Limit theorems for a multi-type Galton-Watson process in random independent environment

Thi da Cam Pham

UNIVERSITÉ DE TOURS

École Doctorale MIPTIS

Institut Denis Poisson

THÈSE présentée par :

Thi Da Cam PHAM

soutenue le : 05 Décembre 2018

pour obtenir le grade de Docteur de l'Université de Tours

Discipline/ Spécialité : Mathématiques

THÉORÈMES LIMITE POUR UN PROCESSUS DE GALTON-WATSON MULTI-TYPE EN ENVIRONNEMENT ALÉATOIRE INDÉPENDANT

THÈSE DIRIGÉE PAR :

PEIGNÉ Marc	Professeur à l'Université de Tours

RAPPORTEURS :

ALSMEYER Gerold	Professeur à l'Université de Münster
DELMAS Jean-François	Professeur à l'École des Ponts ParisTech

JURY :

ALSMEYER Gerold	Professeur à l'Université de Münster
BROFFERIO Sara	Maître de Conférences à l'Université de Paris-Sud
DELMAS Jean-François	Professeur à l'École des Ponts ParisTech
LE PAGE Émile	Professeur à l'Université de Bretagne-Sud
PEIGNÉ Marc	Professeur à l'Université de Tours
RASCHEL Kilian	Chargé de recherches à l'Université de Tours

Talent can take you so far.
Hard work can take you anywhere.

Hiruma Yoichi (Eyeshield 21)

Remerciements

It is so easy to render thanks by saying “thank you” verbally that we do everyday, which becomes even automatic. Even so, on the occasion of my thesis, I would like to at least officially utter my gratitudes to :

- The Laboratory of Tours for its support that has enabled me to carry out my research.

- VIASM to support me during the time I was in Hanoi.

- the Probability group at the laboratory of Tours to be with me during three years.

- Kilian Raschel for inviting many people to the probability seminar so that I could discover a variety of different subjects.

- Vincent Perrollaz for his devoted help and kind patience in teaching me Python, Yann Jullien for his short but memorable communication in C and the mission of using the “giant computer".

- Sara Brofferio for her course in “Stochastic dynamical system” in Halong, Vietnam that encouraged me in different directions of research.

- to Romain Abraham and Jean-François Delmas for their course about “Galton-Watson trees” in Hanoi which gave me more intuitions of my topic.

- Gerold Alsmeyer for having invited me one month in Münster to learn more about renewal theory.

- Emile Le Page for his availability and generosity of giving me ideas to deal with my questions.

- the rapporteurs Gerold Alsmeyer and Jean-François Delmas who gave me important remarks and questions that I can improve my manuscript.

- the faculty thesis committee for spending their time on coming to my thesis defense.

- many people, in particular Florent Malrieu for his caring about the PhD students group, Cedric Lecouvert for his gentle humor, Guy Barles for his welcome helping hand, Elodie and Anouchka for their efficient administrative help and their friendly actions, Sandrine for her supports of resources, etc.

- my friends, who spoke with me, who spent time with me, who listened to me, who were waiting for me silently, who played guitar for me, who introduced me classical music, who taught me to walk everyday, who picked raspberries for me, who climbed with me over mountains, who showed me racehorses, who played boardgames with me, who went to the Christmas markets with me, who played hide-and-seek with me, who gave me flowers, who laughed with me, who danced with me, who cooked for me, who

brought me food, who went with me to the hospital, who always welcomed me, who gave me a part of their garden, who showed me the sky full of stars, who let me pick up eggs of their hens, who appreciate me as who I am ...

- Quy and Xuyen at VIASM for their warm hearts for me, a girl from the south.

- Jin Garcien and Nadine, the staff at my residence Grandmont, who cared about me kindly.

- Germain Rousseau from the International Relations Project Manager at the Euraxess Service Centre, who accompanied me over many troubles that happened to a stranger student in France.

- my family, my parents, my sisters to always trust in me, always be beside me, no matter where I am.

- last but not least, my advisor Marc Peigné, for his strict persistence. I wouldn't be where I am now without him. This is how the steel was tempered!

Thanks for all the smiles I have received during those three years of my PhD.

Résumé

La théorie des processus de branchement multi-type en environnement i.i.d. est considérablement moins développée que dans le cas univarié, et les questions fondamentales ne sont pas résolues en totalité à ce jour. Les réponses exigent une compréhension profonde du comportement des produits des matrices i.i.d. à coefficients positifs.

Sous des hypothèses assez générales et lorsque les fonctions génératrices de probabilité des lois de reproduction sont "linéaire fractionnaires", nous montrons que la probabilité de survie à l'instant n du processus de branchement multi-type en environnement aléatoire est proportionnelle à $\frac{1}{\sqrt{n}}$ lorsque $n \to \infty$.

La démonstration de ce résultat suit l'approche développée pour étudier les processus de branchement uni-variés en environnement aléatoire i. i. d. Il utilise de façon cruciale des résultats récents portant sur les fluctuations des normes de produits de matrices aléatoires i.i.d.

Mots clés : Multi-type branching process, Survival probability, Random environment, Critical case, Exit time, Markov chains, Product of random matrices.

Abstract

The theory of multi-type branching process in i.i.d. environment is considerably less developed than for the univariate case, and fundamental questions are up to date unsolved. Answers demand a deep understanding of the behaviour of products of i.i.d. matrices with non-negative entries.

Under mild assumptions, when the probability generating functions of the reproduction laws are fractional-linear, the survival probability of the multi-type branching process in random environment up to time n is proportional to $\frac{1}{\sqrt{n}}$ as $n \to \infty$.

Techniques for univariate branching processes in random environment and methods from the theory of products of i.i.d. random matrices are required.

Keywords : Multi-type branching process, Survival probability, Random environment, Critical case, Exit time, Markov chains, Product of random matrices.

Table des matières

Liste des tableaux

Table des figures

Introduction

A search of any Mathematical library may reveal a number of volumes devoted to branching processes. Let us mention two books among the most fundamental ones where we found many informations : the ones by T. E. Harris [Harris, 1963] and K. B. Athreya & P. E. Ney [Athreya et Ney, 1972]. Not long after, P. Jagers [Jagers, 2005] introduced general branching processes in a more systematic way.

Branching processes are stochastic processes describing the dynamics of a population of individuals which reproduce and die independently, according to some specific probability distributions. They have numerous applications in population biology and phylogenetics. There are many types of branching processes : discrete time, continuous time, single-type or multi-type (with finitely or infinitely many types), individual reproduction rules may depend on the actual size of the population (population size-dependent branching process), immigration at any generation may disturb the evolution of the size of the population, etc. The method employed in branching processes allow questions about extinction and survival in ecology and evolutionary biology to be addressed.

The study of branching processes began in the 1840's with Irénée-Jules Bienaymé and was advanced in the 1870's with the work of Reverend Henry William Watson and Francis Galton. The simplest and most frequently applied branching process is named after Galton and Watson (or GW-process for short), it is a fundamental example of discrete-time Markov chain.

Informally, we may describe that a population of individuals evolves in discrete time as follows. A Galton-Watson branching process starts with one initial ancestor, sometimes it is possible to have a random number of initial ancestors, in which case it will be explicitly stated. This single ancestor produces a certain number of offspring according to a given probability distribution. All the individuals of the population are assumed to be of the same type and only live for a single unit time period ; but at the end of their life-length, each of them produces a random number (possible 0) of offspring. The number of children born by an individual does not depend on how many other individuals are present. The number of offspring for distinct individuals are mutually independent, and also independent of the offspring of other individuals from earlier generation. Furthermore, they are identically distributed, with common distribution. And the system regenerates.

The following figure allows us to visuallize a Galton-Watson process in the image of a tree, in which case it starts with one founding ancestor and has two children of the first generation, two children of the second generation, four children of the third generation and so on.

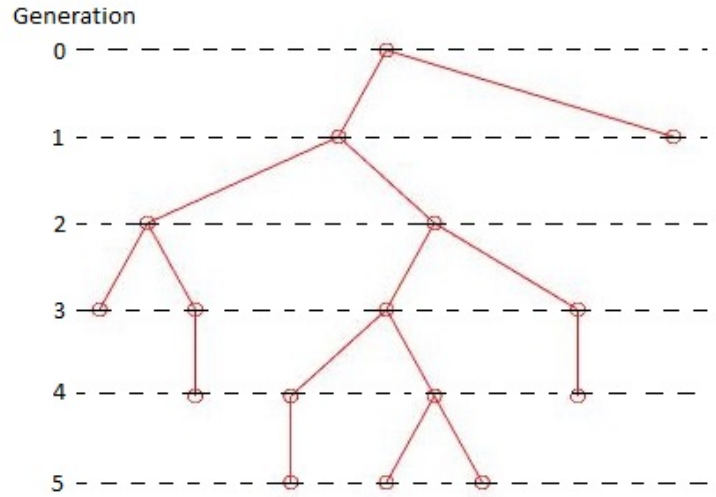

FIGURE 1 – Galton-Watson tree.

The expression "single" type refers to the fact that all individuals are of single type such as the same gender, same cell type, or same geotype or phenotype while in the multi-type process, each parent may have offspring of different types. For instance in genetics, genes can be classified as wild or mutant types and mutation change a wild type into a mutant type.

More rigorously, let us introduce a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ and, for any $n \geq 0$, denote by Z_n the random number of individuals in the n^{th} generation of this population; the numbers Z_n are random variables defined on $(\Omega, \mathfrak{F}, \mathbb{P})$ and may present the population size of animals, plants, cells, or genes at time n or generation n. At each generation n, each individual $i \in \{1, ..., Z_n\}$ has N_i^n children, where $\left(N_i^n, i \geq 1,\ n \geq 0\right)$ are i.i.d. random variables with distribution $\mu = (\mu(k))_{k\geq 0}$. The number of individuals at generation $n+1$ is therefore :

$$Z_{n+1} = N_1^n + N_2^n + ... + N_{Z_n}^n,$$

with the convention if $Z_n = 0$ then $Z_{n+1} = 0$. Since this is a sum of i.i.d. random variables, its distribution depends only on the value of Z_n and neither on past values nor on time n; in other words, the sequence $(Z_n)_{n\geq 0}$ is a time-homogeneous Markov chain on the set of natural numbers $\mathbb{N}$. Having defined the process, we want to know the probability that the random sequence $Z_0, Z_1, Z_2, ..$ eventually goes to zero. In order to answer this question, we take into account the recursive structure of the process which makes it amenable to generating function method. Under some assumptions of the offspring distribution μ, we find some important information about the process, for instance the extinction probability, when $Z_n = 0$ for some $n \geq 1$, or the asymptotic behavior of the process when n tends to infinity. Moreover, according to the mean number m of offspring, the process is dinstinguished into three classes : super-critical ($m > 1$), critical ($m = 1$) and sub-critical

($m < 1$). It is known that for subcritical and critical case, omitting the trivial cases, the population has probability 1 to get extinction while the supercritical process has some positive probability to explode.

In the multi-type case, when the population contains many different type of particles, the process becomes a vector Markov process and the mean number of offspring changes into matrix form M. In this case, the criteria to distinguish three classes depends on the value of the Perron root $\rho(M)$ of the mean matrix M : sub-critical ($\rho(M) < 1$), critical ($\rho(M) = 1$) and super-critical ($\rho(M) > 1$).

