Random Geometry and Reinforced Jump Processes

Nguyen, Tuan-Minh

2017

Document Version:
Publisher's PDF, also known as Version of record

[Link to publication]()

Total number of authors:
1

Random Geometry and Reinforced Jump Processes

TUAN-MINH NGUYEN

Lund University
Faculty of Science
Centre for Mathematical Sciences
Mathematical Statistics

LUND
UNIVERSITY

TUAN-MINH NGUYEN

Random Geometry and Reinforced Jump Processes

2017

Random Geometry and Reinforced Jump Processes

by
Tuan-Minh Nguyen

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisor: Prof. Stanislav Volkov
Faculty opponent: Prof. Ilya Goldsheid,
Queen Mary University of London, United Kingdom

To be, with the permission of Faculty of Science at Lund University,
publicly defended on Friday 08th of December 2017, at 9:00
in lecture hall MH:R, Matematikcentrum, Sölvegatan 18A, Lund.

Organization		Document name	
LUND UNIVERSITY		**DOCTORAL DISSERTATION**	
Centre of Mathematical Sciences Box 118 SE–221 00 LUND Sweden		Date of disputation 2017-12-08	
		Sponsoring organization	
Author(s) Tuan-Minh Nguyen			
Title and subtitle Random Geometry and Reinforced Jump Processes			

Abstract

This thesis comprises three papers studying several mathematical models related to geometric Markov processes and random processes with reinforcements. The main goal of these works is to investigate the dynamics as well as the limiting behaviour of the models as time goes to infinity, the existence of invariant measures and limiting distributions, the speed of convergence and other interesting relevant properties.

In the introduction, we firstly discuss the background: products of random matrices, asymptotic pseudo-trajectories and Markov chains in a general state space. We then outline motivation and overview of the main results in the papers included in this thesis.

In the first paper, we deal with a Markov chain model of convex polygons, which are random consecutive subdivisions of an initial convex polygon. Applying the theory of products of random matrices, we prove the universal convergence of these random convex polygons to a "flat figure". Beside this, we present a discussion about the speed of convergence and the computation of invariant measure in the case of random triangles.

In the second paper, we investigate a model of strongly vertex-reinforced jump processes (VRJP). Using the method of stochastic approximation, we show the connection between strongly VRJP and an asymptotic pseudo-trajectory of a vector field in order to study the dynamics of the model. In particular, we prove that strongly VRJP on a complete graph will almost surely have an infinite local time at one vertex, while the local times at all the remaining vertices remain bounded.

In the last paper, we consider a class of random walks taking values in simplexes and study the existence of limiting distributions. In some special cases of Markov chain models, we prove that the limiting distributions are Dirichlet. In addition, we introduce a related history-dependent random walk model in [0,1] based on Friedman's urn-type schemes and show that this random walk converges in distribution to the arcsine law.

Key words			
random polygons, products of random matrices, vertex-reinforced jump processes, pseudo-trajectories, random walks in simplexes, Markov chains in a general state space			
Classification system and/or index terms (if any)			
Supplementary bibliographical information		Language English	
ISSN and key title ISSN: 1404-0034		ISBN 978-91-7753-505-8 (print) 978-91-7753-506-5 (pdf)	
Recipient's notes	Number of pages 148	Price	
	Security classification		

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____   Date ___2017-10-30___

Random Geometry and Reinforced Jump Processes

by

Tuan-Minh Nguyen

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s).

In the latter case the thesis consists of two parts. An introductory text puts the research work into context and summarizes the main points of the papers. Then, the research publications themselves are reproduced, together with a description of the individual contributions of the authors. The research papers may either have been already published or are manuscripts at various stages (in press, submitted, or in draft).

*Dedicated to my love, Thi Tran
and all my family members*

Contents

Abstract

This thesis comprises three papers studying several mathematical models related to geometric Markov processes and random processes with reinforcements. The main goal of these works is to investigate the dynamics as well as the limiting behaviour of the models as time goes to infinity, the existence of invariant measures and limiting distributions, the speed of convergence and other interesting relevant properties.

In the introduction, we firstly discuss the background: products of random matrices, asymptotic pseudo-trajectories and Markov chains in a general state space. We then outline motivation and overview of the main results in the papers included in this thesis.

In the first paper, we deal with a Markov chain model of convex polygons, which are random consecutive subdivisions of an initial convex polygon. Applying the theory of products of random matrices, we prove the universal convergence of these random convex polygons to a "flat figure". Beside this, we present a discussion about the speed of convergence and the computation of invariant measure in the case of random triangles.

In the second paper, we investigate a model of strongly vertex-reinforced jump processes (VRJP). Using the method of stochastic approximation, we show the connection between strongly VRJP and an asymptotic pseudo-trajectory of a vector field in order to study the dynamics of the model. In particular, we prove that strongly VRJP on a complete graph will almost surely have an infinite local time at one vertex, while the local times at all the remaining vertices remain bounded.

In the last paper, we consider a class of random walks taking values in simplexes and study the existence of limiting distributions. In some special cases of Markov chain models, we prove that the limiting distributions are Dirichlet. In addition, we introduce a related history-dependent random walk model in [0,1] based on Friedman's urn-type schemes and show that this random walk converges in distribution to the arcsine law.

Acknowledgements

This PhD thesis is the first yet crucial milestone in my journey to reach the goal of becoming a professional mathematician. It could not have been successfully completed without the help and support from many influential people in my life.

First and foremost, I would like to express my deep and sincere gratitude to my supervisor Prof. Stanislav Volkov for the patient guidance, positive encouragement and continuous support he has provided throughout my PhD study. His immense knowledge, enthusiasm, sense of humour and kindness are what will stay with me after five years under his supervision.

I am grateful of Dr. Codina Cotar (University College London) and Prof. Olivier Raimond (Université Paris Nanterre). Their great advice and fruitful discussion have been crucial to realising the second paper in this thesis. I am also thankful of Professor Ilya Goldsheid (Queen Mary University of London) and Professor Thomas Kaijser (Linköpings Universitet) for their interesting suggestions about theoretical background on products of random matrices and iterated random functions. Special thanks go to Ngoc P. N. Ngo (Universidade de São Paulo) for his useful comments which help to improve the thesis writing.

I would like to offer my thanks to fellows and colleagues at Mathematical Statistics Division: Debleena, Philip, Vladimir, Dragi, Andreas, Nader, Tatyana, James and many others for creating such a relaxed, friendly and supportive working environment.

Last but not least, I would like to thank all my family members and friends for supporting me spiritually throughout my academic and personal life. Most importantly, I would like to thank my beloved wife, Thi, for all her support and love during the past few years. Her understanding and encouragement were, in the end, part of what made this thesis possible.

Lund, October 2017   *Tuan-Minh Nguyen*

List of papers

This thesis is based on the following papers, referred to by their Latin capitals:

A   **A universal result for consecutive random subdivision of polygons**

Tuan-Minh Nguyen, Stanislav Volkov
Random Structures & Algorithms, Volume 51 (2017), Issue 2, p. 341–371.

B   **Strongly vertex-reinforced jump processes on complete graphs**

Olivier Raimond, **Tuan-Minh Nguyen**
Preprint (2017).

C   **On a class of random walks in simplexes**

Tuan-Minh Nguyen, Stanislav Volkov
Preprint (2017).

All papers are reproduced with permission of their respective publishers.

Introduction

This thesis comprises three papers studying several mathematical models concerned to geometric Markov processes and random processes with reinforcements. The main goal of these works is to investigate the dynamics as well as the limiting behaviour of the models as time goes to infinity, the existence of invariant measures and limiting distributions, the speed of convergence and other interesting relevant properties.

In this introduction, we firstly introduce some notations, basic definitions and main results on several topics: products of random matrices, asymptotic pseudo-trajectories and Markov chains in a general state space which are background material used in this thesis. We then present motivation and review of the included papers.

Background materials

Random walks on real projective spaces

Let $\mathbb{R}^d$ denote the linear space of all d-dimensional real column vectors under the field of real numbers. The real projective space $P(\mathbb{R}^d)$ is defined as the quotient space $(\mathbb{R}^d \setminus \{0\})/\sim$, where $\sim$ is the equivalence relation defined by $x \sim y$, $x, y \in \mathbb{R}^d$ if there exists a real number λ such that $x = \lambda y$. We denote by $\bar{x}$ the equivalence class of x. The projective space $P(\mathbb{R}^d)$ becomes a compact metric space if we consider the following "angular" metric

$$\delta(\bar{x}, \bar{y}) = \sqrt{1 - \frac{(x,y)^2}{||x||^2 . ||y||^2}}, \tag{1}$$

where $||.||$ and $(.,.)$ are respectively the Euclidean norm and the Euclidean scalar product on $\mathbb{R}^d$. For a linear transformation from $\mathbb{R}^d$ to itself, we define

$$A\bar{x} = \overline{Ax}$$

for every $x \in \mathbb{R}^d \setminus \ker(A)$.

Let μ be a probability measure on $GL(d, \mathbb{R})$, which is the group of non-singular $d \times d$ real matrices. Assuming that $(A_n)_{n \geq 1}$ is a sequence of i.i.d. random matrices with common distribution μ, we are interested in the asymptotic behaviour of the random walk $(A_n A_{n-1} \ldots A_2 A_1 \bar{x})_{n \geq 1}$ on $P(\mathbb{R}^d)$ as well as the angular process $(\delta(A_n A_{n-1} \ldots A_1 \bar{x}, A_n A_{n-1} \ldots A_1 \bar{y}))_{n \geq 1}$ for some $\bar{x}, \bar{y} \in P(\mathbb{R}^d)$ as time n goes to infinity.

We first look at the properties for a family of matrices namely irreducible, strongly irreducible and contracting which are defined as follows.

Definition 1.1.   *a. We say a family $\mathcal{H}$ of $d \times d$ matrices is irreducible in $\mathbb{R}^d$ if there exists no proper linear subspace L of $\mathbb{R}^d$ such that $H(L) = L$ for all $H \in \mathcal{H}$.*

   b. We say a family $\mathcal{H}$ of $d \times d$ matrices is strongly irreducible in $\mathbb{R}^d$ if there exists no union L of finite number of proper linear subspaces of $\mathbb{R}^d$ such that $H(L) = L$ for all $H \in \mathcal{H}$.

   c. We say a family $\mathcal{H}$ of $d \times d$ matrices is contracting if there is a sequence of elements $\{h_n\}_{n \geq 1} \subset \mathcal{H}$ such that $||h_n||^{-1} h_n$ converges to a rank-one matrix.

We have the following key result:

Theorem 1.2 (from Theorem III.3.1 and the proof of Proposition III.3.2 in [8], pp. 50–53)**.** *Let $(A_n)_{n \geq 1}$ be a sequence of i.i.d. random matrices in $GL(d, \mathbb{R})$ with common distribution μ. Suppose that the smallest closed semigroup generated by the support of μ is strongly irreducible and contracting. Then, there exists almost surely an one-dimensional space V of $\mathbb{R}^d$ such that any limit point of $\{||A_1 A_2 \ldots A_n||^{-1} A_1 A_2 \ldots A_n, n \geq 1\}$ is a rank-one matrix with range V and any non zero vector $x \in \mathbb{R}^d$,*

$$\mathbb{P}\{x \text{ is orthogonal to } V\} = 0.$$

Furthermore, for any sequence $(x_n)_{n \geq 1} \subset \mathbb{R}^d$ which converges to x,

$$\limsup_{n \to \infty} \frac{||A_n \ldots A_2 A_1||}{||A_n \ldots A_2 A_1 x_n||} \leq ||\zeta_x||^{-1} < \infty$$

almost surely, where ζ_x is the orthogonal projection of x on V.

The next theorem shows that strong irreducibility and contracting property are sufficient conditions for the convergence to 0 of the aforementioned angular process.

Theorem 1.3 (Theorem III.4.3 in [8], p. 56). *Let $(A_n)_{n \geq 1}$ be a sequence of i.i.d. random matrices in $GL(d, \mathbb{R})$ with common distribution μ. Let S_μ be the smallest closed semigroup generated by the support of μ. Suppose that $S_\mu \subset GL(d, \mathbb{R})$ is strongly irreducible and contracting. Then for any $\overline{x}, \overline{y} \in P(\mathbb{R}^d)$*

$$\lim_{n \to \infty} \delta(A_n \ldots A_1 \overline{x}, A_n \ldots A_1 \overline{y}) = 0 \ a.s.$$

Suppose that the semigroup $\mathcal{S}_\mu$ is irreducible. To show that $\mathcal{S}_\mu$ is actually strongly irreducible, one can use the following criterion to give a contradiction assuming that $\mathcal{S}_\mu$ is irreducible but not strongly irreducible.

Proposition 1.4 (see the remark and the equation (2.7) on pp. 121-122 of [14]). *If $\mathcal{S}_\mu$ is irreducible but not strongly irreducible in $\mathbb{R}^d$, then there exist proper linear subspaces $V_1, V_2, ..., V_r$ of $\mathbb{R}^d$ such that*

$$\mathbb{R}^d = \bigoplus_{j=1}^{r} V_j \text{ where } r > 1, V_i \cap V_j = \{0\} \text{ if } i \neq j,$$

where all the subspaces V_j have the same dimension, and

$$M(\cup_{j=1}^{r} V_j) = \cup_{j=1}^{r} V_j,$$

for all $M \in \mathcal{S}_\mu$.

While it is not easy to verify the contracting property of a semigroup of matrices, it suffices to check this property for the Zariski closure which is a larger class of matrices, thanks to the following important statements by Goldsheid and Margulis in [15].

Proposition 1.5. *If a closed semigroup $H \subset GL(d, \mathbb{R})$ is strongly irreducible and its Zariski closure $Zr(H)$ is contracting, then H is also contracting.*

Recall that the Zariski closure $Zr(H)$ of a subset H of an algebraic manifold is defined as the smallest algebraic submanifold that contains H. Furthermore, if H is a closed semigroup of $GL(d, \mathbb{R})$, then the Zariski closure $Zr(H)$ is indeed a group (see [15]).

Taking into account the speed of convergence, we need the following definition of Lyapunov exponents:

3

Definition 1.6. *Let $(A_n)_{n \geq 1}$ be a sequence of i.i.d. random matrices. We define Lyapunov exponents*

$$\mu_j = \lim_{n \to \infty} \mathbb{E} \left(\frac{1}{n} \log \, \sigma_j^{(n)} \right), \quad j = 1, 2, ..., d, \tag{2}$$

where $\sigma_1^{(n)} \geq \sigma_2^{(n)} ... \geq \sigma_d^{(n)}$ are the singular values of $A^{(n)} = A_n A_{n-1} \ldots A_1$, i.e., the square roots of the eigenvalues of $\left(A^{(n)} \right)^{\mathsf{T}} A^{(n)}$.

For $A \in GL(d, \mathbb{R})$, denote

$$\ell(A) = \max(\log^+ (||A||), \log^+ (||A^{-1}||)),$$

where $\log^+(x) := \max\{\log x, 0\}$. The next theorem yields information about the speed of convergence which is connected to Lyapunov exponents.

Theorem 1.7 (from the proof of Proposition III.6.4 in [8], pp. 67–68). *Let A_i be a sequence of i.i.d. random matrices in $GL(d, \mathbb{R})$ with common distribution μ. Let S_μ be the smallest closed semigroup generated by its support. Suppose that $\mathbb{E}\, \ell(A_1) < \infty$, the semigroup $S_\mu \subset GL(d, \mathbb{R})$ is strongly irreducible and contracting then for any $\overline{x}, \overline{y} \in P(\mathbb{R}^d)$*

$$\lim_{n \to \infty} \frac{1}{n} \log \delta(A_n \ldots A_1 \overline{x}, A_n \ldots A_1 \overline{y}) \leq \mu_2 - \mu_1 < 0 \quad a.s. \tag{3}$$

For more discussion about products of random matrices, we refer the readers to excellent monographs written by Bougerol and Lacroix [8], Benoist and Quint [6].

Asymptotic pseudo-trajectories

This section contains a brief discussion of pseudo-trajectories on a metric space studied by Benaim and Hirsch (see [1] and [4]), which have many important applications to asymptotically autonomous differential equations, stochastic differential equations, stochastic approximation processes and related random processes.

Definition 1.8. *A semiflow Φ on metric space (M, d) is a continuous map $\Phi : \mathbb{R}_+ \times M \to M, (t, x) \mapsto \Phi(t, x) = \Phi_t(x)$ such that Φ_0 is the identity map and Φ has the semigroup property, i.e. $\Phi_{t+s} = \Phi_t \circ \Phi_s$ for all $s, t \in \mathbb{R}_+$.*

Definition 1.9. *A continuous function $Z : \mathbb{R}_+ \to M$ is called an asymptotic pseudo-trajectory for the semiflow Φ on metric space (M, d) if*

$$\lim_{t \to \infty} \sup_{0 < s < T} d(Z(t+s), \Phi_s(Z(t))) = 0$$

for all $T > 0$.

Let Φ be a semiflow on a metric space (M, d) and Z be an asymptotic pseudo-trajectory for Φ. We are interested in the limit behaviour of the asymptotic pseudo-trajectory Z. Let $L(Z)$ be the *limit set* of Z, which is defined by

$$L(Z) = \bigcap_{t \geq 0} \overline{X([t, \infty))}$$

or equivalently, $L(X)$ is set of limits of convergent sequences $X(t_k), t_k \to \infty$. A subset $A \subset M$ is called *invariant* of the semiflow Φ if $\Phi_t(A) \subset A$ for all $t \in \mathbb{R}_+$. The asymptotic pseudo-trajectory Z is said *precompact* if its image has compact closure in M.

Definition 1.10. *Let $K \subset M$ be a compact invariant set of the semiflow Φ. A continuous function $H : M \to \mathbb{R}$ is called a Lyapunov function for K if the function $t \in \mathbb{R}_+ \mapsto H(\Phi_t(x)) \in \mathbb{R}$*

  (i) is a constant function for each $x \in K$.

  (ii) is a strictly decreasing function for each $x \in M \setminus K$.

We denote by $\mathcal{C}_\Phi$ the set of all equilibria points for the semiflow Φ, i.e. the set of all points $x \in M$ such that $\Phi_t(x) = x$ for all $t \in \mathbb{R}_+$. It is obvious that $\mathcal{C}_\Phi$ is an invariant set of Φ.

Theorem 1.11 (Theorem 5.7 and Proposition 6.4 in [4]). *Let Φ be a semiflow on a metric space (M, d) and Z be an asymptotic pseudo-trajectory for Φ such that*

  (a) Z is precompact,

  (b) Φ admits a Lyapunov function H for $\mathcal{C}_\Phi$,

  (c) $\mathcal{C}_\Phi$ is a compact subset and $H(\mathcal{C}_\Phi)$ has empty interior in M.

Then the limit set $L(Z)$ is a connected subset of $\mathcal{C}_\Phi$.

In particular, if C_Φ is a set of isolated equilibria then we directly obtain the following important result:

Corollary 1.12. *Let Φ be a semiflow on a metric space (M, d) and Z be an asymptotic pseudo-trajectory for Φ such that the conditions (a) and (b) in Theorem 1.11 are fulfilled, furthermore*

(c') C_Φ has countably many elements.

Then $Z(t)$ converges to an equilibrium $z^* \in C_\Phi$ as $t \to \infty$.

To study speed of convergence of pseudo-trajectories, we need the following definitions:

Definition 1.13. *Assume that Z is an asymptotic pseudo-trajectory for a semiflow Φ on a metric space (M, d). For $\lambda > 0$, Z is called a λ-pseudo-trajectory for Φ if*

$$\sup_{T>0} \limsup_{t\to\infty} \frac{1}{t} \log \left(\sup_{0\leq s\leq T} d(Z(t+s), \Phi_s(Z(t))) \right) \leq -\lambda.$$

Definition 1.14. *Let $K \subset M$ be a compact invariant set of a semiflow Φ and B be a subset of M containing K. For $\alpha > 0$, we say that K attracts B exponentially at rate α if there exists a constant $C > 0$ such that*

$$d(\Phi_t(x), K) \leq Ce^{-\alpha t} d(x, K)$$

for all $x \in B$ and $t \in \mathbb{R}_+$.

Theorem 1.15 (Lemma 8.7 in [4]). *Let Z be a λ-pseudo-trajectory for a semiflow Φ, $K \subset M$ be a compact invariant set of Φ and B be a subset of M containing K. Assume that K attracts B exponentially at rate $\alpha > 0$ such that $Z(t) \in B$ for all $t \in \mathbb{R}_+$. Let $Y(t) \in K$ be the nearest point to $Z(t)$. Then*

$$\limsup_{t\to\infty} \frac{1}{t} \log d(Z(t), Y(t)) \leq -\min(\alpha, \lambda).$$

In particular, if K contains only a single point z^* which is an equilibrium of Φ and B is a neighbourhood of z^* such that z^* attracts exponentially B at rate α and $Z(t) \in B$ for all $t \in \mathbb{R}_+$, then

$$\limsup_{t\to\infty} \frac{1}{t} \log d(Z(t), z^*) \leq -\min(\alpha, \lambda).$$

We now restrict our attention to semiflows generated by vector fields. Let F be a continuous vector field defined on $\mathbb{R}^d$ with unique integral curves. The semi-flow Φ generated by the vector field F is defined such that $\Phi_t(x)$ is the unique solution of the following autonomous differential equation

$$\begin{cases} \frac{d}{dt}\Phi_t(x) & = F(\Phi_t(x)), \\ \Phi_0(x) & = x, \end{cases}$$

for $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Note that the equilibria of Φ coincide with the zeros of F, i.e.

$$\mathcal{C}_\Phi = \{z \in \mathbb{R}^d : F(z) = 0\}.$$

We say an equilibrium $z \in \mathcal{C}_\Phi$ is *hyperbolic* if all eigenvalues of $DF(z)$, the Jacobian matrix of F at z, have nonzero real parts. If all eigenvalues of $DF(z)$ have negative real parts, z is called *linearly stable*. If there exists at least one of its eigenvalues having positive real part, it is called *linearly unstable*. For a hyperbolic unstable equilibrium z, the *stable manifold* of z is the set of initial values z_0 whose forward trajectories $\Phi_t(z_0)$ converge to z as $t \to \infty$.

Let $Z : \mathbb{R}_+ \to \mathbb{R}^d$ be a bounded asymptotic pseudo-trajectory for Φ. Assume that Φ admits a Lyapunov function for $\mathcal{C}_\Phi$ and F has countably many zeros. By Corollary 1.12, there exists an equilibrium $z \in \mathcal{C}_\Phi$ such that $Z_t \to z$ as $t \to \infty$.

In many practical problems, we usually expect that Z_t converges to a stable equilibrium. To imply the convergence to stable equilibrium, one might study the nonconvergence to unstable equilibrium by estimating "distance" from Z_t to stable manifold associated with each unstable equilibrium. This idea was first introduced by Pemantle in [19], Benaim and Hirsch in [2]. More specifically, assume that F is a C^k-vector field and let z^* be a linearly unstable hyperbolic equilibrium of F. We have the orthogonal decomposition

$$\mathbb{R}^d = E^s \oplus E^u,$$

where E^s and E^u are respectively the generalized eigenspaces generated by the eigenvalues of $DF(z^*)$ with positive and negative real parts. Therefore, for each $z \in \mathbb{R}^d$, we have the unique decomposition $z^* = z_s^* + z_u^*$, where $z_s^* \in E^s$ and $z_u^* \in E^u$. By the stable manifold theorem (see, e.g. Theorem 10.1 in [20]), there exist a neighbourhood $\mathcal{N}_0 = \mathcal{N}_0^s \oplus \mathcal{N}_0^u$ of z^*, where $\mathcal{N}_0^s$ (resp. $\mathcal{N}_0^u$) is a neighbourhood of z_s^* in E^s (resp. z_u^* in E^u), and a C^k-function $\Gamma : \mathcal{N}_0^s \to \mathcal{N}_0^u$ such that

- $D\Gamma(z_s^*) = 0$;

- The graph of Γ,

$$\text{Graph}(\Gamma) := \{v + \Gamma(v) : v \in \mathcal{N}_0^s\}$$

equals to the local stable manifold of z^* defined by

$$W_{loc}^s(z^*) = \{z \in \mathbb{R}^d : \forall t \geq 0, \Phi_t(z) \in \mathcal{N}_0 \text{ and } \lim_{t \to \infty} \Phi_t(z) = z^*\};$$

- $W_{loc}^s(z^*)$ is an invariant manifold, i.e. for all $t \in \mathbb{R}$,

$$\Phi_t(W_{loc}^s(z^*)) \cap \mathcal{N}_0 \subset W_{loc}^s(z^*).$$

Let $R : \mathcal{N}_0 \to \mathbb{R}$ be defined by

$$R(z) = \|z - r(z)\|,$$

where the function $r : \mathcal{N}_0 \to W_{loc}^s(z^*)$ is given by

$$r(z) = r(z_s + z_u) = z_s + \Gamma(z_s).$$

In this case, $R(Z_t)$ is defined as the distance from Z_t to the local manifold $W_{loc}^s(z^*)$ in the unstable directions at z^*, which is expected to be not too small so that Z_t can stay away from z^*. For further techniques related to nonconvergence theorems and their applications, we refer the interested readers to [19], [2], [4] and [5].

Let us now consider the speed of convergence to stable equilibrium. Assume that z_* is a linearly stable equilibrium and Z_t is a λ-pseudo-trajectory for the semiflow Φ such that Z_t converges to z_* as $t \to \infty$. Note that there exist $\alpha > 0$ and a neighbourhood B of z_* such that z_* attracts exponentially B at rate α, i.e. for all $z \in B$ and $t \in \mathbb{R}_+$,

$$\|\Phi_t(z) - z_*\| \leq Ce^{-\alpha t}\|z - z_*\|$$

where C is some constant depending only on B and α (see, e.g [20], Theorem 5.1). Therefore, by applying Theorem 1.15, one can conclude that

$$\limsup_{t \to \infty} \frac{1}{t} \log d(Z(t), z_*) \leq -\min(\alpha, \lambda).$$

8

Markov chains in general state spaces

Let $(Z_n)_{n \geq 0}$ be a Markov chain taking values on a measurable state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, where $\mathcal{B}(\mathcal{X})$ is a countably generated σ-algebra (i.e. $\mathcal{B}(\mathcal{X})$ is generated by a countable family of subsets of $\mathcal{X}$). We define the probability transition kernel $P^n : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \to [0, 1]$ by

$$P^n(z, A) = \mathbb{P}(Z_n \in A | Z_0 = z)$$

for each $n \geq 1$, $z \in \mathcal{X}$ and $A \in \mathcal{B}(\mathcal{X})$. We also denote $P(z, A) := P^1(z, A)$. We will always assume that for each $A \in \mathcal{B}(\mathcal{X})$, $P^n(., A)$ is a measurable function while for each $z \in \mathcal{X}$, $P^n(z, .)$ is a probability measure.

As usual, we define the total variation distance $\|\mu - \nu\|$ between two probability measures μ and ν on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ by

$$\|\mu - \nu\| = \sup_{A \in \mathcal{B}(\mathcal{X})} |\mu(A) - \nu(A)|.$$

Definition 1.16. *A measure π on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is called invariant of the Markov chain $(Z_n)_{n \geq 0}$ if*

$$\pi(A) = \int_{\mathcal{X}} P(z, A)\pi(dz)$$

for all $A \in \mathcal{B}(\mathcal{X})$.

Assumption 1. *Assume that there exist a subset $V \in \mathcal{B}(\mathcal{X})$, $q \in (0, 1)$, a probability measure φ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and some positive integer n_0 such that*

  (i) $\mathbb{P}(\tau_V < \infty | Z_0 = z) = 1$ *for all $z \in \mathcal{X}$, where $\tau_V = \inf\{n \geq 1 : Z_n \in V\}$;*

  (ii) $\mathbb{P}(Z_{n_0} \in B | Z_0 = z) \geq q\varphi(B)$ *for all $z \in V$ and measurable set $B \subset V$.*

It is well-known that the Markov chain $(Z_n)_{n \geq 0}$ is *Harris recurrent* if and only if the Assumption 1 is fulfilled (see e.g. [18]). Here the Harris recurrence is in the sense that there exists a σ-additive measure ψ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that $\psi(\mathcal{X}) > 0$ and

$$\mathbb{P}\left(Z_n \in B \text{ for some } n \geq 1 | Z_0 = z\right) = 1, \ \forall z \in \mathcal{X}$$

whenever $B \in \mathcal{B}(\mathcal{X})$ and $\psi(B) > 0$.

The Harris recurrence implies the unique existence of a non-trivial σ-additive invariant measure up to a constant factor while in general it might not be a finite measure and the weak convergence might not occur.

Assume that the condition (ii) in Assumption 1 is fulfilled. We construct an "extended" Markov chain $Z_k^* = (\tilde{Z}_k^{(n_0)}, \xi_k)$ on $\mathcal{X}^* = \mathcal{X} \times \{0,1\}$ such that $\xi_k, k = 0, 1, 2, \ldots$ are i.i.d. random variables taking values on $\{0,1\}$ such that $P(\xi_k = 1) = q$ and the transition distribution of Z_k^* is defined as follows:

- for $z \notin V$,

$$\mathbb{P}(Z_1^* \in (B,j)|Z_0^* = (z,k)) = P^{n_0}(z,B)\mathbb{P}(\xi_1 = j);$$

- for $z \in V$,

$$\mathbb{P}(Z_1^* \in (B,j)|Z_0^* = (z,1)) = \varphi(B)\mathbb{P}(\xi_1 = j);$$
$$\mathbb{P}(Z_1^* \in (B,j)|Z_0^* = (z,0)) = \frac{P^{n_0}(z,B) - q\varphi(B)}{1 - q}\mathbb{P}(\xi_1 = j).$$

Denote
$$\tau^* = \inf\{k \geq 1 : Z_k^* \in (V,1)\}.$$

Proposition 1.17 (Theorem 1.3 in [7]). *Let $(Z_n)_{n\geq 0}$ be a Markov chain corresponding with the measurable state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that (i), (ii) in Assumption 1 and the following conditions*

(iii) $\gcd\{n \geq 1 : \mathbb{P}(Z_n \in B|Z_0 = z) \geq q\varphi(B)\} = 1$ *for all $z \in V$ and measurable subset $B \subset V$;*

(iv) $\mathbb{E}(\tau^*) < \infty$ *for all $z^* \in (V,1)$*

are fulfilled. Then there uniquely exists an invariant probability measure π defined by

$$\pi(B) = \frac{1}{\mathbb{E}\,\tau^*}\,\mathbb{E}\left(\sum_{k=1}^{\tau^*} \mathbf{1}_{\{\tilde{Z}_k^{(n_0)} \in B\}}\right).$$

Moreover, for each $z \in \mathcal{X}$,

$$\|P^n(z,.) - \pi\| \to 0 \quad \text{as} \quad n \to \infty.$$

In practical, it is quite complicated to estimate the expectation $\mathbb{E}(\tau^*)$. Instead of verifying the condition (iv), we might use the following simpler criterion:

Proposition 1.18 (Theorem 2.1 in [7])**.** *Let $Z_n, n = 0, 1, 2, \ldots$ be a Markov chain corresponding with the measurable state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$. Assume that there exist a subset $V \in \mathcal{B}(\mathcal{X})$, $q > 0$, a probability measure φ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ and some positive integer n_0 such that (i) in Assumption 1, (iii) in Proposition 1.17 and the following conditions*

(i') $\sup_{z \in V} \mathbb{E}(\tau_V | Z_0 = z) < \infty$;

(ii') $\mathbb{P}(Z_{n_0} \in B | Z_0 = z) \geq q\varphi(B)$ for all $B \in \mathcal{B}(\mathcal{X})$ and $z \in V$

are fulfilled. Then the conclusion in Proposition 1.17 still holds.

Consecutive random subdivision of convex polygons

Motivation

We consider the Markov chain $(L_n)_{n \geq 0}$ of convex d-polygons (i.e. polygons with d vertices, $d \geq 3$) in the plane recursively defined as follows:

- We begin with an initial convex d-polygon $L_0 = A_1^{(0)} A_2^{(0)} ... A_d^{(0)}$.

- For $n \geq 0$, assume that we obtained the polygon $L_n = A_1^{(n)} A_2^{(n)} ... A_d^{(n)}$. For $j = 1, 2, \ldots, d$, we randomly choose a new point $A_j^{(n+1)}$ inside each edge $A_j^{(n)} A_{j+1}^{(n)}$ such that

$$\frac{\left| A_j^{(n)} A_j^{(n+1)} \right|}{\left| A_j^{(n)} A_{j+1}^{(n)} \right|} := \xi_j^{(n)}, j = 1, 2, \ldots, d,$$

where $\xi_1^{(n)}, \xi_2^{(n)}, \ldots, \xi_d^{(n)}$ are i.i.d. copies of a random variable ξ with support in $[0, 1]$. Here, we use the convention that $A_{d+1}^{(n)} \equiv A_1^{(n)}$. Therefore, we obtain the new convex d-polygon $L_{n+1} = A_1^{(n+1)} A_2^{(n+1)} \ldots A_d^{(n+1)}$.

- Repeating the subdivision process such that $\left(\xi_1^{(n)}, \xi_2^{(n)}, \ldots, \xi_d^{(n)} \right)$, $n = 1, 2, \ldots$ are i.i.d random vectors and have the same distribution with $(\xi_1, \xi_2, \ldots, \xi_d)$, where ξ_i, $i = 1, 2, \ldots, d$ are i.i.d. copies of the random variable ξ, we obtain a sequence of polygons $\left(L_n = A_1^{(n)} A_2^{(n)} \ldots A_d^{(n)} \right)_{n \geq 0}$.

It is trivial to see that the random polygons L_n become smaller and smaller and eventually vanish to a point inside the initial polygon L_0 as time goes to infinity, however the behaviour of the process of their shapes is less clear. To study the process of shapes we may, for example, always place the vertex $A_1^{(n)}$ at the origin $(0, 0)$ of the real plane $\mathbb{R}^2$ and rescale the polygon L_n such that its longest edge has always length 1.

In the case of triangles, i.e. $d = 3$, one can follow the rescaling procedure by Volkov in [25] as follows: for each $n \geq 0$, rescale the triangle $A_1^{(n)} A_1^{(n)} A_3^{(n)}$ to a new triangle $B_1^{(n)} B_1^{(n)} B_3^{(n)}$ such that its longest edge has length 1 and its vertices are reordered in a way that $B_1^{(n)} B_2^{(n)} \geq B_3^{(n)} B_1^{(n)} \geq B_2^{(n)} B_3^{(n)}$; let the

12

Cartesian coordinates of vertices be $B_1^{(n)} = (0,0)$, $B_2^{(n)} = (1,0)$, $B_3^{(n)} = (g_n, h_n)$ as illustrated in Figure 1.
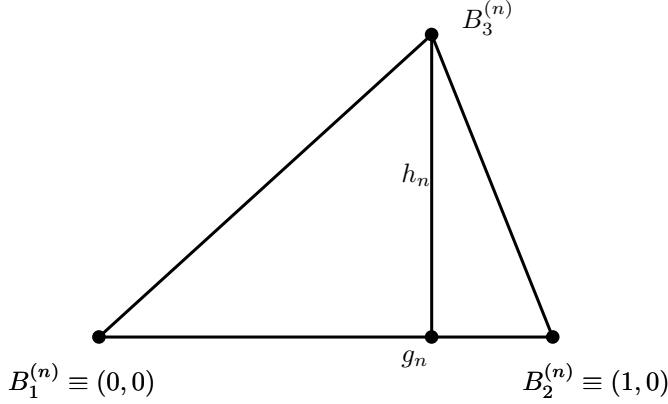
Figure 1: Illustration of the rescaled triangle $B_1^{(n)} B_1^{(n)} B_3^{(n)}$.

The limit shape behaviour of these random triangles therefore is reduced to the limit behaviour of the process $(g_n, h_n)_{n \geq 0}$ in $[1/2, 1] \times \mathbb{R}_+$ as $n \to \infty$. In particular, if ξ is uniformly distributed in $[0,1]$, Volkov showed in [25] that g_n converges almost surely to an uniformly distributed random variable in $[1/2, 1]$ while h_n converges almost surely to 0 at least at exponential rate.

From the above-mentioned example, it is natural to raise a question to determinate some "mild" condition of the distribution of ξ such that the sequence of rescaled convex polygon $(L_n)_n$ converges to a "flat figure", in the sense that all its vertices will be lying approximately along the same line, which is in turn equivalent that the area of the rescaled polygon converges to 0 as n goes to infinity. Formally, we say that the sequence of polygons L_n *converges to a flat figure* as $n \to \infty$ if

$$\lim_{n \to \infty} \frac{S(L_n)}{\left(\max_{j=1,\dots,d} \| l_j^{(n)} \| \right)^2} = 0 \quad \text{a.s.}$$

Here $S(L_n)$ denotes the area of the polygon L_n and $l_j^{(n)} = A_j^{(n)} A_{j+1}^{(n)}$, $j = 1, 2, \dots, d$ is the vector corresponding to the j-th side of L_n.

Let $(x_j^{(n)}, y_j^{(n)})$ denote the Cartesian coordinates of $A_j^{(n)} A_{j+1}^{(n)}$ for $j = 1, 2, \dots, d$. Note that $\sum_{j=1}^{d} x_j^{(n)} = \sum_{j=1}^{d} y_j^{(n)} = 0$. From elementary geometrical calculations

one can obtain the following linear recursion:

$$x^{(n+1)} = T_{n+1}x^{(n)}, \ y^{(n+1)} = T_{n+1}y^{(n)}, \tag{4}$$

where $x^{(n)} = \left(x_1^{(n)}, x_2^{(n)}, \ldots, x_{d-1}^{(n)}\right)^{\mathbb{T}}$, $y^{(n)} = \left(y_1^{(n)}, y_2^{(n)}, \ldots, y_{d-1}^{(n)}\right)^{\mathbb{T}}$ are column vectors and T_n is an i.i.d. copy of the following random $(d-1) \times (d-1)$ matrix

$$T = T(\xi_1, \ldots, \xi_d) = \begin{pmatrix} 1 - \xi_1 & \xi_2 & 0 & \cdots & 0 & 0 \\ 0 & 1 - \xi_2 & \xi_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 - \xi_{d-2} & \xi_{d-1} \\ -\xi_d & -\xi_d & -\xi_d & \cdots & -\xi_d & 1 - \xi_{d-1} - \xi_d \end{pmatrix}. \tag{5}$$

Therefore, to understand the limit shape of L_n one should study the asymptotic behaviour of the backward product of random matrices

$$T^{(n)} := T_n T_{n-1} \ldots T_2 T_1$$

as $n \to \infty$.

Review of Paper A

In Paper A, we apply the theory of products of random matrices to prove the convergence of the random convex polygon L_n to a "flat figure" under very mild non-degenerateness conditions on ξ. In addition, we present a discussion about the speed of convergence and the computation of limiting distribution in the case of random rescaled triangles.

To ensure that L_n is a non-degenerate convex polygon and that the subdivision is genuinely random, we need the following

Assumption 2. $\mathbb{P}(\xi \in \{0,1\}) = 0$ and the support of ξ contains at least two distinct points in $(0,1)$, i.e. the distribution of ξ is non-degenerate.

Assumption 3. If $d \geq 4$ is an even number, we assume that

$$\prod_{i=1}^{d} \xi_i \neq \prod_{i=1}^{d} (1 - \xi_i)$$

almost surely.

Obviously, Assumption 3 always holds for any continuous random variable ξ.

We observe that L_n converges to a flat figure as $n \to \infty$ if

$$\lim_{n\to\infty} \delta(T^{(n)}\bar{x}, T^{(n)}\bar{y}) = 0 \tag{6}$$

almost surely for any $x = (x_1, ..., x_{d-1})^{\mathbb{T}}, y = (y_1, ..., y_{d-1})^{\mathbb{T}} \in \mathbb{R}^{d-1}$, such that $(x_1, y_1), (x_2, y_2), ..., (x_{d-1}, y_{d-1})$ are coordinates of $d-1$ consecutive edges of the convex d-polygon L_0 in the real plane.

Assuming that Assumption 2 and Assumption 3 are fulfilled, we therefore prove the convergence of random polygons L_n to a flat figure by verifying the strong irreducibility and the contracting property for the smallest closed semigroup in $GL(d-1, \mathbb{R})$ generated by the support of the normalized random matrix $\tilde{T} = (|\det T|)^{-1/d} T$. Note that, the convergence to flat figure of the sequence of random polygons with even number of vertices still requires the satisfaction of Assumption 3, which we conjecture it is unnecessary.

To study the speed of convergence, we need to assume that the expectation $\mathbb{E}\,\ell(T) = \mathbb{E}\max(\log^+(||T||), \log^+(||T^{-1}||))$ is finite. This condition is actually equivalent to following:

Assumption 4.

$$\mathbb{E}\log(|\det(T)|) = \mathbb{E}\log\left|\prod_{i=1}^{d}(1-\xi_i) - (-1)^d\prod_{i=1}^{d}\xi_i\right| > -\infty.$$

Applying Theorem 1.7, we deduce that if Assumptions 2, 3 and 4 hold, then the sequence of polygons L_n converges to flatness at least at exponential rate $\mu = \mu_1 - \mu_2 \in (0, \infty)$.

