Mat. Univ. Politec. Torino* **49** (1) (1991) 95–119. MR1218674

[11] R. Dickman, L. T. Rolla and V. Sidoravicius. Activated random walkers: Facts, conjectures and challenges. *J. Stat. Phys.* **138** (1) (2010) 126–142. MR2594894 https://doi.org/10.1007/s10955-009-9918-7
[12] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. *Carus Mathematical Monographs* **22**, xiv + 159. Mathematical Association of America, Washington, DC, 1984. MR0920811
[13] N. Forien. About self-organized criticality. PhD thesis, Université Paris-Saclay, 2020.
[14] N. Gantert and P. Schmidt. Recurrence for the frog model with drift on Z. *Markov Process. Related Fields* **15** (1) (2009) 51–58. MR2509423
[15] S. Helmrich, A. Arias, G. Lochead, T. M. Wintermantel, M. Buchhold, S. Diehl and S. Whitlock. Signatures of self-organized criticality in an ultracold atomic gas. *Nature* **577** (7791) (2020) 481–486.
[16] C. Hoffman, T. Johnson and M. Junge. Recurrence and transience for the frog model on trees. *Ann. Probab.* **45** (5) (2017) 2826–2854. MR3706732 https://doi.org/10.1214/16-AOP1125
[17] C. Hoffman, J. Richey and L. T. Rolla. Active phase for activated random walk on $\mathbb{Z}$. arXiv preprint. Available at arXiv:2009.09491 (2020). MR4010935 https://doi.org/10.1214/18-aihp918
[18] H. J. Jensen. *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*, **10**. Cambridge University Press, Cambridge, 1998. MR1689042 https://doi.org/10.1017/CBO9780511622717
[19] D. Jerison, L. Levine and S. Sheffield. Logarithmic fluctuations for internal DLA. *J. Amer. Math. Soc.* **25** (1) (2012) 271–301. MR2833484 https://doi.org/10.1090/S0894-0347-2011-00716-9
[20] T. Johnson and M. Junge. Stochastic orders and the frog model. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2) (2018) 1013–1030. MR3795075 https://doi.org/10.1214/17-AIHP830
[21] H. Kesten and V. Sidoravicius. The spread of a rumor or infection in a moving population. *Ann. Probab.* **33** (6) (2005) 2402–2462. MR2184100 https://doi.org/10.1214/009117905000000413
[22] L. Levine and V. Silvestri. How far do activated random walkers spread from a single source? *J. Stat. Phys.* **185** (3) (2021) 18. MR4334780 https://doi.org/10.1007/s10955-021-02836-9
[23] R. Lyons and Y. Peres. *Probability on Trees and Networks*. *Cambridge Series in Statistical and Probabilistic Mathematics* **42**, xv + 699. Cambridge University Press, New York, 2016. MR3616205 https://doi.org/10.1017/9781316672815
[24] S. S. Manna. Two-state model of self-organized criticality. *J. Phys. A: Math. Gen.* **24** (7) (1991) L363.
[25] L. T. Rolla. Generalized hammersley process and phase transition for activated random walk models. PhD thesis. Available at arXiv:0812.2473 (2008).
[26] L. T. Rolla. Activated random walks on $\mathbb{Z}^d$. *Probab. Surv.* **17** (2020) 478–544. MR4152668 https://doi.org/10.1214/19-PS339
[27] L. T. Rolla and V. Sidoravicius. Absorbing-state phase transition for driven-dissipative stochastic dynamics on $\mathbb{Z}$. *Invent. Math.* **188** (1) (2012) 127–150. MR2897694 https://doi.org/10.1007/s00222-011-0344-5
[28] L. T. Rolla, V. Sidoravicius and O. Zindy. Universality and sharpness in activated random walks. *Ann. Henri Poincaré* **20** (6) (2019) 1823–1835. MR3956161 https://doi.org/10.1007/s00023-019-00797-0
[29] L. T. Rolla and L. Tournier. Non-fixation for biased activated random walks. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2) (2018) 938–951. MR3795072 https://doi.org/10.1214/17-AIHP827
[30] E. Shellef. Nonfixation for activated random walks. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** (2010) 137–149. MR2651824
[31] V. Sidoravicius and A. Teixeira. Absorbing-state transition for stochastic sandpiles and activated random walks. *Electron. J. Probab.* **22** (2017) 33. MR3646059 https://doi.org/10.1214/17-EJP50
[32] A. Stauffer and L. Taggi. Critical density of activated random walks on transitive graphs. *Ann. Probab.* **46** (4) (2018) 2190–2220. MR3813989 https://doi.org/10.1214/17-AOP1224
[33] L. Taggi. Absorbing-state phase transition in biased activated random walk. *Electron. J. Probab.* **21** (2016) 13. MR3485355 https://doi.org/10.1214/16-EJP4275
[34] L. Taggi. Active phase for activated random walks on $\mathbb{Z}^d$, $d \geq 3$, with density less than one and arbitrary sleeping rate. *Ann. Inst. Henri Poincaré Probab. Stat.* **55** (3) (2019) 1751–1764. MR4010950 https://doi.org/10.1214/18-aihp933

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1215–1245
https://doi.org/10.1214/22-AIHP1343

Hydrodynamics of the t-PNG model via a colored t-PNG model

Hindy Drillick[1,a] and Yier Lin[2,b]

[1]*Department of Mathematics, Columbia University, 2990 Broadway, New York, NY 10027, USA,* [a]*hindy.drillick@columbia.edu*
[2]*Department of Statistics, University of Chicago, 5747 S. Ellis Ave, Chicago, IL 60637, USA,* [b]*ylin10@uchicago.edu*

Abstract. In this paper, we prove the hydrodynamic limit of the t-PNG model using soft techniques. One key element of the proof is the construction of a colored version of the t-PNG model, which allows us to apply the superadditive ergodic theorem and obtain the hydrodynamic limit, albeit without identifying the limiting constant. We then find this constant by proving a law of large numbers for the α-points. Along the way, we construct the stationary t-PNG model and prove a version of Burke's theorem for it.

Résumé. Dans cet article, nous prouvons la limite hydrodynamique du modèle t-PNG en utilisant des méthodes peu techniques. Un élément clé de la preuve est la construction d'une version colorée du modèle t-PNG, qui nous permet d'appliquer le théorème ergodique sur-additif et d'obtenir la limite hydrodynamique, mais sans identifier la constante limite. Nous trouvons ensuite cette constante en démontrant une loi des grands nombres pour les α-points. Ce faisant, nous construisons le modèle stationnaire t-PNG et prouvons une version du théorème de Burke pour celui-ci.

MSC2020 subject classifications: Primary 60F15; secondary 60J25

Keywords: Polynuclear growth model; Hammersley's process; Interacting particle systems

References

[1] A. Aggarwal. Current fluctuations of the stationary ASEP and six-vertex model. *Duke Math. J.* **167** (2) (2018) 269–384. MR3754630 https://doi.org/10.1215/00127094-2017-0029
[2] A. Aggarwal, A. Borodin and M. Wheeler. Colored fermionic vertex models and symmetric functions. Preprint. Available at arXiv:2101.01605 (2021).
[3] A. Aggarwal, A. Borodin and M. Wheeler. Deformed Polynuclear Growth in $(1+1)$ Dimensions. *International Mathematics Research Notices* (2022).
[4] D. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probab. Theory Related Fields* **103** (2) (1995) 199–213. MR1355056 https://doi.org/10.1007/BF01204214
[5] D. Aldous and P. Diaconis. Longest increasing subsequences: From patience sorting to the Baik–Deift–Johansson theorem. *Bull. Amer. Math. Soc.* **36** (4) (1999) 413–432. MR1694204 https://doi.org/10.1090/S0273-0979-99-00796-X
[6] J. Baik, P. Deift and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** (4) (1999) 1119–1178. MR1682248 https://doi.org/10.1090/S0894-0347-99-00307-0
[7] M. Balázs, E. Cator and T. Seppalainen. Cube root fluctuations for the corner growth model associated to the exclusion process. *Electron. J. Probab.* **11** (2006) 1094–1132. MR2268539 https://doi.org/10.1214/EJP.v11-366
[8] A.-L. Basdevant, N. Enriquez, L. Gerin and J.-B. Gouéré. Discrete Hammersley's lines with sources and sinks. *ALEA Lat. Am. J. Probab. Math. Stat.* **13** (2016) 33–52. MR3460869 https://doi.org/10.30757/alea.v13-02
[9] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Elsevier, Amsterdam, 2016. MR0690578
[10] A. Borodin, A. Okounkov and G. Olshanski. Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* **13** (3) (2000) 481–515. MR1758751 https://doi.org/10.1090/S0894-0347-00-00337-4
[11] A. Borodin and L. Petrov. Higher spin six vertex model and symmetric rational functions. *Selecta Math. (N.S.)* **24** (2) (2018) 751–874. MR3782413 https://doi.org/10.1007/s00029-016-0301-7
[12] A. Borodin and M. Wheeler. Coloured stochastic vertex models and their spectral theory. Preprint. Available at arXiv:1808.01866 (2018).
[13] A. Borodin and M. Wheeler. Observables of coloured stochastic vertex models and their polymer limits. *Probab. Math. Phys.* **1** (1) (2020) 205–265. MR4408007 https://doi.org/10.2140/pmp.2020.1.205
[14] P. J. Burke. The output of a queuing system. *Oper. Res.* **4** (6) (1956) 699–704. MR0083416 https://doi.org/10.1287/opre.4.6.699
[15] E. Cator and S. Dobrynin. Behavior of a second class particle in Hammersley's process. *Electron. J. Probab.* **11** (2006) 670–685. MR2242659 https://doi.org/10.1214/EJP.v11-340
[16] E. Cator and P. Groeneboom. Hammersley's process with sources and sinks. *Ann. Probab.* **33** (3) (2005) 879–903. MR2135307 https://doi.org/10.1214/009117905000000053