Furthermore, it is natural that the offspring distribution μ does not stay the same over generation but varies with time, in a deterministic way which is fixed once and for all; several contributions do exist in this direction, in single-type or multi-type cases, let us mention for instance [Church, 1971], [Kersting, 2017] and references therein. It is also natural to consider that these distributions have some random fluctuations, then to replace the distribution μ by a sequence of random distributions $(\mu_n)_{n\geq 0}$. We arrive in a new scheme of a Branching Process in Random Environment (BPRE). The simplest case in the one when the $(\mu_n)_{n\geq 0}$ are assumed to be i.i.d. random probability measures. Others situations may be explored, for instance Markovian random environment; we refer to [Athreya et Karlin, 1971a] and [Athreya et Karlin, 1971b] for general results in these directions.

The case of single-type in random environment is the object of several works over the last 40 years. There are three cases of extinction, the speed of convergence to extinction is smaller with the appearance of some new phenomenons (for instance the sub-critical case is decomposed into three sub-cases).

The goal of this thesis is to find the asymptotic behavior of the survival probability of a critical multi-type Galton-Watson branching process in i.i.d. environment. Before going further to multi-type BPRE, we first pass through single-type BPRE; we present briefly in Chapter 3 the more important contributions about extinction in this case, under quite general and natural hypotheses. In fact, the behavior of the single-type BPRE is mainly determined by the properties of its associated random walk, whose increments are the logarithm of the mean of (random) offspring distributions. Therefore, further studies on random walk interact the evolution on studies on BPRE. Several limit theorems studied over the three classes of BPRE do exist for single-type BPRE [Kozlov, 1977], [Afanasyev *et al.*, 2005], [Geiger *et al.*, 2003]... It is natural to expect a similar result in multi-type case. However, it is not so easy since many useful tools known for ordinary random walks on the real line with i.i.d. increments have no analogues in the case of dependent increments. More precisely (but not exactly), the associated random walk in multi-type case is composed by the logarithms of the norms of products of matrices. Therefore, studies on limit theorems of products of matrices and fluctuations of the logarithm of their norm are required.

In this thesis, we investigate the asymptotic behavior of the probability of non-extinction up to time n of critical multi-type BPREs and obtain an optimal result in the case of linear fractional generating functions. To formulate our main results, we first introduce some standard notations and definitions.

We fix an integer $d \geq 2$ and denote $\mathbb{R}^d$ (resp. $\mathbb{N}^d$) the set of d-dimensional column

and row vectors (with the abuse of notation) with real (resp., non-negative integers) coordinates; for any column vector $x = (x_i)_{1\le i\le d} \in \mathbb{R}^d$, we denote $\tilde{x}$ the row vector $\tilde{x} := (x_1, \dots, x_d)$. Let $\mathbf{1}$ (resp. $\mathbf{0}$) be the column vector of $\mathbb{R}^d$ whose all coordinates equal 1 (resp., 0). We fix a basis $\{e_i, 1 \le i \le d\}$ in $\mathbb{R}^d$ and denote $|.|$ the corresponding L^1 norm defined by $|x| := \sum_{i=1}^{d} |x_i|$ for any column vector $x = (x_i)_{1\le i\le d}$. Let C be the cone of vectors in $\mathbb{R}^d$ with non-negative coordinates

$$C := \{x \in \mathbb{R}^d : x_i \ge 0 \text{ for all } 1 \le i \le d\}$$

and $\mathbb{X}$ be the standard simplex defined by

$$\mathbb{X} := \{x \in C : |x| = 1\}.$$

Let S be the set of $d \times d$ matrices with non-negative entries such that each column contains at least one positive entry; its interior is $S^+ := \{g = (g(i,j))_{1\le i,j\le d} / g(i,j) > 0\}$. Endowed with the standard multiplication of matrices, the set S is a semigroup and S^+ is the ideal of S, more precisely, for any $g \in S^+$ and $h \in S$, it is evident that $gh \in S^+$.

We consider the following actions :

— the left linear action of S on C defined by $(g, x) \mapsto gx$ for any $g \in S$ and $x \in C$,
— the left projective action of S on $\mathbb{X}$ defined by $(g, x) \mapsto g \cdot x := \frac{gx}{|gx|}$ for any $g \in S$ and $x \in \mathbb{X}$.

For any $g = (g(i,j))_{1\le i,j\le d} \in S$, let

$$v(g) := \min_{1\le j\le d} \Big(\sum_{i=1}^{d} g(i,j) \Big) \quad \text{and} \quad |g| := \max_{1\le j\le d} \Big(\sum_{i=1}^{d} g(i,j) \Big),$$

then $|\cdot|$ is a norm on S and for any $x \in C$,

$$0 < v(g)\,|x| \le |gx| \le |g|\,|x|.$$

We set $N(g) := \max\left(\frac{1}{v(g)}, |g|\right)$; notice that $N(g) \ge 1$ for any $g \in S$.

We consider a sequence of i.i.d. S-valued random matrices $(g_n)_{n\ge 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with the same distribution μ on S; let $L_0 = Id$ and $L_n := g_n \dots g_1$ for any $n \ge 1$. The associated "random walk" on $\mathbb{R}$ is defined by $S_0 := a$ and $S_n = S_n(x, a) := a + \log |L_n x|$ for any $n \ge 1$, where $a \in \mathbb{R}$ is fixed. We are interested in $\tau := \min\{n \ge 1 : S_n \le 0\}$, the first time the random process $(S_n)_{n\ge 1}$ becomes non-positive.

Furthermore, from Theorem II.1 in [Hennion et Hervé, 2008], under conditions P1-P3 introduced below, there exists a unique probability measure ν on $\mathbb{X}$ such that for any bounded Borel function φ from $\mathbb{X}$ to $\mathbb{R}$,

$$(\mu * \nu)(\varphi) = \int_S \int_{\mathbb{X}} \varphi(g \cdot x) \nu(dx) \mu(dg) = \int_{\mathbb{X}} \varphi(x) \nu(dx) = \nu(\varphi).$$

Such a measure is said to be μ-invariant. Moreover, the upper Lyapunov exponent (defined in section 2.4) associated with μ is finite and is expressed by

$$\gamma_\mu = \int_S \int_{\mathbb{X}} \rho(g, x) \nu(dx) \mu(dg),$$

where $\rho(g,x) = \log|gx|$ for any $(g,x) \in S \times \mathbb{X}$.

Let $\mathbb{P}_{x,a}$ the probability measure on $(\Omega, \mathcal{F})$ conditioned on the event $[X_0 = x, S_0 = a]$. We specify the main hypotheses concerning the distribution μ of the matrices under which we obtain the main results of this thesis.

HYPOTHESES

P1. *There exists $\epsilon_0 > 0$ such that $\int_S N(g)^{\epsilon_0}\mu(dg) < \infty$.*

P2. *(Strong irreducibility). There exist no affine subspaces $\mathcal{A}$ of $\mathbb{R}^d$ such that $\mathcal{A} \cap C$ is non-empty, bounded and invariant under the action of all elements of the support of μ.*

P3. *There exists $n_0 \geq 1$ such that $\mu^{*n_0}(S^+) > 0$.*

P4. *The upper Lyapunov exponent γ_μ of μ is equal to 0.*

P5. *There exists $\delta > 0$ such that $\mu\{g \in S : \forall x \in \mathbb{X}, \log|gx| \geq \delta\} > 0$.*

Let V be the harmonic function defined in Proposition 1.1, section 1.5. Then, we obtain the following result.

Theorem 1 *Assume P1-P5. Then for any $x \in \mathbb{X}$ and $a \geq 0$,*

$$\mathbb{P}_{x,a}(\tau > n) \sim \frac{2V(x,a)}{\sigma\sqrt{2\pi n}} \text{ as } n \to +\infty,$$

where $\sigma^2 := \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}_x[S_n^2]$ is the variance of the semi-markovian random walk $(S_n)_{n\geq 0}$. Moreover, there exists a constant c such that for any $x \in \mathbb{X}$, $a \geq 0$ and $n \geq 1$,

$$\sqrt{n}\mathbb{P}_{x,a}(\tau > n) \leq cV(x,a).$$

As a direct consequence, we prove that the sequence $(\frac{S_n}{\sigma\sqrt{n}})_{n\geq 1}$, conditioned on the event $\tau > n$, converges in distribution to the Rayleigh law as stated below. It is not of direct interest for our study of critical multi-type BPRE, but it is a natural question once the estimate of the tail of the distribution of τ is obtained.

Theorem 2 *Assume P1-P5. For any $x \in \mathbb{X}$, $a \geq 0$ and $t > 0$,*

$$\lim_{n\to+\infty} \mathbb{P}_{x,a}\left(\frac{S_n}{\sqrt{n}} \leq t \mid \tau > n\right) = 1 - \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

Using these results, we are able to make some significant progess regarding the main topic of this thesis.

Respecting the common notation used in multi-type case, we denote by p the number of types and then the multi-type Galton-Watson process is a temporally homogeneous vector Markov process $(Z_n)_{n\geq 0}$ whose states are row vectors in $\mathbb{N}^p$. We always assume that Z_0 is non-random. For any $1 \leq i \leq p$, the i-th component $Z_n(i)$ of Z_n may be interpreted as the number of objects of type i in the n-th generation.

We consider a measurable function $\xi \mapsto f_\xi$ from $\mathbb{R}$ to the set of multivariate probability generating functions $f_\xi = (f_\xi^{(i)})_{1\leq i\leq p}$ defined by :

$$f_\xi^{(i)}(s) = \sum_{\alpha\in\mathbb{N}^p} p_\xi^{(i)}(\alpha)s^\alpha,$$

for any $s = (s_i)_{1\le i\le p} \in [0,1]^p$, where

(i) $\alpha = (\alpha_i)_i \in \mathbb{N}^p$ and $s^\alpha = s_1^{\alpha_1} \dots s_p^{\alpha_p}$;

(ii) $p_\xi^{(i)}(\alpha) = p_\xi^{(i)}(\alpha_1, \dots, \alpha_p)$ is the probability that an object of type i in environment ξ has α_1 children of type $1, \dots, \alpha_p$ children of type p.

Let $(\xi_n)_{n\ge 0}$ be a sequence of real valued i.i.d. random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The Galton-Watson process with p types of particles in a random environment $(\xi_n)_{n\ge 0}$ describes the evolution of a particle population $Z_n = (Z_n(i))_{1\le i\le p}$ for $n \ge 0$.

We assume that for any $n \ge 0, i = 1, \dots, p$ and $\xi \in \mathbb{R}$, if $\xi_n = \xi$ then each of the $Z_n(i)$ particles of type i existing at time n produces offspring in accordance with the p-dimensional generating function $f_\xi^{(i)}(s)$ independently of the reproduction of all other particles.

If $Z_0 = \tilde{e}_i$ then Z_1 has the generating function :

$$f_{\xi_0}^{(i)}(s) = \sum_{\alpha\in\mathbb{N}^p}^{+\infty} p_{\xi_0}^{(i)}(\alpha) s^\alpha.$$

To simplify notation, we put $f_n = f_{\xi_n}$. In general, if $Z_n = (\alpha_i)_{1\le i\le p}$, then Z_{n+1} is the sum of $\alpha_1 + \dots + \alpha_p$ independent random vectors, where α_i particles of type i have the generating function $f_n^{(i)}$ for $i = 1, \dots, p$. It is obvious that if $Z_n = \tilde{\mathbf{0}}$, then $Z_{n+1} = \tilde{\mathbf{0}}$.