Assumption 4 can be replaced by an "easier" criterion which depends only on the distribution of the random variable ξ. Indeed, for $d = 3, 5, \dots$ is odd, we show that if $\mathbb{E}\,|\log\xi| < \infty$ and $\mathbb{E}\,|\log(1-\xi)| < \infty$ then Assumption 4 is fulfilled. A sufficient condition for these expectations to be finite is

$$\limsup_{v\downarrow 0}\frac{\mathbb{P}(\xi < v)}{v^\alpha} < \infty \text{ and } \limsup_{v\uparrow 1}\frac{\mathbb{P}(\xi > v)}{(1-v)^\alpha} < \infty$$

for some $\alpha > 0$. In particular, in the case that ξ is uniformly distributed on $[0, 1]$, we prove that Assumption 4 is fulfilled for all $d \geq 3$.

Suppose that Assumptions 2, 3 and 4 are fulfilled, we also show that the length of the largest side of L_n converges exponentially to 0 at rate μ_1, i.e.

$$\lim_{n \to \infty} \frac{1}{n} \log(M_n) = \mu_1 \quad \text{a.s.}$$

where

$$M_n = \max_{j=1,\dots,d} \|l_j^{(n)}\|. \tag{7}$$

We now only consider the case of random triangles (i.e. $d = 3$). Let g_n, h_n be respectively the orthogonal projection of $B_3^{(n)}$ on $B_1^{(n)} B_2^{(n)}$ and the height of the triangle $B_1^{(n)} B_2^{(n)} B_3^{(n)}$ drawn from $B_3^{(n)}$ as mentioned in the section Motivation. We firstly prove that g_n converges in distribution to some random variable $\eta \in [1/2, 1]$ if Assumption 2 is fulfilled. Furthermore, assuming that Assumptions 2 and 4 are both fulfilled, we obtain that

$$\lim_{n \to \infty} \frac{1}{n} \log(h_n) = \mathbb{E}(\log(\det(T_1))) - 2 \int_{1/2}^{1} \zeta(x,0) d\mathbb{P}_\eta(x), \quad \text{a.s.} \tag{8}$$

where $\theta_n = (g_n, h_n)$, η is the weak limit of g_n, $\mathbb{P}_\eta$ is its probability measure, and

$$\zeta(x,y) = \mathbb{E}\left(\log(M_1) \mid \theta_0 = (x,y)\right).$$

In particular, in the case where ξ is uniformly distributed on $(0,1)$, we show that

$$\frac{1}{n} \log h_n \to -\frac{\pi^2 - 6}{9} \approx -0.43,$$

thus strengthening the result of Theorem 4 in [25].

Strongly vertex-reinforced jump processes

Motivation

Let us consider a continuous time jump process X on a connected finite (without loops, unoriented) graph $G = (V, E)$ where $V = \{1, 2, ..., d\}$ such that

• at time $t \leq 0$, each vertex $v \in V$ has a local time with a positive initial value $\ell_0^{(v)}$,

• given $\mathcal{F}_t = \sigma\{X_s, s \leq t\}$ and $\{X_t = v\}$, the jumping rate from v to a neighbour $v' \sim v$ at time t is

$$w\left(\ell^{(v')} + \int_0^t \mathbf{1}_{\{X_s = v'\}} ds\right),$$

where $w : \mathbb{R} \to (0, \infty)$ is called a weight function. This type of history-dependent processes is called a *vertex-reinforced jump process (VRJP)*.

Assume $w(t) = t^\alpha$, for some $\alpha > 0$. The jump process X is called

- *strongly vertex-reinforced* if $\alpha > 1$,

- *weakly vertex-reinforced* if $0 < \alpha < 1$ and

- *linearly vertex-reinforced* if $\alpha = 1$.

Denote as $L(v, t) = \ell_0^{(v)} + \int_0^t \mathbf{1}_{\{X_s = v\}} ds$ the *local time* at v up to time t for each $v \in V$ and let

$$Z_t = \left(Z_t^{(1)}, Z_t^{(2)} \ldots, Z_t^{(d)}\right) := \left(\frac{L(1, t)}{t + \ell_0}, \frac{L(2, t)}{t + \ell_0}, ..., \frac{L(d, t)}{t + \ell_0}\right)$$

stand for the *occupation measure process* on V at time t, where $\ell_0 = \ell_0^{(1)} + \ell_0^{(2)} + ... + \ell_0^{(d)}$. Note that Z_t is a continuous process taking values on the $(d - 1)$ dimensional unit simplex Δ, which is defined by

$$\Delta = \{(z_1, z_2, ..., z_d) \in \mathbb{R}^d : \sum_{j=1}^d z_j = 1, z_j \geq 0, j = 1, 2, \cdots, d\}.$$

The model of continuous time vertex-reinforced jump processes was first conceived by Wendelin Werner and later investigated by Davis and Volkov in [9]

and [10]. In particular, these authors considered in [10] the model of linearly VJRP on trees and finite graphs. They showed that linearly VRJP on any finite graph is recurrent, i.e. all local times are almost surely unbounded and the occupation measure process converges almost surely to an element in the interior of Δ as time goes to infinity. In [21], Sabot and Tarrès also obtained the limiting distribution of the centred local times process for linearly VRJP on any finite graph $G = (V, E)$ with d vertices and showed that linearly VRJP is actually a mixture of time-changed Markov jump processes. The connections between linearly VRJP, linearly edge-reinforced random walks and random walks in random environment as well as their applications have been recently investigated in [12], [21], [22], [26] and [17].

From now on, we restrict our attention to the case where G is a complete graph and $w(t) = t^{\alpha}, \alpha > 1$, i.e. X is strongly reinforced.

Let

$$T\Delta = \{(z_1, z_2, ..., z_d) \in \mathbb{R}^d : \sum_{j=1}^{d} z_j = 0\}$$

denote the tangent space of Δ and $e_1 = (1, 0, 0, ..., 0)$, $e_2 = (0, 1, 0, ..., 0),...,$ $e_d = (0, 0, ..., 0, 1)$ be standard unit vectors in $\mathbb{R}^d$ which are also vertices of Δ.

For each $j \in V$, define $w_t^{(j)} = w(L(j, t)) = L(j, t)^{\alpha}$ and $w_t = w_t^{(1)} + w_t^{(2)} + \cdots + w_t^{(d)}$. For each fixed $t \geq 0$, let A_t be an infinitesimal generator matrix such that

$$(A_t)_{i,j} := \begin{cases} \mathbf{1}_{(i,j) \in E} w_t^{(j)}, & i \neq j; \\ - \sum_{k \in V, (k,i) \in E} w_t^{(k)}, & i = j. \end{cases}$$

Note that

$$\pi_t = \left(\frac{w_t^{(1)}}{w_t}, \frac{w_t^{(2)}}{w_t}, \cdots, \frac{w_t^{(d)}}{w_t} \right)$$

is the unique invariant probability measure of A_t in the sense that $\pi_t A_t = \pi_t$. Since π_t can be rewritten as a function of Z_t, we also use the notation $\pi_t = \pi(Z_t)$, where we define the function $\pi : \Delta \to \Delta$, such that for each $z = (z_1, z_2, ..., z_d) \in \Delta$,

$$\pi(z) = \left(\frac{z_1^{\alpha}}{z_1^{\alpha} + \cdots + z_d^{\alpha}}, \cdots, \frac{z_d^{\alpha}}{z_1^{\alpha} + \cdots + z_d^{\alpha}} \right).$$

For $t > 0$, which is not a jumping time of X_t, we have

$$\frac{dZ_t}{dt} = \frac{1}{\ell_0 + t}\left(-Z_t + e_{X_t}\right), \tag{9}$$

We can rewrite (9) as

$$\frac{dZ_t}{dt} = \frac{1}{\ell_0 + t}(-Z_t + \pi(Z_t)) + \frac{1}{\ell_0 + t}(e_{X_t} - \pi_t). \tag{10}$$

Set $t + \ell_0 = e^u$ and $Z_{e^u - \ell_0} = \tilde{Z}_u$, we can rewrite (10) as

$$\frac{d\tilde{Z}_u}{du} = -\tilde{Z}_u + \pi(\tilde{Z}_u) + (e_{X_{e^u - \ell_0}} - \pi_{e^u - \ell_0}).$$

Let $\eta_u := e_{X_{e^u - \ell_0}} - \pi_{e^u - \ell_0}$ and $F : \Delta \to T\Delta$ be a vector field defined by

$$F(z) := -z + \pi(z). \tag{11}$$

We hence obtain the following differential equation

$$\frac{d\tilde{Z}_u}{du} = F(\tilde{Z}_u) + \eta_u.$$

If the noise term

$$\int_t^{t+T} \eta_u du = \int_{e^t - \ell_0}^{e^{t+T}\ell_0} \frac{e_{X_u} - \pi_u}{\ell_0 + u} du$$

converges almost surely to 0 as $t \to \infty$ for each $T > 0$, then one can expect that the dynamics of the process $\tilde{Z}$ is approximately to the dynamics of the autonomous differential equation

$$\frac{dz}{dt} = F(z).$$

For an instance, let $d = 3$ and $\alpha = 2$, we consider the reduced system of differential equations on $\{(z_1, z_2) \in \mathbb{R}^d : 0 \leq z_1 \leq 1, 0 \leq z_2 \leq 1 - z_1\}$ defined as follows

$$\frac{dz_1}{dt} = -z_1 + \frac{z_1^2}{z_1^2 + z_2^2 + (1 - z_1 - z_2)^2}, \tag{12}$$

$$\frac{dz_2}{dt} = -z_2 + \frac{z_1^2}{z_1^2 + z_2^2 + (1 - z_1 - z_2)^2}. \tag{13}$$

As illustrated in Figure 2, it is reasonable to believe that $(\tilde{Z}_t^{(1)}, \tilde{Z}_t^{(2)})$ as well as $(Z_t^{(1)}, Z_t^{(2)})$ might converge to one of the vertices $(0,0)$, $(0,1)$, $(1,0)$, which are actually stable equilibria of the system (12)-(13). It turns out that the limit behaviour of strongly VRJP seems to be completely different from linearly VRJP.
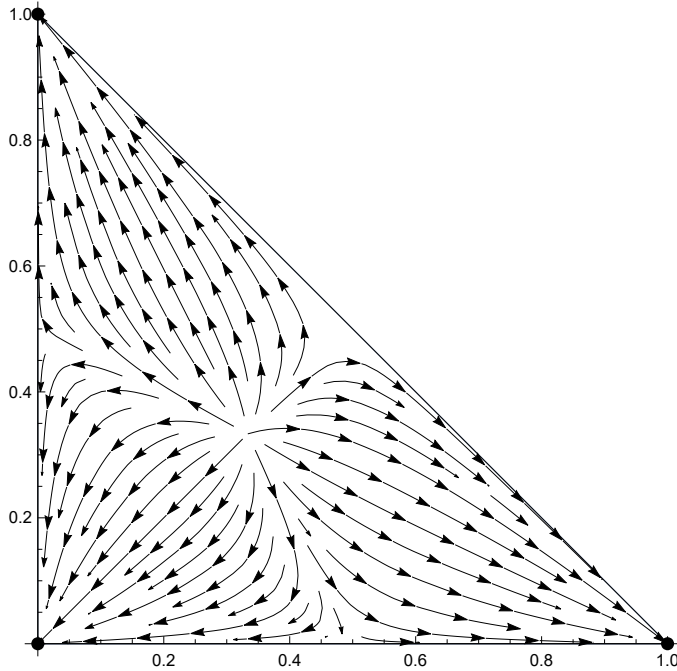
Figure 2: Direction field of the system (12)-(13).

Review of Paper B

In paper B, using the method of stochastic approximation, we show the connection between strongly VRJP and an asymptotic pseudo-trajectory of a vector field in order to study the dynamics of the model. In particular, we prove that strongly VRJP on a complete graph will almost surely have an infinite local time at one vertex, while the local times at all the remaining vertices remain bounded.

Recall that we only consider the case when $w(t) = t^\alpha, \alpha > 1$ and G is a complete graph with the set of vertices $V = \{1, 2, ..., d\}$. Let F be the vector field given by (11) and $\Phi_t(z_0)$ stand for the unique solution of the following initial value problem

$$\begin{cases} \frac{d}{dt}z(t) = & F(z(t)) \quad \text{for } t > 0, \\ z(0) = & z_0. \end{cases}$$

Note that the solution $\Phi_t(z_0)$ can be extended for all $t \in \mathbb{R}_+$. Thus the continuous map $\Phi : \mathbb{R}_+ \times \Delta \to \Delta$ defined by $\Phi(t, z) = \Phi_t(z)$ is a semiflow. We

firstly prove that

$$\lim_{t\to\infty} \sup_{0\le s\le T} \left\| \int_{e^t-\ell_0}^{e^{s+t}-\ell_0} \frac{e_{X_u} - \pi_u}{\ell_0 + u} du \right\| = 0 \ \text{ a.s.}$$

We then show that $\tilde{Z}$ is actually an asymptotic pseudo-trajectory of the semi-flow Φ, i.e. for all $T > 0$,

$$\lim_{t\to\infty} \sup_{0\le s\le T} \left\| \tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t) \right\| = 0 \ \text{ a.s.}$$

Let

$$\mathcal{C} = \{z \in \Delta : F(z) = 0\}$$

stand for the *equilibria set* of F. One can easily obtain that

• $\mathcal{S} = \{e_1 = (1,0,0,...,0), e_2 = (0,1,0,...,0),...,e_d = (0,0,...,0,1)\}$ is the set of stable equilibria, and

• $\mathcal{U} = \{z_{j_1,j_2,...,j_k}; 1 \le j_1 < j_2 < ... < j_k \le d, k = 2,...,d\}$ is the set of unstable equilibria, where $z_{j_1,j_2,...,j_k}$ stands for the point $z = (z_1,....,z_d) \in \Delta$ with $z_{j_1} = z_{j_2} = ... = z_{j_k} = \frac{1}{k}$ and all the remaining coordinates are equal to 0.

For $\alpha > 1$, we notice that the semiflow Φ admits a Lyapunov function $H : \Delta \to \mathbb{R}$ for $\mathcal{C}$, which is defined by

$$H(z) = z_1^\alpha + z_2^\alpha + ... + z_n^\alpha.$$

By using Corollary 1.12, we conclude that Z_t almost surely converges to an equilibrium z^* in $\mathcal{C}$. In addition, we also prove that on the event $\{Z_t \to z^*\}$, where z^* is a stable equilibrium, then

$$\limsup_{t\to\infty} t\|Z_t - z^*\| < \infty \ \text{ a.s.}$$

Beside this, we show that for $\alpha \ge 2$, the probability that Z_t converges to an unstable equilibrium is zero. The above results implies that there exists $j \in V$ such that

$$\limsup_{t\to\infty} L(j,t) \to \infty \text{ a.s.,}$$

while

$$\limsup_{t\to\infty} L(i,t) < \infty \text{ a.s.}$$

for all $i \in V \setminus \{j\}$. The case for $1 < \alpha < 2$ is however still open!

Random walks on simplexes

Motivation

We denote as

$$\mathcal{S}_d = \{(z_1, z_2, \ldots, z_d) \in \mathbb{R}^d : z_1 + z_2 + \cdots + z_d \leq 1, z_j \geq 0, j = 1, 2, \ldots, d\}$$

the standard unit simplex in $\mathbb{R}^d$, $d \geq 1$. Let $E_0 = (0, 0, \ldots, 0)$ be the origin and $E_1 = (1, 0, \ldots, 0)$, $E_2 = (0, 1, 0, \ldots, 0)$, ..., $E_d = (0, \ldots, 0, 1)$ be standard orthonormal basis vectors in $\mathbb{R}^d$, which are also the vertices of $\mathcal{S}_d$. Let $p = (p_1, p_2, ..., p_d)$ be a mapping from $\mathcal{S}_d$ to itself so-called *probability choice function*. Let us consider a random walk $(Z_n)_{n \geq 0}$ on $\mathcal{S}_d$ which is a Markov chain defined as follows:

- At time $n = 0$, the walker stays at some initial point $Z_0 \in \mathcal{S}_d$.

- Assuming that at time $n \geq 0$, the walker stays at Z_n. The walker randomly selects a point Θ_n from vertices $E_0, E_1, \ldots, E_d$, where Θ_n is a discrete random variable independent such that

$$\begin{cases} \mathbb{P}(\Theta_n = E_j | Z_n = z) = p_j(z), j = 1, 2, ..., d; \\ \mathbb{P}(\Theta_n = E_0 | Z_n = z) = 1 - \sum_{j=1}^{d} p_j(z) := p_0(z). \end{cases}$$

- At time $n + 1$, the walker randomly moves from Z_n to a new position Z_{n+1} inside the segment $\Phi_n Z_n$ such that

$$Z_{n+1} = (1 - \xi_n)Z_n + \xi_n \Theta_n,$$

where $\xi_n, n = 0, 1, 2, ...$ are i.i.d copies of a random variable ξ with support in $[0, 1]$.

Note that the above-mentioned settings could be easily generalized to any simplex in $\mathbb{R}^d$. We therefore restrict the problem only on the standard unit simplex $\mathcal{S}_d$ without loss of generality.

Let us consider the case where $d = 1$, $p(z) = \frac{1}{2}$ for all $z \in \mathcal{S}_1 = [0, 1]$ and ξ is uniformly distributed in $[0, 1]$. Diaconis and Freedman in [11] showed that the Markov chain $(Z_n)_{n \geq 0}$ has an unique stationary distribution. If this

stationary distribution has a density f, then it should be a solution of the following equation

$$f(y) = \frac{1}{2} \int_y^1 \frac{f(x)}{x} dx + \frac{1}{2} \int_0^y \frac{f(x)}{1-x} dx$$

for $y \in (0,1)$. Differentiating both sides, we have

$$f'(y) = \frac{1}{2} \left(\frac{1}{1-y} - \frac{1}{y} \right) f(y).$$

It is easy to check that $f(y) = \frac{1}{\pi\sqrt{1-y}}$ is the solution of the above equation, which is also the probability density function of the arcsine law.

A more general one-dimensional case was considered by McKinlay and Borovkov in [16], where $p(z)$ is piecewise continuous on $[0,1]$ such that

$$\sup_{z \in [0,\delta]} \max\{p(z), 1 - p(1-z)\} < 1$$

for some $\delta \in (0, \frac{1}{2})$ and ξ is Beta$(1,\gamma)$ distributed. Recall that we say a distribution with support in $[0,1]$ is Beta(a,b) if its probability density function is given by

$$f(z) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} z^{a-1}(1-z)^{b-1}, z \in (0,1),$$

where Γ is the Gamma function. These authors showed that the Markov chain $(Z_n)_{n\geq 0}$ is ergodic with the stationary density defined by

$$\pi(z) = Cz^\gamma \left(\frac{1}{1-z} + \frac{1}{z} \right) \exp \left(-\gamma \int_{1/2}^z p(u) \left(\frac{1}{1-u} + \frac{1}{u} \right) du \right)$$

for $z \in (0,1)$, where C is the normalizing constant. In particular, if $p(z) = \beta_1(1-z) + (1-\beta_2)z$ then the stationary distribution is Beta$(\beta_1\gamma, \beta_2\gamma)$.

The multidimensional case was also discussed by Diaconis and Freedman in [11], Sethuraman in [23] where $p = (p_1, p_2, \ldots, p_d)$ is a vector of constants. The ergodicity as well as the convergence to stationary distribution of the multidimensional Markov chain $(Z_n)_{n\geq 0}$ in general case, where the probability choice function p depends on Z_n, however, have not been investigated yet.

We are also interested in the modified random walk model $(Y_n)_{n\geq 0}$ in the unit interval $\mathcal{S}_1 = [0,1]$, which is defined as follows

- At time n, there are potentials L_n and R_n at 0 and 1 respectively, and the walker stays at $Y_n \in [0, 1]$.

- At time $n + 1$, with probability $\frac{L_n}{L_n + R_n}$, the potential at 0 will increase a value proportional to Y_n, i.e. the distance from 0 to the current position of the particle, and then the particle is pulled to a new uniformly random position Y_{n+1} inside the interval $(0, Y_n)$. Otherwise, with probability $\frac{R_n}{L_n + R_n}$, the potential at 1 will increase a value proportional to $1 - Y_n$, i.e. the distance from 1 to the current position of the particle, and then the particle is pulled to a new uniformly random position Y_{n+1} inside the interval $(Y_n, 1)$. More precisely, the new position of the walker is defined by

$$Y_{n+1} = Y_n(1 - \xi_n^{(L)})\mathbf{1}_{\{U_n < \frac{L_n}{L_n + R_n}\}} + (Y_n + (1 - Y_n)\xi_n^{(R)})\mathbf{1}_{\{U_n > \frac{L_n}{L_n + R_n}\}},$$

  and the potentials L_n, R_n are defined as follows

$$\begin{aligned} L_{n+1} &= L_n + f(Y_n)\mathbf{1}_{\{U_n < \frac{L_n}{L_n + R_n}\}}, \\ R_{n+1} &= R_n + f(1 - Y_n)\mathbf{1}_{\{U_n > \frac{L_n}{L_n + R_n}\}}, \end{aligned}$$

  where $f : [0, 1] \to [0, +\infty)$ is some function; $\xi_n^{(L)}$ and $\xi_n^{(R)}, n \geq 1$, are independent random variables taking values in $[0, 1]$; $U_n, n \geq 1$ are i.i.d uniformly distributed random variables in $[0, 1]$ and independent of $\xi_n^{(L)}$, $\xi_n^{(R)}$, $Y_n, n \geq 1$.

Since the probabilities jumping to the left or the right of the walker depend on (L_n, R_n), the random walk $\{Y_n\}_{n \geq 0}$ is no longer a Markov chain but depends on its history.

Assume that $\xi_n^{(L)} = \xi_n^{(R)}, n \geq 1$, are independent and uniformly distributed in $[0, 1]$. Let us consider the following simple cases when $f : [0, 1] \to [0, \infty)$ is a constant:

- If $f(x) = 0$ for all $x \in [0, 1]$, the process is reduced to the Markov model with constant choice probability $p = \frac{R_0}{L_0 + R_0}$. Hence, Y_n weakly converges to Beta $\left(\frac{R_0}{R_0 + L_0}, \frac{L_0}{R_0 + L_0}\right)$.

- If $f(x) = \beta > 0$ for all $x \in [0, 1]$, the process $(L_n, R_n)_{n \geq 1}$ is the classical Friedman urn with the matrix

$$\begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix}.$$

24

It is well-known that $L_n/(R_n + L_n)$ converges almost surely to $1/2$ as $n \to \infty$ (see e.g. [13], Corollary 5.2). Therefore, one can show that Y_n converges in distribution to the arcsine law $\text{Beta}(1/2, 1/2)$.

In general case, one can expect that under some conditions for the function f, the random walk Y_n might converge almost surely to some $x^* \in [0,1]$ or weakly converge to some non-trivial distribution.

Review of Paper C

In Paper C, we investigate the existence of stationary distribution and the convergence of the multidimensional Markov chain $(Z_n)_{n \geq 0}$. In the case where ξ is Beta distributed and p is an affine map, we prove that the limiting distributions are Dirichlet. In addition, in a special setting, we show that the history-dependent random walk $(Y_n)_{n \geq 0}$ converges in distribution to the arcsine law.

To ensure the ergodicity of the Markov chain $(Z_n)_{n \geq 0}$, we need the following

Assumption 5. *There are $\delta \in (0, \frac{1}{2^d})$ and $s, t \in (\delta^{1/d}, 1 - \delta^{1/d})$, $s < t$ such that*

(i) $F_\xi(1 - \delta) := 1 - \eta < 1$;

(ii) *there is an $\epsilon > 0$ such that for any $1 \leq k \leq d$ and any $0 \leq j_1 < j_2 < \cdots < j_k \leq d$,*

$$\inf_{z \in \mathcal{S}_d : z_{j_1} + \cdots + z_{j_k} \leq \delta} \sum_{l=1}^{k} p_{j_l}(z) \geq \epsilon;$$

(iii) *there is $c > 0$ such that for all Borel measurable subset $B \subset [s(1-t)^{d-1} - \delta, t] \cup [(1-t)^d - \delta, 1 - s]$,*

$$\mathbb{P}(\xi \in B) > c\lambda(B),$$

where λ is the Lebesgue measure on $[0,1]$.

Let

$$V = \bigcup_{j=0}^{d} V_j,$$

where $V_j = \{z = (z_1, \ldots, z_d) \in \mathcal{S}_d : 1 - \delta \leq z_j \leq 1\}, j = 1, 2, \ldots, d$. We define the probability measure φ such that for each Borel measurable subset $B \subset \mathcal{S}_d$,

$$\varphi(B) = \frac{\lambda_d (B \cap K)}{\lambda_d (K)},$$

where

$$K := \left\{ (u_1, \ldots, u_d) \in \mathcal{S}_d : s \leq \frac{u_j}{1 - \sum_{l=j+1}^d u_l} \leq t, j = 1, 2, \ldots, d \right\}.$$

By verifying all conditions in Proposition 1.18 with $n_0 = d$ and V, φ defined as above, we conclude that the Markov chain $(Z_n)_{n \geq 0}$ converges in distribution if Assumption 5 is fulfilled.

In the case where ξ is Beta$(1, \gamma)$ distributed and the probability choice function $p = (p_1, p_2, \ldots, p_d)$ is defined by

$$p_k(z_1, z_2, \ldots, z_d) = \beta_k(1 - z_k) + \left(1 - \sum_{j=1}^{d+1} \beta_j + \beta_k \right) z_k, \ k = 1, 2, \ldots, d,$$

where $\beta_k > 0$ and $\sum_{j=1}^{d+1} \beta_j - \beta_k < 1$ for $k = 1, 2, \ldots, d + 1$, we prove that Z_n converges in distribution to Dirichlet$(\beta_1\gamma, \beta_2\gamma, \ldots, \beta_d\gamma, \beta_{d+1}\gamma)$. Note that a distribution with support in $\mathcal{S}_d$ is Dirichlet$(\alpha_1, \alpha_2, \ldots, \alpha_{d+1})$ if its probability density function is given by

$$f(z_1, z_2, \ldots, z_d) = \frac{\Gamma\left(\sum_{i=1}^{d+1} \alpha_i \right)}{\prod_{i=1}^{d+1} \Gamma(\alpha_i)} \left(1 - \sum_{i=1}^d z_i \right)^{\alpha_{d+1} - 1} \prod_{i=1}^d z^{\alpha_i - 1}$$

for interior point $(z_1, z_2, \ldots, z_d) \in \mathcal{S}_d$. Note that the one-dimensional model considered by McKinlay and Borovkov in [16] is a special case when $d = 1$, $p(z) = \beta_1(1 - z) + (1 - \beta_2)z$. For the multidimensional case $d \geq 1$, where $\sum_{j=1}^{d+1} \beta_j = 1$, i.e. the choice probabilities are constants, we also obtain Sethuraman's result derived in [23].

We now return to the modified random walk $(Y_n)_{n \geq 0}$. Assume that $f(x) = x$ for all $x \in [0, 1]$ and $\xi_n^{(L)} = \xi_n^{(R)}, n \geq 1$, are uniformly distributed in $[0, 1]$. We show that

$$\zeta_n = \frac{L_n}{L_n + R_n} \to \frac{1}{2} \text{ a.s.}$$

as $n \to \infty$. Moreover, using coupling technique, we prove that Y_n converges in distribution to the arsine law Beta $\left(\frac{1}{2}, \frac{1}{2} \right)$ as $n \to \infty$.

Bibliography

[1] Benaïm, M.; Hirsch, M. W. Asymptotic pseudotrajectories and chain recurrent flows, with applications, J. Dynam. Differential Equations **8** (1996), no. 1, 141–176.

[2] Benaïm, M.; Hirsch, M. W. Dynamics of Morse-Smale urn processes. Ergodic Theory Dynam. Systems **15** (1995), no. 6, 1005–1030.

[3] Benaïm, M. Vertex-reinforced random walks and a conjecture of Pemantle, Ann. Probab. **25** (1997), no. 1, 361–392.

[4] Benaïm, M. Dynamics of stochastic approximation algorithms, in *Séminaire de Probabilités, XXXIII*, 1–68, Lecture Notes in Math., 1709, Springer, Berlin.

[5] Benaïm, M.; Olivier, R. Self-interacting diffusions. III. Symmetric interactions. Ann. Probab. **33** (2005), no. 5, 1717–1759.

[6] Benoist Y. and Quint, J.-F. *Random walks on reductive groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, 62, Springer, Cham, 2016.

[7] Borovkov, A. A. *Ergodicity and stability of stochastic processes*, translated from the 1994 Russian original by V. Yurinsky, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Ltd., Chichester, 1998.

[8] Bougerol, P.; Lacroix, J. *Products of random matrices with applications to Schrödinger operators*. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985.

[9] Davis, B.; Volkov, S. Continuous time vertex-reinforced jump processes, Probab. Theory Related Fields **123** (2002), no. 2, 281–300.

[10] Davis, B.; Volkov, S. Vertex-reinforced jump processes on trees and finite graphs, Probab. Theory Related Fields **128** (2004), no. 1, 42–62.

[11] Diaconis, P.; Freedman, D. Iterated random functions. SIAM Rev. **41** (1999), no. 1, 45–76.

[12] Disertori, M.; Merkl F.; Rolles, S. W. W. Localization for a nonlinear sigma model in a strip related to vertex reinforced jump processes, Comm. Math. Phys. **332** (2014), no. 2, 783–825.

[13] Freedman, D. A. Bernard Friedman's Urn. Ann. Math. Statist. **36** (1965), no. 3, 956–970.

[14] Goldsheid, I. Ya. Guivarc'h, Y. Zariski closure and the dimension of the Gaussian law of the product of random matrices. I, Probab. Theory Related Fields **105** (1996), no. 1, 109–142.

[15] Goldsheid, I. Ya.; Margulis, G. A. Lyapunov exponents of a product of random matrices. Russian Math. Surveys **44** (1989), no. 5, 11–71; translated from Uspekhi Mat. Nauk **44** (1989), no. 5(269), 13–60.

[16] McKinlay, S.; Borovkov, K. On explicit form of the stationary distributions for a class of bounded Markov chains. J. Appl. Probab. **53** (2016), no. 1, 231–243

[17] Merkl F.; Rolles, S. W. W.; Tarrès, P. Convergence of vertex-reinforced jump processes to an extension of the supersymmetric hyperbolic nonlinear sigma model. ArXiv e-prints, arXiv:1612.05409.

[18] Orey, S. *Lecture notes on limit theorems for Markov chain transition probabilities*. Van Nostrand Reinhold Mathematical Studies, No. 34. Van Nostrand Reinhold Co., London-New York-Toronto, Ont., 1971.

[19] Pemantle, R. Nonconvergence to unstable points in urn models and stochastic approximations. Ann. Probab. **18** (1990), no. 2, 698–712.

[20] Robinson, C. *Dynamical systems. Stability, symbolic dynamics, and chaos*. Second edition. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1999.

[21] Sabot, C.; Tarrès, P. Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. J. Eur. Math. Soc. (JEMS) **17** (2015), no. 9, 2353–2378.

[22] Sabot C.; Tarrès P.; Zeng X. The Vertex Reinforced Jump Process and a Random Schrödinger operator on finite graphs. ArXiv e-prints, arXiv:1507.04660.

[23] Sethuraman, J. A constructive definition of Dirichlet priors. Statist. Sinica **4** (1994), no. 2, 639–650.

[24] Volkov, S. Vertex-reinforced random walk on arbitrary graphs, Ann. Probab. **29** (2001), no. 1, 66–91.

[25] Volkov, S. Random geometric subdivisions, Random Structures Algorithms **43** (2013), no. 1, 115–130.

[26] Zeng, X. How vertex reinforced jump process arises naturally. Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 3, 1061–1075.

A

Paper A

A universal result for consecutive random subdivision of polygons

TUAN-MINH NGUYEN AND STANISLAV VOLKOV

Centre for Mathematical Sciences, Lund University

Abstract

We consider consecutive random subdivision of polygons described as follows. Given an initial convex polygon with $d \geq 3$ edges, we choose a point at random on each edge, such that the proportions in which these points divide edges are i.i.d. copies of some random variable ξ. These new points form a new (smaller) polygon. By repeatedly implementing this procedure we obtain a sequence of random polygons. The aim of this paper is to show that under very mild non-degenerateness conditions on ξ, the shapes of these polygons eventually become "flat". The convergence rate to flatness is also investigated; in particular, in the case of triangles ($d = 3$), we show how to calculate the exact value of the rate of convergence, connected to Lyapunov exponents. Using the theory of products of random matrices our paper greatly generalizes the results of [11] which are achieved mostly by using ad hoc methods.

Keywords: Random subdivisions, products of random matrices, Lyapunov exponents.

Introduction

Many problems of consecutive random subdivision of a convex geometrical figure have been investigated by several authors since 1980s. In [13], G. S. Watson introduced the following model: given an initial triangle, one chooses a

point on each edge by keeping the same random proportion ξ and hence obtaining a new triangle. If one repeats the above process with independent identically distributed random proportions $\xi^{(n)}, n = 1, 2, \ldots$ then the limit triangle vanishes to the centroid of the initial triangle. To study the *shapes* of these triangles, let us rescale the newly formed in each step triangle in such a way that the largest side has length 1. It is interesting that the "limit" of these rescaled triangles is non-vanishing and, in fact, random. Veitch and Watson in [12] also gave an extension for a system of points in higher dimensional real space. With the same motivation of random triangles, Mannion in [9] studied the situation where on each step the triangle is formed by choosing three uniformly distributed random points inside the *interior* of the preceding triangle. The sides of these triangles almost surely converge to collinear segments. Diaconis and Miclo [5] considered a triangle split by the three medians such that one of the 6 triangles is chosen at random to replace the original triangle. It turns out that the limiting triangle's shape is flat. Volkov in [11] discovered a similar phenomenon by considering a model where the new triangle is formed by choosing a random point uniformly and independently on each of the sides of the original triangle; he also studied distribution of the "middle" point.

In the present paper, we give a generalization of Volkov's result in [11] for all convex polygons and nearly all non-degenerate distributions of proportions in which the sides of the polygon are split.

Let us now formulate the model rigorously. Fix $d \geq 3$ and a random variable ξ whose support lies on $[0,1]$. Let $L_0 = A_1^{(0)} A_2^{(0)} \ldots A_d^{(0)}$ be a convex d-polygon (i.e., a convex polygon with d sides) in the plane, with edges $A_j^{(0)} A_{j+1}^{(0)}$, $j = 1, 2, \ldots, d$, with the convention $A_{d+1}^{(1)} \equiv A_1^{(1)}$. Randomly choose a point $A_j^{(1)}$ in $A_j^{(0)} A_{j+1}^{(0)}$ such that $|A_j^{(0)} A_j^{(1)}| / |A_j^{(0)} A_{j+1}^{(0)}| = \xi_j$, where ξ_j, $j = 1, \ldots, d$, are i.i.d. copies of the random variable ξ. Thus we obtain new convex polygon $L_1 = A_1^{(1)} A_2^{(1)} \ldots A_d^{(1)}$. Repeating the above process such that the random vectors $\left(\xi_1^{(n)}, \xi_2^{(n)}, \ldots, \xi_d^{(n)} \right)$, $n = 1, 2, \ldots$, are i.i.d., we obtain a Markov chain of polygons $(L_n)_{n \geq 0}$ where $L_n = A_1^{(n)} A_2^{(n)} \ldots A_d^{(n)}$.

It is easy to see that the polygons L_n become smaller and smaller and eventually converge to a point, however the behaviour of their *shapes* is less clear. To study the shapes we may, for example, place one of the vertices at the origin $(0,0)$ and rescale the polygon in such a way that its longest edge has
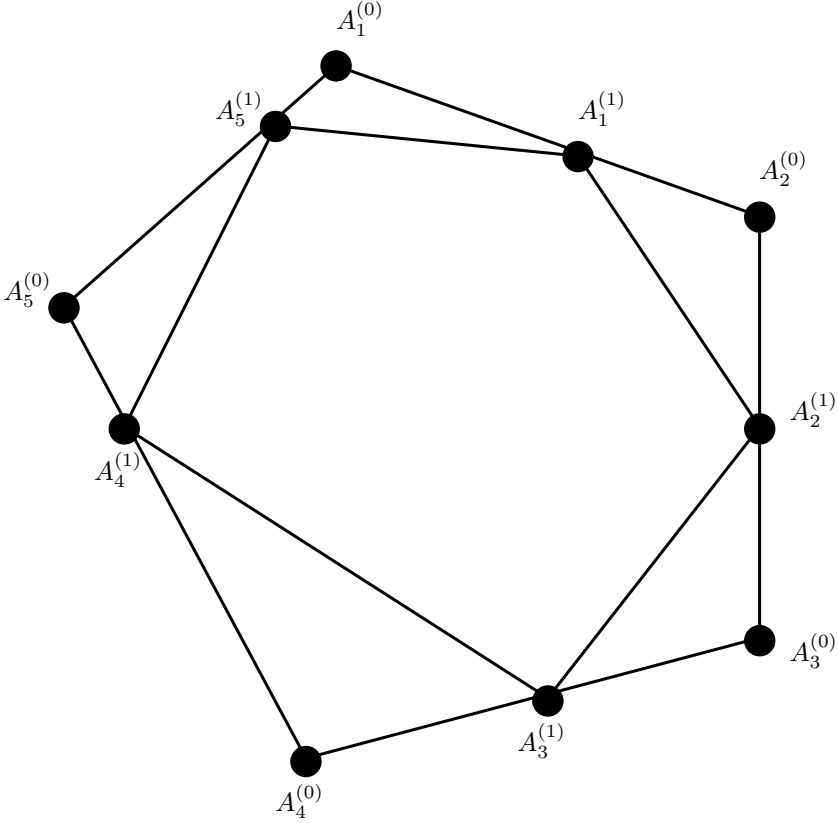
Figure A.1: A new smaller random pentagon L_1 obtaining from the primary pentagon L_0.

always length 1. We will show that under some regularity conditions on the distribution of ξ the rescaled polygon will eventually become degenerate, i.e. flat, in the sense that all of its vertices will be lying approximately along the same line; observe that this is equivalent to the fact that the area of the rescaled polygon converges to 0 as n goes to infinity.

Let $l_j^{(n)} = A_j^{(n)} A_{j+1}^{(n)}$, $j = 1, 2, \ldots, d$, be the vector corresponding to the j-th side of L_n and $(x_j^{(n)}, y_j^{(n)})$ denote its Cartesian coordinates. From elementary geometrical calculations one can obtain the following linear relation:

$$x^{(n+1)} = H_{n+1} x^{(n)}, \; y^{(n+1)} = H_{n+1} y^{(n)} \tag{A.1}$$

where $x^{(n)} = \left(x_1^{(n)}, x_2^{(n)}, \ldots, x_d^{(n)} \right)^{\mathbb{T}}$ and $y^{(n)} = \left(y_1^{(n)}, y_2^{(n)}, \ldots, y_d^{(n)} \right)^{\mathbb{T}}$ are column

vectors, and H_n is an i.i.d. copy of the following random matrix

$$H = H(\xi_1, \ldots, \xi_d) = \begin{pmatrix} 1 - \xi_1 & \xi_2 & 0 & \ldots & 0 \\ 0 & 1 - \xi_2 & \xi_3 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & \xi_d \\ \xi_1 & 0 & 0 & \ldots & 1 - \xi_d \end{pmatrix} \qquad \text{(A.2)}$$

and $\xi_1, \xi_2, \ldots, \xi_d$ are i.i.d. copies of a random variable ξ. Note that $\sum_{j=1}^{d} x_j^{(n)} = 0$ and $\sum_{j=1}^{d} y_j^{(n)} = 0$. In particular,

$$l_j^{(n)} = ((\mathbf{e}_1, \ldots, \mathbf{e}_n)_j H^{(n)} x^{(0)}, (\mathbf{e}_1, \ldots, \mathbf{e}_n)_j H^{(n)} y^{(0)})$$

where $H^{(n)} = H_n H_{n-1} \ldots H_1$ and $(\mathbf{e}_1, \ldots, \mathbf{e}_n)_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ is $1 \times d$ vector with 1 on the j-th place. Note also that if the original polygon is non-degenerate then $H^{(n)} x^{(0)}$ and $H^{(n)} y^{(0)}$ are non-zero vectors for any n.

To ensure that L_n is a non-degenerate convex polygon and that the subdivision is genuinely random, we need the following

Assumption 6. $\mathbb{P}(\xi \in \{0,1\}) = 0$ and the support of ξ contains at least two distinct points in $(0,1)$, i.e. the distribution of ξ is non-degenerate.

We can define "thickness" of a two-dimensional object as the smallest possible ratio between its one-dimensional projections on the two coordinate axes of a Cartesian coordinate system (where we can orient this system arbitrarily); this quantity always lies between 0 and 1; moreover, it equals one for a circle, and it equals zero for any segment. The sequence of L_n converges to a "flat figure", or simply to "flatness", if the sequence of its thicknesses converges to zero. In the case of polygons, this definition is equivalent to

Definition 1.1. We say that the sequence of polygons L_n converges to a flat figure as $n \to \infty$ if

$$\lim_{n \to \infty} \frac{\mathbf{A}(L_n)}{\left(\max_{j=1,\ldots,d} \|l_j^{(n)}\|\right)^2} = 0.$$

Here $\mathbf{A}(L_n)$ denotes the area of the polygon L_n.