[17] E. Cator and P. Groeneboom. Second class particles and cube root asymptotics for Hammersley's process. *Ann. Probab.* **34** (4) (2006) 1273–1295. MR2257647 https://doi.org/10.1214/009117906000000089

[18] E. Cator and L. P. Pimentel. A shape theorem and semi-infinite geodesics for the Hammersley model with random weights. *ALEA Lat. Am. J. Probab. Math. Stat.* **8** (2011) 163–175. MR2783936

[19] H. Chaumont and C. Noack. Characterizing stationary $1+1$ dimensional lattice polymer models. *Electron. J. Probab.* **23** (2018) 1–19. MR3806406 https://doi.org/10.1214/18-EJP163

[20] F. Ciech and N. Georgiou. Order of the variance in the discrete Hammersley process with boundaries. *J. Stat. Phys.* **176** (3) (2019) 591–638. MR3985155 https://doi.org/10.1007/s10955-019-02314-3

[21] I. Corwin and L. Petrov. Stochastic higher spin vertex models on the line. *Comm. Math. Phys.* **343** (2) (2016) 651–700. MR3477349 https://doi.org/10.1007/s00220-015-2479-5

[22] I. Corwin and L.-C. Tsai. KPZ equation limit of higher-spin exclusion processes. *Ann. Probab.* **45** (3) (2017) 1771–1798. MR3650415 https://doi.org/10.1214/16-AOP1101

[23] P. A. Ferrari. TASEP hydrodynamics using microscopic characteristics. *Probab. Surv.* **15** (2018) 1–27. MR3769188 https://doi.org/10.1214/17-PS284

[24] P. A. Ferrari and J. B. Martin. Multiclass Hammersley–Aldous–Diaconis process and multiclass-customer queues. In *Annales de l'IHP Probabilités et statistiques* 250–265, **45**, 2009. MR2500238 https://doi.org/10.1214/08-AIHP168

[25] P. Groeneboom. Ulam's problem and Hammersley's process. *Ann. Probab.* **29** (2) (2001) 683–690. MR1849174 https://doi.org/10.1214/aop/1008956689

[26] P. Groeneboom. Hydrodynamical methods for analyzing longest increasing subsequences. *J. Comput. Appl. Math.* **142** (1) (2002) 83–105. MR1910520 https://doi.org/10.1016/S0377-0427(01)00461-7

[27] L.-H. Gwa and H. Spohn. Six-vertex model, roughened surfaces, and an asymmetric spin Hamiltonian. *Phys. Rev. Lett.* **68** (6) (1992) 725. MR1147356 https://doi.org/10.1103/PhysRevLett.68.725

[28] J. M. Hammersley. A few seedlings of research. In *Proc. Sixth Berkeley Symp. Math. Statist. and Probability* 345–394, **1**, 1972. MR0405665

[29] T. Imamura, M. Mucciconi and T. Sasamoto. Stationary stochastic Higher Spin Six Vertex Model and q-Whittaker measure. *Probab. Theory Related Fields* **177** (3) (2020) 923–1042. MR4126935 https://doi.org/10.1007/s00440-020-00966-x

[30] K. Johansson. Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** (2) (2000) 437–476. MR1737991 https://doi.org/10.1007/s002200050027

[31] K. Johansson. Discrete orthogonal polynomial ensembles and the Plancherel measure. *Annals of Mathematics* (2001) 259–296. MR1826414 https://doi.org/10.2307/2661375

[32] J. Kuan. An algebraic construction of duality functions for the stochastic $U_q(A_n^{(1)})$ vertex model and its degenerations. *Comm. Math. Phys.* **359** (1) (2018) 121–187. MR3781448 https://doi.org/10.1007/s00220-018-3108-x

[33] A. Kuniba, V. V. Mangazeev, S. Maruyama and M. Okado. Stochastic R matrix for $U_q(A_n^{(1)})$. *Nuclear Phys. B* **913** (2016) 248–277. MR3575300 https://doi.org/10.1016/j.nuclphysb.2016.09.016

[34] E. H. Lieb. Residual entropy of square ice. *Phys. Rev.* **162** (1967) 162–172. https://doi.org/10.1103/PhysRev.162.162

[35] T. M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.* **13** (4) (1985) 1279–1285. MR0806224

[36] T. M. Liggett. *Interacting Particle Systems*, **276**. Springer, Berlin, 2012. MR0776231 https://doi.org/10.1007/978-1-4613-8542-4

[37] Y. Lin. KPZ equation limit of stochastic higher spin six vertex model. *Math. Phys. Anal. Geom.* **23** (1) (2020) 1–118. MR4046044 https://doi.org/10.1007/s11040-019-9325-5

[38] Y. Lin. The stochastic telegraph equation limit of the stochastic higher spin six vertex model. *Electron. J. Probab.* **25** (2020) 1–30. MR4193889 https://doi.org/10.1214/20-ejp552

[39] B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Adv. Math.* **26** (2) (1977) 206–222. MR1417317 https://doi.org/10.1016/0001-8708(77)90030-5

[40] J. B. Martin. Limiting shape for directed percolation models. *Ann. Probab.* **32** (4) (2004) 2908–2937. MR2094434 https://doi.org/10.1214/009117904000000838

[41] N. O'Connell and M. Yor. Brownian analogues of Burke's theorem. *Stochastic Process. Appl.* **96** (2) (2001) 285–304. MR1865759 https://doi.org/10.1016/S0304-4149(01)00119-3

[42] Y. Pei. A q-Robinson–Schensted–Knuth algorithm and a q-polymer. Preprint. Available at arXiv:1610.03692 (2016). MR3711039

[43] D. Romik. *The Surprising Mathematics of Longest Increasing Subsequences*. Cambridge University Press, Cambridge, 2015. MR3468738

[44] H. Rost. Non-equilibrium behaviour of a many particle process: Density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete* **58** (1) (1981) 41–53. MR0635270 https://doi.org/10.1007/BF00536194

[45] T. Seppäläinen. A microscopic model for the Burgers equation and longest increasing subsequences. *Electron. J. Probab.* **1** (1996) 1–51. MR1386297 https://doi.org/10.1214/EJP.v1-5

[46] T. Seppäläinen. Increasing sequences of independent points on the planar lattice. *Ann. Appl. Probab.* **7** (4) (1997) 886–898. MR1484789 https://doi.org/10.1214/aoap/1043862416

[47] T. Seppäläinen. Exact limiting shape for a simplified model of first-passage percolation on the plane. *Ann. Probab.* **26** (3) (1998) 1232–1250. MR1640344 https://doi.org/10.1214/aop/1022855751

[48] T. Seppäläinen. Lecture notes on the corner growth model. *Unpublished notes* (2009).

[49] T. Seppäläinen. Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* **40** (1) (2012) 19–73. MR2917766 https://doi.org/10.1214/10-AOP617

[50] T. Seppäläinen. Variational formulas, Busemann functions, and fluctuation exponents for the corner growth model with exponential weights. Preprint. Available at arXiv:1709.05771 (2017). MR3838898 https://doi.org/10.1090/psapm/075

[51] S. M. Ulam. Monte Carlo calculations in problems of mathematical physics. *Modern Mathematics for the Engineers* (1961) 261–281. MR0129165

[52] A. M. Veršik and S. V. Kerov. Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux. In *Dokl. Akad. Nauk SSSR* 1024–1027, **233**, 1977. MR0480398

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1246–1294
https://doi.org/10.1214/23-AIHP1364

Non-stationary KPZ equation from ASEP with slow bonds

Kevin Yang[a]

Department of Mathematics, University of California, Berkeley, [a]*kevin_yang95@berkeley.edu*

Abstract. We prove the height functions for a class of non-integrable and non-stationary particle systems converge to the KPZ equation, thereby making progress on the universality of the KPZ equation. The models herein are ASEP (*Comm. Math. Phys.* **183** (1997) 571–606) with a mesoscopic family of slow bonds, thus we partially extend (*Comm. Math. Phys.* **346** (2016) 801–838) to non-stationary models and add to the almost empty set of non-integrable, non-stationary interacting particle systems for which universality is established. To do this, we develop further the strategy of (Yang (2020); *Probab. Theory Related Fields* **183** (2022) 415–545) introduce a method to establish a novel principle that builds upon the classical hydrodynamic limits of (*Comm. Math. Phys.* **118** (1988) 31–59) and that we call *local hydrodynamics*.

Résumé. Nous prouvons que les fonctions de hauteur pour une classe de systèmes de particules non-intégrables et non-stationnaires convergent vers l'équation KPZ, contribuant ainsi à l'universalité de l'équation KPZ. Les modèles présentés ici sont des modèles ASEP (*Comm. Math. Phys.* **183** (1997) 571–606) avec une famille mésoscopique de liaisons lentes, nous étendons donc partiellement (*Comm. Math. Phys.* **346** (2016) 801–838) aux modèles non-stationnaires et donnons un des rares exemples de systèmes de particules en interaction non-intégrables et non-stationnaires pour lesquels l'universalité est établie. Pour ce faire, nous développons davantage la stratégie de (Yang (2020) ; *Probab. Theory Related Fields* **183** (2022) 415–545) en introduisant une méthode pour établir un nouveau principe qui s'appuie sur les limites hydrodynamiques classiques de (*Comm. Math. Phys.* **118** (1988) 31–59) et que nous appelons *hydrodynamique locale*.