As in the classical one-type case, the asymptotic behavior of the quantity above is controlled by the mean of the offspring distributions. We assume that the offspring distributions have finite first and second moments; the generating functions $f_\xi^{(i)}, \xi \in \mathbb{R}, 1 \le i \le p$, are thus C^2-functions on $[0,1]^p$ and we introduce

(i) the random mean matrices $M_n = (M_n(i,j))_{1\le i,j\le p} = \left(\dfrac{\partial f_n^{(i)}(\mathbf{1})}{\partial s_j}\right)_{i,j}$ taken from the vector-valued random generating function $f_n(s)$ at $s = \mathbf{1}$, namely

$$M_n = \begin{pmatrix} \dfrac{\partial f_n^{(1)}(\mathbf{1})}{\partial s_1} & \cdots & \dfrac{\partial f_n^{(1)}(\mathbf{1})}{\partial s_p} \\ \vdots & & \\ \dfrac{\partial f_n^{(p)}(\mathbf{1})}{\partial s_1} & \cdots & \dfrac{\partial f_n^{(p)}(\mathbf{1})}{\partial s_p} \end{pmatrix}.$$

These matrices M_n belong to the semi-group S of $p \times p$ matrices with non-negative entries. For any $1 \le i, j \le p$, the coefficient $M_n(i,j)$ of the mean matrix M_n is the mean number of offspring of type j produced by individual of type i at generation n.

(ii) the random Hessian matrices $B_n^{(i)} = (B_n^{(i)}(k,l))_{1\le k,l\le p} = \left(\frac{\partial^2 f_n^{(i)}}{\partial s_k \partial s_l}(\mathbf{1})\right)_{k,l}$, $1 \le i \le p$, taken from the real-valued random generating function $f_n^{(i)}(s)$ at $s = \mathbf{1}$.

For any $1 \le i \le p$, the random variables M_n and $B_n^{(i)}$ are i.i.d. in n. The common law of the M_n is denoted by μ.

Recall that $L_n = M_{n-1} \dots M_0$ in this context. By [Furstenberg et Kesten, 1960], if $\mathbb{E}(\max(0, \log|M_0|)) < +\infty$, then the sequence $\left(\frac{1}{n}\log|L_n|\right)_{n\geq 1}$ converges $\mathbb{P}$-almost surely to some constant limit $\gamma_\mu := \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}[\log|L_n|]$. Furthermore, if there exists a constant $B > 0$ such that $\frac{1}{B} \leq M_n(i,j) \leq B$ and $0 \leq B_n^{(i)}(k,l) \leq B$ $\mathbb{P}$-almost surely for any $1 \leq i,j,k,l \leq p$, then the process is positively regular and not singular (see Section 2.2 for the definitions of these notions) ; hence, by [Kaplan, 1974], the process $(Z_n)_{n\geq 0}$ becomes extinct $\mathbb{P}$-almost surely if and only if $\gamma_\mu \leq 0$. These conditions seem to be a bit too strong, see (0.0.1) below.

In the present work, we focus our attention on the so-called **critical case**, that is $\gamma_\mu = 0$, and specify the speed of extinction of the Galton-Watson process. Then we introduce some proper subset of the interior S^+ of the semi-group S which is of interest in the sequel : for any constant $B \geq 1$, let $S^+(B)$ denote the set of $p\times p$ matrices $g = (g(i,j))_{1\leq i,j\leq p}$ with positive coefficients such that for any $1 \leq i,j,k,l \leq p$,

$$\frac{1}{B} \leq \frac{g(i,j)}{g(k,l)} \leq B. \tag{0.0.1}$$

Additionally, we consider "linear-fractional multi-dimensional generating functions" $f = f_{\gamma,M}$ of the form :

$$f(s) = f_{\gamma,M}(s) = \mathbf{1} - \frac{1}{1+\tilde{\gamma}(\mathbf{1}-s)}M(\mathbf{1}-s), \tag{0.0.2}$$

for any $s \in [0,1]^p$, where $\tilde{\gamma} = (\gamma,\dots,\gamma) \in \mathbb{R}^p$ with $\gamma > 0$ and $M \in S^+$. For such a function $f = f_{\gamma,M}$, we set $\gamma = \gamma(f)$ and $M = M(f)$ and notice that $M(f)$ equals the mean matrix $\left(\frac{\partial f^{(i)}}{\partial s_j}(\mathbf{1})\right)_{1\leq i,j\leq p}$.

With the functions f_n being linear-fractional generating functions, we specify hypotheses concerning the distribution μ of the mean matrices $M_n = M(f_n), n \geq 1$ and the random variables $\gamma(f_n), n \geq 0$.

H1. *There exists $\epsilon_0 > 0$ such that $\int_S N(g)^{\epsilon_0}\mu(dg) < \infty$.*

H2. *(Strong irreducibility). There exists no affine subspaces $\mathcal{A}$ of $\mathbb{R}^d$ such that $\mathcal{A} \cap C$ is non-empty, bounded and invariant under the action of all elements of the support of μ.*

H3. *There exists $B \geq 1$ such that $\mu(S^+(B)) = 1$.*

H4. *The upper Lyapunov exponent γ_μ of μ is equal to 0.*

H5. *There exists $\delta > 0$ such that $\mu(E_\delta) > 0$, where*

$$E_\delta := \{g \in S^+ \mid \forall x \in \mathbb{X},\ \log|\tilde{x}g| \geq \delta\}.$$

H6. *There exists $B' \geq 1$ such that $\frac{1}{B'} \leq \gamma(f_n) \leq B' \quad \mathbb{P}-a.s.$ for any $n \geq 1$.*

We recall the result when $p = 1$.

Theorem 3 *[Geiger et Kersting, 2000](Critical case)*

Suppose that $\mathbb{E}\log f_0'(1) = 0$,

$$0 < \mathbb{E}(\log f'(1))^2 < +\infty \quad \textit{and} \quad \mathbb{E}\left(\frac{f''(1)}{f'(1)^2}(1 + \log_+ f'(1))\right) < +\infty.$$

Then, for some $0 < \beta < +\infty$,

$$\mathbb{P}(Z_n > 0) \sim \frac{\beta}{\sqrt{n}} \quad \textit{as} \quad n \to +\infty.$$

Here comes the main results of this thesis.

Theorem 4 *Assume that the random variables* f_n *are linear fractional generating functions and that hypotheses* H1-H6 *hold. Then, for any* $i \in \{1, \dots, p\}$, *there exists a number* $\beta_i \in (0, +\infty)$ *such that*

$$\lim_{n\to+\infty} \sqrt{n}\mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) = \beta_i. \tag{0.0.3}$$

For general generating functions, we obtain the following weaker result.

Theorem 5 *Assume that the random variables* f_n *are* C^2*-functions on* $[0,1]^p$ *such that*

1. *there exists* $A > 0$ *such that for any* $i, k, l \in \{1, \dots, p\}$,

$$\frac{\partial^2 f_n^{(i)}}{\partial s_k \partial s_l}(\mathbf{1}) \leq A \frac{\partial f_n^{(i)}}{\partial s_k}(\mathbf{1}) \quad \mathbb{P} - \text{a.s.}$$

2. *the distribution* μ *of the matrices* $M_n = \left(\frac{\partial f_n^{(i)}}{\partial s_j}(\mathbf{1})\right)_{1\leq i,j\leq p}$ *satisfies hypotheses* H1-H5.

Then, there exist constants $0 < c_1 < c_2 < +\infty$ *such that for any* $i \in \{1, \dots, p\}$, *and* $n \geq 1$,

$$\frac{c_1}{\sqrt{n}} \leq \mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) \leq \frac{c_2}{\sqrt{n}}.$$

Notice that by recent work [Vatutin et Dyakonova, 2017], statement (0.0.3) holds in fact true even when the f_n are not assumed to be linear fractional generating functions. The authors in [Vatutin et Dyakonova, 2017] apply the proof presented in this thesis, adding an extension to the multitype case of Geiger & Kersting's decomposition of the extinction probability and taking into account the residual term which appears in this expression.

This thesis is organised as follow : Chapter 1 provides the conditioned limit theorems on products of matrices, which is an important tool to find the asymptotic behavior of extinction probability in BPRE. We recall general results about fluctuations of 1-dimensional random walk with i.i.d. increments (section 1.1 and 1.2). Then we present general results for products of random matrices and and fix the notations (section 1.3 and 1.4). The main part of Chapter 1 is section 1.5. Chapter 2 concerns in more detail Galton-Watson branching processes in fixed and in random environment. Section 2.4 is the main part of this chapter with the article that concerns the topic of this thesis. Chapter 3 gives some simulations of the considered Galton-Watson process and the thesis ends with conclusions with questions for future research.

Chapitre 1

Random walks on $\mathbb{R}$ and products of matrices

Many limit theorems which arose for the last 60 years, initiated by [Furstenberg et Kesten, 1960], [Guivarc'h et Raugi, 1985], [Le Page, 1982]... and recently [Benoist et Quint, 2016], describe the asymptotic behavior of random walks with i.i.d. increments, for instance the strong law of large numbers, the central limit theorem, the invariant principle... Moreover, the fluctuations of these processes are well studied, for example functional central limit theorems for random walks conditioned on staying positive, or the decay of the probability that the processes stay inside the half real line up to time n or . A vast literature exists on this subject, see for instance [Bolthausen, 1976], [Iglehart, 1974b], [Iglehart, 1974a], [Iglehart, 1975] or [Kaigh, 1976] and references therein. The Wiener-Hopf factorization is usually used in this case and so far, it seems to be impossible to adapt in non-abelian context.

Recently, much efforts are made to apply the analogue results above for the logarithm of the norm of the product of i.i.d. random matrices since it behaves similarly to a sum of i.i.d. random variables. Let us mention also the works by [Hennion, 1984] and [Hennion et Hervé, 2008] for matrices with positive entries. However, the studies on the subject of fluctuation was quite sparse a few years ago. Thanks to the approach of [Denisov et Wachtel, 2015] for random walks in Euclidean spaces and motivated by branching processes, I. Grama, E. Le Page and M. Peigné recently progressed for invertible matrices ([Grama *et al.*, 2014]). Here we propose to develop the same strategy for matrices with positive entries by using [Hennion et Hervé, 2008].

In this chapter we tend to find the asymptotic behavior of the random process $S_n = \log|M_n \dots M_1 x|$, where M_n are independent and identically distributed (i.i.d.) random positive matrices and x is a random vector in $\mathbb{R}^{+d}$. Notice that the increments $S_n - S_{n-1}$ of this random process are not i.i.d. which is the main difference from classical random walks on $\mathbb{R}$.

First, we discuss random walks with i.i.d. increments and second, investigate the Wiener-Hopf method used to obtain those results and the difficulties arising in the case of non-i.i.d. increments, in particular for processes $(S_n)_{n\geq 0}$ over products of random matrices as introduced above. Third, we consider the results obtained for products of random

positive matrices. Last but not least, some brief knowledge about Lyapunov's exponent is recalled for later use in an article put at the end of the present chapter.