The main purpose of our paper is to (partially) establish the following phenomenon.

Conjecture 1. *Suppose that Assumption 6 holds, then the sequence of polygons L_n converges to a flat figure almost surely as $n \to \infty$.*

Further the dynamics of the random subdivisions will be formulated as a certain model related to products of random matrices and its point limit in the projective space. Let $\mathbb{R}^d$ (and $\mathbb{C}^d$) denote the linear space of all d-dimensional real (complex, resp.) column vectors under the field of real (complex) numbers. The real (complex) projective space $P(\mathbb{R}^d)$ is defined as the quotient space $(\mathbb{R}^d \setminus \{0\}) / \sim$, where $\sim$ is the equivalence relation defined by $x \sim y$, $x, y \in \mathbb{R}^d$ if there exists a real (complex) number λ such that $x = \lambda y$. We denote $\overline{x}$ as the equivalence class of x. The projective space $P(\mathbb{R}^d)$ becomes a compact metric space if we consider the following "angular" metric

$$\delta(\overline{x}, \overline{y}) = \sqrt{1 - \frac{(x, y)^2}{||x||^2 \cdot ||y||^2}}. \tag{A.3}$$

where $|| \cdot ||$ and $(\cdot, \cdot)$ are respectively the Euclidean norm and the Euclidean scalar product on $\mathbb{R}^d$. One can see that $\delta(\overline{x}, \overline{y})$ is actually the sinus of the smaller angle between the lines corresponding to $\bar{x}$ and $\bar{y}$.

Next, each linear mapping $A : \mathbb{R}^d \to \mathbb{R}^d$ can be generalized to $P(\mathbb{R}^d)$ by setting

$$A\overline{x} = \overline{Ax}$$

for every $x \in \mathbb{R}^d \setminus \mathrm{Ker}(A)$. Let us also define

$$\mathcal{L} = \{v \in \mathbb{R}^d : v_1 + v_2 + \cdots + v_n = 0\}. \tag{A.4}$$

Observe that since $\sum_{j=1}^d x_j^{(n)} = 0$, $\sum_{j=1}^d y_j^{(n)} = 0$, we have $x^{(n)}, y^{(n)} \in \mathcal{L}$.

Proposition 1.2. *Suppose that*

$$\lim_{n \to \infty} \delta\left(H^{(n)}\overline{x}, H^{(n)}\overline{y}\right) = 0 \tag{A.5}$$

almost surely for every $x, y \in L_n$ such that $(x_1, y_1), (x_2, y_2), \ldots, (x_d, y_d)$ are coordinates of vectors corresponding to consecutive edges of the convex d-polygon in the real plane. Then L_n converges to a flat figure as $n \to \infty$.

Proof. Using the formula for $\delta\left(\overline{x^{(n)}}, \overline{y^{(n)}}\right)$ and omitting the superscript $^{(n)}$ for all $x^{(n)}$ and $y^{(n)}$ for simplicity, we obtain that

$$\delta(\overline{x}, \overline{y})^2 = \frac{\left(\sum_{i=1}^d x_i^2\right)\left(\sum_{i=1}^d y_i^2\right) - \left(\sum_{i=1}^d x_i y_i\right)^2}{\left(\sum_{i=1}^d x_i^2\right)\left(\sum_{i=1}^d y_i^2\right)} = \frac{\sum_{1 \leq i < j \leq d}(x_i y_j - x_j y_i)^2}{\left(\sum_{i=1}^d x_i^2\right)\left(\sum_{i=1}^d y_i^2\right)} =: \delta_n$$

where $\delta_n \to 0$ a.s.

According to a well-know formula for the signed area $\mathbf{A}(L)$ of a planar non-self-intersecting polygon L with vertices $(a_1, b_1), \ldots, (a_d, b_d)$, see [1]

$$2\mathbf{A}(L) = \det \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} + \det \begin{pmatrix} a_2 & a_3 \\ b_2 & b_3 \end{pmatrix} + \cdots + \det \begin{pmatrix} a_d & a_1 \\ b_d & b_1 \end{pmatrix}.$$

Since we know only the coordinates of the vectors forming the edges of polygon (x_i, y_i), $i = 1, 2, \ldots, d$ with the obvious restriction $\sum_{i=1}^d x_i = \sum_{i=1}^d y_i = 0$, we can assume that the polygon's vertices have the coordinates

$$a_i = x_1 + \cdots + x_i,$$
$$b_i = y_1 + \cdots + y_i,$$

$i = 1, 2, \ldots, d$, thus yielding that $a_d = b_d = 0$ so that the last two determinants in the formula for $2\mathbf{A}(L)$ are 0, and hence

$$2\mathbf{A}(L) = \sum_{i=1}^{d-2} \det \begin{pmatrix} a_i & a_{i+1} \\ b_i & b_{i+1} \end{pmatrix} = \sum_{i=1}^{d-2} \det \begin{pmatrix} a_i & a_i + x_{i+1} \\ b_i & b_i + y_{i+1} \end{pmatrix} = \sum_{i=1}^{d-2} (a_i y_{i+1} - b_i x_{i+1})$$

$$= [x_1 y_2 + (x_1 + x_2) y_3 + \cdots + (x_1 + x_2 + \ldots x_{d-2}) y_{d-1}]$$
$$- [y_1 x_2 + (y_1 + y_2) x_3 + \cdots + (y_1 + y_2 + \ldots y_{d-2}) x_{d-1}]$$

$$= \sum_{1 \le i < j \le d-1} \det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix}.$$

Therefore the area $\mathbf{A}(L_n)$ of the polygon L_n satisfies

$$|2\mathbf{A}(L_n)| = \left| \sum_{1 \le i < j \le d-1} \det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix} \right| \le \sum_{1 \le i < j \le d-1} \left| \det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix} \right|$$

$$\le \sqrt{\sum_{1 \le i < j \le d} (x_i y_j - x_j y_i)^2} = \sqrt{\delta_n \left(\sum_{i=1}^d x_i^2 \right) \left(\sum_{i=1}^d y_i^2 \right)}.$$

Consequently,

$$\frac{\mathbf{A}(L_n)}{\left(\max_j \|l_j^{(n)}\| \right)^2} \le \frac{1}{2} \sqrt{\delta_n \frac{\left(\sum_{i=1}^d x_i^2 \right) \left(\sum_{i=1}^d y_i^2 \right)}{\left(\max_{j=1,\ldots,d} \left[x_j^2 + y_j^2 \right] \right)^2}} \le \frac{1}{2} \sqrt{\delta_n \cdot d \cdot d} \to 0$$

since $x_i^2 \le \max_{j=1,\ldots,d} (x_j^2 + y_j^2)$ for each i, and the same holds for y_i. $\qquad \square$

Note that $\mathcal{L}$ defined by (A.4) is an invariant subspace of H. Therefore, we can restrict the linear transformation H to $\mathbb{R}^{d-1}$ by considering only the first $d-1$ coordinates of x and y respectively. One can easily deduce that the restriction of the transformation H can be described by the $(d-1) \times (d-1)$ matrix

$$T = T(\xi_1, \ldots, \xi_d) = \begin{pmatrix} 1 - \xi_1 & \xi_2 & 0 & \cdots & 0 & 0 \\ 0 & 1 - \xi_2 & \xi_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 - \xi_{d-2} & \xi_{d-1} \\ -\xi_d & -\xi_d & -\xi_d & \cdots & -\xi_d & 1 - \xi_{d-1} - \xi_d \end{pmatrix}$$

$$\text{(A.6)}$$

and then the linear relation (A.1) still has the same formulation in $\mathbb{R}^{d-1}$ for T. The condition (A.5) for the matrix (A.6) now can be restated as

Proposition 1.3. *Let $\{T_n\}_{n \geq 1}$ be a sequence of random matrices, which are independent copies of the matrix T in (A.6) and let $T^{(n)} = T_n T_{n-1} \ldots T_2 T_1$. Assume that*

$$\lim_{n \to \infty} \delta(T^{(n)} \bar{x}, T^{(n)} \bar{y}) = 0 \qquad \text{(A.7)}$$

almost surely for any $x = (x_1, \ldots, x_{d-1})^{\mathbb{T}}, y = (y_1, \ldots, y_{d-1})^{\mathbb{T}} \in \mathbb{R}^{d-1}$, such that $(x_1, y_1), (x_2, y_2), \ldots, (x_{d-1}, y_{d-1})$ are coordinates of $d-1$ consecutive edges of a convex d-polygon in the real plane. Then L_n converges to a flat figure as $n \to \infty$.

Proof. Basically, we need to show the following geometric fact. Suppose that $x^{(n)} = (x_1^{(n)}, \ldots, x_{d-1}^{(n)})$ and $y^{(n)} = (y_1^{(n)}, \ldots, y_{d-1}^{(n)})$ are such that $\delta_n := \delta(x^{(n)}, y^{(n)})$ tends to 0 as $n \to \infty$, then $\tilde{\delta}_n := \delta(\tilde{x}^{(n)}, \tilde{y}^{(n)}) \to 0$, where $\tilde{x}^{(n)} = (x_1^{(n)}, \ldots, x_d^{(n)})$ and $\tilde{y}^{(n)} = (y_1^{(n)}, \ldots, y_d^{(n)})$ with $x_d^{(n)} = -\sum_{i=1}^{d-1} x_i^{(n)}, y_d^{(n)} = -\sum_{i=1}^{d-1} y_i^{(n)}$, for all n. Observe that δ_n and $\tilde{\delta}_n$ represent the angular distance on the spaces $P(\mathbb{R}^{d-1})$ and $P(\mathbb{R}^d)$ respectively.

Indeed, suppose that $\delta_n < \epsilon$ for some very small $\epsilon > 0$. Let us from now on also omit the superscript $^{(n)}$ as this does not create a confusion. Without loss of generality we can assume that $\|x\| = \|y\| = 1$, that is, $\sum_{i=1}^{d-1} x_i^2 = 1 = \sum_{i=1}^{d-1} x_i^2$. Denote by $c = (x, y) = \sum_{i=1}^{d-1} x_i y_i = \cos(x, y)$, so that $c^2 + \delta_n^2 = 1$. We have

$$\tilde{\delta}_n^2 = \frac{(1 + x_d^2)(1 + y_d^2) - (\sum_{i=1}^{d} x_i y_i)^2}{(1 + x_d^2)(1 + y_d^2)} = \frac{(1 + x_d^2)(1 - c^2) + (y_d - c x_d)^2}{(1 + x_d^2)(1 + y_d^2)}$$

$$\leq (1 - c^2) + (y_d - c x_d)^2 = \delta_n^2 + \left(\sum_{i=1}^{d-1} u_i \right)^2 \qquad \text{(A.8)}$$

where $u = y - cx = (y_1 - cx_1, \ldots, y_{d-1} - cx_{d-1})$ is the difference between vector y and the projection of y on x. Consequently, u is orthogonal to x and $\|u\|^2 = \|y\|^2 - \|cx\|^2 = 1 - c^2 = \delta_n^2$. By the inequality between the quadratic and arithmetic means $|\sum_{i=1}^{d-1} u_i|^2 \leq (d-1)\|u\|^2$ hence (A.8) implies that $\tilde{\delta}_n^2 \leq [1 + (d-1)]\delta_n^2 \leq d\epsilon^2$. $\qquad\square$

The rest of the paper is organized as follows. In Section 2, by applying the classical Furstenberg's theorem for products of 2×2 invertible random matrices, we will show that (A.7) is fulfilled for $d = 3$ (Theorem 2.2). In a higher dimensional case, it is necessary to show that the closed semigroup generated by the support of the random matrix T is strongly irreducible and contracting. We will show that (A.7) holds for any *odd* number $d > 3$ in Section 3. For the remaining case when $d \geq 4$ is *even*, we will have to require that the random matrix T in (A.6) is invertible almost surely. We actually believe that this extra requirement is not really needed, however we are unable to show the result without this extra condition. The results are summarized in Theorem 3.5. The exponential rate of convergence of random polygons will be considered in Section 4, see Theorems 4.3, 5.4 and 4.10.

Finally, in Section 6 we mention some generalizations of our model, as well as open problems. Also note that throughout the the paper we denote by $GL(d, \mathbb{R})$ the group of $d \times d$ invertible matrices of real numbers and $SL^\pm(d, \mathbb{R})$ the closed subgroup of $GL(d, \mathbb{R})$ containing all matrices with determinant $+1$ or -1.

Random subdivision of triangles ($d = 3$)

Proposition 2.1. (Furstenberg's theorem, Theorem II.4.1 in [2], page 30) *Let μ be a probability measure on $GL(2, \mathbb{R})$ and G_μ be the smallest closed subgroup of $GL(2, \mathbb{R})$ which contains the support of μ. Suppose that the following hold:*

(i) *$G_\mu \subset SL^\pm(2, \mathbb{R})$;*

(ii) *G_μ is not compact;*

(iii) *There does not exist any common invariant finite union of one-dimensional subspaces of $\mathbb{R}^2$ for all matrices of G_μ.*

Let $\{M_n, n \geq 1\}$ be a sequence of independent random matrices with distribution μ and $\bar{x}, \bar{y} \in P(\mathbb{R}^2)$. Then

$$\lim_{n \to \infty} \delta\left(M_n M_{n-1} ... M_1 \bar{x}, M_n M_{n-1} ... M_1 \bar{y}\right) = 0.$$

Note that when M_1 is invertible almost surely and $\det(M_1)$ is possibly not equal to ± 1, it is enough to verify the above conditions for the group $G_{\tilde{\mu}}$ generated by all $\tilde{M} = (\det M)^{-1/2} M$, where M is any invertible matrix in the support of μ.

Theorem 2.2. *Conjecture 1 is fulfilled for $d = 3$.*

Proof. When $d = 3$ the random matrix T equals

$$T = T(\xi_1, \xi_2, \xi_3) = \begin{pmatrix} 1 - \xi_1 & \xi_2 \\ -\xi_3 & 1 - \xi_2 - \xi_3 \end{pmatrix}$$

where ξ_1, ξ_2, ξ_3 are i.i.d. copies of ξ. Let μ be the probability measure associated with the random matrix $T(\xi_1, \xi_2, \xi_3)$. Observe that $\det(T) = \xi_1 \xi_2 \xi_3 + (1 - \xi_1)(1 - \xi_2)(1 - \xi_3) > 0$ as long as $\xi_1, \xi_2, \xi_3 \in (0, 1)$, thus $\tilde{T} = (\det T)^{-1/2} T$ is a.s. well-defined. Let G_μ be the group generated by all the invertible matrices in the support of μ and $G_{\tilde{\mu}}$ be the group generated by all $\tilde{T}$, where $T \in G_\mu$. Since $\det(\tilde{T}(\xi_1, \xi_2, \xi_3)) = 1$ for all possible ξ_1, ξ_2, ξ_3 and the determinant of a product of two matrices equals the product of their determinants, we have $\det(\tilde{T}) = 1$ for all $\tilde{T} \in G_{\tilde{\mu}}$. Consequently, condition (i) of Proposition 2.1 is fulfilled.

Now let us verify condition (ii), i.e. that the group $G_{\tilde{\mu}}$ is not compact. From Assumption 6 it follows that we can choose $a, b \in \text{supp } \xi$ such that $a, b \in (0, 1)$ and $a \neq b$. Let

$$Q = T(a, b, a) \, T(a, b, b)^{-1} \, T(b, a, b) \, T(b, a, a)^{-1} = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \qquad \text{(A.9)}$$

where

$$t = -\frac{(a - b)^2}{2ab + b^2 - a - 2b + 1}.$$

Since $a \neq b$ and $2ab + b^2 - a - 2b + 1 = (a + b - 1)^2 + a(1 - a) > 0$ the quantity t is well-defined and negative. Observe that $Q \in G_{\tilde{\mu}}$ and hence

$$Q^m = \begin{pmatrix} 1 & 0 \\ mt & 1 \end{pmatrix} \in G_{\tilde{\mu}}$$

as well. Since $||Q^m|| \sim m \to \infty$ as $m \to \infty$, the group $G_{\tilde{\mu}}$ is indeed not compact.

Finally, we need to check the condition (iii) of Theorem 2.1, that is, that $G_{\tilde{\mu}}$ is strongly irreducible, or equivalently that G_μ is strongly irreducible. Suppose the contrary, i.e. there is a union L of one-dimensional subspaces of $\mathbb{R}^2$ such that $T(L) = L$ for any $\tilde{T} \in G_{\tilde{\mu}}$. Let $L = V_1 \cup V_2 \cup \cdots \cup V_k$, $k \geq 1$.

First, suppose that L contains a vector of the form $(x, y)^{\mathbb{T}}$ such that $x \neq 0$. Then at least one of V_i is the linear span of $v = (1, r)^{\mathbb{T}}$, $r \in \mathbb{R}$; without loss of generality let this be V_1. Since Q defined by (A.9) belongs to G_μ, for all $m = 1, 2, \ldots$ we must have $Q^m \in G_\mu$ and thus $Q^m L \subseteq L$. The latter implies that $v_m := Q^m v \in L$. However, the slopes of the vectors v_m equal $mt + r$ which take distinct values for different values of m, therefore L cannot be a union of a *finite* number of linear subspaces, leading to a contradiction.

Therefore, the only candidates for V_i can be linear spaces spanned by $(0, 1)^{\mathbb{T}}$. To show that this is not possible either, pick any $a \in (0, 1)^{\mathbb{T}}$ which is in the support of ξ, then

$$T(a, a, a) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a \\ 1 - 2a \end{pmatrix} \in L.$$

Hence there must be a vector in L whose first coordinate is non-zero, which leads to the situation already considered above.

Consequently, the conditions of the Furstenberg's theorem 2.1 are fulfilled, implying a.s. convergence to flatness in case $d = 3$. □

General case ($d \geq 4$)

We start with a few definitions.

Definition 3.1. *We say that a family $\mathcal{H}$ of $d \times d$ matrices is irreducible in $\mathbb{R}^d$ if there exists no proper linear subspace L of $\mathbb{R}^d$ such that $H(L) = L$ for all $H \in \mathcal{H}$.*

Definition 3.2. *We say that a family $\mathcal{H}$ of $d \times d$ matrices is strongly irreducible in $\mathbb{R}^d$ if there exists no union L of finite number of proper linear subspaces of $\mathbb{R}^d$ such that $H(L) = L$ for all $H \in \mathcal{H}$.*

Definition 3.3. *We say a family $\mathcal{H}$ of $d \times d$ matrices has contraction property if there is a sequence of elements $\{A_n\}_{n \geq 1} \subset \mathcal{H}$ such that $||A_n||^{-1} A_n$ converges to a rank one matrix.*

We will make use of the following

Proposition 3.4 (Theorem III.4.3 in [2], p. 56). *Let A_i be a sequence of i.i.d. random matrices in $GL(d, \mathbb{R})$ with common distribution μ. Let S_μ be the smallest closed semigroup generated by its support. Suppose that $S_\mu \subset GL(d, \mathbb{R})$ is strongly irreducible and contracting. Then for any $\overline{x}, \overline{y} \in P(\mathbb{R}^d)$*

$$\lim_{n \to \infty} \delta(A_n \ldots A_1 \overline{x}, A_n \ldots A_1 \overline{y}) = 0 \text{ a.s.}$$

Note that, when A_1 is only invertible almost surely, it is enough to verify the strong irreducibility and contraction condition for the semigroup $\tilde{S}_\mu$ generated by all $\tilde{A} = (|\det A|)^{-1/d} A$, where A is any invertible matrix in the support of μ. In our case the measure μ corresponds to random matrices of type $T = T(\xi_1, ..., \xi_d)$ defined by (A.6). Observe that

$$\det(T) = \prod_{i=1}^{d}(1 - \xi_i) - (-1)^d \prod_{i=1}^{d} \xi_i.$$

Thus we have $|\det(T)| \leq 2$; also obviously $\det(T) > 0$ almost surely for any odd $d \geq 3$; however, if d is an even number, we need the following invertibility

Assumption 7. *If d is an even number, we assume that*

$$\prod_{i=1}^{d} \frac{1 - \xi_i}{\xi_i} \neq 1$$

almost surely.

The main result of this Section is

Theorem 3.5. *Conjecture 1 is fulfilled for all odd $d \geq 3$, and under Assumption 7 also for all even $d \geq 4$.*

From now on we will suppose that Assumption 7 is in fact fulfilled. As a result, we can always choose $a, b \in \text{supp}(\xi)$ such that $a \neq b, a, b \in (0, 1)$ and $T(a_1, a_2, ..., a_d)$ is invertible for all sequences $a_1, a_2, ..., a_d$ where each $a_i \in \{a, b\}$. Let $\mathcal{S}_{a,b}$ stand for the smallest closed semigroup which contain all of the following matrices

$$|\det T(a_1, a_2, ..., a_d)|^{-1/d} T(a_1, a_2, ..., a_d),$$

with $a_1, a_2, ..., a_d \in \{a, b\}$. We will show that $\mathcal{S}_{a,b} \subseteq S_\mu$ is strongly irreducible and contracting, hence so is S_μ itself. Then the result of Theorem 3.5 will immediately follow from Proposition 1.3 and 3.4, provided we check the condition of the latter statement (and this is done in turn in Propositions 3.8 and 3.12 below).

Irreducibility

Proposition 3.6. *Suppose that Assumptions 6 and 7 hold. Then the family of matrices*

$$\{T(a_1, a_2, ..., a_d)\}_{a_1, a_2, ..., a_d \in \{a, b\}}$$

is irreducible in $\mathbb{R}^{d-1}$.

Proof. Observe that, if W is a real proper invariant subspaces of linear operator A then $\tilde{W} = \{w' + iw'' : w', w'' \in W\}$ is also a complex proper invariant subspaces of A. Thus we can complete the proof by proving the irreducibility in $\mathbb{C}^{d-1}$.

From now on, let us denote

$$T_a = T(a, a, ..., a) \text{ and } T_{a,b;k} = T(a_1, a_2, ..., a_d)|_{a_k = b, \ a_j = a, j \neq k}. \tag{A.10}$$

Note that T_a has eigenvectors given by

$$v_1 = \begin{pmatrix} 1 \\ \epsilon \\ \epsilon^2 \\ \vdots \\ \epsilon^{d-2} \end{pmatrix}, \ v_2 = \begin{pmatrix} 1 \\ \epsilon^2 \\ \epsilon^4 \\ \vdots \\ \epsilon^{2(d-2)} \end{pmatrix}, \ \ldots, v_{d-1} = \begin{pmatrix} 1 \\ \epsilon^{d-1} \\ \epsilon^{(d-1)2} \\ \vdots \\ \epsilon^{(d-1)(d-2)} \end{pmatrix} \tag{A.11}$$

where $\epsilon = e^{2\pi i/d}$ is the d−th root of 1; one can easily conclude that these $d - 1$ eigenvectors are linearly independent in $\mathbb{C}^{d-1}$, and correspond to eigenvalues $\lambda_l = 1 - a + a\epsilon^l$, $l = 1, 2, \ldots, d - 1$ respectively.

Let us prove that all complex proper invariant subspaces of T_a are given by the linear spans of $2^n - 2$ non-trivial subsets of $\{v_1, \ldots, v_{d-1}\}$, and only by them. First of all, suppose $V = \text{span}(v_{k_1}, v_{k_2}, \ldots, v_{k_m})$ where $1 \leq k_1 < k_2 < \cdots < k_m \leq d - 1$ and $m \in \{1, 2, \ldots, d - 1\}$. Since $T_a v_{k_l} = \lambda_{k_l} v_{k_l}$ and $\lambda_{k_l} \neq 0$,

$1 \leq l \leq m$, we conclude that $\text{span}(T_a v_{k_1}, \dots, T_a v_{k_m}) = V$ and hence $T_a(V) = V$ and thus V is indeed invariant.

On the other hand, suppose V is an invariant subspace of T_a, that is $T_a(V) = V$. Since $v_1, \dots, v_{d-1}$ form a basis, any vector $w \in V$ can be written as

$$w = q_1 v_{k_1} + q_2 v_{k_2} + \cdots + q_m v_{k_m}$$

where all $q_l \neq 0$. Since V is an invariant subspace, $T_a w \in V$, consequently

$$w' = q_2' v_{k_2} + \cdots + q_m' v_{k_m} = q_2(\lambda_{k_2} - \lambda_{k_1}) v_{k_2} + \dots q_m(\lambda_{k_m} - \lambda_{k_1}) v_{k_m}$$
$$= T_{a;a} w - \lambda_{k_1} w \in V$$

with all $q_l' \neq 0$ since all λ's are distinct. Continuing this by induction, we will obtain that $v_{k_m} \in V$, and hence $v_{k_{m-1}} \in V, \dots, v_{k_1} \in V$. Therefore, V contains all those v_k for which the projection of some vector $w \in V$ on v_k has a non-zero coefficient. At the same time the span of all these v_k will contain all those vectors w, hence V is the span of a subset of $\{v_1, \dots, v_{d-1}\}$.

Next we will show that at the same time no proper invariant subspace $V = \text{span}(v_{k_1}, v_{k_2}, \dots, v_{k_m})$ of T_a can be also an invariant subspace of $T_{a,b;k}$, $k = 1, 2, \dots, d$. First, define the sequence of vectors $u_1, \dots, u_d \in \mathbb{R}^{d-1}$ by

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ u_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ u_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \ \dots, \ u_d = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ -1 \\ 1 \end{pmatrix}. \tag{A.12}$$

We must have $T_{a,b;1} v_r \in V$ for all $r \in \{k_1, k_2, \dots, k_m\}$, hence

$$(a - b)u_1 = T_{a,b;1} v_r - \lambda_r v_r \in V$$

Now, by using the fact that

$$(T_{a,b;k} - T_a)v_r = (a - b)e^{r(k-1)} u_k \in V$$

for $k = 1, 2, \dots, d$ we obtain that $u_1, u_2, \dots, u_d \in V$. Note that $u_2, u_3, \dots, u_d$ are linearly independent, hence $V = \text{span}(u_2, \dots, u_d) \equiv \mathbb{R}^{d-1}$. This contradiction completes the proof. $\qquad \square$

Strong irreducibility

We already know from Proposition 3.6 that $\mathcal{S}_{a,b}$ is irreducible. Now we aim to show its *strong* irreducibility.

Lemma 3.7. *If $\mathcal{S}_{a,b}$ is irreducible but not strongly irreducible in $\mathbb{R}^{d-1}$, then there exist proper linear subspaces $V_1, V_2, ..., V_r$ of $\mathbb{R}^{d-1}$ such that*

$$\mathbb{R}^{d-1} = \bigoplus_{j=1}^{r} V_j \text{ where } r > 1, V_i \cap V_j = \{0\} \text{ if } i \neq j,$$

where all the subspaces V_j have the same dimension, and

$$M(\cup_{j=1}^{r} V_j) = \cup_{j=1}^{r} V_j,$$

for all $M \in \mathcal{S}_{a,b}$.

Proof of Lemma 3.7. See the remark and the equation (2.7) on pp. 121–122 of [6]. $\square$

Proposition 3.8. *Suppose that Assumptions 6 and 7 hold. Then the semigroup $\mathcal{S}_{a,b}$ is strongly irreducible.*

Proof. For a real linear space $W \subset \mathbb{R}^{d-1}$, we define

$$\tilde{W} = \{w' + iw'', w', w'' \in W\} \subset \mathbb{C}^{d-1},$$

which is also a complex linear subspace of $\mathbb{C}^{d-1}$.

We already know that the semigroup $\mathcal{S}_{a,b}$ is irreducible in $\mathbb{R}^{d-1}$. Suppose $\mathcal{S}_{a,b}$ is not strongly irreducible in $\mathbb{R}^{d-1}$. Then it implies from Lemma 3.7 that there exist proper linear space $V_1, V_2, ..., V_r \subset \mathbb{R}^{d-1}$ such that

$$\mathbb{C}^{d-1} = \bigoplus_{j=1}^{r} \tilde{V}_j,$$

where $\tilde{V}_j$ are disjoint linear subspaces of the same dimension, say m, and

$$M(\cup_{j=1}^{r} \tilde{V}_j) = \cup_{j=1}^{r} \tilde{V}_j,$$

for all $M \in \mathcal{S}_{a,b}$.

The rest of the proof is organized as follows. First, we show irreducibility in the case $m > 1$. The case when $m = 1$ is split further in the sub-cases including the one where $k = 2$ and $k \geq 3$, and yet further sub-sub-case where $k = 4$.

Observe also that from Lemma III.4.5.b in [2] it follows that for each $j \in \{1, 2, \ldots, d-1\}$, we have $T_a \tilde{V}_j = \tilde{V}_k$ for some $k = k(j)$. Suppose $k(j) \neq j$ for all j. Let $e_1, \ldots, e_{d-1}$ be the basis $\mathbb{C}^{d-1}$ such that $e_1, \ldots, e_m$ is the basis of V_1, $e_{m+1}, \ldots, e_{m+m}$ is the basis of V_2, etc. In this basis T_a will be a traceless matrix since all the V_j are disjoint. The property of being traceless is invariant with respect to changing the basis as $\mathrm{tr}(PAP^{-1}) = \mathrm{tr}(A)$. However, in the original basis $\mathrm{tr}(T_a) = (1-a)(d-1) - a \neq 0$ unless $a = \frac{d-1}{d}$, but in this case we can replace a by $b \neq a$, so we get a contradiction.

Thus we have established that $k(j) = j$ for some j; w.l.o.g. let us assume that $j = 1$ and consequently $T_a V_1 = V_1$. From the arguments in Proposition 3.6 we know that V_1 is a linear span of some subset of v_k's from (A.11), that is $V_1 = \mathrm{span}\{w_1, \ldots, w_m\}$ where $w_j = v_{r_j}$, for some subset $\{r_1, \ldots, r_m\} \subset \{1, 2, \ldots, d-1\}$. By denoting $\epsilon_j := \epsilon^{r_j}$, some d-th root of 1, we get that $w_j = \left(1, \epsilon_j, \ldots, \epsilon_j^{d-2}\right)^{\mathbb{T}}$. Let u_k be defined as in (A.12). Then

$$T_{a,b;k} w_j = \lambda_{r_j} w_j + (a-b)\epsilon_j^{k-1} u_k.$$

For every k, we must have $T_{a,b;k} V_1 = V_j$ for some $j = j(k)$. Now suppose that there is no k such that $T_{a,b;k} V_1 = V_1$. Recall that $V_1 = \mathrm{span}(w_1, \ldots, w_m)$. Let

$$V_k' = T_{a,b;k} V_1 = \mathrm{span}(\{\lambda_{r_j} w_j + c_k u_k, \ j = 1, \ldots, m\})$$

where $c_k = (a-b)\epsilon_j^{k-1} \neq 0$ for $k = 1, 2, \ldots, d-1-m$. Observe that at the same time $V_k' = V_q$ for some $q = q(k)$, so that the collection V_k', $k = 1, \ldots, d-1-m$, is some subset of $V_1, \ldots, V_r$, possibly with repetitions.

Let us show that $w_1, \ldots, w_m, u_1, \ldots, u_{d-1-m}$ are linearly independent. Indeed, to establish the rank of the matrix of $d-1$ vectors $w_1, \ldots, w_m, u_1, u_2, \ldots, u_{d-1-m}$ observe that

$$
\det \begin{pmatrix}
1 & 1 & \ldots & 1 & 1 & -1 & 0 & \ldots & 0 \\
\epsilon_1 & \epsilon_2 & \ldots & \epsilon_m & 0 & 1 & -1 & \ldots & 0 \\
\epsilon_1^2 & \epsilon_2^2 & \ldots & \epsilon_m^2 & 0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\epsilon_1^{d-m-2} & \epsilon_2^{d-m-2} & \ldots & \epsilon_m^{d-m-2} & 0 & 0 & 0 & \ldots & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\epsilon_1^{d-2} & \epsilon_2^{d-2} & \ldots & \epsilon_m^{d-2} & 0 & 0 & 0 & \ldots & 0
\end{pmatrix}
$$

$$
= \det \begin{pmatrix}
\epsilon_1^{d-m-1} & \epsilon_2^{d-m-1} & \ldots & \epsilon_m^{d-m-1} \\
\vdots & \vdots & \ddots & \vdots \\
\epsilon_1^{d-2} & \epsilon_2^{d-2} & \ldots & \epsilon_m^{d-2}
\end{pmatrix}
$$

$$
= \epsilon_1^{d-m-1} \cdot \cdots \cdot \epsilon_m^{d-m-1} \cdot \det \begin{pmatrix}
1 & 1 & \ldots & 1 \\
\vdots & \vdots & \ddots & \vdots \\
\epsilon_1^{m-1} & \epsilon_2^{m-1} & \ldots & \epsilon_m^{m-1}
\end{pmatrix}
$$

$$
= \prod_{j=1}^{m} \epsilon_j^{d-m-1} \cdot \prod_{1 \leq j < k \leq m} (\epsilon_j - \epsilon_k) \neq 0
$$

since this is a Vandermonde matrix. This, in turn, implies that the subspaces $V_1, V_1', V_2', \ldots, V_{d-m-1}'$ are all pairwise distinct; otherwise there would be a vector which at the same time belongs to $\mathrm{span}(\{\lambda_{r_j} w_j + c_k u_k, \ j = 1, \ldots, m\})$ and $\mathrm{span}(\{\lambda_{r_j} w_j + c_l u_l, \ j = 1, \ldots, m\})$ for $k \neq l$, yielding linear dependence for the set $w_1, \ldots, w_m, u_k, u_l$ which is impossible.

On the other hand, it implies that the dimension of $V_1 \oplus V_1' \oplus \cdots \oplus V_{d-m}'$ is $m \times (d-1-m) > d-1$ unless $m = 1$, yielding a contradiction that this is a subspace of $\mathbb{R}^{d-1}$.

Thus now we have to deal only with the case $m = 1$. In this case, all the spaces $V_1, V_2, \ldots, V_{d-1}$ are one-dimensional, moreover, by letting $v = \epsilon_1$

$$
\begin{aligned}
& w_1 = (1, v, \ldots, v^{d-2})^{\mathbb{T}}, \\
& V_1 = \mathrm{span}(w_1), \\
& V_k' := T_{a,b;k} V_1 = \mathrm{span}(\lambda_{r_1} w_1 + c_k u_k), \ k = 1, 2, \ldots, d-1,
\end{aligned}
$$

and V_k's are some subset of $V_2, \ldots, V_{d-1}$ (if $V_k' = V_1$ for some k then $u_k \in \mathrm{span}(w_1)$ which is impossible for $d \geq 4$). If all the elements of the set V_1, V_1',

..., V'_{d-1} are distinct (we know that then they must be linearly independent since $\mathbb{R}^{d-1} = V_1 \oplus V_2 \oplus \cdots \oplus V_{d-1}$) this would yield a contradiction as our space is only $(d-1)$-dimensional.

Observe that

$$\det(w_1, u_2, u_3, \ldots, u_{d-1}) = \det \begin{pmatrix} 1 & -1 & 0 & \ldots & 0 \\ v & 1 & -1 & \ldots & 0 \\ v^2 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v^{d-2} & 0 & 0 & \ldots & -1 \\ v^{d-1} & 0 & 0 & \ldots & 1 \end{pmatrix}$$

$$= 1 + v + \cdots + v^{d-2} = \frac{1 - v^{d-1}}{1 - v} = \frac{-1}{v} \neq 0$$

since $v^d \equiv \epsilon_1^d = 1$. This implies that the vectors $w_1, u_2, u_3, \ldots, u_{d-1}$ are linearly independent and hence it is impossible that $V'_k = V'_h$ for some $k, h \in \{2, \ldots, d-1\}$ such that $k \neq h$.

So the only not covered case is when V'_1 coincides with some $V'_k, k = 2, \ldots, d-1$, implying a linear dependence between w_1, u_1 and u_k. However, if $k = 2$, then

$$\mathrm{rank}(w_1, u_1, u_k) = \mathrm{rank} \begin{pmatrix} 1 & v & v^2 & \ldots & v^{d-2} \\ 1 & 0 & 0 & \ldots & 0 \\ -1 & 1 & 0 & \ldots & 0 \end{pmatrix}$$

$$= 1 + \mathrm{rank} \begin{pmatrix} v & v^2 & \ldots & v^{d-2} \\ 1 & 0 & \ldots & 0 \end{pmatrix} = 3$$

since $v^2 \neq 0$. Finally, if $k \geq 3$, then

$$\mathrm{rank}(w_1, u_1, u_k) = \mathrm{rank} \begin{pmatrix} 1 & v & \ldots & v^{k-2} & v^{k-1} & v^k & \ldots & v^{d-2} \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & -1 & 1 & 0 & \ldots & 0 \end{pmatrix}$$

$$= 1 + \mathrm{rank} \begin{pmatrix} v & \ldots & v^{k-2} & v^{k-1} & v^k & \ldots & v^{d-2} \\ 0 & \ldots & -1 & 1 & 0 & \ldots & 0 \end{pmatrix} = 3$$

unless simultaneously $d = 4$, $k = 3$ and $v = \epsilon_1 = -1$.

Finally, to deal with the case $d = 4$ and $\epsilon_1 = -1$, observe that

$$T(\xi_1, \xi_2, \xi_3, \xi_4) = \begin{pmatrix} 1 - \xi_1 & \xi_2 & 0 \\ 0 & 1 - \xi_2 & \xi_3 \\ -\xi_4 & -\xi_4 & 1 - \xi_3 - \xi_4 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = e_1 - e_2 + e_3$$

where e_1, e_2, e_3 are the standard basis vectors for $\mathbb{R}^3$. Let us consider

$$\begin{aligned}
w_1^* &:= T(a, a, b, a)\, w_1 = (1 - b - a)w_1 + (b - a)e_1, \\
w_2^* &:= T(a, b, b, a)\, w_1 = (1 - b - a)w_1 + (b - a)e_2, \\
w_3^* &:= T(a, b, a, a)\, w_1 = (1 - b - a)w_1 + (b - a)e_3.
\end{aligned}$$

Then, in the standard Euclidean coordinates,

$$A := [w_1^*, w_2^*, w_3^*] = \begin{pmatrix} 1 - 2a & b + a - 1 & 1 - b - a \\ 1 - b - a & 2b - 1 & 1 - b - a \\ 1 - b - a & b + a - 1 & 1 - 2a \end{pmatrix},$$

and

$$\det(A) = (b - a)^2 (1 - 2a).$$

From Assumptions 6 and 7 it follows that w.l.o.g. we can chooses a and b such that $a \neq 1/2$, $a \neq b$, and $a + b \neq 1$, implying that the above determinant is non-zero. Thus we obtain that the three subspaces span by w_1^*, w_2^*, w_3^* are linearly independent in $\mathbb{R}^3$ again yielding a contradiction. $\qquad \square$

Contracting property

Here we need to show that the semigroup $\mathcal{S}_{a,b}$ is strongly irreducible and contracting. While in general it is not easy to verify the contraction property of a semigroup, thanks to the following important statement by Goldsheid and Margulis in [7], it suffices to check this property for the Zariski closure of $\mathcal{S}_{a,b}$ (which is easier).

Definition 3.9. *Zariski closure of a subset H of an algebraic manifold is the smallest algebraic submanifold that contains H.*

Proposition 3.10 (Lemma 3.3 in [7]). *The Zariski closure $\mathrm{Zr}(H)$ of a closed semigroup of $H \subset GL(d, \mathbb{R})$ is a group.*

Proposition 3.11 (Lemma 6.3 in [7]). *If a closed semigroup $H \subset GL(d, \mathbb{R})$ is strongly irreducible and its Zariski closure $\mathrm{Zr}(H)$ has the contraction property then H also has the contraction property.*

Proposition 3.12. *Suppose that Assumptions 6 and 7 hold. Then the semigroup $\mathcal{S}_{a,b}$ is contracting.*

Proof. According to Proposition 3.11 it is sufficient to show that $\mathrm{Zr}(\mathcal{S}_{a,b})$ is contracting, since we have already established that $\mathcal{S}_{a,b}$ and hence $\mathrm{Zr}(\mathcal{S}_{a,b})$ is strongly irreducible by Proposition 3.8. Note that $T^{-1} \in \mathrm{Zr}(\mathcal{S}_{a,b})$ for any $T \in \mathcal{S}_{a,b}$, since the Zariski closure is necessary a group by Proposition 3.10. We consider two separate cases.

Case $d = 2l + 1$ is odd. Define

$$M = T(a, b, \dots, a, b, \mathbf{a})\, T(a, b, \dots, a, b, \mathbf{b})^{-1} T(b, a, \dots, b, a, \mathbf{b})\, T(b, a, \dots, b, a, \mathbf{a})^{-1}$$
$$\in \mathrm{Zr}(\mathcal{S}_{a,b})$$

After some algebraic computations, one can obtain that

$$M = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ \varphi_1 & \varphi_2 & \dots & \varphi_{2l-1} & 1 \end{pmatrix},$$

where

$$\varphi_{2j-1} = -\frac{(a-b)^2 \left((1-a)(1-b)\right)^{l-j} (ab)^{j-1}}{(1-a)^l (1-b)^{l+1} + a^l b^{l+1}}, \quad \text{and } \varphi_{2j} = 0, \ \ j = 1, 2, \dots, l.$$

Hence

$$M^n = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ n\varphi_1 & n\varphi_2 & \dots & n\varphi_{2l-1} & 1 \end{pmatrix} \in \mathrm{Zr}(\mathcal{S}_{a,b}).$$

It implies that $\|M^n\| \approx Const \cdot n$ hence $\|M^n\|^{-1} M^n$ converges to a matrix whose first $d - 2$ rows are zero rows, and thus $\mathrm{Zr}(\mathcal{S}_{a,b})$ is contracting by definition.