MSC2020 subject classifications: Primary 60K35; secondary 60H17

Keywords: KPZ equation; Universality; Slow bond

References

[1] G. Amir, I. Corwin and J. Quastel. Probability distribution of the free energy of the continuum directed polymer model in $(1+1)$-dimensions. *Comm. Pure Appl. Math.* **64** (2011) 466–537. MR2796514 https://doi.org/10.1002/cpa.20347
[2] R. Basu, S. Sarkar and A. Sly. Invariant Measures for TASEP with a Slow Bond, 2017. Available at arXiv:1704.07799.
[3] R. Basu, V. Sidoravicius and A. Sly. Last passage percolation with a defect line and the solution of the Slow Bond Problem, 2014. Available at arXiv:1408.3464.
[4] L. Bertini and G. Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Comm. Math. Phys.* **183** (3) (1997) 571–606. MR1462228 https://doi.org/10.1007/s002200050044
[5] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1999. MR1700749 https://doi.org/10.1002/9780470316962
[6] L. Bonorino, R. De Paula, P. Goncalves and A. Neumann. Hydrodynamics for the porous medium model with slow reservoirs. *J. Stat. Phys.* **179** (2020) 748–788. MR4099999 https://doi.org/10.1007/s10955-020-02550-y
[7] M. Bramson, T. M. Liggett and T. Mountford. Characterization of stationary measures for one-dimensional exclusion processes. *Ann. Probab.* **30** (4) (2002) 1539–1575. MR1944000 https://doi.org/10.1214/aop/1039548366
[8] I. Corwin. The Kardar–Parisi–Zhang equation and universality class. *Random Matrices Theory Appl.* **1** (1) (2012) 1130001. MR2930377 https://doi.org/10.1142/S2010326311300014
[9] I. Corwin, H. Shen and L.-C. Tsai. ASEP(q,j) converges to the KPZ equation. *Ann. Inst. Henri Poincaré, Probab. Stat.* **54** (2) (2018) 995–1012. MR3795074 https://doi.org/10.1214/17-AIHP829
[10] I. Corwin and L.-C. Tsai. KPZ equation limit of higher-spin exclusion processes. *Ann. Probab.* **45** (3) (2017) 1771–1798. MR3650415 https://doi.org/10.1214/16-AOP1101
[11] A. Dembo and L.-C. Tsai. Weakly asymmetric non-simple exclusion process and the KPZ equation. *Comm. Math. Phys.* **341** (2016) 219–261. MR3439226 https://doi.org/10.1007/s00220-015-2527-1
[12] D. Erhard, T. Franco, P. Goncalves, A. Neumann and M. Tavares. Non-equilibrium fluctuations for SSEP with a slow bond. *Ann. Inst. Henri Poincaré, Probab. Stat.* **56** (2) (2020) 1099–1128. MR4076777 https://doi.org/10.1214/19-AIHP995

[13] C. Erignoux, P. Goncalves and G. Nahum. Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part I, the critical regime and beyond. *J. Stat. Phys.* **181** (2020) 1433–1469. MR4163507 https://doi.org/10.1007/s10955-020-02633-w
[14] C. Erignoux, P. Goncalves and G. Nahum. Hydrodynamics for SSEP with non-reversible slow boundary dynamics: Part II, the critical regime and beyond. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** (2020) 791–823. MR4153070 https://doi.org/10.30757/alea.v17-31
[15] T. Franco, P. Goncalves and A. Neumann. Hydrodynamical behavior of symmetric exclusion with slow bonds. *Ann. Inst. Henri Poincaré, Probab. Stat.* **49** (2) (2013) 402–427. MR3088375 https://doi.org/10.1214/11-AIHP445
[16] T. Franco, P. Goncalves and A. Neumann. Phase transition in equilibrium fluctuations of symmetric slowed exclusion. In *Stochastic Processes and Their Applications* 4156–4185, **123**, 2013. MR3096351 https://doi.org/10.1016/j.spa.2013.06.016
[17] T. Franco, P. Goncalves and A. Neumann. Phase transition of a heat equation with Robin's boundary conditions and exclusion process. *Trans. Amer. Math. Soc.* **367** (9) (2015) 6131–6158. MR3356932 https://doi.org/10.1090/S0002-9947-2014-06260-0
[18] T. Franco, P. Goncalves and A. Neumann. Equilibrium fluctuations for the slow boundary exclusion process. In *PSPDE 2015: From Particle Systems to Partial Differential Equations* 177–197, 2015. MR3746752 https://doi.org/10.1007/978-3-319-66839-0_9
[19] T. Franco, P. Goncalves and A. Neumann. Non-equilibrium and stationary fluctuations of a slowed boundary symmetric exclusion. In *Stochastic Processes and Their Applications* 1413–1442, **129**, 2019. MR3926561 https://doi.org/10.1016/j.spa.2018.05.005
[20] T. Franco, P. Goncalves and M. Simon. Crossover to the stochastic Burgers equation for the WASEP with a slow bond. *Comm. Math. Phys.* **346** (2016) 801–838. MR3537337 https://doi.org/10.1007/s00220-016-2607-x
[21] T. Franco and A. Neumann. Large deviations for the exclusion process with a slow bond. *Ann. Appl. Probab.* **27** (6) (2017) 3547–3587. MR3737932 https://doi.org/10.1214/17-AAP1287
[22] T. Franco, A. Neumann and G. Valle. Hydrodynamic limit for a type of exclusion process with slow bonds in dimension $d \geq 2$. *J. Appl. Probab.* **48** (2) (2011) 333–351. MR2840303 https://doi.org/10.1239/jap/1308662631
[23] P. Goncalves and M. Jara. Scaling limits of additive functionals of interacting particle systems. *Comm. Pure Appl. Math.* **6** (5) (2013) 649–677. MR3028483 https://doi.org/10.1002/cpa.21441
[24] P. Goncalves and M. Jara. Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.* **212** (2014) 597–644. MR3176353 https://doi.org/10.1007/s00205-013-0693-x
[25] P. Goncalves and M. Jara. Stochastic Burgers equation from long range exclusion interactions. In *Stochastic Processes and Their Applications* 4029–4052, **127**, 2017. MR3718105 https://doi.org/10.1016/j.spa.2017.03.022
[26] P. Goncalves, M. Jara and S. Sethuraman. A stochastic Burgers equation from a class of microscopic interactions. *Ann. Probab.* **43** (1) (2015) 286–338. MR3298474 https://doi.org/10.1214/13-AOP878
[27] M. Gubinelli and N. Perkowski. Energy solutions of KPZ are unique. *J. Amer. Math. Soc.* **31** (2018) 427–471. MR3758149 https://doi.org/10.1090/jams/889
[28] M. Z. Guo, G. C. Papnicolaou and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.* **118** (1988) 31–59. MR0954674
[29] M. Hairer. Solving the KPZ equation. *Ann. of Math.* **178** (2) (2013) 559–664. MR3071506 https://doi.org/10.4007/annals.2013.178.2.4
[30] M. Hairer. A theory of regularity structures. *Invent. Math.* **198** (2) (2014) 269–504. MR3274562 https://doi.org/10.1007/s00222-014-0505-4
[31] M. Hairer and J. Quastel. A class of growth models rescaling to KPZ. *Forum Math. Pi* **6** (2018) E3. MR3877863 https://doi.org/10.1017/fmp.2018.2
[32] S. Janowsky and J. Lebowitz. Finite size effects and shock fluctuations in the asymmetric simple exclusion process. *Phys. Rev. A* **45** (1992) 618.
[33] S. Janowsky and J. Lebowitz. Exact results for the asymmetric simple exclusion process with a blockage. *J. Stat. Phys.* **77** (1994) 35–51. MR1300527 https://doi.org/10.1007/BF02186831
[34] M. Kardar, G. Parisi and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Phys. Rev. Lett.* **56** (9) (1986) 889.
[35] C. Kipnis and C. Landim. *Scaling Limits of Interacting Particle Systems*, **320**. Springer-Verlig, Berlin Heidelberg, 1999. MR1707314 https://doi.org/10.1007/978-3-662-03752-2
[36] C. Mueller. On the support of solutions to the heat equation with noise. *Stoch. Stoch. Rep.* **37** (4) (1991) 225–245. MR1149348 https://doi.org/10.1080/17442509108833738
[37] J. Quastel and S. Sarkar. Convergence of exclusion processes and KPZ equation to the KPZ fixed point. *J. Amer. Math. Soc.* **36** (2023) 251–289. MR4495842 https://doi.org/10.1090/jams/999
[38] K. Yang. Kardar–Parisi–Zhang Equation from Long-Range Exclusion Processes. Accepted, *Communications in Mathematical Physics.* (2020). Available at arXiv:2002.05176 [math.PR].
[39] K. Yang. KPZ equation from non-simple variations on open ASEP. *Probab. Theory Related Fields* **183** (2022) 415–545. MR4421178 https://doi.org/10.1007/s00440-022-01133-0
[40] H. T. Yau. Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probab. Theory Related Fields* **109** (1997) 507–538. MR1483598 https://doi.org/10.1007/s004400050140

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1295–1333
https://doi.org/10.1214/22-AIHP1349

Convergence of limit shapes for 2D near-critical first-passage percolation

Chang-Long Yao[a]

Academy of Mathematics and Systems Science, CAS, Beijing, China, [a]*deducemath@126.com*

Abstract. We consider Bernoulli first-passage percolation on the triangular lattice in which sites have 0 and 1 passage times with probability p and $1-p$, respectively. For each $p \in (0, p_c)$, let $\mathcal{B}(p)$ be the limit shape in the classical "shape theorem", and let $L(p)$ be the correlation length. We show that as $p \uparrow p_c$, the rescaled limit shape $L(p)^{-1}\mathcal{B}(p)$ converges to a Euclidean disk. This improves a result of Chayes et al. [*J. Stat. Phys.* **45** (1986) 933–951]. The proof relies on the scaling limit of near-critical percolation established by Garban et al. [*J. Eur. Math. Soc.* **20** (2018) 1195–1268], and uses the construction of the collection of continuum clusters in the scaling limit introduced by Camia et al. [*Springer Proceedings in Mathematics & Statistics*, **299** (2019) 44–89].