1.1 Random walks with i.i.d. increments

The theory and applications of random walks are omnipresent in the modern probability literature, and random walks perhaps present the simplest and the most important examples of stochastic processes. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. If we think of $X_1, X_2, \ldots$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ as the gains or losses of a gambler on successive plays of a gambling game, his cumulative gains or losses on the first n plays equal the partial sums

$$S_n := S_0 + X_1 + \ldots + X_n,$$

where S_0 is the initial fortune of the gambler. The behavior of the "random walk" $(S_n)_n$ describes the evolution of the gambler's fortune. The simplest non-trivial case is to let $X_1, X_2, \ldots$ be independent with distribution μ and the basic operation that forms the nth convolution powers are $\mu,\ \mu^{*2} := \mu * \mu, \ldots,\ \mu^{*n} := \mu * \mu^{*(n-1)}$.

When $S_0 = 0$, one says that the random walk $(S_n)_{n\geq 0}$ starts from 0. One of the main tool to study random walks is the characteristic function of the increments X_n, defined by

$$\hat{\mu}(t) := \mathbb{E}[\exp(itX_n)] = \int e^{itx}\mu(dx);$$

it contains all the information on the distribution μ. Since the increments are i.i.d., we can interprete the characteristic function of S_n as follows

$$\mathbb{E}[\exp(itS_n)] = \int e^{itx}\mu^{*n}(dx) = \hat{\mu}^n(t).$$

Thanks to this expression, the asymptotic properties of $(S_n)_{n\geq 0}$ are closely related to the local behavior near 0 of the function $\hat{\mu}$. The basic dichotomy of transience or recurrence for random walks can be expressed in terms of characteristic functions, which extends to locally compact abelian groups, like how the Fourrier analysis does to the general setting of random walks on groups.

The transience-recurrence dichotomy of random walks depends on the existence of the Green potential,

$$G(x, dy) := \sum_{n=0}^{+\infty} P^n(x, dy).$$

If the Green potential is a Radon measure, that if $G(x, K) = \sum_{n=0}^{+\infty} P^n(x, K) < +\infty$ for every $x \in \mathbb{R}$ and any compact set $K \subset \mathbb{R}$, then the according random walk is transient, otherwise the random walk is recurrent.

furthermore, since random walks are particularly simple and important Markov processes, their potential theory is developed accordingly, where the concept of harmonic function plays a crucial role to illustrate generalized results of classical potential theory.

Formally, a function f is harmonic if it satisfies the equation $Pf = f$, where P is the transition operator of the random walk $(S_n)_{n\geq 0}$; in other words, for every $x \in \mathbb{R}$,

$$f(x) = \int_{\mathbb{R}} f(y)P(x, dy) = \int_{\mathbb{R}} f(x + z)\mu(dz).$$

Nevertheless, while the Fourrier and potential theory are classical tools, the coupling method is more recent : we consider two independent processes on the same probability space with given distributions, and seek to compare those distributions by comparing the processes themselves directly. Proving the convergence theorem for Markov chains and renewal theorems are two of the most successful applications of the coupling method [Lindvall, 1992].

Much of the core of classical probability theory is concerned with the limit theory of a sequence of partial sums S_n, for example the following classical limit theorems : the strong law of large numbers, the central limit theorem, the law of iterated logarithm, the large deviation theorems... Informally, we concern with S_n as $n \to +\infty$. As the number of steps increases, the influence of each step decreases. Thus, we may expect a setting of a stochastic process in continuous time and state-space, which is indeed true. In the simplest case when the X_n have finite variance then the limiting process obtained is a Brownian motion. The necessary tool to handle the passage to the limit is the theory of weak convergence of probability measures, see Billingsley [Billingsley, 1968] for an excellent reference.

It frequently happens that we need to deal with a random walk in which we want to condition on some event by time n, that this event may have small probability and vanishing in the limit. The classical method used to study the rate of convergence of a random walk on Markov chain is spectral theory. If the transition matrix of the chain is P, then the nth step transition matris is P^n. Using Perron-Frobenius theorem, we can decompose the transition matrix as $P = \Pi + R$, where Π is the diagonal matrix of the Perron root 1 and the convergence behavior of P^n is determinded by the second greatest eigenvalue of P. Random walks in random environment is a variant of random walk to model aspects of natural phenomena.

1.2 Wiener-Hopf factorization method

The idea of taking a function which is defined on a strip in the complex plane and expressing it as a product of two functions in which each functions is defined in a half-plane intersecting at this strip and analytic in their interior allows us to use the powerful tools of complex analysis. It may be traced to the work of Wiener-Hopf [Paley et Wiener, 1987] and is known as the Wiener-Hopf factorization [Feller, 1968].

Let $(X_n)_{n\geq 1}$ be a a sequence of real valued i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. We consider the random walk on $\mathbb{R}$ defined by $S_0 = 0$ and $S_n := X_1 + \ldots + X_n$ for every $n \geq 1$, where the X_k are i.i.d. random variables with distribution μ. Let ζ denote the characteristic function of the X_k.

In order to study the first passage of a random walk to some given set A, it is often useful to consider the joint distribution of the couple (N, S_N), where N is the hitting time

of the random walk to A. Thus, to get some information about the distribution of this couple, it is natural to keep in mind the "time" for all the process $(S_n)_{n\geq 0}$ and to consider the process $(n, S_n)_{n\geq 0}$ on $\mathbb{N}\times\mathbb{R}$; the characteristic function of this couple is a function $\Phi(s,t)$ of two variables s and t, defined by $\Phi : (s,t) \longmapsto \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}]$, for every $s \in]-1,1[$ and $t \in \mathbb{R}$. It holds, for every $s \in]-1,1[$ and $t \in \mathbb{R}$,

$$\Phi(s,t) = \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}] = \sum_{n=0}^{+\infty} s^n \hat{\mu}(t)^n = \frac{1}{1-s\zeta(t)}. \tag{1.2.1}$$

In this subsection, we study the first passage of $(S_n)_{n\geq 0}$ to the subset $]0,+\infty[$ and its complement $]-\infty,0]$. Namely, with the convention $\inf\emptyset = +\infty$, we define

$$l^+ := \inf\{n > 0 : S_n > 0\} \qquad \text{and} \qquad l^- := \inf\{n > 0 : S_n \leq 0\}.$$

Set $l_0^+ = l_0^- = 0$ and for every $k \geq 1$, let l_k^+, l_k^- be respectively the (strict) ascending and descending k^{th}-ladder epoch of the walk $(S_n)_{n\geq 0}$, defined inductively by

$$l_k^+ := \inf\{n > l_{k-1}^+ : S_n > S_{l_{k-1}^+}\} \qquad \text{and} \qquad l_k^- := \inf\{n > l_{k-1}^- : S_n \leq S_{l_{k-1}^-}\}.$$

For $k \geq 0$, the random variables l_k^+ and l_k^- are stopping times with respect to the random walk $(S_n)_{n\geq 0}$.

By [Spitzer, 1976], we know that the random walk $(S_n)_{n\geq 0}$ may have three main behavior :.

1. Either $S_n \to +\infty$ $\mathbb{P}$-a.s.; in this case $\mathbb{E}(l^+) < +\infty$ (in particular $\mathbb{P}(l^+ < +\infty) = 1$) and $\mathbb{P}(l^- < +\infty) < 1$.
2. Either $S_n \to -\infty$ $\mathbb{P}$-a.s.; in this case $\mathbb{E}(l^-) < +\infty$ and $\mathbb{P}(l^+ < +\infty) < 1$.
3. Or $(S_n)_{n\geq 0}$ oscillates, namely

$$\liminf_{n\to+\infty} S_n = -\infty \quad \mathbb{P}-\text{a.s.} \quad \text{and} \quad \limsup_{n\to+\infty} S_n = +\infty \quad \mathbb{P}-\text{a.s.}$$

   In this case $\mathbb{P}(l^- < +\infty) = \mathbb{P}(l^+ < +\infty) = 1$ and $\mathbb{E}(l^-) = \mathbb{E}(l^+) = +\infty$.

For instance, assume that $\mathbb{E}(|X_k|) < +\infty$; cases (1), (2) and (3) correspond respectively to the conditions $\mathbb{E}(X_k) > 0$, $\mathbb{E}(X_k) < 0$ and $\mathbb{E}(X_k) = 0$.

From now on, we assume that $(S_n)_{n\geq 0}$ oscillates, that is the stopping times l^- and l^+ are finite $\mathbb{P}$-a.s.; this property holds in particular when the X_n have finite expectation and are centered. In this case, for $k \geq 0$, all the random stopping times l_k^+ and l_k^- are finite $\mathbb{P}$-a.s. By a straightforward argument, one may check that the random variables $l_k^+ - l_{k-1}^+$, for $k \geq 1$, are i.i.d. with the same distribution as l^+ (similarly the random variables $l_k^- - l_{k-1}^-$, for $k \geq 1$, are i.i.d. with the same distribution as l^-). Furthermore, the random sums $S_{l_k^+}$ and $S_{l_k^-}$ are well defined and the random variables $S_{l_k^+} - S_{l_{k-1}^+}$, for $k \geq 1$, (resp., the random variables $S_{l_k^-} - S_{l_{k-1}^-}$, for $k \geq 1$) are also i.i.d. with the same distribution as S_{l^+} (resp., S_{l^-}).

For every s in $[0,1[$ and every $t \in \mathbb{R}$, it holds

$$\begin{aligned}\sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}] &= \mathbb{E}\left[\sum_{n=0}^{l^+-1} s^n e^{itS_n}\right] + \mathbb{E}\left[\sum_{n=l^+}^{+\infty} s^n e^{itS_n}\right] \\ &= \mathbb{E}\left[\sum_{n=0}^{l^+-1} s^n e^{itS_n}\right] + \mathbb{E}[s^{l^+} e^{itS_{l^+}}] \sum_{k=0}^{+\infty} \mathbb{E}[s^k e^{itS_k}],\end{aligned}$$

which implies

$$\mathbb{E}[\sum_{n=0}^{+\infty} s^n e^{itS_n}] = \mathbb{E}[\sum_{n=0}^{l^+-1} s^n e^{itS_n}]\{1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}]\}^{-1}. \tag{1.2.2}$$

The random vectors $(X_1, X_2, \ldots, X_n)$ and $(X_n, X_{n-1}, \ldots, X_1)$ have the same distribution : this is the so-called *duality property*. By using the definition of l^+ and duality property, we obtain

$$\begin{aligned}
\mathbb{E}[\sum_{n=0}^{l^+-1} s^n e^{itS_n}] &= \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}; l^+ > n] \qquad (1.2.3)\\
&= \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}; S_1 \leq 0, \ldots, S_n \leq 0]\\
&= \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}; S_n \leq S_{n-1}, \ldots, S_n \leq 0]\\
&= \sum_{n=0}^{+\infty} \mathbb{E}[s^n e^{itS_n}; \exists k > 0 : n = l_k^-]\\
&= \sum_{n=0}^{+\infty}\sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^-} e^{itS_{l_k^-}}; n = l_k^-]\\
&= \sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^-} e^{itS_{l_k^-}}]. \qquad (1.2.4)
\end{aligned}$$

Notice that

$$\{1 - \mathbb{E}[s^{l^-} e^{itS_{l^-}}]\}^{-1} = \sum_{k=0}^{+\infty} (\mathbb{E}[s^{l^-} e^{itS_{l^-}}])^k = \sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^-} e^{itS_{l_k^-}}] \tag{1.2.5}$$

$$\{1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}]\}^{-1} = \sum_{k=0}^{+\infty} (\mathbb{E}[s^{l^+} e^{itS_{l^+}}])^k = \sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^+} e^{itS_{l_k^+}}]. \tag{1.2.6}$$

From (1.2.2), (1.2.4) and (1.2.5), we obtain

$$\mathbb{E}[\sum_{n=0}^{+\infty} s^n e^{itS_n}] = \{1 - \mathbb{E}[s^{l^-} e^{itS_{l^-}}]\}^{-1}\{1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}]\}^{-1}.$$

Therefore, from (1.2.1), it yields

$$1 - s\zeta(t) = (1 - \mathbb{E}[s^{l^-} e^{itS_{l^-}}])(1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}]), \tag{1.2.7}$$

in which the right side contains two characteristic functions of measures which are supported on $]-\infty, 0]$ and $]0, +\infty[$, respectively.