Case $d = 2l$ is even. Define

$$M = T(a, a, \ldots, a, a, a, \mathbf{a}) \, T(a, a, \ldots, a, a, a, \mathbf{b})^{-1}$$

$$= \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ c_1 & c_2 & \ldots & c_{d-1} & c(a,b) \end{pmatrix}$$

where $c_1 = c_1(a,b), \ldots, c_{d-1} = c_{d-1}(a,b)$ are some constants depending on a and b, and $c(a,b) = \det T(a, \ldots, a, a) / \det T(a, \ldots, a, b)$; observe also that

$$\det T(a, \ldots, a, a) = (1-a)^d - a^d$$
$$\det T(a, \ldots, a, b) = (1-a)^d - a^d + (a-b)[(1-a)^{d-1} + a^{d-1}]$$

Assume initially that $|c(a,b)| > 1$, then

$$M^n = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \\ A_n c_1 & A_n c_2 & \ldots & A_n c_{n-1} & c(a,b)^n \end{pmatrix}$$

where $A_n = 1 + c(a,b) + c(a,b)^2 + \ldots + c(a,b)^{n-1}$, so that $||M^n|| \geq \text{const} \times c(a,b)^n \to \infty$ and thus $||M^n||^{-1} M^n$ converges to a matrix whose first $d-2$ rows are zeros. If $|c(a,b)| < 1$ then we can consider M^{-1} instead of M, which has the form

$$M^{-1} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \ldots & c(a,b)^{-1} \end{pmatrix} \in \mathsf{Zr}(\mathcal{S}_{a,b})$$

and then apply exactly the same arguments as when $|c(a,b)| > 1$. Note that $c(a,b) \neq 1$ since $a \neq b$, so we only have to consider the case when $c(a,b) = -1$.

We have $c(a,b) \neq c(b,a)$ since $a \neq b$. Hence, w.l.o.g. we can assume that $c(a,b) \neq -1$. So in all the cases, either $||M^n||^{-1} M^n$ or $||M^{-n}||^{-1} M^{-n}$ converges to a rank one matrix as $n \to \infty$. $\qquad \square$

Convergence rate of random polygons

Convergence rate of rescaled polygons to flatness

Throughout the rest of the paper we use the notation $\log^+(x) := \max\{\log x, 0\}$.

Let $\ell(T) = \max(\log^+(||T||), \log^+(||T^{-1}||))$. In this section, we suppose that Assumptions 6 and 7 as well as the following condition hold

$$\mathbb{E}\,\ell(T) < \infty. \tag{A.13}$$

Let $T_1, T_2, ...$ be a sequence of random matrices having the same distribution as T. We define Lyapunov exponents

$$\mu_j = \lim_{n\to\infty} \mathbb{E}\left(\frac{1}{n}\log\,\sigma_j^{(n)}\right), \quad j = 1, 2, ..., d-1 \tag{A.14}$$

where $\sigma_1^{(n)} \geq \sigma_2^{(n)}... \geq \sigma_{d-1}^{(n)}$ are the singular values of $T^{(n)} = T_n T_{n-1}\ldots T_1$, i.e., the square roots of the eigenvalues of $\left(T^{(n)}\right)^{\mathbb{T}} T^{(n)}$. Therefore, from the proof of Proposition III.6.4 in [2] (pp. 67–68), for any $x, y \in P(\mathbb{R}^{d-1})$

$$\lim_{n\to\infty}\frac{1}{n}\log\delta(T^{(n)}x, T^{(n)}y) \leq \mu_2 - \mu_1 < 0 \quad \text{a.s.} \tag{A.15}$$

Lemma 4.1. *Let* $\xi_1, \xi_2, \ldots, \xi_d \in [0,1]$. *Then*

$$\prod_{i=1}^{d}\xi_i(1-\xi_i) \leq \xi_1\xi_2\ldots\xi_d + (1-\xi_1)(1-\xi_2)\ldots(1-\xi_d) \leq 1.$$

Proof. The upper bound follows from the fact that it is equal to probability to get either all heads or all tails in an experiment with throwing d independent coins each with probability to turn up head equal to ξ_i, $i = 1, 2, \ldots, d$. To get the lower bound observe that for $d = 1, 2, \ldots$ we have

$$\prod_{i=1}^{d}\xi_i + \prod_{i=1}^{d}(1-\xi_i) \geq \left[\prod_{i=1}^{d-1}\xi_i + \prod_{i=1}^{d-1}(1-\xi_i)\right]\cdot\xi_d(1-\xi_d)$$

and since the statement is true for $d = 1$, we have proved the proposition. □

As it is implied from the following proposition, we can reformulate the requirement (A.13) as

Assumption 8.

$$\mathbb{E}\log\left(|\det(T)|\right) = \mathbb{E}\log\left|\prod_{i=1}^{d}(1-\xi_i) - (-1)^d\prod_{i=1}^{d}\xi_i\right| > -\infty.$$

Proposition 4.2. *Condition* (A.13) *holds if and only if Assumption 8 is fulfilled.*

Proof. Noticing that all the elements of T are bounded, and using the formula for inversion of matrices we obtain that

$$||T|| \le C_1, \quad ||T^{-1}|| \le \frac{C_2}{|\det(T)|} \tag{A.16}$$

where C_i, $i = 1, 2, \dots$ here and further in the text denote some non-random positive constants. Let $\sigma_1 \ge \sigma_2 \ge \dots \sigma_{d-1} > 0$ be the singular values of matrix T, that is, the square roots of the eigenvalues of $T^{\mathsf{T}}T$, arranged in the decreasing order. Then $||T^{-1}|| = 1/\sigma_{d-1}$. On the other hand, using the fact that there is a unitary matrix U such that $U^{\mathsf{T}}(T^{\mathsf{T}}T)U$ is a diagonal matrix with elements σ_i^2, we obtain that

$$\det(T) = \sigma_1\sigma_2\dots\sigma_{d-1} \ge (\sigma_{d-1})^{d-1}$$

so that

$$||T^{-1}|| = \frac{1}{\sigma_{d-1}} \ge \frac{1}{|\det(T)|^{\frac{1}{d-1}}}.$$

On the other hand it is easy that

$$\det(T) = \prod_{i=1}^{d}(1-\xi_i) - (-1)^d\prod_{i=1}^{d}\xi_i$$

which is always non-negative for odd d, but can be positive as well as negative for even d; in both cases $|\det(T)| \le 1$, as it easily follows from Lemma 4.1. Consequently,

$$\log^+\left(||T^{-1}||\right) \le \log^+\left(\frac{C_2}{|\det(T)|}\right) \le \log^+\left(\frac{C_2+1}{|\det(T)|}\right)$$
$$\le \log\left(\frac{1}{|\det(T)|}\right) = -\log\left(|\det(T)|\right),$$
$$\log^+\left(||T^{-1}||\right) \ge \log^+\left(\frac{1}{|\det(T)|^{\frac{1}{d-1}}}\right) \ge -\frac{1}{d-1}\log\left(|\det(T)|\right).$$

Since $\log^+||T||$ is bounded above by some constant, the statement of the proposition follows. $\qquad\square$

Notice that since

$$\mu_1 + \mu_2 + ... + \mu_{d-1} = \mathbb{E}(\log|\det(T)|) \qquad (A.17)$$

all Lyapunov exponents μ_j, $j = 1, 2, ..., d - 1$ are finite if and only if Assumption 8 is fulfilled. Therefore, using (A.15), we can deduce the following

Theorem 4.3. *Suppose that Assumptions 6, 7 and 8 hold. Then the sequence of polygons L_n converges to flatness with at least exponential rate with parameter $\mu = \mu_1 - \mu_2 \in (0, \infty)$*

Now let us give an "easier" sufficient condition for Assumption 8 which depends only on the distribution of one ξ.

Proposition 4.4. *Suppose that $d = 3, 5, \ldots$ is odd. If $\mathbb{E}|\log \xi| < \infty$ and $\mathbb{E}|\log(1 - \xi)| < \infty$ then Assumption 8 is fulfilled. A sufficient condition for these expectations to be finite is*

$$\limsup_{v\downarrow 0} \frac{\mathbb{P}(\xi < v)}{v^\alpha} < \infty \text{ and } \limsup_{v\uparrow 1} \frac{\mathbb{P}(\xi > v)}{(1 - v)^\alpha} < \infty \qquad (A.18)$$

for some $\alpha > 0$.

Remark 4.5. *Note that when d is even we would not be able to bound $|\det(T)|$ from below by the products of $\xi_i(1 - \xi_i)$ as easily as it is done in the following proof. Indeed, if we let all $\xi_i = 1/2$ then $\det(T) = 0$ while all $\xi_i(1 - \xi_i) = 1/4 > 0$.*

Proof of Proposition 4.4. The first part of the statement follows immediately from Lemma 4.1 since

$$\mathbb{E}\log|\det(T)| = \mathbb{E}\log\left[\prod_{i=1}^d \xi_i + \prod_{i=1}^d (1 - \xi_i)\right]$$

$$\geq \mathbb{E}\log\left[\prod_{i=1}^d \xi(1 - \xi_i)\right] = \sum_{i=1}^d \left(\mathbb{E}\log \xi_i + \mathbb{E}\log(1 - \xi_i)\right).$$

To prove the second part, note that

$$\mathbb{E}|\log \xi| \leq 1 + \mathbb{E}\left[|\log \xi| \cdot 1_{\xi<e^{-1}}\right] = 1 + \int_0^\infty \mathbb{P}\left(-(\log \xi) \cdot 1_{\xi<e^{-1}} > u\right) \mathrm{d}u$$

$$= 1 + \int_0^1 \cdots + \int_1^\infty \ldots$$

$$= 1 + \int_0^1 \mathbb{P}\left(e\xi < 1\right) \mathrm{d}u + \int_1^\infty \mathbb{P}\left(-\log \xi > u\right) \mathrm{d}u$$

$$= 1 + \mathbb{P}\left(e\xi < 1\right) + \int_0^{e^{-1}} \frac{\mathbb{P}\left(\xi < v\right)}{v} \mathrm{d}v < \infty$$

since

$$\frac{\mathbb{P}\left(\xi < v\right)}{v} \leq \frac{\text{const}}{v^{1+\alpha}}$$

for sufficiently small v. The expectation $\mathbb{E}\left|\log(1 - \xi)\right|$ is bounded in exactly the same way. $\qquad\square$

An interesting example is when ξ has a uniform distribution, as in the paper [11].

Proposition 4.6. *If the distribution of ξ is uniform on $[0,1]$ then Assumption 8 is fulfilled for all $d \geq 3$.*

Proof. The case when d is odd immediately follows from Proposition 4.4 so we assume that d is even. We have

$$\mathbb{E}\log|\det T| = \int_0^1 \cdots \int_0^1 \log|(1 - x_1)\dots(1 - x_d) - x_1\dots x_d|\; \mathrm{d}x_1 \dots \mathrm{d}x_d$$

$$= \int_0^1 \cdots \int_0^1 \log\left(x_1\dots x_d\right)\; \mathrm{d}x_1 \dots \mathrm{d}x_d$$

$$+ \int_0^1 \cdots \int_0^1 \log\left|1 - \frac{1 - x_1}{x_1} \dots \frac{1 - x_d}{x_d}\right|\; \mathrm{d}x_1 \dots \mathrm{d}x_d$$

$$= -d + \int_0^\infty \cdots \int_0^\infty \frac{\log|1 - u_1\dots u_d|}{(1 + u_1)^2\dots(1 + u_d)^2}\; \mathrm{d}u_1 \dots \mathrm{d}u_d$$

$$= -d + \int_0^\infty \cdots \int_0^\infty \frac{u_1\dots u_{d-1}}{(1 + u_1)^2\dots(1 + u_{d-1})^2}$$

$$\times \left(\int_0^\infty \frac{\log|1 - v|\;\mathrm{d}v}{(u_1\dots u_{d-1} + v)^2}\right)\mathrm{d}u_1 \dots \mathrm{d}u_{d-1}$$

where the inner integral

$$\int_0^\infty \frac{\log|1 - v|\;\mathrm{d}v}{(u_1\dots u_{d-1} + v)^2} = \left(\int_0^{1/2} + \int_{1/2}^{3/2} + \int_{3/2}^2 + \int_2^\infty\right)\frac{\log|1 - v|}{(u_1\dots u_{d-1} + v)^2}\;\mathrm{d}v$$

$$\geq \int_0^{1/2} \frac{-\log 2}{(u_1\dots u_{d-1} + v)^2}\;\mathrm{d}v + \int_{1/2}^{3/2} \frac{\log|1 - v|}{(u_1\dots u_{d-1} + 1/2)^2}\;\mathrm{d}v$$

$$+ \int_{3/2}^2 \frac{-\log 2}{(u_1\dots u_{d-1} + v)^2}\;\mathrm{d}v + 0$$

$$\geq \int_0^\infty \frac{-\log 2}{(u_1\dots u_{d-1} + v)^2}\;\mathrm{d}v + \int_{1/2}^{3/2} \frac{\log|1 - v|}{(u_1\dots u_{d-1} + 1/2)^2}\;\mathrm{d}v$$

$$= -\frac{\log 2}{u_1\dots u_d} + -\frac{1 + \log 2}{(u_1\dots u_{d-1} + 1/2)^2}.$$

Consequently,

$$\mathbb{E} \log | \det T |$$

$$\geq -d - \log 2 \int_0^\infty \cdots \int_0^\infty \frac{du_1 \ldots du_{d-1}}{(1+u_1)^2 \ldots (1+u_{d-1})^2}$$

$$- \int_0^\infty \cdots \int_0^\infty \frac{(1+\log 2)[u_1 \ldots u_{d-1}] \, du_1 \ldots du_{d-1}}{(1+u_1)^2 \ldots (1+u_{d-1})^2 (1/2 + [u_1 \ldots u_{d-1}])^2}$$

$$\geq -d - \left[\log 2 + \frac{1+\log 2}{2} \right] \int_0^\infty \cdots \int_0^\infty \frac{du_1 \ldots du_{d-1}}{(1+u_1)^2 \ldots (1+u_{d-1})^2}$$

$$> -\infty$$

since $a/(1/2+a)^2 \leq 1/2$ for $a \geq 0$. $\qquad\square$

The next statement shows that there are, in fact, examples of distributions for which Assumption 8 is *not fulfilled*.

Proposition 4.7. *Suppose ξ_i are i.i.d. with density*

$$f(x) = \begin{cases} \dfrac{c}{x \log^{1+\delta} x}, & 0 < x \leq 1/2; \\ \dfrac{c}{(1-x) \log^{1+\delta}(1-x)}, & 1/2 < x < 1; \\ 0, & \text{otherwise} \end{cases}$$

where $\delta \in (0, 1/2]$ and $c = c(\delta) \in (0, \infty)$ is the appropriate constant. Then Assumption 8 is <u>not</u> satisfied.

Proof. Assuming d is odd and noticing that $f(1-y) = f(y)$ and that

$$x_1 \ldots x_d + (1-x_1) \ldots (1-x_d) \leq 1,$$

by Lemma 4.1 we have

$$\mathbb{E}\log|\det T|$$

$$= \int_0^1 \cdots \int_0^1 \log\left(x_1 \ldots x_d + (1-x_1)\ldots(1-x_d)\right) f(x_1)\ldots f(x_d)\, dx_1 \ldots dx_d$$

$$\leq \int_0^1 \cdots \int_0^1 \log\left(x_1 x_2 + (1-x_1)(1-x_2)\right) f(x_1)\ldots f(x_d)\, dx_1 \ldots dx_d$$

$$= \int_0^1 \int_0^1 \log(x(1-y) + y(1-x)) f(x) f(y)\, dx\, dy$$

$$\leq \int_0^{1/2} \int_0^{1/2} \log(x + y - xy) f(x) f(y)\, dx\, dy$$

$$\leq \int_0^{1/2} \int_0^{1/2} \log(x + y) f(x) f(y)\, dx\, dy$$

$$= \int_0^{1/2} \int_0^{1/2} \frac{\log(x+y)}{(x\log^{1+\delta} x)(y\log^{1+\delta} y)}\, dx\, dy$$

$$= \int_{\log 2}^{\infty} \int_{\log 2}^{\infty} \frac{\log(e^{-u} + e^{-v})}{u^{1+\delta} v^{1+\delta}}\, du\, dv = 2\int_{\log 2}^{\infty} \int_{\log 2}^{\infty} \frac{\log(e^{-u} + e^{-v})}{u^{1+\delta} v^{1+\delta}} 1_{u>v}\, du\, dv$$

$$\leq 2\int_{\log 2}^{\infty} \left(\int_{\log 2}^{\infty} \frac{\log(2e^{-v})}{u^{1+\delta} v^{1+\delta}} 1_{u>v}\, du \right) dv = \frac{2}{\delta} \int_{\log 2}^{\infty} \frac{\log(2) - v}{v^{1+2\delta}}\, dv = -\infty$$

since $\delta \leq 1/2$. The case when d is even can handled similarly. $\qquad\square$

The next statement shows how quickly the lengths of the largest side of the polygon converge to zero.

Lemma 4.8. *Suppose that Assumptions 6, 7, 8 are fulfilled. Let*

$$M_n = \max_{j=1,\ldots,d} \|l_j^{(n)}\| \tag{A.19}$$

be the length of the largest side of L_n. Then

$$\lim_{n\to\infty} \frac{1}{n} \log(M_n) = \mu_1 \quad a.s.$$

Proof. First of all, observe that by the triangle inequality

$$\max_{j=1,\ldots,d-1} \|l_j\| \leq \max_{j=1,\ldots,d} \|l_j\| \leq \max\left\{ \|l_1\| + \cdots + \|l_{d-1}\|, \max_{j=1,\ldots,d-1} \|l_j\| \right\}$$

$$\leq (d-1) \max_{j=1,\ldots,d-1} \|l_j\|$$

so it suffices to prove the statement of the lemma for the first $d-1$ sides of L_n, i.e., we can redefine just inside of this proof M_n as $\max\{\|l_1^{(n)}\|, \|l_2^{(n)}\|, \ldots, \|l_{d-1}^{(n)}\|\}$. Also, to avoid confusion, in this proof we will denote by $\|\cdot\|_{(k)}$ the Euclidean norm in $\mathbb{R}^k$, while $\|\cdot\|$ is just a Euclidean norm in $\mathbb{R}^2$. By applying Theorem III.7.2.i (pp. 72) in [2], we obtain

$$\lim_{n\to\infty} \frac{1}{n} \log \|T^{(n)}x\|_{(d-1)} = \lim_{n\to\infty} \frac{1}{n} \log \|T_n...T_2 T_1 x\|_{(d-1)} = \mu_1 \qquad \text{(A.20)}$$

for each $x \in \mathbb{R}^{d-1} \setminus \{0\}$. Now recall that the coordinates of $l_j^{(n)} \in \mathbb{R}^2$ are the j-th coordinates of $x^{(n)} = T^{(n)}x^{(0)}$ and $y^{(n)} = T^{(n)}y^{(0)}$ respectively. Omitting the superscript $^{(n)}$, we have

$$\|l_j\|^2 = x_j^2 + y_j^2, \quad \|x\|_{(d-1)}^2 = x_1^2 + \cdots + x_{d-1}^2, \quad \|y\|_{(d-1)}^2 = y_1^2 + \cdots + y_{d-1}^2,$$

so

$$\frac{\|x\|_{(d-1)}^2}{d-1} \le \max_{j=1,\ldots,d-1} x_j^2 \le \max_{j=1,\ldots,d-1} \|l_j\|^2 \le x_1^2 + \cdots + x_{d-1}^2 + y_1^2 + \cdots + y_{d-1}^2$$
$$= \|x\|_{(d-1)}^2 + \|y\|_{(d-1)}^2.$$

Together with (A.20) this immediately implies

$$\limsup_{n\to\infty} \frac{1}{n} \log \max\{\|l_1^{(n)}\|, \|l_2^{(n)}\|, \ldots, \|l_{d-1}^{(n)}\|\} = \mu_1.$$

$\square$

Convergence rate of polygon vertices

The purpose of this Section is to calculate the exact speed of convergence of (not rescaled) polygons L_n to a random point in the plane in the general case $d \ge 3$, under some conditions.

Let $(a_j^{(n)}, b_j^{(n)})$, $j = 1, 2, ..., d$, be the Cartesian coordinates of vertices $A_d^{(n)}$, $A_1^{(n)}$, $A_2^{(n)}, ..., A_{d-1}^{(n)}$ respectively – please note the unusual enumeration of the coordinates, which we do in order to use the same notation for matrix H given by (A.2). We have the following linear relation

$$a^{(n)} = H_{n+1}^{\mathbb{T}} a^{(n-1)}, b^{(n)} = H_{n+1}^{\mathbb{T}} b^{(n-1)}$$

where $a^{(n)} = \left(a_1^{(n)}, a_2^{(n)}, ..., a_d^{(n)}\right), b^{(n)} = \left(b_1^{(n)}, b_2^{(n)}, ..., b_d^{(n)}\right)$. We will make use of the following

Proposition 4.9 (Theorem 4 in [10]). *Let $(X_k)_{k \geq 1}$ be a sequence of i.i.d. random stochastic $d \times d$ matrices such that $X_{n_0} X_{n-1} ... X_2 X_1$ is a positive matrix with a positive probability for some $n_0 < \infty$. Then there exists a random nonnegative vector $W = (w_1, w_2, ..., w_d)$ such that $w_1 + w_2 + ... + w_d = 1$ and*

$$X_n X_{n-1} ... X_2 X_1 \to \mathbf{1}^{\mathbb{T}} W$$

almost surely as $n \to \infty$, where $\mathbf{1} = (1, 1, ..., 1)$. Moreover, if $V = (v_1, ..., v_d)$ is a random non-negative vector such that $v_1 + v_2 + ...v_d = 1$, V is independent of X_1 then $VX_1 = V$ in distribution if and only if $V = W$ in distribution.

Theorem 4.10. *Suppose that Assumptions 6 and 8 hold then the polygon L_n converges almost surely to a random point P inside the initial polygon L_0 such that*

$$\max_{j=1,2,...,d} \|PA_j^{(n)}\| \sim Ce^{\mu_1 n}$$

almost surely as $n \to \infty$, in the sense that $\frac{1}{n} \log \left(\max \|PA_j^{(n)}\| \right) \to \mu_1$ where μ_1 is defined in (A.14).

Proof. By Assumption 6 we have that $H_d^{\mathbb{T}} H_{d-1}^{\mathbb{T}} ... H_2^{\mathbb{T}} H_1^{\mathbb{T}}$ is almost surely a positive stochastic matrix. Therefore, from Proposition 4.9 it follows that there exists a random non-negative vector $\zeta = (\zeta_1, ..., \zeta_d)$ such that $\zeta_1 + ... + \zeta_d = 1$ for which

$$a^{(n)} \to \left(\zeta_1 a_1^{(0)} + ... + \zeta_d a_d^{(0)} \right) \mathbf{1}$$

and

$$b^{(n)} \to \left(\zeta_1 b_1^{(0)} + ... + \zeta_d b_d^{(0)} \right) \mathbf{1}$$

almost surely as $n \to \infty$. It implies that the sequence of polygon L_n converges to a random point P defined by the following vector identity

$$OP = \zeta_1 OA_d^{(0)} + \zeta_2 OA_1^{(0)} + ... + \zeta_d OA_{d-1}^{(0)}$$

where $O = (0,0)$ is the origin of the Cartesian plane. (Observe that if ξ_i is Beta(α, β) distributed on $(0,1)$ then $\zeta = (\zeta_1, ..., \zeta_d)$ is a Dirichlet distributed random vector with parameters $(\alpha + \beta, \alpha + \beta, ..., \alpha + \beta)$.)

Since

$$\|PA_j\| < \|A_d A_1\| + \|A_1 A_2\| + ... + \|A_{d-1} A_d\| \leq d \times \max_{k=1,2,...,d} \|A_k A_{k+1}\|.$$

and on the other hand, for each $k = 1, 2, ..., d$, we have

$$\max_{j=1,2,...,d} \|PA_j\| \geq \frac{1}{2} \left(\|PA_k\| + \|PA_{k+1}\| \right) \geq \frac{1}{2} \|A_k A_{k+1}\|$$

the following inequality inequalities hold:

$$\mathrm{M}_n \leq \max_{j=1,2,...,d} \|PA_j^{(n)}\| \leq d\,\mathrm{M}_n. \tag{A.21}$$

Under the Assumption 8 we have (see our Lemma 4.8 and Proposition III.7.2 in [2])

$$\lim_{n \to \infty} \frac{1}{n} \log(\mathrm{M}_n) = \lim_{n \to \infty} \frac{1}{n} \log \|T_n T_{n-1} ... T_2 T_1\| = \mu_1 \in (-\infty, 0)$$

almost surely. Therefore,

$$\max_{j=1,2,...,d} \|PA_j^{(n)}\| \sim C e^{\mu_1 n}$$

almost surely as $n \to \infty$. $\qquad\qquad\qquad\square$

Random triangles revisited

The goal of this Section is to show that in three-dimensional case the projection of the "middle" vertex on the largest side of the triangle converges in distribution, thus generalizing the result of Theorem 5 in [11]; our main results is presented in Theorems 5.3 which follows later in the Section. We also evaluate the speed of convergence to flatness in Theorem 5.4, as well as study some examples; in particular, we strengthen the result of Theorem 4 in [11].

Since $x \in \mathcal{L}$ defined by (A.4) we can restrict our attention just to the first $d - 1$ coordinates of x. Let us introduce the norm

$$\|x\|_\infty = \max_{j=1,...,d} \|x_j\| = \max\{|x_1|, |x_2|, ..., |x_{d-1}|, |x_1 + ... + x_{d-1}|\}.$$

and for each x in the unit ball $B_\infty = \{(x_1, ..., x_{d-1}) \in \mathbb{R}^{d-1} : \|x\|_\infty = 1\}$ the map $\widehat{T} : B_\infty \to B_\infty$ by

$$\widehat{T}(x) = \frac{1}{\|Tx\|_\infty} Tx.$$

Notice that $\left\{ \widehat{T^{(n)}}(x) \right\}_{n \geq 1}$ is a Markov chain which can be considered as a system of iterated random functions in the sense mentioned in [4, 8]. We will use the following result implied from Lemma 2.5 in [8]:

Lemma 5.1. *Let $D_\epsilon = \{(x,y) : x,y \in B_\infty, \|x - y\|_\infty \leq \epsilon\}$. Assume that*

$$\limsup_{n \to \infty} \sup_{(x,y) \in D_\epsilon} \mathbb{E} \left\| \widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y) \right\|_\infty \to 0 \tag{A.22}$$

as $\epsilon \to 0$ and for each $\alpha > 0$, there exist an integer N and a constant $c > 0$ such that for every $x, y \in B_\infty$

$$\mathbb{P} \left(\left\| \widehat{T'^{(N)}}(x) - \widehat{T''^{(N)}}(y) \right\|_\infty < \alpha \right) \geq c, \tag{A.23}$$

where $\widehat{T'^{(n)}}$ and $\widehat{T''^{(n)}}$ are two independent of copies of $\widehat{T^{(n)}}$. Then, for every $x \in B_\infty$, $\widehat{T^{(n)}}(x)$ weakly converges to some random vector, which is independent of x.

Here is a very important result.

Lemma 5.2. *Assume that Assumption 6 and 7 are fulfilled then*

$$\left(\max_{j=1,\dots,d} \left\| x_j^{(n)} \right\| \right)^{-1} x^{(n)}$$

converges in distribution to some random vector as $n \to \infty$, which is independent of $x^{(0)}$.

Proof. Assume that all the points x, y, etc., belong to B_∞ unless stated otherwise. Next, w.l.o.g. assume that $\|T^{(n)}x\|_\infty \leq \|T^{(n)}y\|_\infty$, then we have

$$\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\|_\infty$$

$$= \frac{1}{\|T^{(n)}x\|_\infty} \left\| T^{(n)} \left(x - \frac{\|T^{(n)}x\|_\infty}{\|T^{(n)}y\|_\infty} y \right) \right\|_\infty \leq \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty} \left\| x - \frac{\|T^{(n)}x\|_\infty}{\|T^{(n)}y\|_\infty} y \right\|_\infty$$

$$\leq \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty} \left(\|x - y\|_\infty + \left(1 - \frac{\|T^{(n)}x\|_\infty}{\|T^{(n)}y\|_\infty} \right) \|y\|_\infty \right)$$

$$\leq \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty} \|x - y\|_\infty + \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty \cdot \|T^{(n)}y\|_\infty} \left(\|T^{(n)}y\|_\infty - \|T^{(n)}x\|_\infty \right)$$

$$\leq \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty} \|x - y\|_\infty + \frac{\|T^{(n)}\|_\infty}{\|T^{(n)}x\|_\infty \cdot \|T^{(n)}y\|_\infty} \left(\|T^{(n)}(y - x)\|_\infty \right)$$

$$\leq 2 \frac{\|T^{(n)}\|_\infty^2}{\|T^{(n)}x\|_\infty \cdot \|T^{(n)}y\|_\infty} \|x - y\|_\infty,$$

where $\|T\|_\infty = \sup_{x\in B_\infty} \|Tx\|_\infty$, $\|T\| = \sup_{\|x\|=1} \|Tx\|$ are the usual operator norms. Therefore, since all the norms on finite dimensional spaces are equivalent, there exists a non random constant $r > 0$ such that

$$\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\|_\infty \leq r \cdot \frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|}\|x - y\|_\infty \tag{A.24}$$

On the other hand, by Theorem III.3.1 in [2], for almost all ω, the exist one-dimensional linear space $V(\omega) \subset \mathbb{R}^{d-1}$ which is the range of limit points of $\|T_1(\omega)\dots T_n(\omega)\|^{-1}T_1(\omega)\dots T_n(\omega)$. By the proof of Proposition III.3.2 in [2] if a sequence $\{x_n\}_{n\geq 1} \subset B_\infty$ converges to x and $\zeta_x(\omega)$ is the orthogonal projection of x onto $V(\omega)$ then

$$\limsup_{n\to\infty} \frac{\|T^{(n)}\|}{\|T^{(n)}x_n\|} \leq \|\zeta_x\|^{-1} \quad \text{a.s.} \tag{A.25}$$

and

$$\mathbb{P}\left(\|\zeta_x\| = 0\right) = 0. \tag{A.26}$$

Therefore, we obtain that

$$\limsup_{n\to\infty} \|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\|_\infty \leq r\|\zeta_x\|^{-1}\|\zeta_y\|^{-1}\|x - y\|_\infty \quad \text{a.s.} \tag{A.27}$$

Let us now verify the condition (A.22). We have

$$\mathbb{E}\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \leq \mathbb{E}\left(\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \mathbf{1}_{\left\{\frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \geq \frac{1}{4r\epsilon}\right\}}\right) +$$

$$+ \mathbb{E}\left(\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \mathbf{1}_{\left\{\frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \leq \frac{1}{4r\epsilon}\right\}}\right)$$

$$=: (\text{I}) + (\text{II})$$

To bound (I), observe that $\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \leq 2$ and therefore (I) $\leq$ $2\,\mathbb{P}\left(\frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \geq \frac{1}{4r\epsilon}\right)$. Suppose

$$\limsup_{n\to\infty} \sup_{(x,y)\in D_\epsilon} \mathbb{P}\left(\frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \geq \frac{1}{4r\epsilon}\right) \not\to 0 \quad \text{as } \epsilon \to 0.$$

Then there exists a $\beta > 0$ and a decreasing sequence $\epsilon_k \downarrow 0$ such that this $\limsup_{n\to\infty} \sup_{(x,y)\in D_{\epsilon_k}} \mathbb{P}(\cdots \geq (4r\epsilon_k)^{-1}) > \beta$ for each k; therefore for each k there is a sequence $n_i^{(k)}$, $i = 1, 2, \ldots$ such that

$$\delta(k) := \mathbb{P}\left(\frac{\|T^{(n_i^{(k)})}\|^2}{\|T^{(n_i^{(k)})}x_{n_i^{(k)}}\| \|T^{(n_i^{(k)})}y_{n_i^{(k)}}\|} \geq \frac{1}{4r\epsilon_k} \right) > \beta \ \text{ for all } i = 1, 2, \ldots. \tag{A.28}$$

Let $m_k = n_k^{(k)}$. Without loss of generality assume that $x_{m_k} \to x_* \in B_\infty$ and $y_{m_k} \to y_* \in B_\infty$; since B_∞ is compact we can always choose a convergence subsequence.

By (A.25) we have

$$\delta_x(k) := \mathbb{P}\left(\frac{\|T^{(m_k)}\|^2}{\|T^{(m_k)}x_{m_k}\|} \geq 2\|\zeta_{x_*}\|^{-1} \right) \to 0,$$

$$\delta_y(k) := \mathbb{P}\left(\frac{\|T^{(m_k)}\|^2}{\|T^{(m_k)}y_{m_k}\|} \geq 2\|\zeta_{y_*}\|^{-1} \right) \to 0,$$

as $k \to \infty$. Hence

$$\delta(k) \leq \delta_x(k) + \delta_y(k) + \mathbb{P}\left(4\|\zeta_{x_*}\|^{-1}\|\zeta_{y_*}\|^{-1} \geq \frac{1}{4r\epsilon_k} \right)$$

$$= \delta_x(k) + \delta_y(k) + \mathbb{P}\left(\|\zeta_{x_*}\| \cdot \|\zeta_{y_*}\| \leq 16r\epsilon_k \right)$$

$$\leq \delta_x(k) + \delta_y(k) + \mathbb{P}\left(\|\zeta_{x_*}\| \leq 4\sqrt{r\epsilon_k} \right) + \mathbb{P}\left(\|\zeta_{x_*}\| \leq 4\sqrt{r\epsilon_k} \right) \to 0$$

by (A.26), leading to a contradiction with (A.28).

On the other hand, if $(x,y) \in D_\epsilon$ and $\frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \leq \frac{1}{4r\epsilon}$ then the inequality (A.27) implies that $\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\|_\infty \leq \frac{1}{4}$, hence

$$\sup_{(x,y)\in D_\epsilon} (\text{II}) = \sup_{(x,y)\in D_\epsilon} \mathbb{E}\left(\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \mathbf{1}_{\left\{ \frac{\|T^{(n)}\|^2}{\|T^{(n)}x\|\|T^{(n)}y\|} \leq \frac{1}{4r\epsilon} \right\}} \right)$$

$$\leq \sup_{(x,y)\in D_\epsilon} \mathbb{E}\left(\left\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\right\|_\infty \mathbf{1}_{\{\|\widehat{T^{(n)}}(x) - \widehat{T^{(n)}}(y)\|_\infty \leq \frac{1}{4}\}} \right)$$

$$\leq \text{Const} \cdot \sup_{x,y\in B_\infty} \mathbb{E}\, \delta\left(T^{(n)}\overline{x}, T^{(n)}\overline{y} \right)$$

64

where the last inequality holds since $\|u - v\|_\infty \leq \text{Const} \cdot \delta(\overline{u}, \overline{v})$ for any vectors u, v such that the angle between u and v is smaller than $\frac{\pi}{2}$. Finally, from the proof of Theorem III.4.3 in [2], we have

$$\limsup_{n \to \infty} \sup_{x,y \in B_\infty} \mathbb{E}\, \delta\left(T^{(n)}\overline{x}, T^{(n)}\overline{y}\right) = 0.$$

Therefore the condition (A.22) is fulfilled.

We now only have to verify the condition (A.23). Note that

$$\mathbb{P}\left(\left\|\widehat{T'^{(n)}}(x) - \widehat{T''^{(n)}}(y)\right\|_\infty < \alpha\right)$$
$$\geq \mathbb{P}\left(\left\|\widehat{T'^{(n)}}(x) - \widehat{T''^{(n)}}(x)\right\|_\infty + \left\|\widehat{T''^{(n)}}(x) - \widehat{T''^{(n)}}(y)\right\|_\infty < \alpha\right)$$
$$\geq \mathbb{P}\left(\left\|\widehat{T'^{(n)}}(x) - \widehat{T''^{(n)}}(x)\right\|_\infty < \alpha/2, \left\|\widehat{T''^{(n)}}(x) - \widehat{T''^{(n)}}(y)\right\|_\infty < \alpha/2\right)$$
$$\geq \mathbb{P}\left(\left\|\widehat{T''^{(n)}}(x) - \widehat{T''^{(n)}}(y)\right\|_\infty < \alpha/2\right) - 1$$
$$+ \mathbb{P}\left(\left\|\widehat{T'^{(n)}}(x) - \widehat{T''^{(n)}}(x)\right\|_\infty < \alpha/2\right). \tag{A.29}$$

W.l.o.g., assume that $\alpha < 1/2$. Observe that

$$\mathbb{P}\left(\left\|\widehat{T''^{(n)}}(x) - \widehat{T''^{(n)}}(y)\right\|_\infty < \alpha/2\right)$$
$$\geq \mathbb{P}\left(\text{Const} \cdot \delta(T''^{(n)}\overline{x}, T''^{(n)}\overline{y}) < \alpha/2\right) \to 1 \tag{A.30}$$

as $n \to \infty$, where the above limit is implied from Theorem III.4.3(i) in [2]. Furthermore, $\widehat{T'^{(n)}}(x)$ and $\widehat{T''^{(n)}}(x)$ are independent identically distributed random variables. Therefore, $\mathbb{P}\left(\left\|\widehat{T'^{(n)}}(x) - \widehat{T''^{(n)}}(x)\right\|_\infty < \alpha/2\right)$ is always positive. This fact together with (A.29) and (A.30) imply the satisfaction of (A.23). $\qquad\square$

From now on assume that $d = 3$. Following [11], for each $n \geq 0$ rescale the triangle $A_1^{(n)} A_1^{(n)} A_3^{(n)}$ to a new triangle $B_1^{(n)} B_1^{(n)} B_3^{(n)}$ such that its longest edge has length 1, its vertices are reordered in a way that $B_1^{(n)} B_2^{(n)} \geq B_3^{(n)} B_1^{(n)} \geq B_2^{(n)} B_3^{(n)}$, and let the Cartesian coordinates of vertices be $B_1^{(n)} = (0,0), B_2^{(n)} = (1,0), B_3^{(n)} = \theta_n = (g_n, h_n)$; formally

$$h_n = \frac{2\mathbf{A}(L_n)}{\mathbf{M}_n^2} \tag{A.31}$$

is the height of the rescaled triangle, corresponding to the largest side with length $\mathbf{M}_n$, formally defined by (A.19). Without loss of generality, we can also assume that $A_1^{(0)} \equiv B_1^{(0)}, A_2^{(0)} \equiv B_2^{(0)}, A_3^{(0)} \equiv B_3^{(0)}$.

Theorem 5.3. *Assume that Assumption 6 is fulfilled then g_n converges in distribution to some random variable $\eta \in [1/2, 1]$.*

Proof. Recall that $(x_i^{(n)}, y_i^{(n)})$, $i = 1, 2, 3$ denote the coordinates of the vectors corresponding to the three sides of the triangle. Since they are asymptotically collinear, g_n has the same limit as

$$g_{x,n} = f(x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$$

where $f(a, b, c)$ is the ratio between the second largest amongst $\{|a|, |b|, |c|\}$ and the largest of them; in fact, $|g_{x,n} - g_n| \leq h_n \to 0$ a.s. from an elementary geometric observation. The only problem which could arise is if the triangle is (nearly) vertical; however this does not happen for large n a.s. by Lemma II.4.2 from [2] which says (equivalently) the the direction of the limiting flat triangle has a continuous distribution.

Since in our case $a + b + c = 0$, we can write f as

$$f(a, b, c) = 1 - \frac{\min\{|a|, |b|, |c|\}}{\max\{|a|, |b|, |c|\}}.$$

Since the function $f(\cdot)$ is continuous, and the vector $x^{(n)}/\|x^{(n)}\|$ converges weakly by Lemma 5.2, the result follows. $\square$

Theorem 5.4. *Suppose that Assumption 8 is fulfilled. Then*

$$\lim_{n \to \infty} \frac{1}{n} \log(h_n) = \mathbb{E}(\log(\det(T_1))) - 2 \int_{1/2}^1 \zeta(x, 0) d\mathbb{P}_\eta(x), \quad a.s.$$

66

where $\theta_n = (g_n, h_n)$ is defined just above (A.31), η is the weak limit of g_n, $\mathbb{P}_\eta$ is its probability measure, and

$$\zeta(x, y) = \mathbb{E}\left(\log(M_1) \mid \theta_0 = (x, y)\right).$$

Proof. We have the following relation

$$h_n = h_{n-1} \cdot \frac{M_{n-1}^2}{M_n^2} \cdot \det(T_n)$$

which implies that

$$\frac{1}{n} \log(h_n) = \frac{1}{n} \sum_{j=1}^{n} \log(\det(T_j)) - \frac{2}{n} \log(M_n) + O\left(\frac{1}{n}\right).$$

Suppose that Assumption 8 is fulfilled. By the strong law of large numbers and equation (A.17) we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} \log(\det(T_j)) = \mathbb{E}(\log(\det(T_1))) = \mu_1 + \mu_2 \quad \text{a.s.}$$

By Lemma 4.8

$$\lim_{n \to \infty} \frac{1}{n} \log(M_n) \to \mu_1 \quad \text{a.s.}$$

so that

$$\lim_{n \to \infty} \frac{1}{n} \log(h_n) = \mu_2 - \mu_1 \quad \text{a.s.}$$

On the other hand, we have

$$\frac{1}{n} \log(h_n) = \frac{1}{n} \sum_{j=1}^{n} \log(\det(T_j)) - \frac{2}{n} \sum_{j=1}^{n} \log\left(\frac{M_j}{M_{j-1}}\right) + \frac{1}{n} \log(h_0).$$

Let $\mathbb{P}_n(d\theta \mid \theta_0)$ be the conditional probability measure of θ_n on θ_0. We have

$$\mathbb{E}\left(\frac{1}{n} \sum_{j=1}^{n} \log\left(\frac{M_j}{M_{j-1}}\right) \mid \theta_0\right) = \sum_{j=1}^{n} \frac{1}{n} \mathbb{E}\left(\log\left(\frac{M_j}{M_{j-1}}\right) \mid \theta_0\right) = \int \zeta(\theta) \tilde{\mathbb{P}}_n(d\theta \mid \theta_0),$$

where we denote $\zeta(\theta) = \mathbb{E}\left(\log(M_1) \mid \theta_0 = \theta\right)$ and $\tilde{\mathbb{P}}_n(d\theta \mid \theta_0) = \frac{1}{n} \sum_{j=1}^{n-1} \mathbb{P}_j(d\theta \mid \theta_0)$.