Résumé. Nous considérons la percolation de premier passage de Bernoulli sur le réseau triangulaire dans lequel les sites ont des temps de passage de 0 et 1 avec une probabilité de p et $1-p$, respectivement. Pour tout $p \in (0, p_c)$, soit $\mathcal{B}(p)$ la forme limite donnée par le "théorème de la forme" classique, et soit $L(p)$ la longueur de corrélation. Nous montrons que lorsque $p \uparrow p_c$, la forme limite renormalisée $L(p)^{-1}\mathcal{B}(p)$ converge vers un disque Euclidien. Ceci améliore un résultat de Chayes et al. [*J. Stat. Phys.* **45** (1986) 933–951]. La preuve repose sur la limite d'échelle de la percolation presque-critique établie par Garban et al. [*J. Eur. Math. Soc.* **20** (2018) 1195–1268], et utilise la construction de l'ensemble de clusters dans le continu dans la limite d'échelle introduite par Camia et al. [*Springer Proceedings in Mathematics & Statistics*, **299** (2019) 44–89].

MSC2020 subject classifications: Primary 60K35; secondary 82B43

Keywords: First-passage percolation; Near-critical percolation; Scaling limit; Correlation length; Shape theorem

References

[1] D. Ahlberg and J. E. Steif. Scaling limits for the threshold window: When does a monotone Boolean function flip its outcome? *Ann. Inst. Henri Poincaré Probab. Stat.* **53** (2017) 2135–2161. MR3729650 https://doi.org/10.1214/16-AIHP786
[2] A. Auffinger, M. Damron and J. Hanson. *50 Years of First-Passage Percolation*. *University Lecture Series* **68**. Am. Math. Soc., Providence, 2017. MR3729447 https://doi.org/10.1090/ulect/068
[3] V. Beffara and P. Nolin. On monochromatic arm exponents for 2D critical percolation. *Ann. Probab.* **39** (2011) 1286–1304. MR2857240 https://doi.org/10.1214/10-AOP581
[4] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, New York, 2006. MR2283880 https://doi.org/10.1017/CBO9781139167383
[5] F. Camia, R. Conijn and D. Kiss. Conformal measure ensembles for percolation and the FK-Ising model. In *Sojourns in Probability Theory and Statistical Physics – II Brownian Web and Percolation, a Festschrift for Charles M. Newman* 44–89. *Springer Proceedings in Mathematics & Statistics* **299**, 2019. MR4044275
[6] F. Camia and C. M. Newman. Critical percolation: The full scaling limit. *Comm. Math. Phys.* **268** (2006) 1–38. MR2249794 https://doi.org/10.1007/s00220-006-0086-1
[7] J. T. Chayes, L. Chayes and R. Durrett. Critical behavior of the two-dimensional first passage time. *J. Stat. Phys.* **45** (1986) 933–951. MR0881316 https://doi.org/10.1007/BF01020583
[8] J. T. Cox and H. Kesten. On the continuity of the time constant of first-passage percolation. *J. Appl. Probab.* **18** (1981) 809–819. MR0633228 https://doi.org/10.1017/s0021900200034161
[9] M. Damron, J. Hanson and W.-K. Lam. Universality of the time constant for 2D critical first-passage percolation. *Ann. Appl. Probab.* To appear. Available at arXiv:1904.12009. MR3706736 https://doi.org/10.1214/16-AOP1129
[10] H. Duminil-Copin. Limit of the Wulff Crystal when approaching criticality for site percolation on the triangular lattice. *Electron. Commun. Probab.* **18** (2013) 93. MR3151749 https://doi.org/10.1214/ECP.v18-3163
[11] R. Durrett. *Probability: Theory and Examples*, 4th edition. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. MR2722836 https://doi.org/10.1017/CBO9780511779398

[12] C. Garban, G. Pete and O. Schramm. Pivotal, cluster and interface measures for critical planar percolation. *J. Amer. Math. Soc.* **26** (2013) 939–1024. MR3073882 https://doi.org/10.1090/S0894-0347-2013-00772-9

[13] C. Garban, G. Pete and O. Schramm. The scaling limits of near-critical and dynamical percolation. *J. Eur. Math. Soc.* **20** (2018) 1195–1268. MR3790067 https://doi.org/10.4171/JEMS/786

[14] C. Garban, G. Pete and O. Schramm. The scaling limits of the minimal spanning tree and invasion percolation in the plane. *Ann. Probab.* **46** (2018) 3501–3557. MR3857861 https://doi.org/10.1214/17-AOP1252

[15] O. Garet and R. Marchand. Large deviations for the chemical distance in supercritical Bernoulli percolation. *Ann. Probab.* **35** (2007) 833–866. MR2319709 https://doi.org/10.1214/009117906000000881

[16] G. Grimmett. *Percolation*, 2nd edition. Springer-Verlag, Berlin, 1999. MR1707339 https://doi.org/10.1007/978-3-662-03981-6

[17] G. Grimmett and H. Kesten. First-passage percolation, network flows and electrical resistances. *Z. Wahrsch. Verw. Gebiete* **66** (1984) 335–366. MR0751574 https://doi.org/10.1007/BF00533701

[18] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif.* 61–110. Springer-Verlag, New York, 1965. MR0198576

[19] N. Holden and X. Sun. Convergence of uniform triangulations under the Cardy embedding. *Acta Math.* To appear. Available at arXiv:1905.13207v3.

[20] J. Jiang and C.-L. Yao. Critical first-passage percolation starting on the boundary. *Stochastic Process. Appl.* **129** (2019) 2049–2065. MR3958423 https://doi.org/10.1016/j.spa.2018.06.008

[21] H. Kesten. Aspects of first passage percolation. In *Lecture Notes in Math.* 125–264, **1180**. Springer, Berlin, 1986. MR0876084 https://doi.org/10.1007/BFb0074919

[22] H. Kesten. Scaling relations for 2D-percolation. *Comm. Math. Phys.* **109** (1987) 109–156. MR0879034

[23] H. Kesten. First-passage percolation. From classical to modern probability. In *Progr. Probab.* 93–143, **54**. Birkhäuser, Basel, 2003. MR2045986

[24] H. Kesten and Y. Zhang. A central limit theorem for "critical" first-passage percolation in two-dimensions. *Probab. Theory Related Fields* **107** (1997) 137–160. MR1431216 https://doi.org/10.1007/s004400050080

[25] G. F. Lawler, O. Schramm and W. Werner. One-arm exponent for critical 2D percolation. *Electron. J. Probab.* **7** (2002) 2. MR1887622 https://doi.org/10.1214/EJP.v7-101

[26] T. M. Liggett. An improved subadditive ergodic theorem. *Ann. Probab.* **13** (1985) 1279–1285. MR0806224

[27] P. Nolin. Near critical percolation in two-dimensions. *Electron. J. Probab.* **13** (2008) 1562–1623. MR2438816 https://doi.org/10.1214/EJP.v13-565

[28] D. Reimer. Proof of the van den Berg–Kesten conjecture. *Combin. Probab. Comput.* **9** (2000) 27–32. MR1751301 https://doi.org/10.1017/S0963548399004113

[29] O. Schramm and S. Smirnov. On the scaling limits of planar percolation (with an appendix by C. Garban). *Ann. Probab.* **39** (2011) 1768–1814. MR2884873 https://doi.org/10.1214/11-AOP659

[30] S. Sheffield. Exploration trees and conformal loop ensembles. *Duke Math. J.* **147** (2009) 79–129. MR2494457 https://doi.org/10.1215/00127094-2009-007

[31] S. Sheffield and W. Werner. Conformal loop ensembles: The Markovian characterization and the loop-soup construction. *Ann. Math.* **176** (2012) 1827–1917. MR2979861 https://doi.org/10.4007/annals.2012.176.3.8

[32] S. Smirnov and W. Werner. Critical exponents for two-dimensional percolation. *Math. Res. Lett.* **8** (2001) 729–744. MR1879816 https://doi.org/10.4310/MRL.2001.v8.n6.a4

[33] N. Sun. Conformally invariant scaling limits in planar critical percolation. *Probab. Surv.* **8** (2011) 155–209. MR2846901 https://doi.org/10.1214/11-PS180

[34] J. van den Berg and R. Conijn. The gaps between the sizes of large clusters in 2D critical percolation. *Electron. Commun. Probab.* **18** (2013) 92. MR3145048 https://doi.org/10.1214/ECP.v18-3065

[35] J. van den Berg, D. Kiss and P. Nolin. Two-dimensional volume-frozen percolation: Deconcentration and prevalence of mesoscopic clusters. *Ann. Sci. Éc. Norm. Supér.* **51** (2018) 1017–1084. MR3861568 https://doi.org/10.24033/asens.2371

[36] J. van den Berg and P. Nolin. Near-critical percolation with heavy-tailed impurities, forest fires and frozen percolation. *Probab. Theory Related Fields* **181** (2021) 211–290. MR4341073 https://doi.org/10.1007/s00440-020-01022-4

[37] W. Werner. Lectures on two-dimensional critical percolation. In *Statistical Mechanics* 297–360. *IAS/Park City Math. Ser.* **16**. Amer. Math. Soc., Providence, RI, 2009. MR2523462 https://doi.org/10.1090/pcms/016/06

[38] C.-L. Yao. Law of large numbers for critical first-passage percolation on the triangular lattice. *Electron. Commun. Probab.* **19** (2014) 18. MR3183571 https://doi.org/10.1214/ECP.v19-3268

[39] C.-L. Yao. Limit theorems for critical first-passage percolation on the triangular lattice. *Stochastic Process. Appl.* **128** (2018) 445–460. MR3739504 https://doi.org/10.1016/j.spa.2017.05.002

[40] C.-L. Yao. Asymptotics for 2D critical and near-critical first-passage percolation. *Probab. Theory Related Fields* **175** (2019) 975–1019. MR4026611 https://doi.org/10.1007/s00440-019-00908-2

[41] C.-L. Yao. Convergence of limit shapes for 2D near-critical first-passage percolation (original preprint version of this paper). Available at arXiv:2104.01211v2.