We take into account some basic formal calculations which may be useful later :

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \quad \text{and} \quad \log(1-x) = -\sum_{n=0}^{+\infty} \frac{x^n}{n}$$

which implies $1-x = \exp(\log(1-x)) = \exp\left(-\sum_{n=1}^{+\infty} \frac{x^n}{n}\right)$. Therefore, we may also decompose formally $1 - s\zeta(t)$ as

$$\begin{aligned} 1 - s\zeta(t) &= \exp\Big(-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}]\Big) \\ &= \exp\left[-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n > 0] - \sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n \leq 0]\right] \\ &= \exp\left[-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n > 0]\right] \exp\left[-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n \leq 0]\right], \quad (1.2.8) \end{aligned}$$

in which the right side contains also two characteristic functions of measures supported on $]-\infty, 0]$ and $]0, +\infty[$, respectively. Some may call these functions the left and right Wiener-Hopf factors. Thus, comparing decompositions (1.2.5), (1.2.7) and (1.2.8) yields

$$1 - \mathbb{E}[s^{l^-} e^{itS_{l^-}}] = \left\{\sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^-} e^{itS_{l_k^-}}]\right\}^{-1} = \exp\left[-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n \leq 0]\right]$$

and

$$1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}] = \left\{\sum_{k=0}^{+\infty} \mathbb{E}[s^{l_k^+} e^{itS_{l_k^+}}]\right\}^{-1} = \exp\left[-\sum_{n=1}^{+\infty} \frac{s^n}{n}\mathbb{E}[e^{itS_n}; S_n > 0]\right].$$

This is the so-called "Wiener-Hopf factorization" which furnishes a simple beautiful but also powerful expression of the quantities $\mathbb{E}[s^{l^-} e^{itS_{l^-}}]$ and $\mathbb{E}[s^{l^+} e^{itS_{l^+}}]$.

To be continued, notice that by (1.2.3) and (1.2.5) it holds

$$\sum_{n=0}^{+\infty} s^n \mathbb{E}[e^{itS_n}; l^+ > n] = \{1 - \mathbb{E}[s^{l^-} e^{itS_{l^-}}]\}^{-1},$$

and similarly,

$$\sum_{n=0}^{+\infty} s^n \mathbb{E}[e^{itS_n}; l^- > n] = \{1 - \mathbb{E}[s^{l^+} e^{itS_{l^+}}]\}^{-1}. \quad (1.2.9)$$

We set $t = 0$ in (1.2.9) and obtain

$$\begin{aligned}
\sum_{n=0}^{+\infty} s^n \mathbb{P}(l^- > n) &= \{1 - \mathbb{E}[s^{l^+}]\}^{-1} \\
&= \exp\left(\sum_{n=1}^{+\infty} \frac{s^n}{n} \mathbb{P}(S_n > 0)\right) \\
&= \exp\left(\sum_{n=1}^{+\infty} \frac{s^n}{n} \left(\mathbb{P}(S_n > 0) - \frac{1}{2}\right)\right) \exp\left(\frac{1}{2} \sum_{n=1}^{+\infty} \frac{s^n}{n}\right) \\
&= \exp\left(\sum_{n=1}^{+\infty} \frac{s^n}{n} \left(\mathbb{P}(S_n > 0) - \frac{1}{2}\right)\right) \exp\left(-\frac{1}{2} \log(1-s)\right) \\
&= \frac{1}{\sqrt{1-s}} \exp\left(\sum_{n=1}^{+\infty} \frac{s^n}{n} \left(\mathbb{P}(S_n > 0) - \frac{1}{2}\right)\right). \qquad (1.2.10)
\end{aligned}$$

With the assumptions $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(X_n^2) < +\infty$, the central limit theorem for random walks on $\mathbb{R}$ yields $\lim_{n\to+\infty} \mathbb{P}(S_n > 0) = \frac{1}{2}$; consequently $\frac{1}{n}\left(\mathbb{P}(S_n > 0) - \frac{1}{2}\right) = \frac{o(n)}{n}$ as $n \to +\infty$. Under a stronger assumption $\mathbb{E}(X_n^{2+\delta}) < +\infty$ for some $\delta \in]0,1[$, a straightforward generalization of the Berry-Essen's theorem [Chow et Teicher, 2012] yields $\left|\mathbb{P}(S_n > 0) - \frac{1}{2}\right| \leq \frac{C}{n^{\delta/2}}$ for some constant $C > 0$, which implies

$$\sum_{n=1}^{+\infty} \frac{1}{n} \left|\mathbb{P}(S_n > 0) - \frac{1}{2}\right| < +\infty. \qquad (1.2.11)$$

Property (1.2.11) holds in fact under the hypothesis $\mathbb{E}(X_n^2) < +\infty$, the proof is quite technical (see [Spitzer, 1976]).

Now, let us briefly explain how to get the asymptotic behavior of the probability $\mathbb{P}(l^- > n)$; indeed, by (1.2.10) and (1.2.11), one may write

$$\sum_{n=0}^{+\infty} s^n \mathbb{P}(l^- > n) = \frac{e^\alpha}{\sqrt{1-s}}(1 + o(s))$$

with $\alpha := \sum_{n=1}^{+\infty} \frac{1}{n}\left(\mathbb{P}(S_n > 0) - \frac{1}{2}\right)$ and $o(s) \to 0$ as $s \to 1$. By a Tauberian theorem for power series [Feller, 1968], it follows that $\frac{1}{n}\sum_{k=1}^{n} \mathbb{P}(l^- > k) \sim \frac{e^\alpha}{\sqrt{\pi n}}$ as $n \to +\infty$; using the fact that the sequence $(\mathbb{P}(l^- > n))_{n\geq 1}$ decreases, one concludes

$$\mathbb{P}(l^- > n) \sim \frac{e^\alpha}{\sqrt{\pi n}} \quad \text{as} \quad n \to +\infty.$$

The same approach leads to a more general statement (we refer to [Le Page et Peigné, 1997] for the details). For every $a > 0$, let τ_a be the first entrance time in $]-\infty, -a]$ of the random walk $(S_n)_{n\geq 0}$ starting from 0 :

$$\tau_a := \inf\{n > 0 : S_n \leq -a\}.$$

For $a = 0$ it holds $\tau_a = l^-$. One may prove that there exists a positive constant c_a such that

$$\mathbb{P}(\tau_a > n) \sim \frac{c_a}{\sqrt{n}} \quad \text{as} \quad n \to +\infty.$$

The constant c_a may be expressed in terms of the renewal function of the increasing ladder random walk $(S_{l_n^+})_{n\geq 0}$ and its behavior at infinity is controlled as follows : there exists a positive constant c such that $c_a \sim ca$ as $a \to +\infty$. Furthermore, the function $a \mapsto c_a$ is harmonic for S_n killed at leaving $]0, +\infty[$.

It is worth mentioning that Wiener-Hopf method is powerful as one does not need to impose any moment conditions on the random walk; indeed, other estimates as above exist for random walks belonging to the domain of attraction of any stable distribution. Nevertheless, it is based on the fact that $(\mathbb{R}, +)$ is abelian and that its bounded characters are the functions $x \mapsto e^{itx}$ for $t \in \mathbb{R}$. The fact that the increments are independent and identically distributed is crucial in order to use the characters of the group as test functions; the decompositions (1.2.7) and (1.2.8) and the pairwise identification of the factors are based on this key point.

In the following sections, we obtain a similar result for the process $(\log|M_n \cdots M_1 x|)_{n\geq 1}$ where M_n are i.i.d. $d \times d$ random matrices with non-negative entries and x is an arbitrary vector in $(\mathbb{R}^{*+})^d$. By classical cocycle properties introduced to study products of random matrices, the random variable $\log|M_n \cdots M_1 x|$ may be decomposed into the sum of real valued random variables Y_k; unfortunately, the "increments" M_k of the product of matrices $M_1 \cdots M_n$ are i.i.d. but this property fails for the increments Y_k. Let us emphasize that the process $(\log|M_n \cdots M_1 x|)_{n\geq 1}$ is not a Markov chain on $\mathbb{R}$. Nevertheless, the matrices M_k act projectively on some compact space $\mathbb{X}$ (see the next paragraph); this action yields to some Markov chain $(X_n)_{n\geq 0}$ on $\mathbb{X}$, starting from x. The key point of this construction is to notice that the process $(X_n, \log|M_n \cdots M_1 x|)_{n\geq 0}$ becomes a Markov chain on $\mathbb{X} \times \mathbb{R}$ and that the distribution of the increment $Y_k = \log|M_k \cdots M_1 x| - \log|M_{k-1} \cdots M_1 x|$ depends only on X_{k-1}. The process $(X_n, \log|M_n \cdots M_1 x|)_{n\geq 0}$ is called a *semi-markovian random walk* on $\mathbb{X} \times \mathbb{R}$; for short, one often say that $(\log|M_n \cdots M_1 x|)_{n\geq 0}$ is a *Markov walk* on $\mathbb{R}$.

This explains why we have to use a different approach from the Wiener-Hopf factorization to study the fluctuations of the process $(\log|M_n \cdots M_1 x|)_{n\geq 0}$. Recently in [Denisov et Wachtel, 2015], D. Denisov and V. Wachtel developed a new strategy to get such results for the first exit time of cones in the context of random walks on the Euclidean space $\mathbb{R}^d, d \geq 2$; it is based on the weak invariance principle for random walks, with a control of the rate of convergence. Grama, E. Le Page and M. Peigné adapted their approach to sums of dependent real valued random variables and applied it to a wide class of Markov walks on $\mathbb{R}$, including the processes $(\log|M_n \cdots M_1 x|)_{n\geq 1}$ mentioned above.

1.3 Product of random matrices

Product of random matrices is the object of many investigations and many limit theorems exist in this context : for instance, the law of large numbers, the central limit theorem, the large deviations principle (see [Bougerol et Lacroix, 1985] and references

therein). Unfortunately, since there is not any type of Wiener-Hopf factorization which is available in the theory of fluctuations of products of random matrices, it remains to be studied. In 2015, Denisov-Wachtel presented a new approach to studies of random walks in Euclidean spaces and their fluctuations with respect to some cones. This approach is totally different from the others before and based on weak functional limit theorems for random walks, with a control of the rate of convergence.