We already know that $h_n \to 0$ almost surely and g_n converges in distribution to some random variable η taking value on $(1/2, 1)$, therefore $\theta_n = (g_n, h_n)$

converges in distribution to $(\eta, 0)$ as $n \to \infty$. Since $\zeta(x, 0)$ is a continuous function of x on $(1/2, 1)$, using Cesàro mean result we have

$$\lim_{n \to \infty} \int \zeta(\theta) \tilde{\mathbb{P}}_n(d\theta|\,\theta_0) = \lim_{n \to \infty} \int \zeta(\theta) \mathbb{P}_n(d\theta|\,\theta_0) = \int_{1/2}^{1} \zeta(x, 0) d\mathbb{P}_\eta(x).$$

where $\mathbb{P}_\eta$ is the probability measure of η. Therefore

$$\lim_{n \to \infty} \frac{1}{n} \log(h_n) = \mathbb{E}(\log(\det(T_1))) - 2 \int_{1/2}^{1} \zeta(x, 0) d\mathbb{P}_\eta(x). \tag{A.32}$$

$\square$

Example 1. Let us consider the case when random variables ξ_1, ξ_2, ξ_3 are uniformly distributed on $(0, 1)$, notice that $\theta_n = (g_n, h_n)$ converges in distribution to $(U, 0)$, where U is uniformly distributed on $(\frac{1}{2}, 1)$, see [11]. We easily obtain that

$$\mathbb{E}\left(\log(\det(T_1))\right) = \mathbb{E}\left(\log\left((1 - \xi_1)(1 - \xi_2)(1 - \xi_3) + \xi_1 \xi_2 \xi_3\right)\right) = \frac{-24 + \pi^2}{9}.$$

and

$$\zeta(x, 0) = \frac{x\left(2x^2 \log(x) - 5x + 5\right) - 2(x-1)^3 \log(1-x)}{6(x-1)x},$$

hence

$$\int_{1/2}^{1} \zeta(x, 0) dx = \frac{-15 + \pi^2}{18}$$

and we can conclude from (A.32) that

$$h_n \sim C e^{-\frac{\pi^2 - 6}{9} n}$$

as $n \to \infty$ in the sense that $\frac{1}{n} \log h_n \to -\frac{\pi^2 - 6}{9} \approx -0.43$, thus strengthening the result of Theorem 4 in [11].

Example 2. Suppose that ξ_1, ξ_2, ξ_3 have a continuous distribution with density symmetric around $\frac{1}{2}$, i.e. $p(1 - x) = p(x)$. Let $x \in (0, 1)$ and set $x_1 = x\xi_1, x_3 = x + (1 - x)\xi_3, x_2 = \xi_2$ and $y_1 \leq y_2 \leq y_3$ be the triple x_1, x_2, x_3 sorted in the increasing order. For $z < x$, we have

$$\mathbb{P}\left(\frac{y_2 - y_1}{y_3 - y_1} < z\right) = I_1(z, x) + I_2(z, x)$$

where

$$I_1(z, x) = \mathbb{P}\left(\frac{y_2 - y_1}{x_3 - y_1} < z; y_1 < y_2 < x < x_3\right)$$

$$= \mathbb{P}\left(y_2 < zx_3 + (1-z)y_1; y_1 < y_2 < x < x_3\right)$$

$$= \mathbb{P}\left(y_1 < y_2 < x; zx_3 + (1-z)y_1 > x\right)$$

$$+ \mathbb{P}\left(y_1 < y_2 < zx_3 + (1-z)y_1; zx_3 + (1-z)y_1 < x\right)$$

$$= \mathbb{P}\left(y_1 < y_2 < x; \frac{x - zx_3}{1 - z} < y_1 < x; \ x < x_3 < 1\right)$$

$$+ \mathbb{P}\left(y_1 < y_2 < zx_3 + (1-z)y_1; zx_3 + (1-z)y_1 < x; \ 0 < y_1 < \frac{x - zx_3}{1 - z}; \ x < x_3 < 1\right)$$

$$= \int_x^1 dx_3 \int_{\frac{x-zx_3}{1-z}}^x dy_1 \int_{y_1}^x \left[\frac{1}{x}p\left(\frac{y_1}{x}\right)p(y_2) + \frac{1}{x}p\left(\frac{y_2}{x}\right)p(y_1)\right] \frac{1}{1-x}p\left(\frac{x_3 - x}{1 - x}\right) dy_2$$

$$+ \int_x^1 dx_3 \int_0^{\frac{x-zx_3}{1-z}} dy_1 \int_{y_1}^{zx_3+(1-z)y_1}$$

$$\left[\frac{1}{x}p\left(\frac{y_1}{x}\right)p(y_2) + \frac{1}{x}p\left(\frac{y_2}{x}\right)p(y_1)\right] \frac{1}{1-x}p\left(\frac{x_3 - x}{1 - x}\right) dy_2,$$

and

$$I_2(z, x) = \mathbb{P}\left(\frac{y_2 - x_1}{y_3 - x_1} < z; x_1 < x < y_2 < y_3\right)$$

$$= \mathbb{P}\left(y_3 > \frac{y_2 - (1-z)x_1}{z}; x_1 < x < y_2 < y_3\right)$$

$$= \mathbb{P}\left(\frac{y_2 - (1-z)x_1}{z} < y_3 < 1; \frac{y_2 - (1-z)x_1}{z} < 1; y_2 > x\right)$$

$$= \mathbb{P}\left(\frac{y_2 - (1-z)x_1}{z} < y_3 < 1; x < y_2 < (1-z)x_1 + z; (1-z)x_1 + z > x\right)$$

$$= \mathbb{P}\left(\frac{y_2 - (1-z)x_1}{z} < y_3 < 1; x < y_2 < (1-z)x_1 + z; \frac{x - z}{1 - z} < x_1 < x\right)$$

$$= \int_{\frac{x-z}{1-z}}^x dx_1 \int_x^{(1-z)x_1+z} dy_2 \int_{\frac{y_2-(1-z)x_1}{z}}^1$$

$$\times \left[\frac{1}{1-x}p\left(\frac{y_3 - x}{1 - x}\right)p(y_2) + \frac{1}{1-x}p\left(\frac{y_2 - x}{1 - x}\right)p(y_3)\right] \frac{1}{x}p\left(\frac{x_1}{x}\right) dy_3.$$

For $z > x$, by the symmetric property, we have

$$\mathbb{P}\left(\frac{y_2 - y_1}{y_3 - y_1} < z\right) = I_1(1 - z, 1 - x) + I_2(1 - z, 1 - x).$$

Let η be the invariant distribution defined in Theorem 5.3. Assume that $2\eta - 1$ has the density $\varphi(x)$, then $\varphi(x)$ is the unique solution of the following integral equation:

$$\int_0^z \varphi(x)dx = \int_0^z \left[I_1(1-z, 1-x) + I_2(1-z, 1-x)\right] \varphi(x)dx$$
$$+ \int_z^1 \left[I_1(z, x) + I_2(z, x)\right] \varphi(x)dx \tag{A.33}$$

since one can look, for example, at the linear projections of the vertices of the triangle, see also [11].

Now fix a positive integer n, and additionally assume that ξ_1, ξ_2, ξ_3 are independent $\text{Beta}(n, n)$ distributed random variables, i.e. their density function is given by

$$p_n(\xi) = \begin{cases} \frac{\xi^{n-1}(1-\xi)^{n-1}}{B(n,n)}, & \xi \in (0,1) \\ 0, & \text{otherwise} \end{cases} \quad \text{where } B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} \, dt$$

is the usual Beta function. Let the corresponding invariant distribution $\varphi_n(x)$ be defined by (A.33).

Using a computer algebra system, e.g. Mathematica$^{\text{TM}}$ or Maple$^{\text{TM}}$, one can check that the solution to (A.33) for $n = 1, 2, 3, 4, 5$ are given by

$$\varphi_1(z) = 1,$$
$$\varphi_2(z) = \frac{6}{7}\left((1-z)z + 1\right),$$
$$\varphi_3(z) = \frac{30}{143}\left(3(1-z)^2z^2 + 4(1-z)z + 4\right),$$
$$\varphi_4(z) = \frac{140}{4199}\left(13(1-z)^3z^3 + 22(1-z)^2z^2 + 25(1-z)z + 25\right),$$
$$\varphi_5(z) = \frac{6174}{7429}\left(\frac{17}{49}(1-z)^4z^4 + \frac{5}{7}(1-z)^3z^3 + \frac{13}{14}(1-z)^2z^2 + (1-z)z + 1\right).$$

We conjecture that in the general case $\varphi_n(z)$ is also a mixture of some Beta distributions, that is, there exist non-negative constants $c_1, c_2, ..., c_n$ summing up to 1 such that

$$\varphi_n(z) = \sum_{j=1}^n c_j \frac{z^{j-1}(1-z)^{j-1}}{B(j, j)}$$
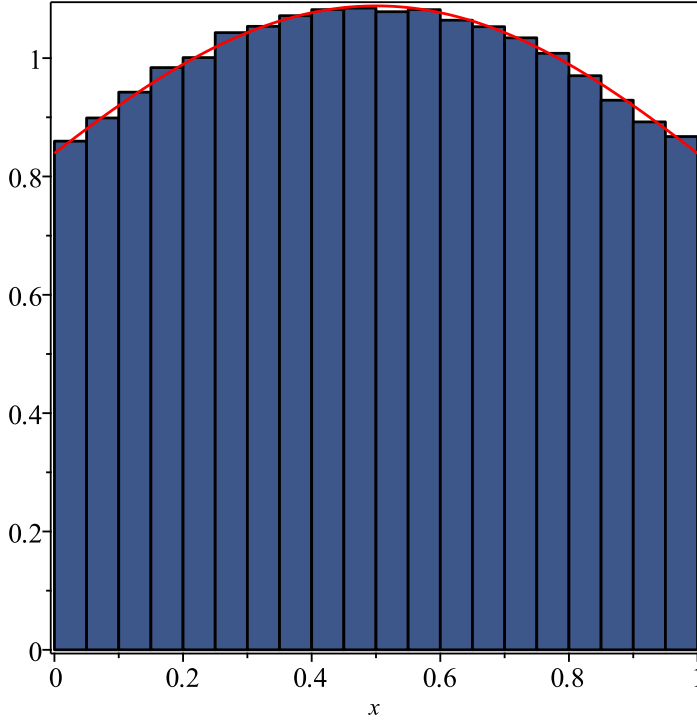
but unfortunately we cannot prove this fact.

Figure A.2: For $\xi \sim \text{Beta}(3,3)$, one can see the similarity between the histogram of $\{2g_j - 1, j = 1, 2, ..., 10^6\}$ obtained from simulation and the plot of $\{\varphi_3(x), x \in [0,1]\}$.

Generalizations and open problems

Let $\xi_1, \xi_2, \ldots, \xi_d$ be the random variables governing how the sides of the d-polygon are split at each iteration. Throughout the paper we have assumed that ξ_j, $j = 1, \ldots, d$ are i.i.d. However, if one looks at the proofs, one can see that the independence assumption can be substantially relaxed without any change in the proofs. Indeed, let $\bar{\xi} = (\xi_1, \xi_2, \ldots, \xi_d)$ be the random variable describing the splitting proportions of the sides of the polygon. Assume that

(i) $\mathbb{P}(0 < \xi_i < 1) = 1$ for all i;

(ii) there are two distinct numbers $a, b \in (0, 1)$ such that all 2^d points of the form $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, where each $x_i = a$ or $= b$, belong to the support of $\bar{\xi}$;

(iii) if d is even then $\xi_1 \xi_2 \ldots \xi_d \neq (1 - \xi_1)(1 - \xi_2) \ldots (1 - \xi_d)$ a.s.

Then Conjecture 1 is fulfilled (observe that we still suppose that random variables $\bar{\xi}$ are drawn in i.i.d. manner for each iteration).

We also strongly feel that assumption (iii) is, in fact, superfluous, so the result will hold even if some matrices are degenerate. Indeed, intuitively, when some of the matrices in the product are not full rank, this should even be helpful for the convergence to lower-dimensional subspaces. However, in this case we would clearly not be able to form a group containing all the matrices in the support of the measure and hence cannot use the standard results from the random matrix theory.

Another possible generalization of our model is to higher dimensional spaces, e.g. random subdivision of tetrahedrons in $\mathbb{R}^3$. We are currently working on showing similar results in this case.

Acknowledgement

Bibliography

[1] Beyer, W. H. (Ed.). CRC Standard Mathematical Tables, 28th ed. Boca Raton, FL: CRC Press, pp. 123–124, 1987.

[2] Bougerol, Philippe; Lacroix, Jean. Products of random matrices with applications to Schrödinger operators. Progress in Probability and Statistics, 8. Birkhäuser Boston, Inc., Boston, MA, 1985.

[3] Davis, Philip J. Circulant Matrices, Wiley, New York, 1970.

[4] Diaconis, Persi; Freedman, David. Iterated random functions. SIAM Rev. 41 (1999), no. 1, 45–76.

[5] Diaconis, Persi; Miclo, Laurent. On Barycentric Subdivision. Combinatorics, Probability and Computing 20 (2011), 213–237.

[6] I. Ya. Goldsheid, Y. Guivarc'h. Zariski closure and the dimension of the Gaussian law of the product of random matrices. Probab. Theory Relat. Fields 105 (1996), 109–142.

[7] Goldsheid, I. Ya.; Margulis, G. A. Lyapunov exponents of a product of random matrices. (Russian) Uspekhi Mat. Nauk 44 (1989), (269), 13–60; translation in Russian Math. Surveys 44 (1989), 11–71.

[8] Kaijser, Thomas. A limit theorem for Markov chains in compact metric spaces with applications to products of random matrices. Duke Math. J. 45 (1978), no. 2, 311–349.

[9] Mannion, D. Convergence to collinearity of a sequence of random triangle shapes. Adv. in Appl. Probab. 22 (1990), no. 4, 831–844.

[10] McKinlay, S. A characterisation of transient random walks on stochastic matrices with Dirichlet distributed limits. J. Appl. Probab. 51 (2014), no. 2, 542–555.

[11] Volkov, S.. Random geometric subdivisions. Random Structures Algorithms 43 (2013), no. 1, 115–130.

[12] Veitch, J. G.; Watson, G. S. Random cyclic transformations of points. Essays in time series and allied processes. J. Appl. Probab. 1986, Special Vol. 23A, 369–382.

[13] Watson, G. S. Random triangles. Proceedings of Second International Workshop on Stereology and Stochastic Geometry (Aarhus, 1983), 183–199, Memoirs, 6, Aarhus Univ., Aarhus, 1983.

B

Paper B

Strongly vertex-reinforced jump processes on complete graphs

Olivier Raimond[a] and Tuan-Minh Nguyen[b]

[a]*Modélisation aléatoire de l'Université Paris Nanterre (MODAL'X)*
[b]*Centre for Mathematical Sciences, Lund University*

Abstract

The aim of our work is to study vertex-reinforced jump processes with superlinear weight function $w(t) = t^\alpha$, for some $\alpha > 1$. On any complete graph $G = (V, E)$, we prove that there is one vertex $v \in V$ such that the total time spent at v almost surely tends to infinity while the total time spent at the remaining vertices is bounded.

Keywords: Vertex-reinforced jump processes; nonlinear reinforcement; random walks with memory; stochastic approximation

Introduction

Let $G = (V, E)$ be a finite connected, unoriented graph without loops, where $V = \{1, 2, ..., d\}$ and E respectively stand for the set of vertices and the set of edges. We consider a continuous time jump process X on the vertices of G such that the law of X satisfies the following condition:

i. at time $t \leq 0$, the local time at each vertex $v \in V$ has a positive initial value $\ell_0^{(v)}$,

ii. at time $t > 0$, given the natural filtration $\mathcal{F}_t$ generated by $\{X_s, s \leq t\}$, the probability that there is a jump from X_t during $(t, t + h]$ to a neighbour

v of X_t (i.e. $(v, X_t) \in E$) is given by

$$w \left(\ell_0^{(v)} + \int_0^t \mathbf{1}_{\{X_s = v\}} ds \right) \cdot h + o(h),$$

where $w : [0, \infty) \to (0, \infty)$ is called weight function.

For each vertex $v \in V$, we also denote as $L(v, t) = \ell_0^{(v)} + \int_0^t \mathbf{1}_{\{X_s = v\}} ds$ the local time at v up to time t and let

$$Z_t = \left(\frac{L(1, t)}{\ell_0 + t}, \frac{L(2, t)}{\ell_0 + t}, ..., \frac{L(d, t)}{\ell_0 + t} \right)$$

stand for the (normalized) occupation measure on V at time t, where $\ell_0 = \ell_0^{(1)} + \ell_0^{(2)} + \cdots + \ell_0^{(d)}$.

In our work, we consider the weight function $w(t) = t^\alpha$, for some $\alpha > 0$. The jump process X is called *strongly vertex-reinforced* if $\alpha > 1$, *weakly vertex-reinforced* if $\alpha < 1$ or *linearly vertex-reinforced* if $\alpha = 1$.

The model of discrete time edge-reinforced random walks (ERRW) was first studied by Coppersmith and Diaconis in their unpublished manuscripts [7] and later the model of discrete time vertex-reinforced random walks (VRRW) was introduced by Pemantle in [11] and [12]. Several remarkable results about localization of ERRW and VRRW were obtained in [20], [17], [21], [6] and [22]. Following the idea about discrete time reinforced random walks, Wendelin Werner conceived a model in continuous time so-called vertex reinforced jump processes (VRJP) whose linear case was first investigated by Davis and Volkov in [8] and [9]. In particular, these authors showed in [9] that linearly VRJP on any finite graph is recurrent, i.e. all local times are almost surely unbounded and the normalized occupation measure process converges almost surely to an element in the interior of the $(d - 1)$ dimensional standard unit simplex as time goes to infinity. The relation between VRJP, ERRW and random walks in random environment as well as its applications were studied in [10], [13], [14], [18] and [19].

In this paper, we prove that strongly VRJP on a complete graph $G = (V, E)$ will almost surely have an infinite local time at some vertex v, while the local times at the remaining vertices remain bounded. The main technique of our proofs is based on stochastic approximation (see, e.g. [1, 2, 3, 4]).

Preliminary notations and remarks

Throughout this paper, we denote as Δ and $T\Delta$ respectively the $(d-1)$ dimensional standard unit simplex in $\mathbb{R}^d$ and its tangent space, which are defined by

$$\Delta = \{z = (z_1, z_2, ..., z_d) \in \mathbb{R}^d : z_1 + z_2 + \cdots + z_d = 1, z_j \geq 0, j = 1, 2, \cdots, d\},$$

$$T\Delta = \{z = (z_1, z_2, ..., z_d) \in \mathbb{R}^d : z_1 + z_2 + \cdots + z_d = 0\}.$$

Also, let $\|.\|$ and $\langle .,. \rangle$ denote the Euclidean norm and the Euclidean scalar product on $\mathbb{R}^d$.

For a càdlàg process Y, we denote as $Y_{t-} = \lim_{s \to t-} Y_t$ and $\Delta Y_t = Y_t - Y_{t-}$ respectively the left limit and the size of the jump of Y at time t. Let $[Y]$ be as usual the quadratic variation of the process Y. Note that, for a càdlàg finite variation process Y, we have $[Y]_t - [Y]_s = \sum_{s < u \leq t} (\Delta Y_u)^2$. In the next sections, we will use the following useful well-known results of stochastic calculus (see e.g. [15]):

1. (Change of variables formula) Let A be a càdlàg finite variation process and let f be a C^1 function. Then

$$f(A_t) - f(A_s) = \int_s^t f'(A_{u-}) dA_u + \sum_{s < u \leq t} \left(\Delta f(A_u) - f'(A_{u-}) \Delta A_u \right).$$

2. Let M be a martingale such that $M_t = M_0 + I_t + \int_0^t K_s ds$ with K being some adapted process. Suppose that I is a pure jump process i.e. $I_t = \sum_{0 < s \leq t} \Delta I_s$. Then, for f a C^1 function,

$$f(M_t) = f(M_0) + \int_0^t f'(M_s) K_s ds + \sum_{0 < s \leq t} \Delta f(M_s).$$

In particular,

$$M_t^2 = M_0^2 + 2 \int_0^t M_s K_s ds + \sum_{0 < s \leq t} \Delta M_s^2.$$

If $N_t = \int_0^t H_s dM_s$, with H a bounded predictable process, then

$$N_t = \sum_{0 < s \leq t} H_s \Delta I_s + \int_0^t H_s K_s ds.$$

It is also a martingale,

$$[N]_t = \int_0^t H_s^2 d[M]_s$$

and

$$\mathbb{E}[N_t^2] = \mathbb{E}\left[\sum_{0<s\leq t}(\Delta N_s)^2\right] = \mathbb{E}\left[\sum_{0<s\leq t}H_s^2(\Delta I_s)^2\right] = \mathbb{E}\left[\int_0^t H_s^2 d[M]_s\right] = \mathbb{E}[[N]_t],$$

where we recall that $[M]_t = \sum_{0<s\leq t}(\Delta I_s)^2$.

3. (Integration by part formula) Let X and Y be two càdlàg finite variation processes. Then

$$X_t Y_t - X_s Y_s = \int_s^t X_{u-} dY_u + \int_s^t Y_{u-} dX_u + [X,Y]_t - [X,Y]_s.$$

4. (Doob's Inequality) Let X be a càdlàg martingale. Then for any $p > 1$,

$$\mathbb{E}[\sup_{s\leq t}|X_s|^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_t|^p].$$

5. (Burkholder-Davis-Gundy Inequality) Let X be a martingale. For each $1 \leq p < \infty$ there exist positive constants c_p and C_p depending on only p such that

$$c_p\mathbb{E}\left[[X]_t^{p/2}\right] \leq \mathbb{E}\left[\sup_{s\leq t}|X_s|^p\right] \leq C_p\mathbb{E}\left[[X]_t^{p/2}\right].$$

Dynamics of occupation measure process

For $t > 0$ which is not a jumping time of X_t, we have

$$\frac{dZ_t}{dt} = \frac{1}{\ell_0 + t}\left(-Z_t + I[X_t]\right), \tag{B.1}$$

where for each matrix M, $M[j]$ is the j-th row vector of M and I is as usual the identity matrix. Observe that the process Z_t always takes values in the interior of the standard unit simplex Δ.

For fixed $t \geq 0$, let A_t be an infinitesimal generator matrix such that the (i,j) element is defined by

$$
(A_t)_{i,j} := \begin{cases} \mathbf{1}_{(i,j)\in E} w_t^{(j)}, & i \neq j; \\ -\displaystyle\sum_{k\in V, (k,i)\in E} w_t^{(k)}, & i = j, \end{cases}
$$

where we define $w_t^{(j)} = w(L(j,t)) = L(j,t)^\alpha$ for each $j \in V$. Also, let $w_t = w_t^{(1)} + w_t^{(2)} + \cdots + w_t^{(d)}$. Note that

$$
\pi_t := \left(\frac{w_t^{(1)}}{w_t}, \frac{w_t^{(2)}}{w_t}, \cdots, \frac{w_t^{(d)}}{w_t} \right)
$$

is the unique invariant probability measure of A_t in the sense that $\pi_t A_t = \pi_t$. Since π_t can be rewritten as a function of Z_t, we will also use the notation $\pi_t = \pi(Z_t)$, where we define the function $\pi : \Delta \to \Delta$, such that for each $z = (z_1, z_2, ..., z_d) \in \Delta$,

$$
\pi(z) = \left(\frac{z_1^\alpha}{z_1^\alpha + \cdots + z_d^\alpha}, \cdots, \frac{z_d^\alpha}{z_1^\alpha + \cdots + z_d^\alpha} \right).
$$

Now we can rewrite the equation (B.1) as

$$
\frac{dZ_t}{dt} = \frac{1}{\ell_0 + t}(-Z_t + \pi_t) + \frac{1}{\ell_0 + t}(I[X_t] - \pi_t). \tag{B.2}
$$

Changing variable $\ell_0 + t = e^u$ and denoting $\tilde{Z}_u = Z_{e^u - \ell_0}$ for $u > 0$, we can transform the equation (B.2) as

$$
\frac{d\tilde{Z}_u}{du} = -\tilde{Z}_u + \pi(\tilde{Z}_u) + (I[X_{e^u - \ell_0}] - \pi_{e^u - \ell_0}).
$$

Taking integral of both sides, we obtain that

$$
\tilde{Z}_{t+s} - \tilde{Z}_t = \int_t^{t+s} \left(-\tilde{Z}_u + \pi(\tilde{Z}_u) \right) du + \int_{e^t - \ell_0}^{e^{t+s} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du. \tag{B.3}
$$

Lemma 3.1. *For $t \geq 0$, the process $M_t = (M_t^1, M_t^2, \cdots, M_t^d)$ defined by*

$$
M_t = I[X_t] - \int_0^t A_s[X_s] ds
$$

is a martingale in $\mathbb{R}^d$.

Proof. For small $h > 0$, we have

$$\mathbb{E}\left(I[X_{t+h}] - I[X_t] \mid \mathcal{F}_t\right) = \sum_{j \sim X_t} (I[j] - I[X_t])\, \mathbb{P}(X_{t+h} = j \mid \mathcal{F}_t)$$

$$= \sum_{j \sim X_t} (I[j] - I[X_t])\, w_t^j.h + o(h)$$

$$= A_t[X_t].h + o(h).$$

Let us fix $0 < s < t$ and define $t_j = s + j(t-s)/n$ for $j = 0, 1, ..., n$. Note that

$$\mathbb{E}\left(I[X_t] - I[X_s] \mid \mathcal{F}_s\right) = \mathbb{E}\left(\sum_{j=1}^{n} \mathbb{E}(I[X_{t_j}] - I[X_{t_{j-1}}] \mid \mathcal{F}_{t_{j-1}}) \mid \mathcal{F}_s\right)$$

$$= \mathbb{E}\left(\sum_{j=1}^{n} A_{t_{j-1}}[X_{t_{j-1}}](t_j - t_{j-1}) + n \cdot o\left(\frac{t-s}{n}\right) \mid \mathcal{F}_s\right).$$

Since the left hand side is independent on n, by using Lebesgue's dominated convergence theorem and taking the limit of the random sum under the expectation sign on the right hand side, we obtain that

$$\mathbb{E}\left(I[X_t] - I[X_s] \mid \mathcal{F}_s\right) = \mathbb{E}\left(\int_s^t A_u[X_u]du \mid \mathcal{F}_s\right).$$

Thus, $\mathbb{E}\left(M_t \mid \mathcal{F}_s\right) = M_s$. $\qquad\square$

Lemma 3.2. *Let $w(t) = t^\alpha$. If $\alpha > 0$ and $G = (V, E)$ is a complete graph, then a.s.*

$$\lim_{t \to \infty} \sup_{1 \le c \le C} \left\| \int_{t-\ell_0}^{ct-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0 + s} ds \right\| = 0 \tag{B.4}$$

for each $C > 1$.

Proof. Note that, for $t \ge 0$,

$$\pi_t - I[X_t] = \frac{1}{w_t} A_t[X_t].$$

By taking integration by part, we obtain the following identity for each $c \in$

$[1, C]$

$$\int_{t-\ell_0}^{ct-\ell_0} \frac{\pi_s - I[X_s]}{\ell_0 + s} ds = \int_{t-\ell_0}^{ct-\ell_0} A_s[X_s] \frac{ds}{(\ell_0 + s)w_s}$$

$$= \left(\frac{I[X_{ct-\ell_0}]}{ctw_{ct-\ell_0}} - \frac{I[X_{t-\ell_0}]}{tw_{t-\ell_0}} \right)$$

$$- \int_{t-\ell_0}^{ct-\ell_0} I[X_s] \frac{d}{ds} \left(\frac{1}{(s+\ell_0)w_s} \right) ds$$

$$- \int_{t-\ell_0}^{ct-\ell_0} \frac{dM_s}{(s+\ell_0)w_s}.$$

Observe that for some positive constant k, $w_s \geq ks^\alpha$ (which is easy to prove, using the fact that $L(1,t) + L(2,t) + \cdots + L(d,t) = \ell_0 + t$). We now estimate the terms in the right hand side of the above-mentioned identity. In the following, the positive constant k may change from lines to lines and only depends on C and ℓ_0. First,

$$\left\| \frac{I[X_{ct-\ell_0}]}{ctw_{ct-\ell_0}} - \frac{I[X_{t-\ell_0}]}{tw_{t-\ell_0}} \right\| \leq k/t^{\alpha+1}. \tag{B.5}$$

Second,

$$\frac{d}{ds} \left(\frac{1}{(\ell_0 + s)w_s} \right) = - \left(\frac{1}{(\ell_0 + s)^2 w_s} + \frac{1}{(\ell_0 + s)w_s^2} \frac{dw_s}{ds} \right).$$

When s is not a jump time, it is easy to check that $\left| \frac{dw_s}{ds} \right| \leq \alpha(\ell_0 + s)^{\alpha-1}$. Therefore, for $s \in [t, ct]$,

$$\left\| \frac{d}{ds} \left(\frac{1}{(\ell_0 + s)w_s} \right) \right\| \leq k/s^{2+\alpha}$$

and thus,

$$\left\| \int_{t-\ell_0}^{ct-\ell_0} I[X_s] \frac{d}{ds} \left(\frac{1}{(\ell_0 + s)w_s} \right) ds \right\| \leq k/t^{\alpha+1}. \tag{B.6}$$

And at last (using first Doob's inequality), for $i, j \in \{1, 2, \cdots, d\}$,

$$\mathbb{E} \left[\sup_{1 \leq c \leq C} \left| \int_{t--\ell_0}^{ct-\ell_0} \frac{dM_s^i}{(\ell_0 + s)w_s} \right|^2 \right] \leq 4 \, \mathbb{E} \left[\left(\int_{t-\ell_0}^{Ct-\ell_0} \frac{dM_s^i}{(\ell_0 + s)w_s} \right)^2 \right].$$

Observe that in our setting, for $i \in \{1, 2, \cdots, d\}$, $(\Delta I_s^i)^2 = 1$ if s is a jump time between i and another vertex. Thus $[M^1]_t + [M^2]_t + \cdots + [M^d]_t$ is just twice

the number of jumps up to time t of X. So, for $i, j \in \{1, 2, \cdots, d\}$,

$$
\begin{aligned}
\mathbb{E}\left[\left(\int_{t-\ell_0}^{Ct-\ell_0} \frac{dM_s^i}{(\ell_0+s)w_s}\right)^2\right] &= \mathbb{E}\left[\int_{t-\ell_0}^{Ct-\ell_0} \frac{d[M^i]_s}{(\ell_0+s)^2 w_s^2}\right] \\
&\leq \frac{k}{t^{2(\alpha+1)}} \mathbb{E}\left[[M^i]_{Ct-\ell_0} - [M^i]_{t-\ell_0}\right] \\
&\leq \frac{k}{t^{2(\alpha+1)}}(Ct)^\alpha (C-1)t,
\end{aligned}
$$

where in the last inequality, we have used the fact that the number of jumps in $[t-\ell_0, Ct-\ell_0]$ is dominated by the number of jumps of a Poisson process with constant intensity $(Ct)^\alpha$ in $[t-\ell_0, Ct-\ell_0]$. Therefore,

$$
\mathbb{E}\left[\sup_{1 \leq c \leq C}\left\|\int_{t-\ell_0}^{ct-\ell_0} \frac{dM_s}{(\ell_0+s)w_s}\right\|^2\right] \leq \frac{k}{t^{\alpha+1}}. \tag{B.7}
$$

From (B.5), (B.6), (B.7) and by using Markov's inequality, we have

$$
\mathbb{P}\left(\sup_{1 \leq c \leq C}\left\|\int_{t-\ell_0}^{ct-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| \geq \frac{1}{t^\gamma}\right) \leq \frac{k}{t^{\alpha+1-2\gamma}} \tag{B.8}
$$

for every $0 < \gamma \leq \frac{\alpha+1}{2}$. By Borel-Cantelli lemma,

$$
\limsup_{n \to \infty} \sup_{1 \leq c \leq C}\left\|\int_{C^n-\ell_0}^{cC^n-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| = 0.
$$

Moreover, for $C^n \leq t \leq C^{n+1}$, we have

$$
\begin{aligned}
\sup_{1 \leq c \leq C}\left\|\int_{t-\ell_0}^{ct-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| &\leq \left\|\int_{C^n-\ell_0}^{t-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| \\
&+ \sup_{1 \leq c \leq C}\left\|\int_{C^n-\ell_0}^{\min(ct,C^{n+1})-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| \\
&+ \sup_{1 \leq c \leq C}\left\|\int_{C^{n+1}-\ell_0}^{\max(ct,C^{n+1})-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| \\
&\leq 2\sup_{1 \leq c \leq C}\left\|\int_{C^n-\ell_0}^{cC^n-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\| + \sup_{1 \leq c \leq C}\left\|\int_{C^{n+1}-\ell_0}^{cC^{n+1}-\ell_0} \frac{I[X_s] - \pi_s}{\ell_0+s}ds\right\|.
\end{aligned}
$$

This inequality immediately implies (B.4). $\qquad\square$

84

From now on, we always assume that $w(t) = t^\alpha, \alpha > 1$ and $G = (V, E)$ is a complete graph. Let us define the vector field $F : \Delta \to T\Delta$ such that $F(z) = -z + \pi(z)$ for each $z \in \Delta$. We also remark that for each $z = (z_1, z_2, \cdots, z_d) \in \Delta$,

$$F(z) = \left(-z_1 + \frac{z_1^\alpha}{z_1^\alpha + \cdots + z_d^\alpha}, \cdots, -z_d + \frac{z_d^\alpha}{z_1^\alpha + \cdots + z_d^\alpha} \right). \qquad \text{(B.9)}$$

A continuous map $\Phi : \mathbb{R}_+ \times \Delta \to \Delta$ is called a *semi-flow* if $\Phi(0, .) : \Delta \to \Delta$ is the identity map and Φ has the semi-group property, i.e. $\Phi(t + s, .) = \Phi(t, .) \circ \Phi(s, .)$ for all $s, t \in \mathbb{R}_+$.

Now for each $z^0 \in \Delta$, let $\Phi_t(z^0)$ be the solution of the differential equation

$$\begin{cases} \dfrac{d}{dt} z(t) = & F(z(t)), \ t > 0; \\ z(0) = & z^0. \end{cases} \qquad \text{(B.10)}$$

Note that F is Lipschitz. Thus the solution $\Phi_t(z^0)$ can be extended for all $t \in \mathbb{R}_+$ and $\Phi : \mathbb{R}_+ \times \Delta \to \Delta$ defined by $\Phi(t, z) = \Phi_t(z)$ is a semi-flow.

Theorem 3.3. $\tilde{Z}$ *is an asymptotic pseudo-trajectory of the semi-flow* Φ*, i.e. for all* $T > 0$*, a.s.*

$$\lim_{t \to \infty} \sup_{0 \leq s \leq T} \left\| \tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t) \right\| = 0. \qquad \text{(B.11)}$$

Proof. Indeed, by the definition of Φ

$$\Phi_s(\tilde{Z}_t) - \tilde{Z}_t = \int_0^s F(\Phi_u(\tilde{Z}_t)) du.$$

Moreover, from (B.3)

$$\tilde{Z}_{t+s} - \tilde{Z}_t = \int_0^s F(\tilde{Z}_{t+u}) du + \int_{e^t - \ell_0}^{e^{t+s} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du.$$

Subtracting both sides of the two above identities, we obtain that

$$\tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t) = \int_0^s \left(F(\tilde{Z}_{t+u}) - F(\Phi_u(\tilde{Z}_t)) \right) du + \int_{e^t - \ell_0}^{e^{t+s} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du.$$

Observe that F is Lipschitz, hence

$$\| \tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t) \| \leq K \int_0^s \| \tilde{Z}_{t+u} - \Phi_u(\tilde{Z}_t) \| du + \left\| \int_{e^t - \ell_0}^{e^{s+t} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\|,$$

85

where K is the Lipschitz constant of F. Therefore, by the Grönwall's inequality,

$$\|\tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t)\| \leq \sup_{0 \leq s \leq T} \left\| \int_{e^t - \ell_0}^{e^{s+t} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| e^{Ks}. \qquad \text{(B.12)}$$

Finally, (B.11) follows from the inequality (B.12) and Lemma 1 that a.s.

$$\lim_{t \to \infty} \sup_{0 \leq s \leq T} \left\| \int_{e^t - \ell_0}^{e^{s+t} - \ell_0} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| = 0.$$

$\square$

Convergence to equilibria

Let
$$\mathcal{C} = \{z \in \Delta : F(z) = 0\}$$

stand for the *equilibria set* of the vector field F defined in (B.9). We say an equilibrium $z \in \mathcal{C}$ is (linearly) *stable* if all the eigenvalues of $DF(z)$, the Jacobian matrix of F at z, have negative real parts. If there is one of its eigenvalues having a positive real part, then it is called (linearly) *unstable*.

Observe that $\mathcal{C} = \mathcal{S} \cup \mathcal{U}$, where we define

$$\mathcal{S} = \{e_1 = (1, 0, 0, \cdots, 0), e_2 = (0, 1, 0, \cdots, 0), \cdots, e_d = (0, 0, \cdots, 0, 1)\}$$

as the set of all stable equilibria and

$$\mathcal{U} = \{z_{j_1, j_2, \cdots, j_k} : 1 \leq j_1 < j_2 < \cdots < j_k \leq d, k = 2, \cdots, d\}$$

as the set of all unstable equilibria, where $z_{j_1, j_2, \cdots, j_k}$ stands for the point $z = (z_1, \cdots, z_d) \in \Delta$ such that $z_{j_1} = z_{j_2} = \cdots = z_{j_k} = \frac{1}{k}$ and all the remaining coordinates are equal to 0.

Indeed, for each $z \in \mathcal{S}$, we note that $DF(z) = -I$. Moreover,

$$DF\left(\frac{1}{d}, \frac{1}{d}, \cdots, \frac{1}{d}\right) = \begin{pmatrix} \frac{(d-1)\alpha}{d} - 1 & -\frac{\alpha}{d} & \cdots & -\frac{\alpha}{d} \\ -\frac{\alpha}{d} & \frac{(d-1)\alpha}{d} - 1 & \cdots & -\frac{\alpha}{d} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{\alpha}{d} & -\frac{\alpha}{d} & \cdots & \frac{(d-1)\alpha}{d} - 1 \end{pmatrix}.$$

and $DF(z_{j_1, j_2, \cdots, j_k}) = (D_{m,n})$ where

$$
D_{m,n} = \begin{cases}
\frac{(k-1)\alpha}{k} - 1, & m = n = j_i, i = 1, \cdots, k; \\
-1, & m = n, m \neq j_i, i = 1, \cdots, k; \\
0, & m \neq n, m \text{ or } n \neq j_i, i = 1, \cdots, k; \\
-\frac{\alpha}{k}, & \text{otherwise.}
\end{cases}
$$

Therefore, we can easily compute that for each $z \in \mathcal{U}$, the eigenvalues of $DF(z)$ are $-1, \alpha - 1$.

Theorem 4.1. Z_t converges almost surely to a point in $\mathcal{C}$ as $t \to \infty$.