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1334–1355
https://doi.org/10.1214/22-AIHP1352

Random walks in Dirichlet environments on $\mathbb{Z}$ with bounded jumps

Daniel J. Slonim[a]

Department of Mathematics, Purdue University, 150 N University St., West Lafayette, IN 47907, USA , [a]*dslonim@virginia.edu*

Abstract. We examine a class of random walks in random environments on $\mathbb{Z}$ with bounded jumps, a generalization of the classic one-dimensional model. The environments we study have i.i.d. transition probability vectors drawn from Dirichlet distributions. For the transient case of this model, we characterize ballisticity—nonzero limiting velocity. We do this in terms of two parameters, κ_0 and κ_1. The parameter κ_0 governs finite trapping effects. The parameter κ_1, which already is known to characterize directional transience, also governs repeated traversals of arbitrarily large regions of the graph. We show that the walk is ballistic if and only if $\min(\kappa_0, |\kappa_1|) > 1$. We prove some stronger results regarding moments of the quenched Green function and other functions that the quenched Green function dominates. These results help us to better understand the phenomena and parameters affecting ballisticity.

Résumé. Nous considérons une classe de marches aléatoires dans des environnements aléatoires sur $\mathbb{Z}$ avec des sauts bornés, une généralisation du modèle classique unidimensionnel. Les environnements que nous étudions ont des vecteurs de probabilité de transition i.i.d. tirés selon des lois de Dirichlet. Pour le cas transient de ce modèle, nous caractérisons la balisticité (vitesse limite non nulle). Nous le faisons en fonction de deux paramètres, κ_0 et κ_1. Le paramètre κ_0 régit les effets de pièges finis. Le paramètre κ_1, qui est déjà connu pour caractériser la transience directionnelle, contrôle également les traversées répétées de régions arbitrairement grandes du graphe. Nous montrons que la marche est balistique si et seulement si $\min(\kappa_0, \kappa_1) > 1$. Nous prouvons des résultats plus forts concernant les moments de la fonction de Green et d'autres fonctions que la fonction de Green domine. Ces résultats nous aident à mieux comprendre les phénomènes et les paramètres qui influent sur la balisticité.

MSC2020 subject classifications: 60G50; 60J10; 60K37

Keywords: Random walk; Random environment; Dirichlet environments; Bounded jumps; Ballisticity

References

[1] E. Bolthausen and I. Goldsheid. Recurrence and transience criteria for random walk in a random environment. *Comm. Math. Phys.* **214** (2000) 429–447. MR1796029 https://doi.org/10.1007/s002200000279
[2] É. Bouchet. Sub-ballistic random walk in Dirichlet environment. *Electron. J. Probab.* **18** (2013). MR3068389 https://doi.org/10.1214/EJP.v18-2109
[3] J. Brémont. Random walks in random medium odn Z and Lyapunov spectrum. *Ann. Inst. Henri Poincaré Probab. Stat.* **40** (3) (2004) 309–336. MR2060456 https://doi.org/10.1016/j.anihpb.2003.10.006
[4] J. Brémont. One-dimensional finite range random walk in random medium and invariant measure equation. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** (1) (2009) 70–103. MR2500229 https://doi.org/10.1214/07-AIHP150
[5] N. Enriquez and C. Sabot. Edge oriented reinforced random walks and RWRE. *C. R. Math. Acad. Sci. Paris* **335** (2002) 941–946. MR1952554 https://doi.org/10.1016/S1631-073X(02)02580-3
[6] M. S. Keane and S. W. W. Rolles. Tubular recurrence. *Acta Math. Hungar.* **97** (2002) 207–221. MR1933730 https://doi.org/10.1023/A:1020855011898
[7] E. S. Key. Recurrence and transience criteria for random walk in a random environment. *Ann. Probab.* **12** (2) (1984) 529–560. MR0735852 https://doi.org/10.1214/aop/1176993304
[8] A. Roitershtein. Transient random walks on a strip in a random environment. *Ann. Probab.* **36** (6) (2008) 2354–2387. MR2478686 https://doi.org/10.1214/08-AOP393
[9] C. Sabot. Random walks in random Dirichlet environment are transient in dimension $d \geq 3$. *Probab. Theory Related Fields* **151** (1) (2011) 297–317. MR2834720 https://doi.org/10.1007/s00440-010-0300-0
[10] C. Sabot and L. Tournier. Reversed Dirichlet environment and directional transience of random walks in Dirichlet environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** (1) (2011) 1–8. MR2779393 https://doi.org/10.1214/09-AIHP344

[11] C. Sabot and L. Tournier. Random walks in Dirichlet environment: An overview. *Ann. Fac. Sci. Univ. Toulouse Sci. Math. Sci. Phys.* **26** (2016) 463–509. MR3640900 https://doi.org/10.5802/afst.1542
[12] D. J. Slonim Directional transience of random walks in Dirichlet environments with bounded jumps. Preprint, submitted, 2021. Available at https://arxiv.org/abs/2108.11424.
[13] D. J. Slonim. On total weight exiting finite, strongly connected sets in shift-invariant weighted directed graphs on $\mathbb{Z}$. Preprint, submitted, 2022. Available at https://arxiv.org/abs/2205.07414.
[14] D. J. Slonim. Ballisticity of random walks in random environments on $\mathbb{Z}$ with bounded jumps. *Markov Process. Relat. Fields* **28** (5) (2022) 659–671.
[15] F. Solomon. Random walks in a random environment. *Ann. Probab.* **3** (1) (1975) 1–31. MR0362503 https://doi.org/10.1214/aop/1176996444
[16] L. Tournier. Integrability of exit times and ballisticity for random walks in Dirichlet environment. *Electron. J. Probab.* **14** (2009) 431–451. MR2480548 https://doi.org/10.1214/EJP.v14-609
[17] L. Tournier. Asymptotic direction of random walks in Dirichlet environment. *Ann. Inst. Henri Poincaré Probab. Stat.* **51** (2) (2015) 716–726. MR3335022 https://doi.org/10.1214/13-AIHP582

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1356–1386
https://doi.org/10.1214/22-AIHP1354

Fractal properties of Aldous–Kendall random metric

Guillaume Blanc[a]

Université Paris-Saclay, Orsay, France, [a]*guillaume.blanc1@universite-paris-saclay.fr*

Abstract. Investigating a model of scale-invariant random spatial network suggested by Aldous, Kendall constructed a random metric T on $\mathbb{R}^d$, for which the distance between points is given by the optimal connection time, when travelling on the road network generated by a Poisson process of lines with a speed limit. In this paper, we look into some fractal properties of that random metric. In particular, although almost surely the metric space $(\mathbb{R}^d, T)$ is homeomorphic to the usual Euclidean $\mathbb{R}^d$, we prove that its Hausdorff dimension is given by $(\gamma - 1)d/(\gamma - d) > d$, where $\gamma > d$ is a parameter of the model; which confirms a conjecture of Kahn. We also find that the metric space $(\mathbb{R}^d, T)$ equipped with the Lebesgue measure exhibits a multifractal property, as some points have untypically big balls around them.

Résumé. En étudiant un modèle de "scale-invariant random spatial network" suggéré par Aldous, Kendall a construit une métrique aléatoire T sur $\mathbb{R}^d$, pour laquelle la distance entre les points est donnée par le temps de trajet optimal, lorsqu'on se déplace sur le réseau de routes engendré par un processus de Poisson de droites avec une limitation de vitesse. Dans cet article, nous nous intéressons aux propriétés fractales de cette métrique aléatoire. En particulier, bien que presque sûrement l'espace métrique $(\mathbb{R}^d, T)$ soit homéomorphe à l'espace euclidien $\mathbb{R}^d$, nous montrons que sa dimension de Hausdorff est donnée par $(\gamma - 1)d/(\gamma - d) > d$, où $\gamma > d$ est un paramètre du modèle ; cela confirme une conjecture de Kahn. Nous montrons par ailleurs que l'espace métrique $(\mathbb{R}^d, T)$ muni de la mesure de Lebesgue est multifractal, puisque certains points se trouvent être au centre de boules atypiquement grosses.

MSC2020 subject classifications: 60D05

Keywords: Random geometry; Poisson process; Hausdorff dimension

References

[1] D. Aldous. Scale-invariant random spatial networks. *Electron. J. Probab.* **19** (2014) 1–41. MR3164768 https://doi.org/10.1214/EJP.v19-2920
[2] S. N.Chiu, D. Stoyan, W. S. Kendall and J. Mecke. *Stochastic Geometry and Its Applications*, 3rd edition. *Wiley Series in Probability and Statistics*. Wiley, New York, 2013. MR3236788 https://doi.org/10.1002/9781118658222
[3] R. Durrett. *Probability: Theory and Examples*, 5th edition. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. MR3930614 https://doi.org/10.1017/9781108591034
[4] J. B.Hough, M. Krishnapur, Y. Peres and B. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*. *University Lecture Series* **51**. American Mathematical Society, Providence, 2009. MR2552864 https://doi.org/10.1090/ulect/051
[5] J. Kahn. Improper Poisson line process as SIRSN in any dimension. *Ann. Probab.* **44** (2016) 2694–2725. MR3531678 https://doi.org/10.1214/15-AOP1032
[6] O. Kallenberg. *Foundations of Modern Probability*, 2nd edition. *Probability and Its Applications.* Springer, Berlin, 2002. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8
[7] W. S. Kendall. From random lines to metric spaces. *Ann. Probab.* **45** (2017) 469–517. MR3601654 https://doi.org/10.1214/14-AOP935
[8] W. S. Kendall. Rayleigh random flights on the Poisson line SIRSN. *Electron. J. Probab.* **25** (2020) 1–36. MR4161134 https://doi.org/10.1214/20-ejp526
[9] G. Last and M. Penrose. *Lectures on the Poisson Process*. *Institute of Mathematical Statistics Textbooks*. Cambridge University Press, Cambridge, 2017. MR3791470
[10] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.* **41** (2013) 2880–2960. MR3112934 https://doi.org/10.1214/12-AOP792
[11] J.-F. Le Gall. Brownian geometry. *Jpn. J. Math.* **14** (2019) 135–174. MR4007665 https://doi.org/10.1007/s11537-019-1821-7
[12] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*. *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. MR1333890 https://doi.org/10.1017/CBO9780511623813
[13] P. Mörters and Y. Peres. *Brownian Motion*. *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. MR2604525 https://doi.org/10.1017/CBO9780511750489
[14] J. Munkres. *Topology*, 2nd edition. Pearson, Upper Saddle River, 2000. MR3728284