In the following subsections, we introduce some notations used in the article presented in this chapter and present some elements of proofs.

1.3.1 On the semi-group of matrices with non-negative entries

We fix an integer $d \geq 2$ and denote $\mathbb{R}^d$ the set of d-dimensional column vectors with real coordinates; for every column vector $x = (x_i)_{1\leq i\leq d} \in \mathbb{R}^d$, we denote $\tilde{x}$ the row vector $\tilde{x} := (x_1, \ldots, x_d)$. Let $\mathbf{1}$ (resp. $\mathbf{0}$) be the column vector of $\mathbb{R}^d$ whose all coordinates equal 1 (resp. 0). We fix a basis $\{e_i, 1 \leq i \leq d\}$ in $\mathbb{R}^d$ and denote $|\cdot|$ the corresponding L^1-norm and $\langle\ ,\ \rangle$ the usual scalar product on $\mathbb{R}^d$.

Let S be the general linear semi -group of $d\times d$ matrices with non-negative coefficients such that each column contains at least one positive element; we denote S^+ the interior of S, that is the set of matrices with all strictly positive coefficients. Let us endow S with the L^1-norm : for every matrix M in S, where $M = (M(i,j))_{i,j}$, then the norm of matrix M is defined as follows

$$|M| := \max_{1\leq j\leq d} \sum_{i=1}^{d} M(i,j).$$

For any $M, N \in S$, it holds $|MN| \leq |M|\,|N|$. Set $C := (\mathbb{R}^+)^d$ and $\mathbb{X} := \{x \in C : |x| = 1\}$.

Matrices in S act linearly on the cone C, this property is crucial in the sequel; they also act on $\mathbb{X}$ as follows :

$$\forall M \in S, \forall x \in \mathbb{X} \qquad M \cdot x := \frac{1}{|Mx|} Mx.$$

We mention [Hennion, 1997] for further information about a distance denoted by d that we endow $\mathbb{X}$ with. This distance d is a variant of the Hilbert metric, bounded on $\mathbb{X}$ and every element g in S acts on $(\mathbb{X}, d)$ as a contraction. We summarise its construction and its major properties as follows. For every $x,\ y$ in $\mathbb{X}$, we write :

$$\begin{aligned} \mathsf{m}(x,y) &= \sup\{\lambda \in \mathbb{R}^+ | \forall i \in \{1,\ldots,d\}, \lambda y_i \leq x_i\} \\ &= \min\{\frac{x_i}{y_i} | \forall i = 1,\ldots,d \ \text{ such that } \ y_i > 0\} \end{aligned}$$

It follows that $0 \leq \mathsf{m}\left(x,y\right) \leq 1$ and we define

$$d\left(x,y\right) := \varphi\left[\mathsf{m}\left(x,y\right)\mathsf{m}\left(y,x\right)\right],$$

where φ is the one-to-one function on $[0,1]$ defined by $\varphi\left(s\right) := \dfrac{1-s}{1+s}$.

For $M \in S$, let $c\left(M\right) := \sup\left\{d\left(M\cdot x, M\cdot y\right), x, y \in \mathbb{X}\right\}$.

Proposition 1.3.1 *[Hennion, 1997] The function d is a distance on $\mathbb{X}$ which satisfies the following properties.*

1. $\sup\{d(x,y) \mid x,\ y \in \mathbb{X}\} = 1$.
2. *For every $M = (M(i,j))_{i,j} \in S$,*

$$c(M) = \max_{i,j,k,l \in \{1,\dots,d\}} \frac{|M(i,j)M(k,l) - M(i,l)M(k,j)|}{M(i,j)M(k,l) + M(i,l)M(k,j)}.$$

   In particular $c(M) < 1$ for every $M \in S^+(B)$.
3. $d(M \cdot x, M \cdot y) \le c(M)d(x,y) \le c(M)$ *for every* $x,\ y \in \mathbb{X}$ *and* $M \in S$.
4. $c(MN) \le c(M)c(N)$ *for every* $M, N \in S$.

1.3.2 On some sub-semigroup of S^+ and its contraction properties

We consider some proper subset of S^+ which is of interest in the sequel : for every constant $B \ge 1$, let $S^+(B)$ denote the set of $d \times d$ matrices $M = (M(i,j))_{1 \le i,j \le d}$ with positive coefficients such that for any $1 \le i, j, k, l \le d$,

$$\frac{1}{B} \le \frac{M(i,j)}{M(k,l)} \le B.$$

By Proposition 1.3.1, there exists $\kappa_B \in [0,1[$ such that $c(M) \le \kappa_B$ for every $M \in S^+(B)$. Indeed, we consider the function $D : (x,y,z,t) \mapsto \frac{|xy-zt|}{|xy+zt|}$ on the compact set $[\frac{1}{B}, B]^4$. Since D is continuous on the given compact set, D attains its maximum here, that is the constant $\kappa_B := \max_{(x,y,z,t) \in [\frac{1}{B},B]^4} \frac{|xy-zt|}{|xy+zt|}$ belongs to $[0,1[$ and depends on B.

Denote by $T_{S^+(B)}$ the semigroup generated by the elements in $S^+(B)$. The following lemma is the key argument to control the asymptotic behavior of the norm of products of matrices in $S^+(B)$.

Notation. *Let $c > 0$ and ϕ, ψ be two functions of some variable x; we shall write $\phi \overset{c}{\preceq} \psi$ (or simply $\phi \preceq \psi$) when $\phi(x) \le c\psi(x)$ for any value of x. The notation $\phi \overset{c}{\asymp} \psi$ (or simply $\phi \asymp \psi$) means $\phi \overset{c}{\preceq} \psi \overset{c}{\preceq} \phi$.*

Since $|M| \le |M|' \le d|M|$ for any matrix M in S, where $|M|' := \sum_{i,j=1}^{d} M(i,j)$, then without loss of generality, we locally use the norm $|\cdot|'$ in Lemma 1.3.2 and the same notation $|\cdot|$ for the new norm.

Lemma 1.3.2 *There exists $c \ge 1$ which depends on B such that for every $M, N \in T_{S^+(B)}$ and $x,\ y \in \mathbb{X}$,*

1. $|Mx| \overset{c}{\asymp} |M|$ *and* $|\tilde{y}M| \overset{c}{\asymp} |M|$,
2. $|\tilde{y}Mx| \overset{c}{\asymp} |M|$,
3. $|MN| \overset{c}{\asymp} |M||N|$.

We begin with the proof of Furstengberg-Kesten [Furstenberg et Kesten, 1960] and then present an alternate proof which is based on spectral theory of matrices and contains ideas from geometric arguments; this new proof may be used in work in progress on stochastic dynamical systems generated by affine maps in $\mathbb{R}^d$.

The proof of Furstengberg-Kesten is based on the following simple result.

Lemma 1.3.3 *For every $n \geq 1$, any matrices $M_1, \ldots, M_n$ in $S^+(B)$, letting $L_n := M_n \ldots M_1$, it holds :*

$$\forall i, j, k, l = 1, \ldots, d \qquad \frac{1}{B^2} \leq \frac{L_n(i,j)}{L_n(k,l)} \leq B^2.$$

Proof. For $n = 1$, it is obvious. For $n \geq 2$, setting $L_{n-1,2} = M_{n-1} \ldots M_2$ (with the convention $L_{1,2} = I$), it holds

$$\begin{aligned}
\frac{L_n(i,j)}{L_n(k,l)} &= \frac{\sum_{r,s=1}^n M_n(i,r) L_{n-1,2}(r,s) M_1(s,j)}{\sum_{r,s=1}^n M_n(k,r) L_{n-1,2}(r,s) M_1(s,l)} \\
&\leq \frac{\sum_{r,s=1}^n \max\limits_{1\leq i',j'\leq n} M_n(i',j') L_{n-1,2}(r,s) \max\limits_{1\leq i',j'\leq n} M_1(i',j')}{\sum_{r,s=1}^n \min\limits_{1\leq i',j'\leq n} M_n(i',j') L_{n-1,2}(r,s) \min\limits_{1\leq i',j'\leq n} M_1(i',j')} \\
&\leq B^2.
\end{aligned}$$

Similarly,

$$\frac{L_n(i,j)}{L_n(k,l)} \geq \frac{\sum_{r,s=1}^n \min\limits_{1\leq i',j'\leq n} M_n(i',j') L_{n-1,2}(r,s) \min\limits_{1\leq i',j'\leq n} M_1(i',j')}{\sum_{r,s=1}^n \max\limits_{1\leq i',j'\leq n} M_n(i',j') L_{n-1,2}(r,s) \max\limits_{1\leq i',j'\leq n} M_1(i',j')} \geq \frac{1}{B^2}.$$

□

Proof of Lemma 1.3.2. By Lemma 1.3.3, for every $M \in T_{S^+(B)}$ and $1 \leq i, j, k, l \leq d$, it holds

$$M(i,j) \overset{B^2}{\asymp} M(k,l),$$

which yields

$$|M| = \sum_{i,j=1}^d M(i,j) \overset{d^2B^2}{\asymp} M(k,l). \tag{1.3.1}$$

Then, further properties can be easily deduced from (1.3.1). Indeed, the assertions we need are obvious by noticing that for any $x, y \in \mathbb{X}$ and for any $M, M' \in S^+(B)$,

$$|Mx| = \sum_{i,j=1}^d M(i,j) x_j \overset{d^3B^2}{\asymp} |M|,$$

$$\tilde{y} M x = \sum_{i,j=1}^d y_i M(i,j) x_j \overset{d^2B^2}{\asymp} |M|,$$

$$|MM'| = \sum_{i,j,k=1}^{d} M(i,j)M'(j,k) \overset{d^7B^4}{\asymp} |M||M'|.$$

□

We present an alternative proof of Lemma 1.3.2 which is based on the analysis of the spectrum of elements of $S^+(B)$; we expect that this approach may be useful for a larger class of matrices. We only get into the details for the assertion $|Mx| \asymp |M|$, the other assertions are easy consequences.

Before getting into the proof, let us state a useful lemma.

Lemma 1.3.4 *For $d \geq 2$, there exists a compact set included in the interior of $\mathbb{X}$, denoted by $\mathbb{X}_B$, such that $M \cdot x \in \mathbb{X}_B$ for every $M \in T_{S^+(B)}$ and $x \in \mathbb{X}$.*

Proof. *Step 1 :* Assume that the matrix M belongs to $S^+(B)$, where $M = (M(i,j))_{1\leq i,j\leq d}$. For every $i, j = 1, \dots, d$,

$$\langle M \cdot e_j, e_i \rangle = \frac{M(i,j)}{\sum_{k=1}^d M(k,j)} = \frac{1}{\sum_{k=1}^d \frac{M(k,j)}{M(i,j)}},$$

which implies that $\langle M \cdot e_j, e_i \rangle \in \left[\frac{1}{dB}; \frac{1}{1+\frac{1}{B}}\right]$. Thus $M \cdot e_j \in \left[\frac{1}{dB}; \frac{1}{1+\frac{1}{B}}\right]^d$. Fix $\epsilon_B > 0$ such that $\left[\frac{1}{dB}; \frac{1}{1+\frac{1}{B}}\right] \subset [\epsilon_B; 1-\epsilon_B]$ and set $\mathbb{X}_B := \mathbb{X} \cap [\epsilon_B; 1-\epsilon_B]^d$; it holds $M \cdot x \in \mathbb{X}_B$ for every $x \in \mathbb{X}$ and $M \in S^+(B)$.