Proof. Consider the the map $H : \Delta \to \mathbb{R}$ such that

$$
H(z) = z_1^\alpha + z_2^\alpha + \cdots + z_n^\alpha.
$$

Note that H is a strict Lyapunov function of F, i.e $\nabla H(z).F(z)^T$ is positive for all $z \in \Delta \setminus \mathcal{C}$. Indeed, we have

$$
\nabla H(z).F(z)^T = \sum_{i=1}^d \alpha z_i^{\alpha-1} \left(-z_i + \frac{z_i^\alpha}{\sum_{j=1}^d z_j^\alpha} \right)
$$

$$
= \alpha \left(-\sum_{i=1}^d z_i^\alpha + \frac{\sum_{i=1}^d z_i^{2\alpha-1}}{\sum_{i=1}^d z_i^\alpha} \right)
$$

$$
= \frac{\alpha}{H(z)} \left(-\left(\sum_{i=1}^d z_i^\alpha \right)^2 + \sum_{i=1}^d z_i^{2\alpha-1} \sum_{i=1}^d z_i \right)
$$

$$
= \frac{\alpha}{H(z)} \sum_{1 \leq i < j \leq d} z_i z_j \left(z_i^{\alpha-1} - z_j^{\alpha-1} \right)^2.
$$

For $z \in \Delta \setminus \mathcal{C}$, there exist distinct indexes $j_1, j_2 \in \{1, 2, ..., d\}$ such that z_{j_1}, z_{j_2} are positive and $z_{j_1} \neq z_{j_2}$. Therefore,

$$
\nabla H(z).F(z)^T \geq \frac{\alpha}{H(z)} z_{j_1} z_{j_2} \left(z_{j_1}^{\alpha-1} - z_{j_2}^{\alpha-1} \right)^2 > 0.
$$

Let

$$
L(Z) = \bigcap_{t \geq 0} \overline{Z([t, \infty))}
$$

be limit set of Z. Since $\tilde{Z}$ is an asymptotic pseudo-trajectory of Φ, by Theorem 5.7 and Proposition 6.4 in [4], we can conclude that $L(Z) = L(\tilde{Z})$ is a connected subset of $\mathcal{C}$. Moreover, $\mathcal{C}$ is actually an isolated set and this fact implies the almost sure convergence of Z_t toward an equilibrium $z \in \mathcal{C}$ as $t \to \infty$. $\qquad \square$

Theorem 4.2. *Assume that $\alpha > 1$, $G = (V, E)$ is a complete graph and $z^* = e_j \in \mathcal{S}$ is a stable equilibrium, where $j \in \{1, 2, ..., d\}$. Then, on the event $\{Z_t \to z^*\}$,*

$$\limsup_{t \to \infty} t\|Z_t - z_*\| < \infty,$$

thus implies that the local time at j is unbounded, while the local times at the remaining vertices remain bounded.

Proof. We observe that

$$F(z) = (z - z^*).DF(z^*) + R(z - z^*),$$

where

$$R(y) = y.\left(\int_0^1 DF(ty + z^*)dt - DF(z^*)\right).$$

Note that $\|R(y)\| \leq k\|y\|^{1+\beta}$, where $\beta = \min(1, \alpha - 1)$ and k is some positive constant. Therefore, we can transform the differential equation (B.10) to the following integral form

$$z(t) - z^* = (z(0) - z^*)e^{DF(z^*)t} + \int_0^t R(z(s) - z^*)e^{DF(z^*)(t-s)}ds.$$

Note that for $z^* \in \mathcal{S}$, we have $DF(z^*) = -I$. Therefore,

$$\|z(t) - z^*\| \leq e^{-t}\|z(0) - z^*\| + \int_0^t e^{-(t-s)}\|R(z(s) - z^*)\|ds.$$

For each small $\epsilon > 0$, if $\|z(s) - z^*\| \leq \left(\frac{\epsilon}{k}\right)^{1/\beta}$ for all $0 \leq s \leq t$, then

$$e^t\|z(t) - z^*\| \leq \|z(0) - z^*\| + \epsilon \int_0^t e^s\|z(s) - z^*\|ds.$$

Thus, by Gronwall inequality, if $\|z(s) - z^*\| \leq \left(\frac{\epsilon}{k}\right)^{1/\beta}$ for all $0 \leq s \leq t$, then

$$\|z(t) - z^*\| \leq \|z(0) - z^*\|e^{-(1-\epsilon)t}.$$

But this is also implies that if $\|z(0) - z^*\| \leq \left(\frac{\epsilon}{k}\right)^{1/\beta}$ then $\|z(t) - z^*\| \leq \left(\frac{\epsilon}{k}\right)^{1/\beta}$ for all $t \geq 0$. Hence, for all $t \geq 0$ and any small $\epsilon > 0$ and $z(0)$ such that $\|z(0) - z^*\| \leq \left(\frac{\epsilon}{k}\right)^{1/\beta}$, we have

$$\|z(t) - z^*\| \leq e^{-(1-\epsilon)t}\|z(0) - z^*\|. \tag{B.13}$$

On the other hand, from the inequality (B.12) in the proof of Lemma 3, we obtain that for $T > 0$,

$$\sup_{0 \le s \le T} \|\tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t)\| \le k \sup_{0 \le s \le T} \left\| \int_{e^t}^{e^{s+t}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\|, \tag{B.14}$$

where k is some positive constant depending on T and may change from lines to lines. From (B.8), we have

$$\mathbb{P}\left(\sup_{0 \le s \le T} \left\| \int_{e^t}^{e^{s+t}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| \ge e^{-\gamma t} \right) \le k e^{-(\alpha + 1 - 2\gamma)t},$$

for every $0 < \gamma \le \frac{\alpha+1}{2}$. By Borel-Cantelli lemma, it implies that

$$\limsup_{n \to \infty} \frac{1}{nT} \log \left(\sup_{0 \le s \le T} \left\| \int_{e^{nT}}^{e^{s+nT}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| \right) \le -\gamma \quad \text{a.s.}$$

Taking $\gamma \to \frac{\alpha+1}{2}$, by the continuity from the right of cumulative distribution functions, we obtain that

$$\limsup_{n \to \infty} \frac{1}{nT} \log \left(\sup_{0 \le s \le T} \left\| \int_{e^{nT}}^{e^{s+nT}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| \right) \le -\frac{\alpha+1}{2} \quad \text{a.s.}$$

Note that for $nT \le t \le (n+1)T$ and $0 \le s \le T$,

$$\left\| \int_{e^t}^{e^{s+t}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| \le 2 \sup_{0 \le s \le T} \left\| \int_{e^{nT}}^{e^{s+nT}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\|$$

$$+ \sup_{0 \le s \le T} \left\| \int_{e^{(n+1)T}}^{e^{s+(n+1)T}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\|.$$

Therefore,

$$\limsup_{t \to \infty} \frac{1}{t} \log \left(\sup_{0 \le s \le T} \left\| \int_{e^t}^{e^{s+t}} \frac{I[X_u] - \pi_u}{\ell_0 + u} du \right\| \right) \le -\frac{\alpha+1}{2} \quad \text{a.s.} \tag{B.15}$$

From (B.14) and (B.15), we obtain that

$$\limsup_{t \to \infty} \frac{1}{t} \log \left(\sup_{0 \le s \le T} \|\tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t)\| \right) \le -\frac{\alpha+1}{2} \quad \text{a.s.}$$

Moreover, from (B.13), for each $\epsilon > 0$ there exist $\delta(\epsilon) > 0$ such that

$$\|\Phi_s(z) - z^*\| \le e^{-(1-\epsilon)s}\|z - z^*\|$$

for all $s > 0$ and $z \in B_{\delta(\epsilon)}(z^*) = \{z \in \Delta : \|z - z^*\| < \epsilon\}$. Therefore, we can conclude that on the event $\{Z_t \to z^*\}$ there exists $t(\epsilon)$ such that $\tilde{Z}_t \in B_{\delta(\epsilon)}(z^*)$ for all $t \geq t(\epsilon)$ and furthermore,

$$
\begin{aligned}
\|\tilde{Z}_{t+s} - z^*\| &\leq \|\tilde{Z}_{t+s} - \Phi_s(\tilde{Z}_t)\| + \|\Phi_s(\tilde{Z}_t) - z^*\| \\
&\leq e^{-(\frac{\alpha+1}{2} - \epsilon)t} + e^{-(1-\epsilon)s} \|\tilde{Z}_t - z^*\| \\
&\leq e^{-(1-\epsilon)t} + e^{-(1-\epsilon)s} \|\tilde{Z}_t - z^*\|
\end{aligned}
$$

for all $s > 0$. It implies that

$$
\limsup_{t \to \infty} \frac{1}{t} \log \|\tilde{Z}_t - z^*\| \leq -1 + \epsilon
$$

for arbitrary $\epsilon > 0$. Taking $\epsilon \to 0$, we obtain that

$$
\limsup_{t \to \infty} \frac{1}{t} \log \|\tilde{Z}_t - z^*\| \leq -1.
$$

This also implies that on the event $\{Z_t \to z^*\}$, we have

$$
\|Z_t - z^*\| = O\left(\frac{1}{t}\right)
$$

as $t \to \infty$. Therefore, if Z_t converges to a stable equilibrium, then there is only one vertex which eventually has infinite local time. $\qquad\square$

Nonconvergence to unstable equilibria

On graph with two vertices

Let us consider strongly VRJP on the graph $G = (V, E)$, where $V = \{1, 2\}$, $E = \{(1, 2)\}$ and assume that $w(t) = t^\alpha, \alpha > 1$. Note that, $z^* = (1/2, 1/2)$ is the unique unstable equilibrium of the vector field F defined in (B.9) for $d = 2$.

Throughout this subsection, we use the notation Const for the existence of some positive constants and they may vary from line to line. We will use the next results, which are inspired by the work of Brandière and Duflo in [1] for discrete time processes.

Lemma 5.1. *Assume that φ_t is a $\mathcal{F}_t$-adapted process satisfying the following integral equation*

$$\varphi_t = \eta + \int_p^t \left(\lambda \gamma(u) \varphi_{u-} du + c(u) dA_u \right), \ t \geq p,$$

where $\lambda > 0$; p is not a jump time; η is a $\mathcal{F}_p$-measurable random variable; γ and c are deterministic continuous real functions such that $\lim_{t\to\infty} \int_p^t \gamma(u) du = +\infty$; A_t are càdlàg finite variation $\mathcal{F}_t$-adapted processes. Assume also that a.s. $\lim_{t\to\infty} \varphi_t = \varphi^* < +\infty$. Then

$$\int_p^\infty c(t) e^{-\lambda \int_p^t \gamma(u) du} dA_t = -\eta \ a.s.$$

Proof. Applying the integration by part formula, we note that the process

$$\varphi_t = \left(\int_p^t c(s) e^{-\lambda \int_p^s \gamma(u) du} dA_s + \eta \right) e^{\lambda \int_p^t \gamma(u) du}$$

is the unique solution to the above-mentioned stochastic integral equation. Observe that, $e^{\lambda \int_p^t \gamma(u) du} \to +\infty$ and $\varphi_t \to \varphi^* < +\infty$ almost surely as $t \to \infty$. Thus,

$$\int_p^t c(s) e^{-\lambda \int_p^s \gamma(u) du} dA_s + \eta \to 0 \text{ a.s.}$$

as $t \to \infty$. □

Theorem 5.2. *Let F be a differentiable bounded function and z^* be an unstable equilibrium of F, i.e $F(z^*) = 0$ and $F'(z^*) > 0$. Suppose that F' is Lipschitz continuous in a neighbourhood of z^*. Let us consider the finite variation process Z_t satisfying the following equation*

$$Z_t - Z_s = \int_s^t \gamma(u) F(Z_u) du + \int_s^t c(u)(G_u du + dN_u)$$

where N_t is a martingale w.r.t $\mathcal{F}_t = \sigma(Z_s, s \leq t)$ and G_t is a $\mathcal{F}_t$-adapted process; γ and c are deterministic positive continuous functions such that $\int_0^\infty \gamma(t) dt = \infty$, $\int_0^\infty \gamma(t)^2 dt < \infty$, $\int_0^\infty c(t)^2 dt < \infty$.

Assume that on the event $\Gamma = \{\lim_{t\to\infty} Z_t = z^*\}$, for every $t \geq 0$

$$\frac{\gamma(t)}{c(t)} |\Delta Z_t| \leq \text{Const}|G_t|, \tag{B.16}$$

$$\int_0^\infty G_t^2 dt < \infty, \tag{B.17}$$

$$\mathbb{E} \left(\mathbf{1}_\Gamma \sum_{t < u \leq t+1} |\Delta Z_u|^2 \right) \leq \text{Const.} c^2(t), \tag{B.18}$$

$$\limsup_{t \to \infty} \mathbb{E} \left(\mathbf{1}_\Gamma |N_{t+h} - N_t|^2 \mid \mathcal{F}_t \right) < \text{Const.} h, \tag{B.19}$$

and

$$\liminf_{t \to \infty} \mathbb{E} \left(\mathbf{1}_\Gamma |N_{t+h} - N_t|^2 \mid \mathcal{F}_t \right) > \text{Const.} h \tag{B.20}$$

for every $h > 0$. Then $\mathbb{P}(\Gamma) = 0$.

Proof. 1. Define $\varphi(z) = z e^{\psi(z)}$, where

$$\psi(z) = \int_0^z \frac{\lambda - \int_0^1 F'(tu + z^*) dt}{u \int_0^1 F'(tu + z^*) dt} du.$$

Note that φ is differentiable and its derivative φ' is Lipschitz in a neighbourhood W of 0. Moreover, $\varphi(0) = 0, \varphi'(0) = 1$. We denote $W(z^*) = z^* + W$ and thus $W(z^*)$ is a neighbourhood of z^*.

2. By the change of variables formula for finite variation processes, we have

$$\varphi(Z_t - z^*) - \varphi(Z_s - z^*) = \int_s^t \varphi'(Z_{u-} - z^*) dZ_u$$
$$+ \sum_{s < u \leq t} \left(\Delta \varphi(Z_u - z^*) - \varphi'(Z_{u-} - z^*) \Delta Z_u \right).$$

Define $\varphi_t = \varphi(Z_t - z^*)$, we obtain that

$$\varphi_t - \varphi_s = \int_s^t \varphi'(Z_{u-} - z^*) \left[\gamma(u) F(Z_u) du + c(u)(G_u du + dN_u) \right] + c(u) dH_u,$$

where

$$H_t = \sum_{0 < u \leq t} \frac{1}{c(u)} \left(\Delta \varphi(Z_u - z^*) - \varphi'(Z_{u-} - z^*) \Delta Z_u \right).$$

Therefore, on the event $\Gamma_p = \{ Z_t \in W(z^*) \text{ for all } t \geq p \}$, for $p \leq s < t$,

$$\varphi_t - \varphi_s = \lambda \int_s^t \gamma(u) \varphi_{u-} du + c(u)(\tilde{G}_u du + \varphi'(Z_{u-} - z^*) dN_u + dH_u),$$

where

$$\tilde{G}_t = -\frac{\gamma(t)}{c(t)} \varphi'(Z_{t-} - z^*) \Delta F(Z_t) + \varphi'(Z_{t-} - z^*) G_t.$$

Since F is Lipschitz on $W(z^*)$, on the event Γ_p, we remark from (B.16) that

$$|\tilde{G}_t| \leq \text{Const} \frac{\gamma(t)}{c(t)} |\Delta Z_t| + \text{Const}|G_t| \leq \text{Const}|G_t|.$$

Therefore, from (B.17)

$$\mathbb{E}\left(\mathbf{1}_{\Gamma_p} \int_t^\infty \tilde{G}_u^2 du\right) \to 0, \text{ as } t \to \infty. \tag{B.21}$$

By Lemma 5.1, on the event Γ_p we obtain that

$$\int_p^\infty k(t)(\tilde{G}_t dt + \varphi'(Z_{t-} - z^*)dN_t + dH_t) = -\varphi_p, \tag{B.22}$$

where $k(t) = e^{-\lambda \int_p^t \gamma(u)du} c(t)$. We also note that

$$\mathbb{E}\left|\mathbf{1}_{\Gamma_p} \int_t^\infty k(u)dH_u\right|$$
$$= \mathbb{E}\left|\mathbf{1}_{\Gamma_p} \sum_{t<u<\infty} e^{-\lambda \int_p^u \gamma(s)ds} \left(\Delta\varphi(Z_u - z^*) - \varphi'(Z_{u-} - z^*)\Delta Z_u\right)\right|.$$

Note that φ' is a Lipschitz function on W, therefore

$$\left|\Delta\varphi(Z_u - z^*) - \varphi'(Z_{u-} - z^*)\Delta Z_u\right|$$
$$= \left|\int_0^1 \varphi'(Z_{u-} - z^* + t\Delta Z_u)dt - \varphi'(Z_{u-} - z^*)\right|.|\Delta Z_u|$$
$$\leq \text{Const.}|\Delta Z_u|^2.$$

On the other hand, it implies from (B.18) that

$$\mathbb{E}\left(\mathbf{1}_{\Gamma_p} \sum_{t<u<\infty} e^{-\lambda \int_p^u \gamma(s)ds}|\Delta Z_u|^2\right)$$
$$= \sum_{n\geq 0} e^{-\lambda \int_p^{t+n} \gamma(s)ds} \mathbb{E}\left(\mathbf{1}_{\Gamma_p} \sum_{t+n<u\leq t+n+1} |\Delta Z_u|^2\right)$$
$$\leq \text{Const.} \sum_{n\geq 0} e^{-\lambda \int_p^{t+n} \gamma(s)ds} c^2(t+n) \leq \text{Const.} \int_t^\infty e^{-\lambda \int_p^u \gamma(s)ds} c(u)^2 du$$
$$\leq \text{Const.} \left(\int_t^\infty \left(e^{-\lambda \int_p^u \gamma(s)ds} c(u)\right)^2 du\right)^{1/2} \left(\int_t^\infty c(u)^2 du\right)^{1/2}.$$

Since $\int_t^\infty c(u)^2 du \to 0$ as $t \to \infty$, we obtain that

$$\mathbb{E} \left| \mathbf{1}_{\Gamma_p} \int_t^\infty k(u) dH_u \right| = o \left(\int_t^\infty k^2(u) du \right)^{1/2}. \tag{B.23}$$

2. Denote

$$S_t = \varphi_p + \int_p^t k(u)(\varphi'(Z_{u-} - z^*) dN_u + \tilde{G}_u du + dH_u),$$

$$T_t = \int_t^\infty k(u)(\varphi'(Z_{u-} - z^*) dN_u + \tilde{G}_u du + dH_u)$$

and

$$\tau_t = \int_t^\infty k(u) \left(\tilde{G}_u du + dH_u + (\varphi'(Z_{u-} - z^*) - 1) dN_u \right), \rho_t = \int_t^\infty k(u) dN_u.$$

We claim that

$$\mathbb{E} \left(|\rho_t|^2 \mid \mathcal{F}_t \right) \leq \text{Const.} \alpha_t^2 \tag{B.24}$$

and

$$\mathbb{E} \left(|\rho_t| \mid \mathcal{F}_t \right) \geq \text{Const.} \alpha_t, \tag{B.25}$$

where $\alpha_t = \left(\int_t^\infty k^2(u) du \right)^{1/2}$. Indeed, the inequality (B.24) immediately follows from the assumption (B.19). On the other hand, by Burkholder-Davis-Gundy inequality, Doob's inequality and the assumption (B.20)

$$\mathbb{E} \left(|\rho_t|^{3/2} \mid \mathcal{F}_t \right) \geq \frac{1}{\sqrt{9}} \mathbb{E} \left(\sup_{s \leq t} |\rho_s|^{3/2} \mid \mathcal{F}_t \right) \geq \text{Const.} \mathbb{E} \left([\rho]_t^{3/4} \mid \mathcal{F}_t \right)$$

$$= \text{Const.} \mathbb{E} \left(\left(\int_t^\infty k(u)^2 d[N]_u \right)^{3/4} \mid \mathcal{F}_t \right)$$

$$\geq \text{Const.} \mathbb{E} \left(\left(\int_t^\infty k(u)^2 du \right)^{3/4} \mid \mathcal{F}_t \right) = \text{Const.} \alpha_t^{3/2}.$$

Finally, by Hölder inequality we can conclude that

$$\mathbb{E} \left(|\rho_t| \mid \mathcal{F}_t \right) \geq \frac{\left(\mathbb{E} \left(|\rho_t|^{3/2} \mid \mathcal{F}_t \right) \right)^2}{\mathbb{E} \left(|\rho_t|^2 \mid \mathcal{F}_t \right)} \geq \text{Const.} \alpha_t.$$

3. From (B.22), we note that on the event Γ_p,

$$T_t = -S_t.$$

Therefore,

$$||\rho_t| + \text{sign}(S_t)\rho_t| = ||\rho_t| - \text{sign}(T_t)\rho_t| = ||\rho_t| - |T_t| + \text{sign}(T_t)(T_t - \rho_t)|$$
$$\leq 2|(T_t - \rho_t)| = 2|\tau_t|.$$

Denote $\Gamma^{(p)} = \{\lim_{t\to\infty} Z_t = z^*\} \cap \{Z_t \in W(z^*), \text{for all } t \geq p\}$ and $\Gamma_t^{(p)} = \{\mathbb{P}(\Gamma^{(p)}|\mathcal{F}_t) > 1/2\}$. By the inequality (B.25), we observe that

$$\mathbb{P}(\Gamma_t^{(p)}) \leq \text{Const}\alpha_t^{-1} \mathbb{E}\left[\mathbf{1}_{\Gamma_t^{(p)}} \mathbb{E}\left(|\rho_t| \mid \mathcal{F}_t\right)\right]$$
$$= \text{Const}\alpha_t^{-1} \mathbb{E}\left[\mathbf{1}_{\Gamma_t^{(p)}} \mathbb{E}\left(|\rho_t| + \text{sign}(S_t)\rho_t \mid \mathcal{F}_t\right)\right],$$

where the later equality follows from the fact that S_t is $\mathcal{F}_t$-measurable and $\mathbb{E}(\rho_t|F_t) = 0$. Thus,

$$\mathbb{P}(\Gamma_t^{(p)}) \leq \text{Const}\alpha_t^{-1} \mathbb{E}\left[\mathbf{1}_{\Gamma^{(p)}} \mathbb{E}(2|\tau_t| \mid \mathcal{F}_t)\right] + \text{Const}\alpha_t^{-1} \mathbb{E}\left[\left|\mathbf{1}_{\Gamma^{(p)}} - \mathbf{1}_{\Gamma_t^{(p)}}\right| \mathbb{E}(2|\rho_t|) |\mathcal{F}_t\right]$$
$$\leq \text{Const}\alpha_t^{-1} \mathbb{E}\left[\mathbf{1}_{\Gamma^{(p)}}|\tau_t|\right] + \text{Const}\alpha_t^{-1} \mathbb{E}\left[|\mathbf{1}_{\Gamma^{(p)}} - \mathbf{1}_{\Gamma_t^{(p)}}|^2\right]^{1/2} \mathbb{E}\left[|\rho_t|^2\right]^{1/2}.$$

Observe that $\mathbb{E}\left[|\mathbf{1}_{\Gamma^{(p)}} - \mathbf{1}_{\Gamma_t^{(p)}}|^2\right]^{1/2} \to 0$ as $t \to \infty$ and $\mathbb{E}\left[|\rho_t|^2\right] \leq \text{Const.}\alpha_t^2$. On the other hand,

$$\mathbb{E}\left[\mathbf{1}_{\Gamma^{(p)}}|\tau_t|\right] \leq \text{Const.}\alpha_t \left(\mathbb{E}\left(\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty \tilde{G}_u^2 du\right)\right)^{1/2} + \text{Const.} \mathbb{E}\left|\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty k(u)dH_u\right|$$

$$+ \text{Const.} \mathbb{E}\left|\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty k(u)(\varphi'(Z_{u-} - z^*) - 1)dN_u\right|.$$

The first and the second term have order $o(\alpha_t)$ as $t \to \infty$ by (B.21) and (B.23). Moreover,

$$\mathbb{E}\left|\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty k(u)\left(\varphi'(Z_{u-} - z^*) - 1\right)dN_u\right|$$
$$\leq \mathbb{E}\left(\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty k^2(u)\left(\varphi'(Z_{u-} - z^*) - 1\right)^2 d[N]_u\right)^{1/2}$$
$$\leq \text{Const} \mathbb{E}\left(\mathbf{1}_{\Gamma^{(p)}} \int_t^\infty k^2(u)\left(\varphi'(Z_{u-} - z^*) - 1\right)^2 du\right)^{1/2} = o(\alpha_t),$$

where the last estimation is implied from the fact that, $\lim_{t\to\infty}\varphi'(Z_{t-} - z^*) = 1$ on the event $\Gamma^{(p)}$. Thus,

$$\mathbb{E}\left[\mathbf{1}_{\Gamma^{(p)}}|\tau_t|\right] = o(\alpha_t).$$

Therefore, we can conclude that $\mathbb{P}(\Gamma^{(p)}) \leq \lim_{t\to\infty} P(\Gamma_t^{(p)}) = 0$. It implies that $\mathbb{P}(\Gamma) = \mathbb{P}(\cup_p \Gamma^{(p)}) = 0$.

$\square$

We now return to the occupation measure process $Z_t = (Z_t^1, Z_t^2)$. For $t \geq 0$, define

$$\bar{Z}_t = Z_t^1 + \frac{\mathbf{1}_{\{X_t=1\}}}{(t+\ell_0)^{\alpha+1}}$$

and for each $z \in \mathbb{R}$,

$$\tilde{F}(z) = F_1(z, 1-z) = -z + \frac{z^\alpha}{z^\alpha + (1-z)^\alpha}.$$

By the integration by part formula, the equation (B.2) can be rewritten as the following form for $0 \leq s < t$

$$\bar{Z}_t - \bar{Z}_s = \int_s^t \frac{1}{u+\ell_0} \tilde{F}(\bar{Z}_u) du + \int_s^t \frac{dE_u + R_u du}{(u+\ell_0)^{\alpha/2+1}},$$

where we define the martingale

$$E_t = \int_0^t \frac{1}{(u+\ell_0)^{\alpha/2}} dM_u^1$$

$$= \sum_{0<u\leq t} \frac{1}{(u+\ell_0)^{\alpha/2}} \left(\mathbf{1}_{\{X_u=1\}} - \mathbf{1}_{\{X_{u-}=1\}}\right) - \int_0^t \frac{(A_u)_{X_u,1}}{(u+\ell_0)^{\alpha/2}} du$$

and the $\mathcal{F}_t$-adapted process

$$R_t = (\tilde{F}(Z_t^1) - \tilde{F}(\bar{Z}_t))(t+\ell_0)^{\alpha/2} - \frac{\alpha/2+1}{(t+\ell_0)^{\alpha/2+1}} \mathbf{1}_{\{X_{t-}=1\}}.$$

Note that

$$|Z_t^1 - \bar{Z}_t| \leq \frac{1}{(t+\ell_0)^{\alpha+1}} \to 0$$

as $t \to \infty$. Furthermore,

$$|(\tilde{F}(Z_t^1) - \tilde{F}(\bar{Z}_t))| \leq \text{Const.} |Z_t^1 - \bar{Z}_t| \leq \frac{\text{Const}}{(t+\ell_0)^{\alpha+1}}.$$

It implies that

$$|R_t| \leq \frac{\text{Const}}{(t+\ell_0)^{\alpha/2+1}}.$$

Therefore
$$\int_0^\infty R_t^2 dt < +\infty.$$

For $h > 0$ and large t, we observe that

$$\mathbb{E}\left(|E_{t+h} - E_t|^2 \mid \mathcal{F}_t\right)$$

$$= \mathbb{E}\left(\sum_{t<u\leq t+h} \frac{(\mathbf{1}_{\{X_u=1\}} - \mathbf{1}_{\{X_{u-}=1\}})^2}{(u+\ell_0)^\alpha} \mid \mathcal{F}_t\right)$$

$$< \frac{1}{(t+\ell_0)^\alpha} \times \mathbb{E}\,(\text{number of jumps between 1 and 2 during } (t, t+h))$$

$$< \text{Const.}h,$$

where in the last inequality, we have used the fact that the number of jumps between 1 and 2 during $(t, t+h)$ is dominated by the number of jumps of a Poisson process with constant intensity $(t+h)^\alpha$ during $(t, t+h)$.

Beside this, on the event $\Gamma = \{Z_t \to z^* = (1/2, 1/2)\}$, $L(1, t) = t/2 + o(t)$ and $L(2, t) = t/2 + o(t)$ as $t \to \infty$. We thus have

$$\mathbb{E}\left(\mathbf{1}_\Gamma |E_{t+h} - E_t|^2 \mid \mathcal{F}_t\right)$$

$$> \frac{1}{(t+\ell_0+h)^\alpha} \times \mathbb{E}\,(\mathbf{1}_\Gamma \times \text{number of jumps from 1 to 2 during } (t, t+h))$$

$$> \text{Const.}h,$$

where in the last inequality, we have used the fact that the number of jumps from 1 to 2 during $(t, t+h)$ is greater than the number of jumps of a Poisson process with constant intensity $L(2, t)^\alpha = (\frac{t}{2} + o(t))^\alpha$ during $(t, t+h)$.

Furthermore, for each $t > 0$,

$$\mathbb{E}\left(\sum_{t<s\leq t+1} |\bar{Z}_s|^2\right) = \mathbb{E}\left(\sum_{t<s\leq t+1} \frac{(\mathbf{1}_{\{X_s=1\}} - \mathbf{1}_{\{X_{s-}=1\}})^2}{(s+\ell_0)^{2\alpha+2}}\right)$$

$$\leq \frac{\text{Const}}{(t+\ell_0)^{2\alpha+2}} \cdot \mathbb{E}\left(\sum_{t<s\leq t+1} (\mathbf{1}_{\{X_s=1\}} - \mathbf{1}_{\{X_{s-}=1\}})^2\right)$$

$$\leq \frac{\text{Const}}{(t+\ell_0)^{2\alpha+2}} \cdot (t+\ell_0)^\alpha = \frac{\text{Const}}{(t+\ell_0)^{\alpha+2}}.$$

Applying Theorem 5.2 for $\gamma(t) := \frac{1}{t+\ell_0}, c(t) := \frac{1}{(t+\ell_0)^{\alpha/2+1}}$, $N_t := E_t$ and $G_t := R_t$, we can conclude that $\mathbb{P}\{Z_t \to z^* = (1/2, 1/2)\} = \mathbb{P}\{\bar{Z}_t \to 1/2\} = 0$.

Multivariate nonconvergence theorem

In this subsection, we consider a càdlàg stochastic process Z_t taking values in $\mathbb{R}^d$ such that

$$Z_t - Z_0 = \int_0^t \gamma_s F(Z_s) ds + \int_0^t G_s ds + \int_0^t H_{s-} dM_s \tag{B.26}$$

and the following assumptions are fulfilled

Assumption 9. *1. F is a C^2-vector field;*

2. *G and H are cadlag adapted processes taking their values respectively in $\mathbb{R}^d$ and in $\mathbb{R}^{d \times d}$, and there are decreasing deterministic continuous functions $t \mapsto g_t$ and $t \mapsto h_t$ with $\lim_{t \to \infty} g_t = \lim_{t \to \infty} h_t = 0$ such that for all $t \geq 0$, $\sup_i |G_t^i| \leq g_t$ and $\sup_{i,j} |H_t^{i,j}| \leq h_t$;*

3. *M is a martingale of the form $M_t = N_t - \int_0^t \Lambda_s ds$ such that for all $i \in \{1, \ldots, d\}$,*

   * *N^i is an increasing cadlag process taking its values in $\mathbb{N}$ such that $\Delta N_t^i := N_t^i - N_{t-}^i \in \{0, 1\}$ and if $i \neq j$, $\Delta N_t^i \Delta N_t^j = 0$ for all $t \geq 0$,*
   * *Λ^i is a nonnegative cadlag adapted process.*

Let z^* be a non-degenerate unstable equilibrium of F. Then $\mathbb{R}^d = E^s \oplus E^u$, where E^s and E^u are respectively the generalized eigenspaces generated by the eigenvalues of $DF(z^*)$ with positive and negative real parts. To simplify the notation, we suppose that E^u is spanned by $e_1, \ldots, e_m$ and E^s by $e_{m+1}, \ldots, e_d$. Since z^* is unstable, there exist constants $C, \lambda > 0$ such that

$$\|D_v \Phi_t(z^*)\| = \|e^{DF(z^*)t} v\| \geq C e^{\lambda t} \|v\|$$

for all $v \in E^u$ and $t \geq 0$.

Let $(z_s^*, z_u^*) \in E^s \times E^u$ be such that $z^* = z_s^* + z_u^*$. By the stable manifold theorem (see, e.g [16], Theorem 10.1), there exist a neighbourhood $\mathcal{N}_0 = \mathcal{N}_0^s \oplus \mathcal{N}_0^u$ of z^*, with $\mathcal{N}_0^s$ (resp. $\mathcal{N}_0^u$) a ball around z_s^* in E^s (resp. around z_u^* in E^u) and a C^2-function $\Gamma : \mathcal{N}_0^s \to \mathcal{N}_0^u$ such that

(a) $D\Gamma(z_s^*) = 0$.

(b) The graph of Γ:
$$\text{Graph}(\Gamma) := \{v + \Gamma(v) : v \in \mathcal{N}_0^s\}$$

equals to the local stable manifold of z^*:
$$W_{loc}^s(z^*) = \{z \in \mathbb{R}^d : \forall t \geq 0, \Phi_t(z) \in \mathcal{N}_0 \text{ and } \lim_{t \to \infty} \Phi_t(z) = z^*\}.$$

(c) $W_{loc}^s(z^*)$ is an invariant manifold, i.e. for all $t \in \mathbb{R}$,
$$\Phi_t(W_{loc}^s(z^*)) \cap \mathcal{N}_0 \subset W_{loc}^s(z^*).$$

Let $r : \mathcal{N}_0 \to W_{loc}^s(z^*)$ and $R : \mathcal{N}_0 \to \mathbb{R}$ be defined by
$$r(z_s + z_u) = z_s + \Gamma(z_s)$$

and
$$R(z) = \|z - r(z)\|^2.$$
Then r and R are C^2 and R vanishes on $W_{loc}^s(z^*)$.

Following Lemma 6.7 in [5], there exists $t_1 > 0$ and a neighbourhood $\mathcal{N}_1 \subset \mathcal{N}_0$ of z^* such that for all $z \in \mathcal{N}_1$, $\Phi_{t_1}(z) \in \mathcal{N}_0$ and
$$R(\Phi_{t_1}(z)) \geq R(z). \tag{B.27}$$

Let $\mathcal{N}_2 \subset \mathcal{N}_1$ be a neighbourhood of z^* such that $\Phi_{-t}(z) \in \mathcal{N}_1$ for every $t \in [0, t_1]$. For $z \in \mathcal{N}_2$, set
$$\eta(z) = \int_0^{t_1} R(\Phi_{-s}(z)) ds.$$

Then η is satisfies the following

Lemma 5.3.   (i) $\eta(z) = 0$, for every $z \in \mathcal{N}_2 \cap W_{loc}^s(z^*)$,

(ii) η is C^2 on $\mathcal{N}_2$,

(iii) For all $z \in \mathcal{N}_2$, $D\eta(z)F(z) \geq 0$,

(iv) There is a constant $C_\eta \in [1, \infty)$ such that for all $z \in \mathcal{N}_2$ and $u \in \mathbb{R}^d$,
$$\|D\eta(z)\| \leq C_\eta \eta^{\frac{1}{2}}(z),$$
$$\|D^2\eta(z)\| \leq C_\eta,$$
$$2\eta(z)D_{u,u}^2\eta(z) - (D_u\eta(z))^2 \geq -C_\eta\|u\|^2\eta(z)^{3/2}.$$

(v) For every $\epsilon > 0$ there exists $\mathcal{N}_2^\epsilon \subset \mathcal{N}_2$ such that for all $z, z' \in \mathcal{N}_2^\epsilon$ and $v = z' - z$,

$$\eta(z') - \eta(z) - D_v\eta(z) \geq \frac{1}{2}D^2_{v,v}\eta(z^*) - \epsilon\|v\|^2,$$

(vi) $D^2_{u,v}\eta(z^*) = 0$ if $v \in E^s$ and there is a positive constant ρ such that for all $u \in E^u$, $D^2_{u,u}\eta(z^*) \geq 2\rho\|u\|^2$.

Proof. (i) and (ii) are clear. For $z \in \mathcal{N}_2$,

$$
\begin{aligned}
D\eta(z)F(z) &= \lim_{\delta \to 0} \frac{1}{\delta}\left(\eta(\Phi_\delta(z)) - \eta(z)\right) \\
&= \lim_{\delta \to 0} \frac{1}{\delta}\left(\int_{-\delta}^0 R(\Phi_{-s}(z))ds - \int_{t_1-\delta}^{t_1} R(\Phi_{-s}(z))ds\right) \\
&= R(z) - R(\Phi_{-t_1}(z)) \geq 0.
\end{aligned}
$$

Therefore, (iii) is implied from (B.27).

Set $z_t = \Phi_{-t}(z)$, and for $u \in \mathbb{R}^d$, $u_t = D_u z_t = D_u\Phi_{-t}(z)$. Then

$$D_u\eta(z) = D_u\int_0^{t_1}\|z_t - r(z_t)\|^2 dt = 2\int_0^{t_1}\langle u_t - D_{u_t}r(z_t), z_t - r(z_t)\rangle dt.$$

The first inequality in (iv) can be proved by using Cauchy-Schwartz inequality and the fact that r and Φ are C^2. The second inequality in (iv) simply follows from (ii). The inequality in (v) is a simple consequence of the Taylor expansion of η. We now prove the third inequality in (iv). Using Cauchy-Schwartz inequality, one obtains that

$$
\begin{aligned}
2\eta(z)D^2_{u,u}\eta(z) - (D_u\eta(z))^2 \geq &-4\eta(z)\int_0^{t_1}\langle D^2_{u_t,u_t}r(z_t), z_t - r(z_t)\rangle dt \\
&+ 4\eta(z)\int_0^{t_1}\langle w_t - D_{w_t}r(z_t), z_t - r(z_t)\rangle dt,
\end{aligned}
$$

where we set $w_t = D_{u,u}\Phi_{-t}(z)$. Therefore, the third inequality in (iv) is obtained by using again Cauchy-Schwartz inequality.