[15] W. Rudin. *Real and Complex Analysis*, 3rd edition. McGraw-Hill, New York, 1987. MR0924157

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

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1387–1417
https://doi.org/10.1214/22-AIHP1358

The Seneta–Heyde scaling for supercritical super-Brownian motion

Haojie Hou[1,a], Yan-Xia Ren[2,b] and Renming Song[3,c]

[1]*School of Mathematical Sciences, Peking University, Beijing, 100871, P.R. China,* [a]*houhaojie@pku.edu.cn*
[2]*LMAM School of Mathematical Sciences & Center for Statistical Science, Peking University, Beijing, 100871, P.R. China,* [b]*yxren@math.pku.edu.cn*
[3]*Department of Mathematics, University of Illinois, Urbana, IL 61801, U.S.A.,* [c]*rsong@math.uiuc.edu*

Abstract. We consider the additive martingale $W_t(\lambda)$ and the derivative martingale $\partial W_t(\lambda)$ for one-dimensional supercritical super-Brownian motions with general branching mechanism. In the critical case $\lambda = \lambda_0$, we prove that $\sqrt{t} W_t(\lambda_0)$ converges in probability to a positive limit, which is a constant multiple of the almost sure limit $\partial W_\infty(\lambda_0)$ of the derivative martingale $\partial W_t(\lambda_0)$. We also prove that, on the survival event, $\limsup_{t\to\infty} \sqrt{t} W_t(\lambda_0) = \infty$ almost surely.

Résumé. Nous considérons la martingale additive $W_t(\lambda)$ et la martingale dérivée $\partial W_t(\lambda)$ pour les super-mouvements browniens surcritiques unidimensionnels avec mécanisme général de branchement. Dans le cas critique où $\lambda = \lambda_0$, nous prouvons que $\sqrt{t} W_t(\lambda_0)$ converge en probabilité vers une limite positive, qui est un multiple constant de la limite presque sûre $\partial W_\infty(\lambda_0)$ de la martingale dérivée $\partial W_t(\lambda_0)$. Nous prouvons également que, dans l'événement de survie, $\limsup_{t\to\infty} \sqrt{t} W_t(\lambda_0) = \infty$ presque sûrement.

MSC2020 subject classifications: 60J68; 60F05; 60F15

Keywords: Seneta–Heyde scaling; Super-Brownian motion; Spine decomposition; Skeleton decomposition; Additive martingale; Derivative martingale

References

[1] E. Aïdékon and Z. Shi. The Seneta–Heyde scaling for the branching random walk. *Ann. Probab.* **42** (2014) 959–993. MR3189063 https://doi.org/10.1214/12-AOP809
[2] J. Berestycki, A. E. Kyprianou and A. Murillo-Salas. The prolific backbone for supercritical superprocesses. *Stochastic Process. Appl.* **121** (2011) 1315–1331. MR2794978 https://doi.org/10.1016/j.spa.2011.02.004
[3] J. D. Biggins and A. E. Kyprianou. Seneta–Heyde norming in the branching random walk. *Ann. Probab.* **25** (1997) 337–360. MR1428512 https://doi.org/10.1214/aop/1024404291
[4] J. D. Biggins and A. E. Kyprianou. Measure change in multitype branching. *Adv. in Appl. Probab.* **36** (2004) 544–581. MR2058149 https://doi.org/10.1239/aap/1086957585
[5] J. D. Biggins and A. E. Kyprianou. Branching random walk: Seneta–Heyde norming. In *Trees* 31–49. B. Chauvin et al (Eds). *Versailles*, *1995*. *Progr. Probab.* **40**. Birkhäuser, Basel, 1996. MR1439971
[6] B. Chauvin. Multiplicative martingales and stopping lines for branching Brownian motion. *Ann. Probab.* **30** (1991) 1195–1205. MR1112412
[7] T. Duquesne and M. Winkel. Growth of Lévy trees. *Probab. Theory Related Fields* **139** (2007) 313–371. MR2322700 https://doi.org/10.1007/s00440-007-0064-3
[8] R. Durrett. *Probability: Theory and Examples*, 4th edition. *Cambridge Series in Statistical and Probabilistic Mathematics* **31**. Cambridge University Press, Cambridge, 2010. MR2722836 https://doi.org/10.1017/CBO9780511779398
[9] E. B. Dynkin. Branching particle systems and superprocesses. *Ann. Probab.* **19** (3) (1991) 1157–1194. MR1112411
[10] E. B. Dynkin. Superprocesses and partial differential equations. *Ann. Probab.* **21** (1993) 1185–1262. MR1235414
[11] E. B. Dynkin. Branching exit Markov systems and superprocesses. *Ann. Probab.* **29** (2001) 1833–1858. MR1880244 https://doi.org/10.1214/aop/1015345774
[12] E. B. Dynkin. *Diffusions, Superdiffusions and Partial Differential Equations*. AMS, Providence, R.I., 2002. MR1883198 https://doi.org/10.1090/coll/050
[13] E. B. Dynkin and S. E. Kuznetsov. ℕ-Measures for branching exit Markov systems and their applications to differential equations. *Probab. Theory Related Fields* **130** (1) (2004) 135–150. MR2092876 https://doi.org/10.1007/s00440-003-0333-8
[14] M. Eckhoff, A. E. Kyprianou and M. Winkel. Spine, skeletons and the strong law of large numbers. *Ann. Probab.* **43** (2015) 2594–2659. MR3395469 https://doi.org/10.1214/14-AOP944
[15] H. He, J.-N. Liu and M. Zhang. On Seneta–Heyde scaling for a stable branching random walk. *Adv. in Appl. Probab.* **50** (2) (2018) 565–599. MR3832885 https://doi.org/10.1017/apr.2018.25

[16] C. C. Heyde. Extension of a result of Seneta for the super-critical Galton–Watson process. *Ann. Math. Stat.* **41** (1970) 739–742. MR0254929 https://doi.org/10.1214/aoms/1177697127

[17] Y. Hu and Z. Shi. Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* **37** (2009) 742–789. MR2510023 https://doi.org/10.1214/08-AOP419

[18] J.-P. Imhof. Density factorizations for Brownian motion, meander and the three dimensional Bessel process, and applications. *J. Appl. Probab.* **21** (1984) 500–510. MR0752015 https://doi.org/10.2307/3213612

[19] A. E. Kyprianou. Travelling wave solutions to the K-P-P equation: Alternatives to Simon Harris' probabilistic analysis. *Ann. Inst. Henri Poincaré Probab. Stat.* **40** (2004) 53–72. MR2037473 https://doi.org/10.1016/S0246-0203(03)00055-4

[20] A. E. Kyprianou, R.-L. Liu, A. Murillo-Salas and Y.-X. Ren. Supercritical super-Brownian motion with a general branching mechanism and travelling waves. *Ann. Inst. Henri Poincaré Probab. Stat.* **48** (2012) 661–687. MR2976558 https://doi.org/10.1214/11-AIHP448

[21] Z. Li. *Measure Valued Branching Markov Processes*. Springer, Berlin, 2011. MR2760602 https://doi.org/10.1007/978-3-642-15004-3

[22] P. Maillard and M. Pain. 1-stable fluctuations in branching Brownian motion at critical temperature I: The derivative martingale. *Ann. Probab.* **47** (5) (2019) 2953–3002. MR4021242 https://doi.org/10.1214/18-AOP1329

[23] J. Neveu. Multiplicative martingales for spatial branching processes. In *Seminar on Stochastic Processes, 1987* 223–241. *Progr. Probab. Statist.* **15**. Birkhaüser, Boston, 1988. MR1046418 https://doi.org/10.1007/978-1-4684-0550-7_10

[24] Y.-X. Ren, R. Song and R. Zhang. The extremal process of super-Brownian motion. *Stochastic Process. Appl.* **137** (2021) 1–34. MR4242982 https://doi.org/10.1016/j.spa.2021.03.007

[25] E. Seneta. On recent theorems concerning the supercritical Galton–Watson process. *Ann. Math. Stat.* **39** (1968) 2098–2102. MR0234530 https://doi.org/10.1214/aoms/1177698037

[26] T. Shiga and S. Watanabe. Bessel diffusions as a one-parameter family of diffusion processes. *Z. Wahrsch. Verw. Gebiete* **27** (1973) 37–46. MR0368192 https://doi.org/10.1007/BF00736006

[27] T. Yang and Y.-X. Ren. Limit theorem for derivative martingale at criticality w.r.t. branching Brownian motion. *Statist. Probab. Lett.* **81** (2011) 195–200. MR2748182 https://doi.org/10.1016/j.spl.2010.11.007

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1418–1457
https://doi.org/10.1214/23-AIHP1361

A branching process with deletions and mergers that matches the threshold for hypercube percolation

Laura Eslava[1,a], Sarah Penington[2,b] and Fiona Skerman[3,c]

[1] *Instituto de investigacidades en matematicas aplicadas y en sistemas, Universidad Nacional Autonoma Mexico, Mexico,* [a]*laura@sigma.iimas.unam.mx*
[2] *Department of Mathematical Sciences, University of Bath, Bath, UK,* [b]*sp2355@bath.ac.uk*
[3] *Department of Mathematics, Uppsala University, Uppsala, Sweden,* [c]*fiona.skerman@math.uu.se*

Abstract. We define a graph process $\mathcal{G}(p, q)$ based on a discrete branching process with deletions and mergers, which is inspired by the 4-cycle structure of both the hypercube Q_d and the lattice $\mathbb{Z}^d$ for large d. Individuals have Poisson offspring distribution with mean $1 + p$ and certain deletions and mergers occur with probability q; these parameters correspond to the mean number of edges discovered from a given vertex in an exploration of a percolation cluster and to the probability that a non-backtracking path of length four closes a cycle, respectively.