Step 2 : If $M \in T_{S^+(B)}$, then there exists some $k \geq 1$ such that $M = M_1 \dots M_k$ with $M_1, \dots, M_k \in S^+(B)$. For $k = 1$, the vector $M \cdot x$ belongs to $\mathbb{X}_B$ by Step 1. For $k > 1$, the vector $M_2 \dots M_k \cdot x$ belongs to $\mathbb{X}$ which yields $M \cdot x = M_1 \cdot (M_2 \dots M_k \cdot x) \in \mathbb{X}_B$. □

Remark. First, we recall general results on spectral theory of matrices. Let $M \in T_{S^+(B)}$, then its characteristic polynomial is $P_M(\lambda)$, expressed by

$$P_M(\lambda) = \det(M - \lambda I) = (\lambda - \lambda_1)^{\alpha_1} \dots (\lambda - \lambda_r)^{\alpha_m},$$

where $\lambda_1, \dots, \lambda_m$ are eigenvalues of M with respective algebraic multiplicities $\alpha_1, \dots, \alpha_m$. For $1 \leq j \leq m$, we denote by H_{λ_j} the real part of the characteristic space of λ_j :

$$\begin{aligned} H_{\lambda_j} &= \mathrm{Re}\Big(\mathrm{Ker}(M - \lambda_j I)^{\alpha_j}\Big) \\ &= \Big\{x \in \mathbb{R}^d \mid \exists y \in \mathbb{R}^d, x + iy \in \mathrm{Ker}(M - \lambda_j I)^{\alpha_j}\Big\} \\ &= \left\{v + \bar{v} \mid v \in \mathrm{Ker}(M - \lambda_j I)^{\alpha_j}\right\}. \end{aligned}$$

It is noticable that H_{λ_j} are M-invariant subspaces of $\mathbb{R}^d$. By Cayley-Hamilton's theorem, it holds that $\mathbb{C}^d = \oplus_{j=1}^m \mathrm{Ker}(M - \lambda_j I)^{\alpha_j}$, which implies

$$\mathbb{R}^d = \bigoplus_{j=1}^{m} H_{\lambda_j}.$$

Since M has positive entries, by Perron-Frobenius' theorem, its spectral radius is the simple real dominant eigenvalue, denoted by λ_M. Without loosing generality, we suppose that $\lambda_M = \lambda_1$, then $\alpha_1 = 1$. Corresponding with λ_M is the eigenvector u_M in $\mathbb{X}$. Since λ_M is dominant, all the other eigenvalues λ_k have modulus stricly less than λ_M, for $2 \leq k \leq m$.

Set $H_M := \oplus_{j=2}^{m} H_{\lambda_j}$ and notice that $H_M \subset C^c \cup \{\mathbf{0}\}$. Indeed, if there exists some $u \in C$ such that $u \in H_M$ and $|u| = 1$, then $d(M^n \cdot u, M^n \cdot u_M) = c(M^n)d(u, u_M) \leq \kappa_B^n$, so that $M^n \cdot u \to u_M$ as $n \to +\infty$. This contradicts the fact that $u \in H_M$ and H_M is an M-invariant subspace of $\mathbb{R}^d$.

Therefore, by the inclusion $H_M \subset C^c \cup \{\mathbf{0}\}$, there exists a positive constant δ_B (which only depends on B) such that $|\cos(\widehat{u, v})| = |\langle u, v\rangle| \leq 1 - \delta_B$ for every $u \in H_M$ and $v \in \mathbb{X}_B$ such that $|u| = 1$.

Alternative proof of Lemma 1.3.2. Recall that $\mathbb{X}_B$ is defined in Lemma 1.3.4. We begin to prove a part of the first assertion of Lemma 1.3.2, that there exists a positive constant C_B such that $|Mx| \overset{C_B}{\asymp} |M|$ for every $M \in T_{S^+(B)}$ and $x \in \mathbb{X}$.

Let us fix $x \in \mathbb{X}$. The inequality $|Mx| \leq |M|$ is obvious; let us prove the converse inequality, up to a multiplicative constant. If $M \in S^+(B)$, the conclusion follows from the definition of $S^+(B)$ ($|Mx| \geq \frac{1}{B^2}|M|$). Assume that $M = M_1 \dots M_k$, where $M_1, \dots, M_k \in S^+(B)$ for $k \geq 2$; set $M_{1,k-1} = M_1 \dots M_{k-1}$, then $M = M_{1,k-1}M_k$, which yields $|Mx| = \left|M_{1,k-1}\frac{M_k x}{|M_k x|}\right| \times |M_k x|$ with $\frac{M_k x}{|M_k x|} \in \mathbb{X}_B$ by Lemma 1.3.4. The conclusion follows if we can prove that

$$\forall N \in T_{S^+(B)}, \forall y \in \mathbb{X}_B \qquad |Ny| \geq \frac{1}{C_B}|N|. \tag{1.3.2}$$

Indeed, the assertion follows by taking into account that (1.3.2) implies

$$\begin{aligned} |Mx| &= \left|M_{1,k-1}\frac{M_k x}{|M_k x|}\right| \times |M_k x| \\ &\geq \frac{1}{C_B}|M_{1,k-1}| \times \frac{1}{B^2}|M_k| \\ &\geq \frac{1}{B^2 C_B}|M|, \end{aligned}$$

where the last inequality follows from the fact that $|AB| \geq |A||B|$ for two matrices A and B.

It remains to prove inequality (1.3.2); we fix $N \in T_{S^+(B)}$ and $y \in \mathbb{X}_B$. With the convention $\mathbf{1} = (1, \dots, 1)$, for $\lambda \in \mathbb{R}$, notice that $(\lambda\mathbf{1} + C^c) \cap \mathbb{X}_B \neq \emptyset$ if and only if there exists a positive constant ϵ_B that depends on B such that $\epsilon_B \leq \lambda \leq 1 - \epsilon_B$. The vector y can be decomposed as $y = \lambda\mathbf{1} + y'$, where $\epsilon_B \leq \lambda \leq 1 - \epsilon_B$ and $y' \in H_N \subset C^c \cup \{\mathbf{0}\}$; consequently $Ny = \lambda N\mathbf{1} + Ny'$ with $Ny' \in H_N \subset C^c \cup \{\mathbf{0}\}$. The L^2-norm of the vector Ny satisfies the following inequalities :

$$\begin{aligned} ||Ny||_2^2 &= \lambda^2||N\mathbf{1}||_2^2 + ||Ny'||_2^2 + 2\lambda\langle N\mathbf{1}, Ny'\rangle \\ &= \lambda^2||N\mathbf{1}||_2^2 + ||Ny'||_2^2 + 2\lambda\cos(\widehat{N\mathbf{1}, Ny'})||N\mathbf{1}||_2||Ny'||_2 \\ &\geq \lambda^2||N\mathbf{1}||_2^2 + ||Ny'||_2^2 - 2|\lambda|(1-\delta_B)||N\mathbf{1}||_2||Ny'||_2 \\ &= \lambda^2\Big(1 - (1-\delta_B)^2\Big)||N\mathbf{1}||_2^2 + \Big(|\lambda|(1-\delta_B)||N\mathbf{1}||_2 - ||Ny'||_2\Big)^2 \\ &\geq \epsilon_B^2\Big(1 - (1-\delta_B)^2\Big)||N\mathbf{1}||_2^2. \end{aligned}$$

Hence, for every $x \in C$, the double inequality $\|x\|_2 \leq |x| \leq \sqrt{d}\|x\|_2$ and the fact that $|M| = |M\mathbf{1}|$ yield

$$|Ny| \geq \|Ny\|_2 \geq \epsilon_B \sqrt{1-(1-\delta_B)^2}\|N\mathbf{1}\|_2 \geq \epsilon_B \sqrt{\frac{1-(1-\delta_B)^2}{d}}|N|,$$

which is the inequality (1.3.2) with $C_B^{-1} := \epsilon_B \sqrt{\frac{1-(1-\delta_B)^2}{d}}$.

□

1.4 Lyapunov exponent for product of matrices in S^+

In this section, we give the definition of the upper Lyapunov exponent γ which gives the exponential rate of growth of the norm of products of i.i.d. matrices. We only speak of Lyapunov's exponent for positive matrices. For the case of invertible matrices, readers can find in [Bougerol et Lacroix, 1985].

Let L_n denote the left product $L_n := M_n \dots M_1$, where $M_1, M_2, \dots$ are i.i.d. random matrices in S^+ with common distribution μ, defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It is well known that $|L_n| \leq |M_n| \dots |M_1|$. If $\mathbb{E}(\log^+ |M_1|) < +\infty$ then $\log^+ |L_n|$ is integrable. Furthermore for any $n, p \geq 1$,

$$\begin{aligned} \mathbb{E}[\log^+ |L_{n+p}] &= \mathbb{E}[\log |M_{n+p}...M_{n+1}...M_1|] \\ &= \mathbb{E}[\log |M_{n+p}...M_{n+1}| + \log |M_n...M_1|] \\ &\leq \mathbb{E}[\log |L_p|] + \mathbb{E}[\log |L_n|]. \end{aligned}$$

Therefore, the sequence $(\mathbb{E}[\log |L_n|]_{n\geq 1})_{n\geq 1}$ is sub-additive and

$$\frac{1}{n}\mathbb{E}[\log |L_n|] \longrightarrow \gamma_\mu := \inf_{m\geq 1} \frac{1}{m}\mathbb{E}[\log |L_m|]$$

as $n \to +\infty$ with $\gamma_\mu \in \mathbb{R} \cup \{-\infty\}$.

Definition 1.4.1 *If $\mathbb{E}[\log^+ |M_1|] < +\infty$, then the upper Lyapunov exponent associated with μ is the element γ_μ of $\mathbb{R} \cup \{-\infty\}$ defined by $\gamma_\mu = \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}[\log |L_n|]$.*

The convergence of the quantity $\frac{1}{n}\mathbb{E}[\log |L_n|]$ towards γ_μ may be strengthened as follows.

Theorem 1.4.2 *[Bougerol et Lacroix, 1985] [Benoist et Quint, 2016] If $\mathbb{E}[\log^+ |M_1|] < +\infty$, then*

$$\lim_{n\to+\infty} \frac{1}{n} \log |L_n| = \gamma_\mu \quad \mathbb{P}-\text{a.s.}$$

The upper Lyapunov exponent of μ has an explicit expression, under stronger assumptions on the semi-group generated by the support of μ (namely, conditions P1-P3

introduced below); from Theorem II.1 in [Hennion et Hervé, 2008], under these conditions, there exists a unique probability measure ν on $\mathbb{X}$ such that for every bounded Borel function $\varphi : \mathbb{X} \to \mathbb{R}$,

$$(\mu * \nu)(\varphi) = \int_S \int_{\mathbb{X}} \varphi(g \cdot x)\nu(dx)\mu(dg) = \int_{\mathbb{X}} \varphi(x)\nu(dx) = \nu(\varphi).$$

Such a measure is said to be μ-invariant. The Lyapunov exponent γ_μ is expressed by

$$\gamma_\mu = \int_S \int_{\mathbb{X}} \rho(g, x)\nu(dx)\mu(dg),$$

where $\rho(g, x) = \log |gx|$ for any $(g, x) \in S \times \mathbb{X}$.