To prove (vi), we remark that for $u, v \in \mathbb{R}^d$,

$$D^2_{u,v}\eta(z^*) = 2\int_0^{t_1}\langle u_t - D_{u_t}r(z^*), v_t - D_{v_t}r(z^*)\rangle dt,$$

where $u_t = D_u\Phi_{-t}(z^*)$ and $v_t = D_v\Phi_{-t}(z^*)$. Since for $v \in E^s$, $v_t \in E^s$, we have $D_{v_t}r(z^*) = v_t$. This shows the first assertion of (vi). Following the proof of the

claim (iv), page 54 in [4], there is a positive constant c_0 such that for all $u \in E^u$ and $0 \le t \le t_1$,

$$\|u_t - D_{u_t} r(z_t)\| \ge c_0 \|u\|.$$

This prove the last statement of (vi), with $\rho = c_0^2 t_1$. $\qquad\square$

A consequence of the first inequality in Lemma 5.3(iv) is that for all $z, z' \in \mathcal{N}_2$, if $u = z' - z$, it holds that

$$\sqrt{\eta(z')} \ge \sqrt{\eta(z)} - \frac{C_\eta \|u\|}{2}.$$

This with the third inequality of (iv) implies that if $\|u\| < \frac{2\sqrt{\eta(z)}}{C_\eta}$ (which implies that for all $t \in [0,1]$, $\eta(z + tu) > 0$), then

$$
\sqrt{\bar{\eta}(z')} - \sqrt{\bar{\eta}(z)} - \frac{D_u \eta(z)}{2\sqrt{\eta(z)}}
$$
$$
= \frac{1}{4} \int_0^1 \int_0^t \frac{2\eta(z+tu)D_{u,u}^2\eta(z+su) - (D_u\eta(z+su))^2}{\eta(z+su)^{3/2}} ds dt
$$
$$
\ge -\frac{C_\eta}{8} \times \|u\|^2. \tag{B.28}
$$

Suppose moreover that

Assumption 10.   • *There are a constant c^* and a positive decreasing deterministic positive function $t \mapsto c_t$ with $\lim_{t\to\infty} c_t = 0$ such that when $Z_t \in \mathcal{N}_0$,*

$$
c_t \le \sum_{i=1}^m \sum_k (H_t^{i,k})^2 \Lambda_t^k \qquad and \qquad \sum_{i=1}^d \sum_k (H_t^{i,k})^2 \Lambda_t^k \le c^* c_t. \tag{B.29}
$$

• *Set $\eta_t = \eta(Z_t) 1_{Z_t \in \mathcal{N}_2^c}$. Let $t \mapsto \alpha_t$ be a positive decreasing function such that $\lim_{t\to\infty} \alpha_t = 0$. We suppose also that there is $T < \infty$ and $C < \infty$ such that for all $t \ge T$,*

$$
\frac{4\alpha_t^2}{\rho C} \le \int_t^\infty c_u du \le \left(\frac{\alpha_t}{C_\eta c^*}\right) \wedge \left(\frac{\alpha_t^2}{8C_\eta^2 c^*}\right); \tag{B.30}
$$

$$
4C_\eta \int_t^\infty \alpha_u g_u du \le \rho \int_t^\infty c_u du; \tag{B.31}
$$

$$
4C_\eta \int_t^\infty g_u du \le \alpha_t; \tag{B.32}
$$

$$
h_t \le \left(\frac{4}{\sqrt{d}}\right) \wedge \left(\frac{\alpha_t}{2C_\eta}\right). \tag{B.33}
$$

For $\epsilon < \frac{\rho}{2c^*}$ we will choose later on, and $t > 0$, define the following stopping times

$$S_t = \inf\{s \geq t : \eta_s > \alpha_s^2\}, \tag{B.34}$$
$$U_t = \inf\{s \geq t : Z_s \notin \mathcal{N}_2^\epsilon\}. \tag{B.35}$$

Lemma 5.4. *For all $t > T$,*

$$\mathbb{P}[S_t \wedge U_t < \infty | \mathcal{F}_t] \geq p = \frac{1}{2(1+C)}. \tag{B.36}$$

Proof. Applying the change of variable formula to (B.26), we have

$$\eta_s - \eta_t = \int_t^s \gamma_u D\eta(Z_u) F(Z_u) du + \sum_{i,k} \int_t^s \partial_i \eta(Z_{u-}) H_{u-}^{i,k} dM_u^k$$
$$+ \int_t^s D\eta(Z_u) G_u du + \sum_{t < u \leq s} (\Delta\eta(Z_u) - D\eta(Z_{u-})\Delta Z_u).$$

Using Lemma 5.3(iii-iv-v), the definitions (B.34), (B.35) and Assumption 9(2), for $t \leq s \leq S_t \wedge U_t$, we obtain that

$$\eta_s - \eta_t \geq \sum_{i,k} \int_t^s \partial_i \eta(Z_{u-}) H_{u-}^{i,k} dM_u^k - C_\eta \int_t^s \alpha_u g_u du$$
$$+ \frac{1}{2} \sum_{t < u \leq s} \sum_{i,j} \partial_{i,j}^2 \eta(z^*) \Delta Z_u^i \Delta Z_u^j - \epsilon \sum_{t < u \leq s} \|\Delta Z_u\|^2.$$

Since $\Delta Z_u^i \Delta Z_u^j = \sum_{k,\ell} H_{u-}^{ik} H_{u-}^{j\ell} \Delta N_u^k \Delta N_u^\ell = \sum_k H_{u-}^{i,k} H_{u-}^{j,k} \Delta N_u^k$, we get for $t \leq s \leq S_t \wedge U_t$,

$$\eta_s - \eta_t \geq \sum_{i,k} \int_t^s \partial_i \eta(Z_{u-}) H_{u-}^{i,k} dM_u^k - C_\eta \int_t^s \alpha_u g_u du$$
$$+ \frac{1}{2} \sum_{t < u \leq s} \sum_{i,j,k} \left(\partial_{i,j}^2 \eta(z^*) - 2\epsilon\delta_{i,j} \right) H_{u-}^{i,k} H_{u-}^{j,k} \Delta N_u^k$$
$$= \sum_k \int_t^s K_{u-}^k dM_u^k - C_\eta \int_t^s \alpha_u g_u du$$
$$+ \frac{1}{2} \int_t^s \sum_{i,j,k} \left(\partial_{i,j}^2 \eta(z^*) - 2\epsilon\delta_{i,j} \right) H_u^{i,k} H_u^{j,k} \Lambda_u^k du,$$

where

$$K_u^k := \sum_i \partial_i \eta(Z_u) H_u^{i,k} + \frac{1}{2} \sum_{i,j} (\partial_{i,j}^2 \eta(z^*) - 2\epsilon\delta_{i,j}) H_u^{i,k} H_u^{j,k}.$$

Note that K^k is bounded on $[t, S_t \wedge U_t]$. For $1 \leq k \leq d$, Lemma 5.3(vi) and the assumption (B.29) imply that

$$\frac{1}{2}\sum_{i,j,k}(\partial_{i,j}^2\eta(z^*) - 2\epsilon\delta_{i,j})H_u^{i,k}H_u^{j,k}\Lambda_u^k \geq \rho\sum_{i=1}^{m}(H_u^{i,k})^2\Lambda_u^k - \epsilon\sum_i(H_u^{i,k})^2\Lambda_u^k$$

$$\geq (\rho - \epsilon c^*)c_u \geq \frac{\rho}{2}c_u$$

(since we have chosen $\epsilon \leq \frac{\rho}{2c^*}$). We thus have

$$\eta_{S_t \wedge U_t} - \eta_t \geq \sum_k \int_t^{S_t \wedge U_t} K_{u-}^k dM_u^k - C_\eta \int_t^\infty \alpha_u g_u du + \frac{\rho}{2}\int_t^{S_t \wedge U_t} c_u du. \quad \text{(B.37)}$$

Note that when $S_t < U_t$, from the definitions (B.34), (B.35), we have $\eta(Z_{S_t-}) \leq \alpha_{S_t}^2$ and $Z_{S_t}, Z_{S_t-} \in \mathcal{N}_2^\epsilon \subset \mathcal{N}_2$. Furthermore, from Assumption 9(2-3), $\|\Delta Z_{S_t}\| = \|H_{S_t-}\Delta N_{S_t}\| \leq d.h_{S_t}$. Therefore, from the change of variables formula, the two first inequalities in Lemma 5.3(iv) and the assumption (B.33), we obtain that

$$\eta(Z_{S_t}) - \eta(Z_{S_t-}) \leq D\eta(Z_{S_t-})\Delta Z_{S_t} + \frac{1}{2}C_\eta\|\Delta Z_{S_t}\|^2$$

$$\leq C_\eta[\sqrt{\eta(Z_{S_t-})}\|\Delta Z_{S_t}\| + \frac{1}{2}\|\Delta Z_{S_t}\|^2]$$

$$\leq C_\eta[d\alpha_{S_t}h_{S_t} + \frac{1}{2}d^2h_{S_t}^2] \leq \alpha_{S_t}^2.$$

Thus we get that $\eta_{S_t} = \eta_{S_t-} + (\eta_{S_t} - \eta_{S_t-}) \leq 2\alpha_{S_t}^2$. On the other hand, when $U_t \leq S_t$, $\eta_{U_t} = 0$ and when $S_t \wedge U_t = \infty$, $\lim_{s\to\infty}\eta_s = 0$. Therefore, from (B.37) and the fact that $\{M_t\}_{t\geq 0}$ is a martingale, we have

$$2\mathbb{E}[\alpha_{S_t \wedge U_t}^2|\mathcal{F}_t] \geq \mathbb{E}[\eta_{S_t \wedge U_t}|\mathcal{F}_t]$$

$$\geq -C_\eta \int_t^\infty \alpha_u g_u du + \frac{\rho}{2}\mathbb{E}\left[\int_t^{S_t \wedge U_t} c_u du \,\middle|\, \mathcal{F}_t\right]$$

$$\geq -C_\eta \int_t^\infty \alpha_u g_u du + \frac{\rho}{2}\left(\int_t^\infty c_u du\right)\mathbb{P}[S_t \wedge U_t = \infty|\mathcal{F}_t].$$

Using that $\alpha_{S_t \wedge U_t}^2 \leq \alpha_t^2 1_{S_t \wedge U_t < \infty}$, we obtain from the above inequality that

$$\mathbb{P}[S_t \wedge U_t < \infty|\mathcal{F}_t] \geq \left(1 - \frac{2C_\eta \int_t^\infty \alpha_u g_u du}{\rho \int_t^\infty c_u du}\right)\left(1 + \frac{4\alpha_t^2}{\rho \int_t^\infty c_u du}\right)^{-1}.$$

In the last estimation, using the assumptions (B.30) and (B.31), we easily obtain (B.36). $\qquad\square$

Lemma 5.5. *Let $H = \{\lim_{t\to\infty} Z_t = z^*\}$ and for $t \geq 0$, $H_t = H \cap \{U_t = \infty\}$. Then, for all $t > T$, on the event $\{S_t < U_t\}$,*

$$\mathbb{P}[H_t|\mathcal{F}_{S_t}] \leq \frac{1}{2}. \tag{B.38}$$

Proof. Set $T_t = \inf\{s \geq S_t : \eta_s = 0\}$. On the event $\{S_t < U_t\}$, for $s \in [S_t, T_t \wedge U_t]$,

$$\sqrt{\eta_s} = \sqrt{\eta_{S_t}} + \int_{S_t}^s \gamma_u \frac{D\eta(Z_u)F(Z_u)}{2\sqrt{\eta_u}}du + \sum_{i,j} \int_{S_t}^s \frac{\partial_i \eta(Z_{u-})}{2\sqrt{\eta(Z_{u-})}} H_{u-}^{i,j} dM_u^j$$

$$+ \int_{S_t}^s \frac{D\eta(Z_u)G_u}{2\sqrt{\eta_u}}du + \sum_{S_t < u \leq s} \left(\Delta\sqrt{\eta}(Z_u) - \frac{D\eta(Z_{u-})\Delta Z_u}{2\sqrt{\eta(Z_{u-})}}\right). \tag{B.39}$$

We remark from the definition (B.34) that

$$\sqrt{\eta_{S_t}} \geq \alpha_{S_t}. \tag{B.40}$$

From the second inequality in Lemma 5.3(iv), Assumption 9(2) and (B.32), we have

$$\int_{S_t}^s \frac{D\eta(Z_u)G_u}{2\sqrt{\eta_u}}du \geq -\frac{C_\eta}{2}\int_{S_t}^s g_u du \geq -\frac{1}{8}\alpha_{S_t}. \tag{B.41}$$

From (B.28),

$$\sum_{S_t < u \leq s} \left(\Delta\sqrt{\eta}(Z_u) - \frac{D\eta(Z_{u-})\Delta Z_u}{2\sqrt{\eta(Z_{u-})}}\right)$$

$$\geq -\frac{C_\eta}{8}\|\Delta Z_u\|^2$$

$$\geq -\frac{C_\eta}{8}\sum_{i,k}\sum_{S_t < u \leq s}(H_{u-}^{i,k})^2\Delta N_u^k$$

$$\geq -\frac{C_\eta}{8}\sum_{i,k}\int_{S_t}^s(H_{u-}^{i,k})^2 dM_u^k - \frac{C_\eta}{8}\sum_{i,k}\int_{S_t}^s(H_{u-}^{i,k})^2\Lambda_u^k du$$

$$\geq -\frac{C_\eta}{8}\sum_{i,k}\int_{S_t}^s(H_{u-}^{i,k})^2 dM_u^k - \frac{c^* C_\eta}{8}\int_{S_t}^s c_u du \tag{B.42}$$

$$\geq -\frac{C_\eta}{8}\sum_{i,k}\int_{S_t}^s(H_{u-}^{i,k})^2 dM_u^k - \frac{1}{8}\alpha_{S_t}, \tag{B.43}$$

where (B.42) is implied from the assumption (B.29) and (B.43) is implied from (B.41).

Set

$$M_t^\eta = \sum_{i,k} \int_{S_t}^s \left(\frac{\partial_i \eta(Z_{u-})}{2\sqrt{\eta(Z_{u-})}} H_{u-}^{i,k} - \frac{C_\eta}{8} (H_{u-}^{i,k})^2 \right) dM_u^k \qquad (B.44)$$

and

$$I_t = \inf_{s \in [S_t, U_t \wedge T_t]} M_s^\eta. \qquad (B.45)$$

Then, combining (B.39) together with the inequalities (B.40), (B.41), (B.43), the definitions (B.44) and (B.45), we have

$$\sqrt{\eta_s} \geq \alpha_{S_t} + I_t - \frac{1}{8}\alpha_{S_t} - \frac{1}{8}\alpha_{S_t} = \frac{3}{4}\alpha_{S_t} + I_t$$

on the event $\{S_t < U_t\}$, for $s \in [S_t, T_t \wedge U_t]$. Therefore,

$$\inf_{s \in [S_t, U_t \wedge T_t]} \sqrt{\eta_s} \geq \frac{1}{4}\alpha_{S_t}$$

on the event $\{S_t < U_t\} \cap \{I_t \geq -\frac{1}{2}\alpha_{S_t}\}$. Thus on the event $\{S_t < U_t\}$,

$$\mathbb{P}[H_t | \mathcal{F}_{S_t}] \leq \mathbb{P}\left[I_t \leq -\frac{1}{2}\alpha_{S_t} \,\middle|\, \mathcal{F}_{S_t} \right]. \qquad (B.46)$$

Using Doob's inequality, we have

$$\mathbb{P}\left[I_t \leq -\frac{1}{2}\alpha_{S_t} \,\middle|\, \mathcal{F}_{S_t} \right] \leq \frac{4}{\alpha_{S_t}^2} \left[\int_{S_t}^{U_t \wedge T_t} \sum_k (L_u^k)^2 \Lambda_u^k du \,\middle|\, \mathcal{F}_{S_t} \right], \qquad (B.47)$$

where, for $u \in [S_t, U_t \wedge T_t]$,

$$L_u^k = \sum_i \frac{\partial_i \eta(Z_u)}{2\sqrt{\eta(Z_u)}} H_u^{i,k} - \frac{C_\eta}{8} \sum_i (H_u^{i,k})^2.$$

From Assumption 9(2), note that $\sum_i (H_u^{i,k})^2 \leq d.h_u^2 \leq 1$ as u is large enough. Using the first inequality in Lemma 5.3(iv), we have

$$\begin{aligned}
|L_u^k| &\leq \frac{C_\eta}{2} \sqrt{\sum_i (H_u^{i,k})^2} + \frac{C_\eta}{8} \sum_i (H_u^{i,k})^2 \\
&\leq \frac{C_\eta}{2} \sqrt{\sum_i (H_u^{i,k})^2} \left[1 + \frac{\sqrt{d}h_u}{4} \right] \leq C_\eta \sqrt{\sum_i (H_u^{i,k})^2}, \qquad (B.48)
\end{aligned}$$

where the last inequality of (B.48) is implied from the assumption (B.33). The assumption (B.29) and (B.48) imply that, for $u \in [S_t, U_t \wedge T_t]$,

$$\sum_k |L_u^k|^2 \Lambda_u^k \leq C_\eta^2 \sum_k \sum_i (H_u^{i,k})^2 \Lambda_u^k$$

$$\leq C_\eta^2 c^* c_u. \tag{B.49}$$

Therefore, from (B.47), (B.49) and the assumption (B.30)

$$\mathbb{P}\left[I_t \leq -\frac{1}{2} \alpha_{S_t} \,\middle|\, \mathcal{F}_{S_t} \right] \leq \frac{4 C_\eta^2 c^*}{\alpha_{S_t}^2} \int_{S_t}^\infty c_u du \leq \frac{1}{2}.$$

This inequality and (B.46) imply the lemma. $\qquad\square$

Theorem 5.6. $\mathbb{P}[\lim_{t\to\infty} Z_t = z^*] = 0.$

Proof. Set $A := \{\exists t; U_t = \infty\}$. From (B.36) and (B.38), for $t > T$,

$$\mathbb{P}[H_t^c | \mathcal{F}_t] \geq \mathbb{E}[\mathbb{P}[H_t^c | \mathcal{F}_{S_t}] 1_{\{S_t < U_t\}} | \mathcal{F}_t]$$

$$\geq \frac{1}{2} \times \mathbb{P}[S_t < U_t | \mathcal{F}_t]$$

$$\geq \frac{1}{2} \times (p - \mathbb{P}[U_t < \infty | \mathcal{F}_t]).$$

Since $H = \{\lim_{t\to\infty} Z_t = z^*\} \subset A$, for $t > T$, a.s.

$$\mathbb{P}[H^c | \mathcal{F}_t] = \lim_{s\to\infty} \mathbb{P}[H_s^c | \mathcal{F}_t]$$

$$= \lim_{s\to\infty} \mathbb{E}[\mathbb{P}[H_s^c | \mathcal{F}_s] | \mathcal{F}_t]$$

$$\geq \lim_{s\to\infty} \frac{1}{2} \times (p - \mathbb{E}[\mathbb{P}[U_s < \infty | \mathcal{F}_s] | \mathcal{F}_t])$$

$$\geq \lim_{s\to\infty} \frac{1}{2} \times (p - \mathbb{P}[U_s < \infty | \mathcal{F}_t])$$

$$\geq \frac{1}{2} \times (p - \mathbb{P}[A^c | \mathcal{F}_t]).$$

Since, a.s., $\lim_{t\to\infty} \mathbb{P}[H^c | \mathcal{F}_t] = 1_{H^c}$ and $\lim_{t\to\infty} \mathbb{P}[A^c | \mathcal{F}_t] = 1_{A^c}$, we obtain that

$$1_{H^c} \geq \frac{1}{2}(p - 1_{A^c}) \qquad \text{a.s.}$$

This implies that a.s., $A \subset H^c$, which is possible only if $\mathbb{P}(A) = 0$ or $\mathbb{P}(H^c) = 1$. We thus have proved (since $H \subset A$) that $\mathbb{P}(H) = 0$. $\qquad\square$

Application to strongly VRJP on complete graphs

Theorem 5.7. *Assume that $w(t) = t^\alpha, \alpha \geq 2$ and $G = (V, E)$ is a complete graph. Then for every unstable equilibrium z^*,*

$$\mathbb{P}\{Z_t \to z^*\} = 0.$$

Proof. Without lost of generality, we only need to consider the event $\{Z_t \to z^*\}$, where z^* is a unstable equilibrium such that $z_i^* = 1/m$ if $1 \leq i \leq m$ and $z_i^* = 0$ otherwise, where $m \in \{2, ..., d\}$. We apply the previous theorem to $\bar{Z}_t = Z_t + \frac{I[X_t]}{(t + \ell_0)^{\alpha+1}}$. Then $\bar{Z}$ satisfies (B.26), by taking

$$\gamma_t := \frac{1}{t + \ell_0},$$

$$G_t := \frac{F(Z_t) - F(\bar{Z}_t)}{t + \ell_0} - \left(\frac{\alpha}{2} + 1\right) I[X_{t-}],$$

$$H_t^{i,j} := \delta_{i,j} \frac{\epsilon_t^i}{(t + \ell_0)^{\alpha+1}},$$

$$N_t^i := \sum_{0 < u \leq t} |\Delta I^i[X_u]|,$$

$$\Lambda_t^i := \epsilon_t^i A_t^i[X_t],$$

where

$$\epsilon_t^i := 1_{\{X_t \neq i\}} - 1_{\{X_t = i\}}.$$

Then N^i counts the number of jumps of X from i or towards i.

For some sufficiently large constant g, one can take

$$g_t = \frac{g}{t^{\alpha+2}} \text{ and } h_t = \frac{1}{t^{\alpha+1}}.$$

We thus have

$$\sup_i |G_t^i| \leq g_t, \text{ and } \sup_{i,j} |H_t^{ij}| \leq h_t.$$

Note that

$$\bar{M}_t^i := \sum_{0 < u \leq t} \Delta N_t^i - \int_0^t \Lambda_u^i du = \int_0^t \epsilon_{u-}^i dM_u^i$$

is a martingale. Note also that

$$\Lambda_t^i = \sum_{j:j \sim i} 1_{\{X_t \in \{i,j\}\}} w_t^{(j)}.$$

If one assumes that G is a complete graph, then by taking c a sufficiently small constant, $c_t = ct^{-(\alpha+2)}$ and c^* sufficiently large, one can check that (B.29) is satisfied.

We then have $\int_t^\infty c_u du = \frac{c}{\alpha+1}t^{-(\alpha+1)}$.

Set $\alpha_t = at^{-\frac{\alpha+1}{2}}$. Then if one takes a such that $a^2 \geq \frac{8C_\eta^2 c^* c}{\alpha+1}$ and $C \geq \frac{4a^2(\alpha+1)}{\rho c}$, T sufficiently large, then (B.30) is satisfied.

We calculate $\int_t^\infty \alpha_u g_u du = \frac{2ag}{3(\alpha+1)}t^{-\frac{3(\alpha+1)}{2}} = o\left(\int_t^\infty c_u du\right)$, which implies that if T is taken sufficiently large, (B.31) holds for all $t \geq T$.

If now one takes T sufficiently large, then (B.32) and (B.33) also hold.

All this proves that $\mathbb{P}[Z_t \to z^*] = 0$.

$\square$

Remark. Since Lemma 5.3 requires that F is a C^2-vector field, the case of subsquare weight functions, i.e. $1 < \alpha < 2$, is still not solved for complete graphs with $d \geq 3$ vertices.

Bibliography

[1] Brandière, Odile; Duflo, Marie. Les algorithmes stochastiques contournent-ils les pièges? (French). Ann. Inst. H. Poincaré Probab. Statist. **32** (1996), no. 3, 395–427

[2] Benaïm, Michel. A dynamical system approach to stochastic approximations. SIAM J. Control Optim. **34** (1996), no. 2, 437–472.

[3] Benaïm, Michel. Vertex-reinforced random walks and a conjecture of Pemantle. Ann. Probab. **25** (1997), no. 1, 361–392.

[4] Benaïm, Michel. Dynamics of stochastic approximation algorithms. Séminaire de Probabilités, XXXIII, 1–68, Lecture Notes in Math., 1709, Springer, Berlin, 1999.

[5] Benaïm, Michel; Raimond, Olivier. Self-interacting diffusions. III. Symmetric interactions. Ann. Probab. **33** (2005), no. 5, 1717–1759.

[6] Benaïm, Michel; Raimond, Olivier; Schapira, Bruno. Strongly vertex-reinforced-random-walk on a complete graph. ALEA Lat. Am. J. Probab. Math. Stat. **10** (2013), no. 2, 767–782.

[7] Coppersmith, D.; Diaconis, P. Random walks with reinforcement. Unpublished manuscript, 1986.

[8] Davis, Burgess; Volkov, Stanislav. Continuous time vertex-reinforced jump processes. Probab. Theory Related Fields **123** (2002), no. 2, 281–300.

[9] Davis, Burgess; Volkov, Stanislav. Vertex-reinforced jump processes on trees and finite graphs, Probab. Theory Related Fields **128** (2004), no. 1, 42–62.

[10] Disertori, M.; Merkl, F.; Rolles S. W. W. Localization for a nonlinear sigma model in a strip related to vertex reinforced jump processes, Comm. Math. Phys. **332** (2014), no. 2, 783–825.

[11] Pemantle, Robin. Phase transition in reinforced random walk and RWRE on trees. Ann. Probab. **16** (1988), no. 3, 1229–1241.

[12] Pemantle, Robin. Vertex-reinforced random walk, Probab. Theory Related Fields **92** (1992), no. 1, 117–136.

[13] Sabot, Christophe; Tarrès, Pierre. Edge-reinforced random walk, vertex-reinforced jump process and the supersymmetric hyperbolic sigma model. J. Eur. Math. Soc. (JEMS) **17** (2015), no. 9, 2353–2378.

[14] Sabot, Christophe; Tarrès, Pierre; Zeng, Xiaolin. The Vertex Reinforced Jump Process and a Random Schrödinger operator on finite graphs. ArXiv e-prints, arXiv:1507.04660.

[15] Protter, Philip. Stochastic integration and differential equations. Second edition. Stochastic Modelling and Applied Probability, 21. Springer-Verlag, Berlin, 2005.

[16] Robinson, Clark. Dynamical systems. Stability, symbolic dynamics, and chaos. Second edition. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1999.

[17] Tarrès, Pierre. Vertex-reinforced random walk on $\mathbb{Z}$ eventually gets stuck on five points, Ann. Probab. **32** (2004), no. 3B, 2650–2701.

[18] Zeng, Xiaolin. How vertex reinforced jump process arises naturally. Ann. Inst. Henri Poincaré Probab. Stat. **52** (2016), no. 3, 1061–1075.

[19] Merkl, Franz; Rolles, Silke W.W.; Tarrès, Pierre. Convergence of vertex-reinforced jump processes to an extension of the supersymmetric hyperbolic nonlinear sigma model. ArXiv e-prints, arXiv:1612.05409.

[20] Volkov, Stanislav. Vertex-reinforced random walk on arbitrary graphs, Ann. Probab. **29** (2001), no. 1, 66–91.

[21] Volkov, Stanislav. Phase transition in vertex-reinforced random walks on $\mathbb{Z}$ with non-linear reinforcement. J. Theoret. Probab. **19** (2006), no. 3, 691–700.

[22] Cotar, Codina; Thacker, Debleena, Edge- and vertex-reinforced random walks with super-linear reinforcement on infinite graphs, Ann. Probab. **45** (2017), no. 4, 2655–2706.

C

Paper C

On a class of random walks in simplexes

Tuan-Minh Nguyen and Stanislav Volkov

Centre for Mathematical Sciences, Lund University

Abstract

In this paper, we will study the limit behaviour of a class of Markov chain models taking values in the d-dimensional unit standard simplex, $d \geq 1$, defined as follows: from an interior point z, the chain will choose one from $d + 1$ vertices of the simplex with probabilities depending on z and then it will randomly jump to a new point z' in the segment connecting z and the chosen vertex. In some specific cases using Beta distribution, we prove that the limiting distributions of the Markov chain are Dirichlet. We also consider a related history-dependent random walk model in [0,1] based on Friedman's urn-type schemes. We will show that this random walk converges in distribution to the arcsine law.

Keywords: Random walks in simplexes, iterated random functions, Dirichlet distribution, stick-breaking process.

Introduction

We denote as

$$\mathcal{S}_d = \{(z_1, z_2, \ldots, z_d) \in \mathbb{R}^d : z_1 + z_2 + \cdots + z_d \leq 1, z_j \geq 0, j = 1, 2, \ldots, d\}$$

the standard unit simplex in $\mathbb{R}^d$, also let $\mathcal{B}(\mathcal{S}_d)$ and λ_d stand for the Borel σ-algebra and Lebesgue measure on $\mathcal{S}_d$ respectively. Let $E_0 = (0, 0, \ldots, 0)$ be the origin and $E_1 = (1, 0, \ldots, 0)$, $E_2 = (0, 1, 0, \ldots, 0)$, ..., $E_d = (0, \ldots, 0, 1)$ be standard orthonormal basis vectors in $\mathbb{R}^d$, which are also the vertices of $\mathcal{S}_d$.

113

Let $p = (p_1, p_2, ..., p_d)$ be a mapping from $\mathcal{S}_d$ to itself so-called *probability choice function*. For some initial point $Z_0 \in \mathcal{S}_d$, we consider the following random iteration

$$Z_{n+1} = (1 - \xi_n)Z_n + \xi_n \Theta_n,$$

for $n \geq 0$, where $\xi_n, n = 0, 1, 2, ...$ are i.i.d copies of a random variables ξ with support in $[0, 1]$; Θ_n is a discrete random variable independent from $\{\xi_n\}_{n \geq 0}$ such that

$$\begin{cases} \mathbb{P}(\Theta_n = E_j | Z_n = z) = p_j(z), j = 1, 2, ..., d; \\ \mathbb{P}(\Theta_n = E_0 | Z_n = z) = 1 - \sum_{j=1}^{d} p_j(z). \end{cases}$$

The aforementioned model was first introduced by Sethuraman in [10], in the case the choice probabilities $p_1, p_2, ..., p_d$ are positive constants. Sethuraman proved that if $\xi \sim \text{Beta}(1, \gamma)$, Θ is a discrete random variable such that $\mathbb{P}(\Theta = E_j) = p_j$ for $j = 1, 2, ..., d$, $p_0 = 1 - p_1 - p_2 - ... - p_d$ and $Z \sim \text{Dirichlet}(p_1\gamma, p_2\gamma, ..., p_d\gamma, p_0\gamma)$, then

$$Z \sim (1 - \xi)Z + \xi\Theta,$$

where $\text{Beta}(a, b)$ denotes the beta distribution with the probability density function

$$g(\xi) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \xi^{a-1}(1 - \xi)^{b-1}, \ 0 < \xi < 1,$$

Γ being the Gamma function and $\text{Dirichlet}(\alpha_1, \alpha_2, ..., \alpha_d, \alpha_{d+1})$ denotes the Dirichlet distribution with the probability density function

$$f(z_1, z_2, ..., z_d) = \frac{\Gamma\left(\sum_{i=1}^{d+1} \alpha_i\right)}{\prod_{i=1}^{d+1} \Gamma(\alpha_i)} \left(1 - \sum_{i=1}^{d} z_i\right)^{\alpha_{d+1}-1} \prod_{i=1}^{d} z^{\alpha_i - 1}$$

defined for each $(z_1, z_2, ..., z_d)$ in the interior of $\mathcal{S}_d$.

This identity is often used for the construction of Dirichlet distribution in $\mathcal{S}_d$, which has been intensively applied to Bayesian nonparametric statistics. Further extensions when $\xi \sim \text{Beta}(k, \gamma)$, k is a positive integer, and Θ has quasi-Bernoulli distributions were studied by Hitczenco and Letac in [5].

In [4], Diaconis and Freedman reconsidered Sethuraman's model from the point of view of random iterated functions and also discussed the case in which $p(z)$ depends on $z \in \mathcal{S}_1 = [0, 1]$. Other models in $\mathcal{S}_1$ with various specific cases of $p(z)$ and ξ were studied in [8], [9], and [7]. Inspired by the work of Diaconis and Freedman, Ladjimi and Peigné in their recent work [6]

studied iterated random functions with place dependent choice probabilities and demonstrated several applications to the one-dimensional model where $\xi \sim \text{Uniform}[0,1]$ and $p(z)$ is Hölder continuous in $[0,1]$.

In [7], McKinlay and Borovkov gave a general condition for the ergodicity of the one-dimensional Markov chain $\{Z_n\}_{n \geq 0}$ in $\mathcal{S}_1$. By solving integral equations, they derived a closed-form expression for the stationary density function in the case where $\xi \sim \text{Beta}(1, \gamma)$ and $p(z)$ is a piecewise continuous function on $[0,1]$. In particular, if $p(z) = (1-c)z + b(1-z), b, c \in (0,1)$, then the stationary distribution is $\text{Beta}(b\gamma, c\gamma)$.

The model, also known in the literature as stick-breaking process, stochastic give-and-take (see [2], [7]) or Diaconis-Friedman's chain (see [6]) has numerous applications to other fields such as human genetics, robot coverage algorithm, random search, etc. For further discussions on these, we refer the reader to [2], [9] and [7].

The aim of our work is to study the limit behaviour of the d-dimensional Markov chain $\{Z_n\}_{n \geq 0}$ in $\mathcal{S}_d$ and a related history-dependent random walk model in $[0,1]$ which is not a Markov chain. In Section 2, we give an extension of the work of MacKinlay and Borovkov in higher dimensional simplexes under some certain assumptions for $p(z)$ and ξ. In the case where ξ is Beta distributed and the probability choice function $p(z)$ is linearly dependent on z, we prove that the limiting distributions of these Markov chains are Dirichlet in Section 3. In Section 4, we also consider a history-dependent random walk model in [0,1] based on Friedman's urn-type schemes. Using martingales and coupling techniques, we will show that the random walk converges in distribution to the arcsine law.

Existence of the limiting distribution

To prove the ergodicity of the Markov chain $\{Z_n\}_{n \geq 0}$, we will use the following result (see [1], Theorem 2.1)

Proposition 2.1. *Let $Z_n, n = 0, 1, 2, \ldots$ be a Markov chain corresponding with the measurable state space $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ such that for $n \geq 1$, $\mathbb{P}(Z_n \in A | Z_0 = z)$ is a measurable function of $z \in \mathcal{X}$ when $A \in \mathcal{B}(\mathcal{X})$ is fixed, while it is a probability measure of A when z is fixed.*

Then Z_n is ergodic if there exist a subset $V \in \mathcal{B}(\mathcal{X})$, $q > 0$, a probability measure φ

on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *and some positive integer* n_0 *such that*

(a) $\mathbb{P}(\tau_V < \infty | Z_0 = z) = 1$ *for all* $z \in \mathcal{X}$, *where* $\tau_V = \inf\{n \geq 1 : Z_n \in V\}$;

(b) $\sup_{z \in V} \mathbb{E}\,(\tau_V | Z_0 = z) < \infty$;

(c) $\mathbb{P}(Z_{n_0} \in B | Z_0 = z) \geq q\varphi(B)$ *for all* $B \in \mathcal{B}(\mathcal{X})$ *and* $z \in V$;

(d) $\gcd\{n : \mathbb{P}(Z_n \in B | Z_0 = z) \geq q\varphi(B)\} = 1$ *for* $z \in V$.

Moreover, if the above conditions are fulfilled, then there exists an invariant measure μ *such that the distribution of* Z_n *converges to* μ *in total variation norm.*

For $z = (z_1, z_2 \ldots, z_d) \in \mathcal{S}_d$, we define $z_0 = 1 - z_1 - z_2 - \cdots - z_d$ and $p_0(z) = 1 - p_1(z) - p_2(z) - \cdots - p_d(z)$.

Assumption 11. *There are* $\delta \in (0, \frac{1}{2^d})$ *and* $s, t \in (\delta^{1/d}, 1 - \delta^{1/d})$, $s < t$ *such that*

(i) $F_\xi(1 - \delta) := 1 - \eta < 1$;

(ii) *there is an* $\epsilon > 0$ *such that for any* $1 \leq k \leq d$ *and any* $0 \leq j_1 < j_2 < \cdots < j_k \leq d$,

$$\inf_{z \in \mathcal{S}_d : z_{j_1} + \cdots + z_{j_k} \leq \delta} \sum_{l=1}^{k} p_{j_l}(z) \geq \epsilon;$$

(iii) *there is* $c > 0$ *such that for all* $B \in \mathcal{B}([0,1])$, $B \subset [s(1-t)^{d-1} - \delta, t] \cup [(1-t)^d - \delta, 1 - s]$,

$$\mathbb{P}(\xi \in B) > c\lambda(B),$$

where λ *is the Lebesgue measure on* $[0,1]$.

Remark. The condition (i) is quite natural to avoid the absorption of Z_n at the boundary of $\mathcal{S}_d$. For $d = 1$, the above conditions are mostly similar to the assumptions (E1-E2-E3) of McKinlay and Borovkov in [7]. However, in comparison with the condition (iii), they required that ξ has a density on $[s - \delta, t]$ and $[1 - t - \delta, 1 - s]$.
Also, observe that in condition (iii) the intervals are properly defined (but may overlap).

For $j = 0, 1, \ldots, d$, let

$$V_j = \left\{z = (z_1, \ldots, z_d) \in \mathcal{S}_d : 1 - \delta \leq z_j \leq 1\right\}.$$

In particular,

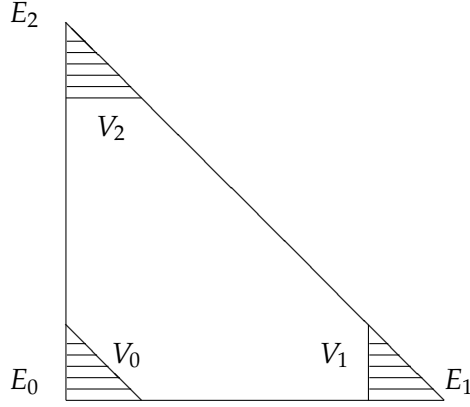$$V_0 = \left\{ z = (z_1, \ldots, z_d) \in \mathcal{S}_d : \sum_{j=1}^{d} z_j \leq \delta \right\}.$$

Figure C.1: Illustration of V_j, $j = 0, 1, 2$ in case $d = 2$.

Let $T : [0,1]^d \to \mathcal{S}^d$ defined as follows, for each $x = (x_1, x_2, \ldots, x_d) \in [0,1]^d$,

$$T(x) = \left(x_1 \prod_{1 < j \leq n} (1 - x_j), x_2 \prod_{2 < j \leq n} (1 - x_j), \ldots, x_{d-1}(1 - x_d), x_d \right).$$

Note that T is a homeomorphism from $(0,1)^d$ to the interior of $\mathcal{S}^d$. Moreover, the inverse T^{-1} is defined such that, for each $z = (z_1, \ldots, z_d)$ in the interior of $\mathcal{S}_d$,

$$T^{-1}(z) = \left(\frac{z_1}{1 - \displaystyle\sum_{2 \leq j \leq n} z_j}, \frac{z_2}{1 - \displaystyle\sum_{3 \leq j \leq n} z_j}, \ldots, \frac{z_{d-1}}{1 - z_d}, z_d \right).$$

For $j = 1, 2, \ldots, d, z = (z_1, \ldots, z_d), u = (u_1, \ldots, u_d) \in \mathcal{S}_d, z_0 = 1 - \sum_{k=1}^{d} z_k, u_0 = 1 - \sum_{k=1}^{d} u_k$, we define the following affine operators

$$R_j(z_1, z_2, \ldots, z_d) = (z_0, z_1, \ldots, z_{j-1}, z_{j+1}, z_{j+2}, \ldots, z_d),$$

$$G_z(u) = (u_0 z_1 + u_1, u_0 z_2 + u_2, \ldots, u_0 z_d + u_d).$$

Note that provided $z_0 \neq 0$, the map G_z is invertible and its inverse can be computed as

$$G_z^{-1}(u) = \left(u_1 - \frac{z_1 u_0}{z_0}, u_2 - \frac{z_2 u_0}{z_0}, \ldots, u_d - \frac{z_d u_0}{z_0} \right).$$

We also define

$$K := \left\{ (u_1, \ldots, u_d) \in \mathcal{S}_d : s \leq \frac{u_j}{1 - \sum_{l=j+1}^{d} u_l} \leq t, j = 1, 2, \ldots, d \right\} = T([s, t]^d).$$

The proof of the following Lemma is given in the Appendix.

Lemma 2.2. *(a) If $z \in V_0$ then*

$$T^{-1} \circ G_z^{-1}(K) \subset [s(1-t)^{d-1} - \delta, t]^d.$$

(b) If $z \in V_k$ with $k \in \{1, 2, \ldots, d\}$ then

$$T^{-1} \circ G_{R_k(z)}^{-1} \circ R_k(K) \subset [(1-t)^d - \delta, 1 - s] \times [s(1-t)^{d-1} - \delta, t]^{d-1}.$$

Theorem 2.3. *Assume that all conditions in Assumption 1 are fulfilled with a common δ. Then the Markov chain $\{Z_n\}_{n \geq 0}$ converges in distribution.*

Proof.
Step 1. We define

$$V = \bigcup_{j=0}^{d} V_j.$$

Observe from the assumption (i) that $\mathbb{P}(Z_1 \in V | Z_0 = z) \geq \eta = 1 - F_{\xi}(1 - \delta) > 0$ for all $z \in \mathcal{S}_d$. Therefore, for all $z \in \mathcal{S}_d$, given $Z_0 = z$, $\tau_V = \inf\{n \geq 1 : Z_n \in V\}$ stochastically dominates a geometric random variable with parameter η. Thus,

$$\mathbb{P}(\tau_V > n | Z_0 = z) \leq (1 - \eta)^n.$$

Hence, the conditions (a) and (b) are easily verified.