We prove survival and extinction under certain conditions on p and q that heuristically match the known expansions of the critical probabilities for bond percolation on the lattice $\mathbb{Z}^d$ and the hypercube Q_d. These expansions have been rigorously established by Hara and Slade in 1995, and van der Hofstad and Slade in 2006, respectively. We stress that our method does not constitute a branching process proof for the percolation threshold. However, it can provide a conjecture for other high-dimensional, odd-cycle free transitive graphs such as the body-centered cubic lattice.

The analysis of the graph process survival is considerably more challenging than for branching processes in discrete time, due to the interdependence between the descendants of different individuals in the same generation. In fact, it is left open whether the survival probability of $\mathcal{G}(p, q)$ is monotone in p or q; we discuss this and some other open problems regarding the new graph process.

Résumé. Nous définissons un processus de graphes $\mathcal{G}(p, q)$ à partir d'un processus de branchement discret avec suppressions et fusions, qui s'inspire de la structure à 4 cycles de l'hypercube Q_d et du réseau $\mathbb{Z}^d$ pour des valeurs élevées de d. Les individus ont une loi de reproduction de Poisson avec une moyenne de $1 + p$ et certaines suppressions et fusions se produisent avec une probabilité q ; ces paramètres correspondent respectivement au nombre moyen d'arêtes découvertes à partir d'un sommet donné dans une exploration d'un amas de percolation et à la probabilité qu'un chemin sans retour de longueur quatre ferme un cycle.

Nous prouvons la survie et l'extinction sous certaines conditions sur p et q qui correspondent heuristiquement aux expansions connues des probabilités critiques de percolation des liaisons sur le réseau $\mathbb{Z}^d$ et l'hypercube Q_d. Ces expansions ont été rigoureusement établies par Hara et Slade en 1995, et van der Hofstad et Slade en 2006, respectivement. Nous soulignons que notre méthode ne constitue pas une preuve pour le seuil de percolation qui utiliserait les processus de branchement.

L'analyse de la survie du processus de graphes est considérablement plus difficile que pour les processus de branchement en temps discret, en raison de l'interdépendance entre les descendants de différents individus dans la même génération. Par exemple, le fait que la probabilité de survie de $\mathcal{G}(p, q)$ est monotone en p ou q n'est pas clair ; nous discutons de ceci et de quelques autres problèmes ouverts concernant le nouveau processus de graphes.

MSC2020 subject classifications: Primary 60J80; secondary 60C05; 60K35; 05C80

Keywords: Branching process; Survival threshold; Percolation threshold; Hypercube; Graph exploration

References

[1] M. Ajtai, J. Komlós and E. Szemerédi. Largest random component of a k-cube. *Combinatorica* **2** (1) (1982) 1–7. MR0671140 https://doi.org/10.1007/BF02579276
[2] K. B. Athreya and P. E. Ney. *Branching Processes*. *Math. Surveys* **7**. Springer-Verlag, Berlin, 1972. MR0373040
[3] B. Bollobás. The evolution of sparse graphs. In *Graph Theory and Combinatorics: Proc. Cambridge Combinatorial Conf. in Honour of Paul Erdős* 35–57. Academic Press, New York, 1984. MR0777163
[4] B. Bollobás. The evolution of the cube. In *Combinatorial Mathematics (Marseille-Luminy, 1981)*. *North-Holland Math. Stud.* **75** 91–97. North-Holland, Amsterdam. MR0841284

[5] B. Bollobás, Y. Kohayakawa and T. Łuczak. The evolution of random subgraphs of the cube. *Random Structures Algorithms* **3** (1) (1992) 55–90. MR1139488 https://doi.org/10.1002/rsa.3240030106

[6] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, Cambridge, 2006. MR2283880 https://doi.org/10.1017/CBO9781139167383

[7] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs: I. The scaling window under the triangle condition. *Random Structures Algorithms* **27** (2) (2005) 137–184. MR2155704 https://doi.org/10.1002/rsa.20051

[8] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs: II. The lace expansion and the triangle condition. *Ann. Probab.* **33** (5) (2005) 1886–1944. MR2165583 https://doi.org/10.1214/009117905000000260

[9] C. Borgs, J. T. Chayes, R. van der Hofstad, G. Slade and J. Spencer. Random subgraphs of finite graphs: III. The phase transition for the n-cube. *Combinatorica* **26** (4) (2006) 395–410. MR2260845 https://doi.org/10.1007/s00493-006-0022-1

[10] P. Erdős and J. Spencer. Evolution of the n-cube. *Comput. Math. Appl.* **5** (1) (1979) 33–39. MR0534014 https://doi.org/10.1016/0898-1221(81)90137-1

[11] L. Federico, R. van der Hofstad, F. den Hollander and T. Hulshof. Expansion of percolation critical points for Hamming graphs. *Combin. Probab. Comput.* **29** (1) (2020) 68–100. MR4052928 https://doi.org/10.1017/s0963548319000208

[12] D. Gaunt and H. Ruskin. Bond percolation processes in d dimensions. *J. Phys. A: Math. Gen.* **11** (7) (1978) 1369–1380.

[13] S. Handa, Y. Kamijima and A. Sakai. A survey on the lace expansion for the nearest-neighbor models on the BCC lattice. *Taiwanese J. Math.* **24** (3) (2020) 723–784. MR4100717 https://doi.org/10.11650/tjm/190904

[14] T. Hara and G. Slade. The self-avoiding-walk and percolation critical points in high dimensions. *Combin. Probab. Comput.* **4** (3) (1995) 197–215. MR1356575 https://doi.org/10.1017/S0963548300001607

[15] M. Heydenreich and K. Matzke. Critical site percolation in high dimension. *J. Stat. Phys.* **3** (181) (2020) 816–853. MR4160912 https://doi.org/10.1007/s10955-020-02607-y

[16] M. Heydenreich and K. Matzke. Expansion for the critical point of site percolation: The first three terms. *Random Structures Algorithms* **31** (3) (2022) 430–454. MR4410719 https://doi.org/10.1017/s0963548321000365

[17] M. Heydenreich and R. van der Hofstad. *Progress in High-Dimensional Percolation and Random Graphs*. *CRM Short Courses*. Springer, Cham, 2017. MR3729454

[18] S. Janson and L. Warnke. On the critical probability in percolation. *Electron. J. Probab.* **23** (2018) Paper No. 1. MR3751076 https://doi.org/10.1214/17-EJP52

[19] S. Mertens and C. Moore. Series expansion of the percolation threshold on hypercubic lattices. *J. Phys. A: Math. Theor.* **51** (47) (2018) 475001. MR3876576 https://doi.org/10.1088/1751-8121/aae65c

[20] M. Molloy. Cores in random hypergraphs and Boolean formulas. *Random Structures Algorithms* **27** (1) (2005) 124–135. MR2150018 https://doi.org/10.1002/rsa.20061

[21] A. Nachmias and Y. Peres. Critical random graphs: Diameter and mixing time. *Ann. Probab.* **36** (4) (2008) 1267–1286. MR2435849 https://doi.org/10.1214/07-AOP358

[22] B. Pittel, J. Spencer and N. Wormald. Sudden emergence of a giant k-core in a random graph. *J. Combin. Theory Ser. B* **67** (1) (1996) 111–151. MR1385386 https://doi.org/10.1006/jctb.1996.0036

[23] O. Riordan. The k-core and branching processes. *Combin. Probab. Comput.* **17** (1) (2008) 111–136. MR2376426 https://doi.org/10.1017/S0963548307008589

[24] R. van der Hofstad and A. Nachmias. Unlacing hypercube percolation: A survey. *Metrika* **77** (1) (2014) 23–50. MR3152019 https://doi.org/10.1007/s00184-013-0473-5

[25] R. van der Hofstad and G. Slade. Asymptotic expansions in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$. *Random Structures Algorithms* **27** (3) (2005) 331–357. MR2162602 https://doi.org/10.1002/rsa.20074

[26] R. van der Hofstad and G. Slade. Expansion in n^{-1} for percolation critical values on the n-cube and $\mathbb{Z}^n$: The first three terms. *Combin. Probab. Comput.* **15** (5) (2006) 695–713. MR2248322 https://doi.org/10.1017/S0963548306007498

Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2024, Vol. 60, No. 2, 1458–1509
https://doi.org/10.1214/23-AIHP1367

Generalized range of slow random walks on trees

Pierre Andreoletti[a] and Alexis Kagan[b]

Institut Denis Poisson, UMR CNRS 7013, Université d'Orléans, Orléans, France, [a]*pierre.andreoletti@univ-orleans.fr,* [b]*alexis.kagan@univ-orleans.fr*

Abstract. In this work, we are interested in the set of visited vertices of a tree $\mathbb{T}$ by a randomly biased random walk $\mathbb{X} := (X_n, n \in \mathbb{N})$. The aim is to study a generalized range, that is to say the volume of the trace of $\mathbb{X}$ with both constraints on the trajectories of $\mathbb{X}$ and on the trajectories of the underlying branching random potential $\mathbb{V} := (V(x), x \in \mathbb{T})$. Focusing on slow regime's random walks (see Hu and Shi (2016); Andreoletti and Chen in Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018) 466–513), we prove a general result and detail examples. These examples exhibit many different behaviors for a wide variety of ranges, showing the interactions between the trajectories of $\mathbb{X}$ and the ones of $\mathbb{V}$.