Step 2. Throughout this proof, we let Const stand for an existence of some positive constant. From the definition of $\{Z_n\}_{n \geq 0}$, we observe that

$$Z_d = \zeta_0 Z_0 + \zeta_1 \Theta_0 + \zeta_2 \Theta_1 + \cdots + \zeta_d \Theta_{d-1},$$

where

$$(\zeta_1, \ldots, \zeta_d) := T(\xi_0, \ldots, \xi_{d-1}), \quad \zeta_0 := \prod_{j=0}^{d-1}(1 - \xi_j) = 1 - \sum_{j=1}^{d} \zeta_j.$$

For $1 \leq k \leq d$ and $0 \leq j_1 < j_2 < \cdots < j_k \leq d$, define

$$U_{j_1 j_2 \ldots j_k} := \left\{ z = (z_1, z_2, \ldots, z_d) \in \mathcal{S}_d : z_{j_1} + z_{j_2} + \cdots + z_{j_k} \leq \delta \right\}.$$

Let $B \in \mathcal{B}(\mathcal{S}_d)$. Then, if $Z_0 = z \in V_0$ and $\Theta_0 = E_1, \Theta_1 = E_2, \ldots, \Theta_{d-1} = E_d$, then $Z_j \in U_{j+1,j+2,\ldots,d}$ for $j = 0, 1, \ldots, d$. Therefore, by assumption (ii), we note that

$$
\begin{aligned}
&\mathbb{P}\left(Z_d \in B | Z_0 = z\right) \\
&\geq \mathbb{P}\left(Z_d \in B \cap K, (\Theta_0, \Theta_1, \ldots, \Theta_{d-1}) = (E_1, E_2, \ldots, E_d) \mid Z_0 = z\right) \\
&\geq \prod_{l=1}^{d} \inf_{z \in U_{l,l+1,\ldots,d}} \left(\sum_{j=l}^{d} p_j(z)\right) \times \mathbb{P}\left(\zeta_0 z + \zeta_1 E_1 + \cdots + \zeta_d E_d \in B \cap K\right) \\
&\geq \epsilon^d \, \mathbb{P}((\zeta_0 z_1 + \zeta_1, \zeta_0 z_2 + \zeta_2, \ldots, \zeta_0 z_d + \zeta_d) \in B \cap K) \\
&= \epsilon^d \, \mathbb{P}\left((\zeta_1, \zeta_2, \ldots, \zeta_d) \in G_z^{-1}(B \cap K)\right) \\
&= \epsilon^d \, \mathbb{P}\left((\xi_0, \xi_1, \ldots, \xi_{d-1}) \in T^{-1} \circ G_z^{-1}(B \cap K)\right).
\end{aligned} \tag{C.1}
$$

Step 3. For $B \in \mathcal{B}(\mathcal{S}_d)$ and $z \in V_0$, from the assumption (iii) and Lemma 2.2, we have

$$
\mathbb{P}\left((\xi_0, \xi_1, \ldots, \xi_{d-1}) \in T^{-1} \circ G_z^{-1}(B \cap K)\right) \geq \lambda_d\left(T^{-1} \circ G_z^{-1}(B \cap K)\right). \tag{C.2}
$$

We will demonstrate that,

$$
\lambda_d\left(T^{-1} \circ G_z^{-1}(B \cap K)\right) \geq \lambda_d(B \cap K). \tag{C.3}
$$

Indeed, the inequality (C.3) is implied from the fact that for any differentiable mapping $Q : \mathcal{S}_d \to [0,1]^d$ and $A \in \mathcal{B}(\mathcal{S}_d)$,

$$
\lambda_d(Q(A)) \geq \inf_{u \in \mathcal{S}_d} \left|\det\left(\frac{\partial}{\partial u} Q(u)\right)\right| \lambda_d(A)
$$

and by setting the Jacobian $Q = \frac{\partial}{\partial u} G_z^{-1}(u)$ to get

$$
\det\left(\frac{\partial}{\partial u} G_z^{-1}(u)\right) = \det \begin{pmatrix} 1 + \frac{z_1}{z_0} & \frac{z_1}{z_0} & \frac{z_1}{z_0} & \cdots & \frac{z_1}{z_0} \\ \frac{z_2}{z_0} & 1 + \frac{z_2}{z_0} & \frac{z_2}{z_0} & \cdots & \frac{z_2}{z_0} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ \frac{z_{d-1}}{z_0} & \cdots & \frac{z_{d-1}}{z_0} & 1 + \frac{z_{d-1}}{z_0} & \frac{z_{d-1}}{z_0} \\ \frac{z_d}{z_0} & \cdots & \frac{z_d}{z_0} & \frac{z_d}{z_0} & 1 + \frac{z_d}{z_0} \end{pmatrix} = \frac{1}{z_0} \geq 1,
$$

and

$$\det\left(\frac{\partial}{\partial v}T^{-1}(v)\right) = \det \begin{pmatrix} \frac{1}{1-\sum_{j=2}^{d}v_j} & \frac{v_1}{\left(1-\sum_{j=2}^{d}v_j\right)^2} & \frac{v_1}{\left(1-\sum_{j=2}^{d}v_j\right)^2} & \cdots & \frac{v_1}{\left(1-\sum_{j=2}^{d}v_j\right)^2} \\ 0 & \frac{1}{1-\sum_{j=3}^{d}v_j} & \frac{v_2}{\left(1-\sum_{j=3}^{d}v_j\right)^2} & \cdots & \frac{v_2}{\left(1-\sum_{j=3}^{d}v_j\right)^2} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & \frac{1}{1-v_d} & \frac{v_{d-1}}{(1-v_d)^2} \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

$$= \left(\prod_{j=1}^{d}\left(1-\sum_{l=j+1}^{d}v_l\right)\right)^{-1} \geq 1.$$

From (C.1), (C.2)) and (C.3), for each $B \in \mathcal{B}(\mathcal{S}_d)$ and $z \in V_0$, we obtain that

$$\mathbb{P}\left(Z_d \in B | Z_0 = z\right) \geq \text{Const } \lambda_d\left(B \cap K\right). \tag{C.4}$$

Step 4. For each $k \in \{1, 2, ..., d\}$, $B \in \mathcal{B}(\mathcal{S}_d)$ and $z \in V_k$, we have

$\mathbb{P}(Z_d \in B | Z_0 = z)$
$\geq \mathbb{P}\left(Z_d \in B \cap (K), (\Theta_0, \Theta_1, \ldots, \Theta_{d-1}) = (E_0, E_1, \ldots, E_{k-1}, E_{k+1}, \ldots, E_d) \mid Z_0 = z\right)$
$\geq \text{Const } \mathbb{P}\left(R_k^{-1} \circ G_{R_k(z)}(\zeta_1, \zeta_2, ..., \zeta_d) \in B \cap K\right)$
$= \text{Const } \mathbb{P}\left((\zeta_1, \ldots, \zeta_d) \in G_{R_k(z)}^{-1} \circ R_k(B \cap K)\right)$
$= \text{Const } \mathbb{P}\left((\xi_0, \xi_1, \ldots, \xi_{d-1}) \in T^{-1} \circ G_{R_k(z)}^{-1} \circ R_k(B \cap K)\right),$

where we remark that, for $u \in \mathcal{S}_d$ and $z \in V_k$,

$$R_k^{-1}(G_{R_k(z)}(u)) = (u_0 z_1 + u_2, \ldots, u_0 z_{k-1} + u_k, u_0 z_k, u_0 z_{k+1} + u_{k+1}, ..., u_0 z_d + u_d).$$

From the assumption (iii), we have

$$\mathbb{P}\left((\xi_0, \xi_1, \ldots, \xi_{d-1}) \in T^{-1} \circ G_{R_k(z)}^{-1} \circ R_k(B \cap K)\right) \geq \lambda_d(B \cap K).$$

It implies that for each $B \in \mathcal{B}(\mathcal{S}_d), k = 1, 2, ..., d$ and $z \in V_k$,

$$\mathbb{P}\left(Z_d \in B | Z_0 = z\right) \geq \text{Const } \lambda_d\left(B \cap K\right). \tag{C.5}$$

We define the probability measure φ as follows, for each $B \in \mathcal{B}(\mathcal{S}_d)$,

$$\varphi(B) = \frac{\lambda_d\left(B \cap K\right)}{\lambda_d\left(K\right)}.$$

From (C.4) and (C.5), we can conclude that the condition (c) is verified.

Step 5. For each $B \in \mathcal{B}(\mathcal{S}_d)$ and $z \in V$,

$$
\begin{aligned}
\mathbb{P}\left(Z_{d+1} \in B | Z_0 = z\right) & \geq \mathbb{P}\left(Z_{d+1} \in B, Z_1 \in V | Z_0 = z\right) \\
& = \mathbb{P}\left(Z_{d+1} \in B | Z_1 \in V, Z_0 = z\right) \mathbb{P}(Z_1 \in V | Z_0 = z) \\
& \geq \eta \, \mathbb{P}\left(Z_{d+1} \in B | Z_1 \in V\right) \\
& \geq \text{Const } \lambda_d \left(B \cap K\right).
\end{aligned}
$$

Since $\gcd(d, d+1) = 1$, the condition (d) is now clearly fulfilled. $\qquad\square$

Beta walks with linearly place-dependent probabilities

Assume that Z_n converges in distribution to a random variable Z. Then

$$
Z \sim (1 - \xi)Z + \xi\Theta,
$$

where Θ is a discrete random variable independent from ξ such that

$$
\begin{cases}
\mathbb{P}(\Theta = E_j | Z = z) = p_j(z), j = 1, 2, ..., d; \\
\mathbb{P}(\Theta = E_0 | Z = z) = 1 - \sum_{j=1}^{d} p_j(z).
\end{cases}
$$

Lemma 3.1. *Assume that ξ and Z respectively have the probability density functions $g(u)$ for $u \in (0,1)$ and $f(z_1, z_2, \ldots, z_d)$ for $(z_1, z_2, \ldots, z_d)$ in the interior of $\mathcal{S}_d$. Then $Z \sim (1 - \xi)Z + \xi\Theta$ if and only if f and g satisfy the following equation*

$$
f(z_1, z_2, \ldots, z_d) = \sum_{j=0}^{d} T_j, \tag{C.6}
$$

where

$$
T_0 = \int_{z_1 + z_2 + \cdots + z_d}^{1} \frac{1}{u^d} f\left(\frac{z_1}{u}, \frac{z_2}{u}, \ldots, \frac{z_d}{u}\right) p_0 \left(\frac{z_1}{u}, \frac{z_2}{u}, \ldots, \frac{z_d}{u}\right) g(1 - u) du
$$

and

$$
\begin{aligned}
T_j = \int_{1 - z_j}^{1} \frac{1}{u^d} f &\left(\frac{z_1}{u}, \ldots \frac{z_{j-1}}{u}, \frac{z_j - 1 + u}{u}, \frac{z_{j+1}}{u}, \ldots \frac{z_d}{u}\right) \\
&\times p_j \left(\frac{z_1}{u}, \ldots \frac{z_{j-1}}{u}, \frac{z_j - 1 + u}{u}, \frac{z_{j+1}}{u}, \ldots \frac{z_d}{u}\right) g(1 - u) du
\end{aligned}
$$

for $j = 1, 2, \ldots, d$.

Proof. Denote $\tilde{Z} = (1 - \xi)Z + \xi\Theta$. We have

$$\mathbb{P}\left(\tilde{Z} < z\right) = \sum_{j=0}^{d} \int_0^1 \mathbb{P}\left(Z + \frac{1-u}{u}\Theta < \frac{z}{u}, \Theta = E_j\right) g(1-u)du, \qquad \text{(C.7)}$$

where for $y = (y_1, y_2, ..., y_d), z = (z_1, z_2, ..., z_d) \in \mathbb{R}^d$, we write $z < y$ if $z_j < y_j, j = 1, 2, ..., d$. Note that

$$\mathbb{P}\left(Z + \frac{1-u}{u}\Theta < \frac{z}{u}, \Theta = E_j\right) = \int_{\mathcal{S}_d} \mathbf{1}_{\left\{y < \frac{1}{u}(z-(1-u)E_j)\right\}} f(y) p_j(y) dy.$$

Observe that for $j = 0, 1, ..., d$,

$$\frac{\partial^d}{\partial z_1 \partial z_2 ... \partial z_d} \int_{\mathcal{S}_d} \mathbf{1}_{\left\{y < \frac{1}{u}(z-(1-u)E_j)\right\}} f(y) p_j(y) dy$$

$$= \frac{1}{u^d} f\left(\frac{1}{u}(z - (1-u)E_j)\right) p_j\left(\frac{1}{u}(z - (1-u)E_j)\right) \mathbf{1}_{\left\{\frac{1}{u}(z-(1-u)E_j \in \mathcal{S}_d\right\}}$$

$$= \frac{1}{u^d} f\left(\frac{1}{u}(z - (1-u)E_j)\right) p_j\left(\frac{1}{u}(z - (1-u)E_j)\right) \mathbf{1}_{\left\{1-z_j \leq u \leq 1\right\}}.$$

Therefore, the density function $\tilde{f}$ of $\tilde{Z}$ is computed as follows

$$\tilde{f}(z_1, z_2, ..., z_d) = \sum_{j=0}^{d} \int_{z_j}^1 \frac{1}{u^d} f\left(\frac{1}{u}(z - (1-u)E_j)\right) p\left(\frac{1}{u}(z - (1-u)E_j)\right) g(1-u)du$$

$$= \sum_{j=0}^{d} T_j.$$

The lemma is implied from the fact that Z and $\tilde{Z}$ have the same distribution if and only if $f(z) = \tilde{f}(z)$ for all z in the interior of $\mathcal{S}_d$. $\qquad \square$

Theorem 3.2. *Assume that*

(a) $\xi \sim \text{Beta}(1, \gamma)$;

(b) $p = (p_1, p_2, \ldots, p_d) : \mathcal{S}_d \to \mathcal{S}_d$ is defined by

$$p_k(z_1, z_2, \ldots, z_d) = \beta_k(1 - z_k) + \left(1 - \sum_{j=1}^{d+1} \beta_j + \beta_k\right) z_k, \ k = 1, 2, ..., d,$$

where $\beta_k > 0$ and $\sum_{j=1}^{d+1} \beta_j - \beta_k < 1$ for $k = 1, 2, ..., d+1$;

(c) $Z \sim \text{Dirichlet}(\beta_1\gamma, \beta_2\gamma, \ldots, \beta_d\gamma, \beta_{d+1}\gamma)$.

Then $Z \sim (1-\xi)Z + \xi\Theta$ yielding by Lemma 3.1 that Z_n converges to a Dirichlet distribution.

Proof. Let f and g be respectively the probability density functions of $\xi \sim \text{Beta}(1, \gamma)$ and $Z \sim \text{Dirichlet}(\beta_1\gamma, \beta_2\gamma, \ldots, \beta_d\gamma, \beta_{d+1}\gamma)$. We only have to verify that f and g satisfy the integral equation (C.6).

We have

$$
T_0 = \frac{\Gamma\left(\gamma\sum_{j=1}^{d+1}\beta_j\right)}{\prod_{j=1}^{d+1}\Gamma(\beta_j\gamma)} \prod_{j=1}^{d} z_j^{\beta_j\gamma-1} \int_{\sum_{j=1}^{d} z_j}^{1} \left(u - \sum_{j=1}^{d} z_j\right)^{\gamma\beta_{d+1}-1} u^{-\gamma\left(\sum_{j=1}^{d+1}\beta_j-1\right)-1}
$$

$$
\times \left[\gamma\beta_{d+1}u - \gamma\left(\sum_{j=1}^{d+1}\beta_j - 1\right)\left(u - \sum_{j=1}^{d} z_j\right)\right] du
$$

$$
= \frac{\Gamma\left(\gamma\sum_{j=1}^{d+1}\beta_j\right)}{\prod_{j=1}^{d+1}\Gamma(\beta_j\gamma)} \left(1 - \sum_{j=1}^{d} z_j\right)^{\beta_{d+1}\gamma} \prod_{j=1}^{d} z_j^{\beta_j\gamma-1},
$$

where we note that

$$
\int_z^1 u^{-b-1}(u-z)^{a-1}[au - b(u-z)]\, du = (1-z)^a.
$$

Similarly, for $k = 1, 2, \ldots, d$, we also obtain that

$$
T_k = \frac{\Gamma\left(\gamma\sum_{j=1}^{d+1}\beta_j\right)}{\prod_{j=1}^{d+1}\Gamma(\beta_j\gamma)} \left(1 - \sum_{j=1}^{d} z_j\right)^{\beta_{d+1}\gamma-1} z_k \prod_{j=1}^{d} z_j^{\beta_j\gamma-1}.
$$

Therefore,

$$
\sum_{k=0}^{d} T_k = \frac{\Gamma\left(\gamma\sum_{j=1}^{d+1}\beta_j\right)}{\prod_{j=1}^{d+1}\Gamma(\beta_j\gamma)} \left(1 - \sum_{j=1}^{d} z_j\right)^{\beta_{d+1}\gamma-1} \prod_{j=1}^{d} z_j^{\beta_j\gamma-1} = f(z_1, z_2, \ldots, z_d).
$$

Remarks.

- For $d = 1$, $p_1(z) = \beta_1(1 - z) + (1 - \beta_2)z$. This is the one-dimensional case considered by McKinlay and Borovkov in [7].

- For $d \geq 1$, if $\sum_{j=1}^{d+1} \beta_j = 1$ then we obtain the model considered by Sethuraman in [10].

Random walks in [0,1] based on Friedman's urn-type schemes

In this section, we are interested in the following random walk model in the unit interval $\mathcal{S}_1 = [0, 1]$ with the following assumptions

1. At time n, there are potentials L_n and R_n at 0 and 1 respectively, and we denote as Z_n the position of the particle at this time.

2. At time n, with probability $\frac{L_n}{L_n + R_n}$, the potential at 0 will increase a value proportional to Z_n, i.e. the distance from 0 to the current position of the particle, and then the particle is pulled to a new uniformly random position Z_{n+1} inside the interval $(0, Z_n)$. Otherwise, with probability $\frac{R_n}{L_n + R_n}$, the potential at 1 will increase a value proportional to $1 - Z_n$, i.e. the distance from 1 to the current position of the particle, and then the particle is pulled to a new uniformly random position Z_{n+1} inside the interval $(Z_n, 1)$.

For more general, we can consider the following random recursion

$$Z_{n+1} = Z_n(1 - \xi_n^{(L)})\mathbf{1}_{\left\{U_n < \frac{L_n}{L_n + R_n}\right\}} + (Z_n + (1 - Z_n)\xi_n^{(R)})\mathbf{1}_{\left\{U_n > \frac{L_n}{L_n + R_n}\right\}},$$

together with two "urns" L_n, R_n respectively at 0 and 1, defined as follows

$$\begin{aligned}
L_{n+1} &= L_n + f(Z_n)\mathbf{1}_{\left\{U_n < \frac{L_n}{L_n + R_n}\right\}}, \\
R_{n+1} &= R_n + f(1 - Z_n)\mathbf{1}_{\left\{U_n > \frac{L_n}{L_n + R_n}\right\}},
\end{aligned}$$

where $f : [0, 1] \to [0, +\infty)$ is some function; $\xi_n^{(L)}$ and $\xi_n^{(R)}, n \geq 1$, are independent random variables taking values in $[0, 1]$; $U_n, n \geq 1$ are i.i.d uniformly

distributed random variables in $[0,1]$ and independent of $\xi_n^{(L)}$, $\xi_n^{(R)}$, $Z_n, n \geq 1$. Since the probabilities jumping to the left or the right depend on (L_n, R_n), the random walk $\{Z_n\}_{n \geq 1}$ is no longer a Markov chain but depends on its history.

We assume from now on that $\xi_n^{(L)} = \xi_n^{(R)} = \xi_n, n \geq 1$, are uniformly distributed in $[0,1]$. Note that in the case where $f(x) = \beta > 0$ for all $x \in [0,1]$, the process $(L_n, R_n)_{n \geq 1}$ is reduced to the classical Friedman urn w.r.t the matrix

$$\begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix}.$$

It is well-known in this case that $L_n / (R_n + L_n)$ converges almost surely to $1/2$ as $n \to \infty$ (see e.g. [3], Corollary 5.2). Therefore, one can show that Z_n converges in distribution to the arcsine law Beta$(1/2, 1/2)$.

In the remaining part of this paper, for simplicity, we restrict to the case where $f(x) = x$ for all $x \in [0,1]$.

Lemma 4.1. $L_n + R_n \to \infty$ almost surely as $n \to \infty$.

Proof. There are the following cases that might occur:

- Case 1: There is $\epsilon \in (0, 1/2)$ such that

$$\limsup_{n \to \infty} Z_n > \epsilon \quad \text{and} \quad \liminf_{n \to \infty} Z_n < 1 - \epsilon.$$

  In this case, we will show that there exist infinitely many ns such that $Z_n \in (\epsilon, 1 - \epsilon)$; therefore $L_{n+1} + R_{n+1} - (L_n + R_n) > \epsilon$ for all such n implying that $R_n + L_n \to \infty$. Indeed, assume that $Z_n \in [0, \epsilon] \cup [1 - \epsilon, 1]$ for all large n; moreover the assumption on lim inf and lim sup imply that Z_n visit both segments infinitely often; hence it must make infinitely many moves to the right when $Z_n \in [0, \epsilon]$. However, each time it does so, Z_{n+1} will land in $(\epsilon, 1 - \epsilon)$ with probability at least $1 - 2\epsilon > 0$ so eventually this will happen a.s. contradicting the assumption.

- Case 2: $\lim_{n \to \infty} Z_n = 0$. In this case, for each $\epsilon > 0$, there exists N large enough such that for all $n \geq N$, $Z_n \in (0, \epsilon)$. If $Z_{n+1} \in (Z_n, 1)$ infinitely many times for $n \geq N$, then, since $\mathbb{P}(Z_{n+1} \in (\epsilon, 1) | Z_{n+1} \in (Z_n, 1)) \geq 1 - \epsilon$, there exists $N' > N$ such that $Z_{N'} \in (\epsilon, 1)$ which is against our assumption. This contradiction implies that there exists N_0 such that $Z_{n+1} \in (0, Z_n)$ for all $n \geq N_0$. Hence, we have

$$L_n = L_{N_0}(1 + \eta_1 + \eta_1 \eta_2 + \ldots + \eta_1 \eta_2 \ldots \eta_{n-N_0}),$$

where $\eta_1, \eta_2, \ldots$ are independent uniformly distributed random variable in [0,1]. It is well-known that $1 + \eta_1 + \eta_1\eta_2 + \ldots + \eta_1\eta_2\ldots\eta_n$ converges almost surely to a finite random variable as $n \to \infty$ (see e.g. [11]). On the other hand, for all $n \geq N_0$

$$\mathbb{P}\left(Z_{n+1} \in (Z_n, 1) | Z_j, j = 1, 2, \ldots, n\right) = \frac{R_n}{L_n + R_n} \geq \frac{R_{N_0}}{R_{N_0} + L_\infty} > 0, \text{ a.s.,}$$

where $L_\infty := \lim_{n \to \infty} L_n$. The above contradiction implies that this case cannot happen.

- Case 3: $\lim_{n \to \infty} Z_n = 1$, this case is similar to Case 2.

$\square$

Lemma 4.2.
$$\frac{1}{12} \leq \liminf_{n \to \infty} \frac{L_n}{R_n} \leq \limsup_{n \to \infty} \frac{L_n}{R_n} \leq 12.$$

Proof. The proof is based on the arguments from the previous statement. For simplicity rewrite the recursion equivalently as

$$Z_{n+1} = \xi_n Z_n \mathbf{1}_{\left\{U_n < \frac{L_n}{L_n + R_n}\right\}} + (Z_n + \xi_n(1 - Z_n))\mathbf{1}_{\left\{U_n > \frac{L_n}{L_n + R_n}\right\}}.$$

We know that Z_n makes a.s. infinitely many steps to the left as well as to the right. Hence there exists a sequence of stopping times

$$\tau_1 < \eta_1 < \tau_2 < \eta_2 < \ldots$$

going to infinity, such that

$$Z_{n+1} < Z_n, \text{ if } n \in [\tau_i, \eta_i) \text{ for some } i,$$
$$Z_{n+1} > Z_n, \text{ if } n \in [\eta_i, \tau_{i+1}) \text{ for some } i.$$

Consequently,

$$L_{\eta_i} - L_{\tau_i} = Z_{\eta_i}(1 + \xi_{\tau_i+1} + \xi_{\tau_i+1}\xi_{\tau_i+2} + \cdots + \xi_{\tau_i+1}\xi_{\tau_i+2}\cdots\xi_{\eta_i-1})$$
$$\leq 1 + \xi_{\tau_i+1} + \xi_{\tau_i+1}\xi_{\tau_i+2} + \cdots + \xi_{\tau_i+1}\xi_{\tau_i+2}\cdots\xi_{\eta_i-2} + 1,$$
$$R_{\eta_i} - R_{\tau_i} = 0.$$

At the same time with probability $1/2$ we have $\xi_{\eta_i-1} < 1/2$. Assuming this, we get that $Z_{\eta_i} \leq \xi_{\eta_i-1} < 1/2$. Since at the time η_i the walk moves to the right, we have $R_{\eta_i+1} - R_{\eta_i} = 1 - Z_{\eta_i} > 1/2$. Consequently, we obtain that

$$L_{\eta_n+1} \leq \sum_{j=1}^{n} \nu_n, \quad R_{\eta_n+1} \geq \sum_{j=1}^{n} \tilde{\nu}_n$$

where all ν_n and $\tilde{\nu}_n$ are independent and ν_n has the distribution of $2 + \sum_{k=1}^{\infty} \prod_{i=1}^{k} \xi_i$ yielding $\mathbb{E}\,\nu_n = 3$ while $\tilde{\nu}_n$ equals $1/2$ and 0 with equal probabilities $1/2$, so that $\mathbb{E}\,\tilde{\nu}_n = 1/4$. Now the strong law of large numbers together with monotonicity of L_n and R_n imply that

$$\limsup_{n\to\infty} \frac{L_n}{R_n} \leq \frac{3}{1/4} = 12.$$

The complimentary inequality can be proved identically. $\qquad\square$

Lemma 4.3. $\zeta_n := \dfrac{L_n}{L_n + R_n}$ *converges almost surely to* $\zeta_\infty \in (0,1)$ *as* $n \to \infty$.

Remark. From the previous lemma it follows that

$$1/13 \leq \liminf \zeta_n \leq \limsup \zeta_n \leq 12/13.$$

Proof. Let $\mathcal{F}_n$ stand for the σ-algebra generated by $Z_1, ..., Z_n$. Let us consider the quantity

$$W_n = \left(\frac{1}{2} - \frac{L_n + Z_n}{L_n + R_n}\right)^2 + \frac{1}{L_n + R_n}.$$

We have

$$\mathbb{E}\left(W_{n+1}|\mathcal{F}_n\right)$$

$$= \frac{L_n}{L_n + R_n}\mathbb{E}\left(W_{n+1}\,\middle|\,U_n < \frac{L_n}{L_n + R_n},\mathcal{F}_n\right)$$

$$+ \frac{R_n}{L_n + R_n}\mathbb{E}\left(W_{n+1}\,\middle|\,U_n > \frac{L_n}{L_n + R_n},\mathcal{F}_n\right)$$

$$= \frac{L_n}{L_n + R_n}\int_0^{Z_n}\frac{1}{Z_n}\left[\left(\frac{1}{2} - \frac{L_n + Z_n + u}{L_n + R_n + Z_n}\right)^2 + \frac{1}{L_n + R_z + Z_n}\right]du$$

$$+ \frac{R_n}{L_n + R_n}\int_{Z_n}^1\frac{1}{1 - Z_n}\left[\left(\frac{1}{2} - \frac{L_n + u}{L_n + R_n + 1 - Z_n}\right)^2 + \frac{1}{L_n + R_n + 1 - Z_n}\right]du$$

$$= \frac{L_n}{L_n + R_n}\frac{3\left(L_n - R_n\right)^2 + 12\left(L_n - R_n\right)Z_n + 12\left(L_n + R_n + Z_n\right) + 13Z_n^2}{12\left(L_n + R_n + Z_n\right)^2}$$

$$+ \frac{R_n}{L_n + R_n}\frac{3\left(L_n - R_n\right)^2 + 12\left(L_n - R_n\right)Z_n + 12\left(L_n + R_n + Z_n\right) + 13(1 - Z_n)^2}{12\left(L_n + R_n + 1 - Z_n\right)^2}.$$

Substituting $\zeta_n = \dfrac{L_n}{L_n + R_n}$ and $\varepsilon_n = \dfrac{1}{L_n + R_n}$, or equivalently, $L_n = \dfrac{\zeta_n}{\epsilon_n}, R_n = \dfrac{1 - \zeta_n}{\epsilon_n}$, to the above identity, we obtain that

$$\mathbb{E}\left(W_{n+1} - W_n|\mathcal{F}_n, Z_n = z\right) = \frac{\epsilon_n\left[r_0(\zeta_n, z) + r_1(\zeta_n, z)\epsilon_n + \cdots + r_5(\zeta_n, z)\epsilon_n^5\right]}{6(\epsilon_n z + 1)^2(1 + \epsilon_n(1 - z))^2}, \tag{C.8}$$

where

$$\begin{aligned}
r_0(\zeta, z) &= -24z\zeta^3 + 36z\zeta^2 + 12\,\zeta^3 - 18z\zeta - 24\zeta^2 + 3z + 15\zeta - 3\\
&= -3(2\zeta - 1)^2(\zeta z + (1 - z)(1 - \zeta)),\\
r_1(\zeta, z) &= -30\,z^2\zeta^2 - 12\,z\zeta^3 + 30\,z^2\zeta + 24\,z\zeta^2\\
&\quad + 6\,\zeta^3 - 7\,z^2 - 38\,z\zeta - 12\,\zeta^2 + 14\,z + 13\,\zeta - 7,\\
r_2(\zeta, z) &= 10\,z^3\zeta - 12\,z^2\zeta^2 - 5\,z^3 - 9\,z^2\zeta + 7\,z^2 - 19\,z\zeta + 4\,z + 6\,\zeta - 6,\\
r_3(\zeta, z) &= -z\left(6\,z^3\zeta^2 - 6\,z^3\zeta - 12\,z^2\zeta^2 - 17\,z^3\right.\\
&\quad \left. -12\,z^2\zeta + 6\,z\zeta^2 + 28\,z^2 + 30\,z\zeta - 23\,z - 6\,\zeta + 12\right),\\
r_4(\zeta, z) &= -6\,z^2\left(1 - z\right)\left(-2\,z^2\zeta + z^2 + 2\,z\zeta + 1\right),\\
r_5(\zeta, z) &= -6\,z^4\left(z - 1\right)^2.
\end{aligned}$$

One can show that $\max_{x,y\in[0,1]} r_i(x,y) \leq 0$ for $i = 0,2,3,4,5$. From the proof of Lemma 4.1, $\epsilon_n \to 0$, and one can show that

$$\max_{x,y\in[0,1]} (r_0(x,y) + r_1(x,y))\epsilon \leq 0$$

for $0 \leq \epsilon < 0.5$. Hence W_n is a supermartingale. Therefore, by Doob's martingale convergence theorem, there exists $W_\infty := \lim_{n\to\infty} W_n$ almost surely.

Observe that $\zeta_n \in \left\{ \frac{1}{2} - \sqrt{W_n - \frac{1}{L_n+R_n} - \frac{Z_n}{L_n+R_n}}, \frac{1}{2} + \sqrt{W_n - \frac{1}{L_n+R_n} - \frac{Z_n}{L_n+R_n}} \right\}$. On the other hand, note that

$$|\zeta_{n+1} - \zeta_n| \leq \max\left\{ \frac{L_n}{L_n + R_n} \cdot \frac{1 - Z_n}{L_{n+1} + R_{n+1}}, \frac{R_n}{L_n + R_n} \cdot \frac{Z_n}{L_{n+1} + R_{n+1}} \right\} \to 0$$

as $n \to \infty$. It implies that $\limsup_{n\to\infty} \zeta_n = \liminf_{n\to\infty} \zeta_n := \zeta_\infty$ almost surely.

$\square$

Lemma 4.4. $\zeta_\infty = \frac{1}{2}$ *almost surely.*

Proof. Suppose $\mathbb{P}(W_\infty = 0) < 1$. Then, there exists $\epsilon > 0$ such that $\mathbb{P}(W_\infty > \epsilon) > 0$. Let us denote the stopping time

$$\tau_m = \inf\left\{ n \geq m \ : \ W_n < \frac{\epsilon}{2} \right\}.$$

Since $\mathbb{P}(W_\infty > \epsilon) > 0$, there exists m such that $\mathbb{P}(\tau_m = \infty) > 0$. Let us consider $Y_n = W_{n\wedge\tau_m}$. Y_n is also a supermartingale, hence, there exists $Y_\infty = \lim_{n\to\infty} Y_n$. From (C.8), observe that for $N > 0$,

$$\mathbb{E}(W_{m+N}) - \mathbb{E}(W_m) =$$
$$-\mathbb{E}\left(\sum_{n=m}^{m+N} \frac{1}{2} (2\zeta_n - 1)^2 (Z_n\zeta_n + (1 - Z_n)(1 - \zeta_n)) \varepsilon_n(1 + O(\varepsilon_n)) \right).$$

Therefore,

$$\mathbb{E}(Y_\infty) - \mathbb{E}(Y_m) \tag{C.9}$$
$$= -\mathbb{E}\left(\sum_{n=m}^{\tau_m} \frac{1}{2} (2\zeta_n - 1)^2 (Z_n\zeta_n + (1 - Z_n)(1 - \zeta_n)) \varepsilon_n(1 + O(\varepsilon_n)) \right).$$

Note that $W_n \geq \frac{\epsilon}{2}$, for all $n \leq \tau_m$. Hence, combining with the remark in Lemma 4.3, one can show that there exists $\gamma > 0$ such that on the event $\{\tau_m = \infty\}$,

$$\frac{1}{2} (2\zeta_n - 1)^2 (Z_n \zeta_n + (1 - Z_n)(1 - \zeta_n)) > \gamma,$$

for large enough n. Since $\mathbb{P}(\tau_m = \infty) > 0$, the LHS of (C.9) is finite while the RHS is divergent. This contradiction proves the lemma. $\qquad\square$

Theorem 4.5. *As $n \to \infty$, Z_n converges in distribution to the arsine law* Beta $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Proof. Using Lemma 4.4, for small $\epsilon > 0$, there exists (random) N such that

$$\frac{1}{2} - \epsilon \leq \frac{L_n}{L_n + R_n} \leq \frac{1}{2} + \epsilon$$

for all $n \geq N$. We couple $\{Z_n\}_{n \geq 0}$ with two random walks $\{\check{Z}_n\}_{n \geq 0}$ and $\{\hat{Z}_n\}_{n \geq 0}$ defined as follows:

- For $0 \leq n \leq N$, set $\check{Z}_n = \hat{Z}_n = Z_n$.

- For $n \geq N$, set

$$\check{Z}_{n+1} = \begin{cases} \xi_n \check{Z}_n & \text{if } U_n \leq \frac{1}{2} - \epsilon; \\ \check{Z}_n + \xi_n(1 + \check{Z}_n) & \text{if } U_n > \frac{1}{2} - \epsilon, \end{cases}$$

   and

$$\hat{Z}_{n+1} = \begin{cases} \xi_n \hat{Z}_n & \text{if } U_n \leq \frac{1}{2} + \epsilon; \\ \hat{Z}_n + \xi_n(1 - \hat{Z}_n) & \text{if } U_n > \frac{1}{2} + \epsilon. \end{cases}$$

Assume that for some $n \geq N$, $\hat{Z}_n \leq Z_n \leq \check{Z}_n$. We observe that:

- When $\check{Z}_n$ chooses left, Z_n also chooses left since $U_n \leq \frac{1}{2} - \epsilon < \frac{L_n}{L_n + R_n}$. In this case, $Z_{n+1} = \xi_n Z_n \leq \xi_n \check{Z}_n = \check{Z}_{n+1}$. When $\check{Z}_n$ chooses right, Z_n might choose left or right, but we still have $Z_{n+1} \leq Z_n + \xi_n(1 - Z_n) \leq \check{Z}_n + \xi_n(1 - \check{Z}_n) = \check{Z}_{n+1}$.

- When Z_n chooses left, $\hat{Z}_n$ also chooses left since $U_n \leq \frac{L_n}{L_n + R_n} < \frac{1}{2} + \epsilon$. In this case, $Z_{n+1} = \xi_n Z_n \geq \xi_n \hat{Z}_n = \hat{Z}_{n+1}$. When Z_n chooses right, $\hat{Z}_n$ might choose left or right, but we still have $\hat{Z}_{n+1} \leq \hat{Z}_n + \xi_n(1 - \hat{Z}_n) \leq Z_n + \xi_n(1 - Z_n) = Z_{n+1}$.

By induction, we obtain that for all $n \geq 0$, $\hat{Z}_n \leq Z_n \leq \tilde{Z}_n$. Therefore, we have

$$\mathbb{P}(\tilde{Z}_n \leq x) \leq \mathbb{P}(Z_n \leq x) \leq \mathbb{P}(\hat{Z}_n \leq x)$$

for all $n \geq 0$ and $x \in [0,1]$. On the other hand, by Theorem 3.2, $\tilde{Z}_n$ and $\hat{Z}_n$ weakly converge respectively to Beta $\left(\frac{1}{2} + \epsilon, \frac{1}{2} - \epsilon\right)$ and Beta $\left(\frac{1}{2} - \epsilon, \frac{1}{2} + \epsilon\right)$ as $n \to \infty$. Since ϵ is arbitrarily small, the theorem is proved. $\qquad\square$

Appendix

Proof of Lemma 2.2.

(a) For $u = (u_1, ..., u_d) \in K$ and $z = (z_1, ..., z_d) \in V_0$, we will consider

$$v = (v_1, v_2, \ldots, v_d) := G_z^{-1}(u).$$

Note that $u_0 \leq 1 - u_d \leq 1 - \delta \leq z_0$, $z_j \leq 1 - z_0 \leq \delta$ and thus

$$s(1-t)^{d-1} - \delta \leq u_j - z_j \leq v_j = u_j - \frac{u_0}{z_0} z_j \leq u_j \leq t$$

for $j = 1, 2, ..., d$. Therefore, for $j = 1, 2, \ldots, d$, we have

$$s(1-t)^{d-1} - \delta \leq v_j \leq \frac{v_j}{1 - \sum\limits_{l=j+1}^{d} v_l} = \frac{u_j - z_j \dfrac{u_0}{z_0}}{1 - \sum\limits_{l=j+1}^{d} u_l + \sum\limits_{l=j+1}^{d} z_l \dfrac{u_0}{z_0}} \leq \frac{u_j}{1 - \sum\limits_{l=j+1}^{d} u_l} \leq t.$$

It implies that $v = G_z^{-1}(u) \in K_0$ for each $u \in K$ and $z \in V_0$, where we denote

$$K_0 = \left\{ (v_1, \ldots, v_d) \in \mathcal{S}_d : s(1-t)^{d-1} - \delta \leq \frac{v_j}{1 - \sum_{l=j+1}^{d} v_l} \leq t, j = 1, 2, ..., d \right\}.$$

(C.10)

Observe that $T^{-1}(K_0) = [s(1-t)^{d-1} - \delta, t]^d$. Thus, $T^{-1} \circ G_z^{-1}(K) \subset [s(1-t)^{d-1} - \delta, t]^d$.

(b) For $u \in K, z \in V_k$, let

$$v = (v_1, v_2, \ldots, ..., v_k) := G_{R_k(z)}^{-1}(R_k(u))$$

$$= \left(u_0 - z_0 \frac{u_k}{z_k}, u_1 - z_1 \frac{u_k}{z_k}, \ldots, u_{k-1} - z_{k-1} \frac{u_k}{z_k}, u_{k+1} - z_{k+1} \frac{u_k}{z_k}, \ldots, u_d - z_d \frac{u_k}{z_k} \right).$$

Note that $z_l \leq 1 - z_k \leq \delta$ for $l \in \{0, 2, ..., d\} \setminus \{k\}$ and $u_k \leq \max\{u_d, 1 - u_d\} \leq 1 - \delta \leq z_k$. Therefore, we observe that:

(i) for $k + 1 \leq j \leq d$,

$$\frac{v_j}{1 - \sum\limits_{l=j+1}^{d} v_l} = \frac{u_j - z_j \dfrac{u_k}{z_k}}{1 - \sum\limits_{l=j+1}^{d} u_l + \sum\limits_{l=j+1}^{d} z_l \dfrac{u_k}{z_k}} \leq \frac{u_j}{1 - \sum\limits_{l=j+1}^{d} u_l} \leq t$$

and

$$\frac{v_j}{1 - \sum\limits_{l=j+1}^{d} v_l} \geq v_j = u_j - z_j \frac{u_k}{z_k} \geq u_j - z_j \geq s(1-t)^{d-1} - \delta.$$

(ii) for $j = 1$, we have

$$\frac{v_1}{1 - \sum\limits_{l=2}^{d} v_l} = \frac{u_0 - z_0 \dfrac{u_k}{z_k}}{u_0 + \sum\limits_{l=1}^{d} z_l \dfrac{u_k}{z_k}} = 1 - \frac{u_k}{\left(1 - \sum\limits_{l=1}^{d} u_l\right) z_k + u_k \sum\limits_{l=1}^{d} z_l}$$

$$\leq 1 - \frac{u_k}{\left(1 - \sum\limits_{l=k}^{d} u_l\right) z_k + u_k} \leq 1 - \frac{u_k}{1 - \sum\limits_{l=k+1}^{d} u_l} \leq 1 - s$$

and

$$\frac{v_1}{1 - \sum\limits_{l=2}^{d} v_l} \geq v_1 = u_0 - z_0 \frac{u_k}{z_k} \geq (1-t)^d - \delta.$$

(iii) for $2 \leq j \leq k$,

$$s(1-t)^{d-1} - \delta \leq v_j \leq \frac{v_j}{1 - \sum\limits_{l=j+1}^{d} v_l} = \frac{u_{j-1} - z_{j-1} \dfrac{u_k}{z_k}}{1 - \sum\limits_{l=j}^{d} u_l + \sum\limits_{l=j}^{d} z_l \dfrac{u_k}{z_k}} \leq \frac{u_{j-1}}{1 - \sum\limits_{l=j}^{d} u_l} \leq t.$$

Therefore,

$$v \in T\left([(1-t)^d - \delta, 1 - s] \times [s(1-t)^{d-1} - \delta, t]^{d-1}\right).$$

132

Bibliography

[1] Borovkov, A. A. Ergodicity and Stability of Stochastic Processes (1998). Wiley, New York.

[2] DeGroot, M. H.; Rao, M. M. Stochastic give-and-take. J. Math. Anal. Appl. **7** (1963), 489–498.

[3] Freedman, D. A. Bernard Friedman's Urn. Ann. Math. Statist. **36** (1965), no. 3, 956–970.

[4] Diaconis, P.; Freedman, D. Iterated random functions. SIAM Rev. **41** (1999), no. 1, 45–76.

[5] Hitczenko, P.; Letac, G. Dirichlet and quasi-Bernoulli laws for perpetuities. J. Appl. Probab. **51** (2014), no. 2, 400–416.

[6] Ladjimi F.; Peigné. M. Iterated function systems with place dependent probabilities and application to the Diaconis-Friedman's chain on [0,1]. ArXiv e-prints, arXiv:1707.07237.

[7] McKinlay, S.; Borovkov, K. On explicit form of the stationary distributions for a class of bounded Markov chains. J. Appl. Probab. **53** (2016), no. 1, 231–243

[8] Pacheco-González, C. G. Ergodicity of a bounded Markov chain with attractiveness towards the centre. Statist. Probab. Lett. **79** (2009), no. 20, 2177–2181.

[9] Ramli, M. A.; Leng, G. The stationary probability density of a class of bounded Markov processes. Adv. in Appl. Probab. **42** (2010), no. 4, 986–993.

[10] Sethuraman, J. A constructive definition of Dirichlet priors. Statist. Sinica **4** (1994), no. 2, 639–650.

[11] Vervaat, W. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. Adv. in Appl. Probab. **11** (1979), no. 4, 750–783.