Résumé. Nous nous intéressons aux sommets d'un arbre de Galton–Watson $\mathbb{T}$ visités par une marche biaisée aléatoirement $\mathbb{X} := (X_n, n \in \mathbb{N})$. Plus particulièrement, nous étudions une trace généralisée, c'est à dire le volume des points visités par $\mathbb{X}$ avec des containtes à la fois sur $\mathbb{X}$ et sur l'environnement aléatoire branchant $\mathbb{V} := (V(x), x \in \mathbb{T})$. Nous nous concentrons sur le régime lent (voir Hu et Shi (2016) ; Andreoletti et Chen dans Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018) 466–513), en montrant un résultat général et en détaillant des exemples caractéristiques. Ces exemples font apparaître une grande variété de comportements asymptotiques pour ce type de traces mettant en avant les interactions fortes entre $\mathbb{X}$ et $\mathbb{V}$.

MSC2020 subject classifications: Primary 60K37; 60J80; secondary 60G50

Keywords: Randomly biased random walks; Branching random walks; Range

References

[1] E. Aïdékon. Transient random walks in random environment on a Galton–Watson tree. *Probab. Theory Related Fields* **142** (3) (2008) 525–559. MR2438700 https://doi.org/10.1007/s00440-007-0114-x
[2] E. Aïdékon. Tail asymptotics for the total progeny of the critical killed branching random walk. *Electron. Commun. Probab.* **15** (2010) 522–533. MR2737710 https://doi.org/10.1214/ECP.v15-1583
[3] E. Aïdékon. Convergence in law of the minimum of a branching random walk. *Ann. Probab.* **41** (2013) 1362–1426. MR3098680 https://doi.org/10.1214/12-AOP750
[4] E. Aïdékon and L. de Raphélis. Scaling limit of the recurrent biased random walk on a Galton–Watson tree. *Probab. Theory Related Fields* **169** (3) (2017) 643–666. MR3719055 https://doi.org/10.1007/s00440-016-0739-8
[5] E. Aidekon and Z. Shi. The Seneta–Heyde scaling for the branching random walk. *Ann. Probab.* **42** (3) (2014) 959–993. MR3189063 https://doi.org/10.1214/12-AOP809
[6] P. Andreoletti and X. Chen. Range and critical generations of a random walk on Galton–Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (1) (2018) 466–513. MR3765897 https://doi.org/10.1214/16-AIHP812
[7] P. Andreoletti and R. Diel. The heavy range of randomly biased walks on trees. *Stochastic Process. Appl.* **130** (2) (2020) 962–999. MR4046527 https://doi.org/10.1016/j.spa.2019.04.004
[8] E. Bolthausen. On a functional central limit theorem for random walks conditioned to stay positive. *Ann. Probab.* **4** (1976) 480–485. MR0415702 https://doi.org/10.1214/aop/1176996098
[9] F. Caravenna. A local limit theorem for random walks conditioned to stay positive. *Probab. Theory Related Fields* **133** (2005) 508–530. MR2197112 https://doi.org/10.1007/s00440-005-0444-5
[10] F. Caravenna and L. Chaumont. An invariance principle for random walk bridges conditioned to stay positive. *Electron. J. Probab.* **18** (2013) 32 pp. MR3068391 https://doi.org/10.1214/EJP.v18-2362
[11] X. Chen. Heavy range of the randomly biased walk on Galton–Watson trees in the slow movement regime. *Stochastic Process. Appl.* **150** (2022) 446–509. Available at https://www.sciencedirect.com/science/article/pii/S0304414922001077. MR4426162 https://doi.org/10.1016/j.spa.2022.04.018
[12] L. de Raphélis. Scaling limit of the random walk in random environment in the subdiffusive case. *Ann. Probab.* **50** (2022) 339–396. MR4385828 https://doi.org/10.1214/21-aop1535

[13] P. Erdös and S. Taylor. Some problems concerning the structure of random walk paths. *Acta Math. Hungar.* **03** (1960) 137–162. MR0121870 https://doi.org/10.1007/BF02020631

[14] G. Faraud. A central limit theorem for random walk in a random environment on marked Galton–Watson trees. *Electron. J. Probab.* **16** (6) (2011) 174–215. MR2754802 https://doi.org/10.1214/EJP.v16-851

[15] G. Faraud, Y. Hu and Z. Shi. Almost sure convergence for stochastically biased random walks on trees. *Probab. Theory Related Fields* **154** (2011) 621–660. MR3000557 https://doi.org/10.1007/s00440-011-0379-y

[16] W. Feller. *An Introduction to Probability Theory, Vol. 1*, 3rd edition. Wiley, New York, NY, 1968. MR0228020

[17] A. O. Golosov. Localization of random walks in one-dimensional random environments. *Comm. Math. Phys.* **92** (1984) 491–506. MR0736407

[18] Y. Hu and Z. Shi. The slow regime of randomly biased walks on trees. *Ann. Probab.* **44** (6) (2016) 3893–3933. MR3572327 https://doi.org/10.1214/15-AOP1064

[19] Y. Hu and Z. Shi. The potential energy of biased random walks on trees, 2016. Available at arXiv:1403.6799.

[20] M. V. Kozlov. On the asymptotic behavior of the probability of non-extinction for critical branching processes in a random environment. *Theory Probab. Appl.* **21** (1976) 791–804.

[21] J. Rosen. A random walk proof of the Erdös–Taylor conjecture. *Period. Math. Hungar.* **50** (2005) 223–245. MR2162811 https://doi.org/10.1007/s10998-005-0014-8

[22] Z. Shi. Branching Random Walks. *École d'été de Probabilités de Saint-Flour XLII – 2012* (2015). MR3444654 https://doi.org/10.1007/978-3-319-25372-5

[23] Y. G. Sinai. The limit behaviour of a one-dimensional random walk in a random medium. *Theory Probab. Appl.* **27** (2) (1982) 256–268. MR0657919

ANNALES DE L'INSTITUT HENRI POINCARÉ PROBABILITÉS ET STATISTIQUES

Recommandations aux auteurs
Instructions to authors

Les *Annales de l'I.H.P., Probabilités et Statistiques*, sont une revue internationale publiant des articles originaux en français et en anglais. Le journal publie des articles de qualité reflétant les différents aspects des processus stochastiques, de la statistique mathématique et des domaines contigus.

A compter du 14 avril 2008, les Annales ont adopté le système EJMS pour soumettre et traiter les articles. Les auteurs sont encouragés à utiliser ce système et peuvent y accéder à l'adresse https://www.e-publications.org/ims/submission/. Des informations complémentaires se trouvent à https://imstat.org/journals-and-publications/annales-de-linstitut-henri-poincare/annales-de-linstitut-henri-poincare-manuscript-submission/. Les articles peuvent être écrits en français ou en anglais. Sous le titre, les auteurs indiqueront leurs prénoms et noms ainsi qu'une désignation succincte de leur laboratoire – notamment l'adresse. Afin de faciliter la communication, il est aussi souhaitable que les auteurs fournissent une adresse de courrier électronique.

Les articles doivent être accompagnés d'un résumé précisant clairement les points essentiels développés dans l'article. Pour les articles en français, l'auteur est invité à fournir la traduction en anglais de ce résumé. Pour les articles en anglais, l'auteur est invité à fournir un résumé en français. Le comité éditorial pourra effectuer ces traductions le cas échéant.

Les références seront numérotées continûment, renvoyant à la liste bibliographique indiquant l'initiale du prénom + le nom de l'auteur, le titre de la publication, le titre de la Revue, l'année, la tomaison ou le cas échéant le numéro, les pages de début et de fin d'article et, dans le cas d'un livre, l'éditeur, le lieu et l'année d'édition.

Les articles acceptés pour publication sont considérés comme ne varietur. Les auteurs recevront une seule épreuve de leur article : celle-ci devra être retournée à l'éditeur dans le délai maximal d'une semaine. Toutes modifications ou corrections excessives autres que celles provenant d'erreurs typographiques peuvent être facturées aux auteurs. La publication des articles ou mémoires est gratuite, la facturation des pages est optionnelle. L'auteur correspondant recevra un fichier pdf de leur article final par courrier electronique.

Les auteurs sont encouragés à préparer leurs manuscrits au moyen de l'un des logiciels Plain TeX, LaTeX ou AMS TeX. Un support LATEX se trouve à l'adresse https://www.e-publications.org/ims/support/

The *Annales de l'I.H.P., Probabilites et Statistiques* is an international Journal publishing original articles in French or in English. The Journal publishes papers of high quality representing different aspects of the theory of stochastic processes, mathematical statistics and related fields.

Effective April 14, 2008, the Annales has adopted an Electronic Journal Management System (EJMS) for submission and handling of papers by the editorial board and reviewers for its journal. Authors are encouraged to use this system and should access EJMS at https://www.e-publications.org/ims/submission/. Please see https://imstat.org/journals-and-publications/annales-de-linstitut-henri-poincare/annales-de-linstitut-henri-poincare-manuscript-submission/ for additional details. Papers may be written in French or in English. The authors give their name below the title, together with their current affiliation and address. In order to facilitate communication, the authors should also provide an e-mail address.

Articles should begin with a summary explaining the basic points developed in the article. For papers in English, the authors should provide the translation in French of this summary. The Editorial committee will supply the translation if necessary. The authors of articles in French should provide the translation in English of this summary.

References should be numbered, referring to the bibliography giving for each reference the initial and name of the author followed by the title of the publication, the name of the Journal, the volume, the year, the pages of the article, and in the case of a book, the editor, the place and year of edition.

Papers accepted for publication are considered as ne variatur. The corresponding author will receive email regarding proofs, which should be sent back to the publisher within one week.

A charge may be made for any excessive corrections other than those due to typographical errors.

The publication of articles is free, page charges are optional. Every corresponding author will receive a pdf file via email of the final article. Paper offprints may be purchased by using the IMS Offprint Purchase Order Forms below.

The authors are encouraged to prepare their manuscripts using Plain TeX, LaTeX or AMS TeX. A LaTeX support page is available at https://www.e-publications.org/ims/support/

Indexations: *Current Contents (PC&ES), Zentralblatt für Mathematik, Inspec, Current Index to statistics, Pascal (INIST), Science Citation Index, SciSearch®, Research Alert®, Compu Math Citation Index®, MathSciNet. Also covered in the abstract and citation database SCOPUS®.*