1.5 Conditioned limit theorems for products of positive random matrices

This section presents the article published in 2018 in the journal "Latin American Journal of Probability and Mathematical Statistics". For the sake of unification, we change the notation $\mathring{S}$ in the following article into S^+.

Alea 15, 67–100 (2018)

Conditioned limit theorems for products of positive random matrices

Thi Da Cam Pham

Institut Denis-Poisson, Université de Tours, Université d'Orléans, CNRS, Tours, France.
Parc de Grandmont,
37 200 Tours, France.
E-mail address: Thi-Da-Cam.Pham@lmpt.univ-tours.fr

Abstract. Inspired by a recent paper of I. Grama, E. Le Page and M. Peigné, we consider a sequence $(g_n)_{n\geq 1}$ of i.i.d. random $d\times d$-matrices with non-negative entries and study the fluctuations of the process $(\log|g_n\cdots g_1x|)_{n\geq 1}$ for any non-zero vector x in $\mathbb{R}^d$ with non-negative coordinates. Our method involves approximating this process by a martingale and studying harmonic functions for its restriction to the upper half line. Under certain conditions, the probability for this process to stay in the upper half real line up to time n decreases as $\frac{c}{\sqrt{n}}$ for some positive constant c.

1. Introduction

Many limit theorems describe the asymptotic behaviour of random walks with i.i.d. increments, for instance the strong law of large numbers, the central limit theorem, the invariant principle... Besides, the fluctuations of these processes are well studied, for example the decay of the probability that they stay inside the half real line up to time n or functional central limit theorems for random walks conditioned to stay positive. A vast literature exists on this subject, see for instance Bolthausen (1976), Iglehart (1974a), Iglehart (1974b), Iglehart (1975), Kaigh (1976) or Shimura (1983), and references therein. The Wiener-Hopf factorization is usually used in this case and so far, it seems to be impossible to adapt in non-abelian context. Recently, much efforts are made to apply the results above for the logarithm of the norm of the product of i.i.d. random matrices since it behaves similarly to a sum of i.i.d. random variables. Many limit theorems arose for the last 60 years, initiated by Furstenberg and Kesten (1960), Guivarc'h and Raugi (1985), Le Page (1982)... and recently Benoist and Quint (2016). Let us mention also the works by Hennion (1984) and Hennion and Hervé (2008) for matrices with positive entries. However, the studies on the subject of fluctuation was quite sparse a few years ago. Thanks to the approach of Denisov and Wachtel (2015) for random walks in Euclidean spaces and motivated by branching processes, I. Grama, E. Le Page

2000 *Mathematics Subject Classification.* 60J80, 60F17, 60K37.
Key words and phrases. Exit time, Markov chains, Product of random matrices.

and M. Peigné recently progressed for invertible matrices (Grama et al. (2014)). Here we propose to develop the same strategy for matrices with positive entries by using Hennion and Hervé (2008).

We endow $\mathbb{R}^d$ with the norm $|\cdot|$ defined by $|x| := \sum_{i=1}^{d} |x_i|$ for any column vector $x = (x_i)_{1\le i\le d}$. Let $\mathcal{C}$ be the cone of vectors in $\mathbb{R}^d$ with non-negative coordinates

$$\mathcal{C} := \{x \in \mathbb{R}^d : \forall 1 \le i \le d, x_i \ge 0\}$$

and $\mathbb{X}$ be the standard simplex defined by

$$\mathbb{X} := \{x \in \mathcal{C}, |x| = 1\}.$$

Let S be the set of $d \times d$ matrices with non-negative entries such that each column contains at least one positive entry; its interior is $S^+ := \{g = (g(i,j))_{1\le i,j\le d} / g(i,j) > 0\}$. Endowed with the standard multiplication of matrices, the set S is a semigroup and S^+ is the ideal of S, more precisely, for any $g \in S^+$ and $h \in S$, it is evident that $gh \in S^+$.

We consider the following actions:

- the left linear action of S on $\mathcal{C}$ defined by $(g,x) \mapsto gx$ for any $g \in S$ and $x \in \mathcal{C}$,
- the left projective action of S on $\mathbb{X}$ defined by $(g,x) \mapsto g \cdot x := \frac{gx}{|gx|}$ for any $g \in S$ and $x \in \mathbb{X}$.

For any $g = (g(i,j))_{1\le i,j\le d} \in S$, without confusion, let

$$v(g) := \min_{1\le j\le d}\Big(\sum_{i=1}^{d} g(i,j)\Big) \quad \text{and} \quad |g| := \max_{1\le j\le d}\Big(\sum_{i=1}^{d} g(i,j)\Big),$$

then $|\cdot|$ is a norm on S and for any $x \in \mathcal{C}$,

$$0 < v(g)\ |x| \le |gx| \le |g|\ |x|. \tag{1.1}$$

We set $N(g) := \max\left(\frac{1}{v(g)}, |g|\right)$; notice that $N(g) \ge 1$ for any $g \in S$.

On some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a sequence of i.i.d. S-valued matrices $(g_n)_{n\ge 0}$ with the same distribution μ on S. Let $L_0 = Id$ and $L_n := g_n \dots g_1$ for any $n \ge 0$. For any fixed $x \in \mathbb{X}$, we define the $\mathbb{X}$-valued Markov chain $(X_n^x)_{n\ge 0}$ by setting $X_n^x := L_n \cdot x$ for any $n \ge 0$ (or simply X_n if there is no confusion). We denote by P the transition probability of $(X_n)_{n\ge 0}$, defined by: for any $x \in \mathbb{X}$ and any bounded Borel function $\varphi : \mathbb{X} \to \mathbb{C}$,

$$P\varphi(x) := \int_S \varphi(g \cdot x)\mu(dg) = \mathbb{E}[\varphi(L_1 \cdot x)].$$

Hence, for any $n \ge 1$,

$$P^n\varphi(x) = \mathbb{E}[\varphi(L_n \cdot x)].$$

We assume that with positive probability, after finitely many steps, the sequence $(L_n)_{n\ge 1}$ reaches S^+. In mathematical term, it is equivalent to writing as

$$\mathbb{P}\left(\bigcup_{n\ge 1} [L_n \in S^+]\right) > 0.$$

On the product space $S \times \mathbb{X}$, we define the function ρ by setting for any $(g,x) \in S \times \mathbb{X}$,

$$\rho(g,x) := \log |gx|.$$

Notice that $gx = e^{\rho(g,x)} g \cdot x$; in other terms, the linear action of S on $\mathcal{C}$ corresponds to the couple $(g \cdot x, \rho(g,x))$. This function ρ satisfies the cocycle property $\rho(gh,x) = \rho(g, h \cdot x) + \rho(h,x)$ for any $g, h \in S$ and $x \in \mathbb{X}$ and implies the basic decomposition for any $x \in \mathbb{X}$,

$$\log |L_n x| = \sum_{k=1}^{n} \rho(g_k, X_{k-1}^x).$$

For any $a \in \mathbb{R}$ and $n \geq 1$, let $S_0 := a$ and $S_n = S_n(x,a) := a + \sum_{k=1}^{n} \rho(g_k, X_{k-1})$. Then the sequence $(X_n, S_n)_{n \geq 0}$ is a Markov chain on $\mathbb{X} \times \mathbb{R}$ with transition probability $\widetilde{P}$ defined by: for any $(x,a) \in \mathbb{X} \times \mathbb{R}$ and any bounded Borel function $\psi : \mathbb{X} \times \mathbb{R} \to \mathbb{C}$,

$$\widetilde{P}\psi(x,a) = \int_S \psi(g \cdot x, a + \rho(g,x)) \mu(dg).$$

For any $(x,a) \in \mathbb{X} \times \mathbb{R}$, we denote by $\mathbb{P}_{x,a}$ the probability measure on $(\Omega, \mathcal{F})$ conditioned to the event $[X_0 = x, S_0 = a]$ and by $\mathbb{E}_{x,a}$ the corresponding expectation; for the sake of brevity, by $\mathbb{P}_x$ we denote $\mathbb{P}_{x,a}$ when $S_0 = 0$ and by $\mathbb{E}_x$ the corresponding expectation. Hence for any $n \geq 1$,

$$\widetilde{P}^n \psi(x,a) = \mathbb{E}[\psi(L_n \cdot x, a + \log |L_n x|)] = \mathbb{E}_{x,a}[\psi(X_n, S_n)].$$

Now we consider the restriction $\widetilde{P}_+$ to $\mathbb{X} \times \mathbb{R}^+$ of $\widetilde{P}$ defined by: for any $(x,a) \in \mathbb{X} \times \mathbb{R}^+$ and any bounded function $\psi : \mathbb{X} \times \mathbb{R} \to \mathbb{C}$,

$$\widetilde{P}_+ \psi(x,a) = P(\psi \mathbf{1}_{\mathbb{X} \times \mathbb{R}_*^+})(x,a).$$

Let us emphasize that $\widetilde{P}_+$ may not be a Markov kernel on $\mathbb{X} \times \mathbb{R}^+$.

Let $\tau := \min\{n \geq 1 : S_n \leq 0\}$ be the first time the random process $(S_n)_{n \geq 1}$ becomes non-positive; for any $(x,a) \in \mathbb{X} \times \mathbb{R}^+$ and any bounded Borel function $\psi : \mathbb{X} \times \mathbb{R} \to \mathbb{C}$,

$$\widetilde{P}_+ \psi(x,a) = \mathbb{E}_{x,a}[\psi(X_1, S_1); \tau > 1] = \mathbb{E}[\psi(g_1 \cdot x, a + \rho(g_1, x)); a + \rho(g_1, x) > 0].$$

A positive $\widetilde{P}_+$-harmonic function V is any function from $\mathbb{X} \times \mathbb{R}^+$ to $\mathbb{R}^+$ satisfying $\widetilde{P}_+ V = V$. We extend V by setting $V(x,a) = 0$ for $(x,a) \in \mathbb{X} \times \mathbb{R}_*^-$. In other words, the function V is $\widetilde{P}_+$-harmonic if and only if for any $x \in \mathbb{X}$ and $a \geq 0$,

$$V(x,a) = \mathbb{E}_{x,a}[V(X_1, S_1); \tau > 1]. \tag{1.2}$$

From Theorem II.1 in Hennion and Hervé (2008), under conditions P1-P3 introduced below, there exists a unique probability measure ν on $\mathbb{X}$ such that for any bounded Borel function φ from $\mathbb{X}$ to $\mathbb{R}$,

$$(\mu * \nu)(\varphi) = \int_S \int_{\mathbb{X}} \varphi(g \cdot x) \nu(dx) \mu(dg) = \int_{\mathbb{X}} \varphi(x) \nu(dx) = \nu(\varphi).$$

Such a measure is said to be μ-invariant. Moreover, the upper Lyapunov exponent associated with μ is finite and is expressed by

$$\gamma_\mu = \int_S \int_{\mathbb{X}} \rho(g,x) \nu(dx) \mu(dg). \tag{1.3}$$

Now we state the needed hypotheses for later work.

HYPOTHESES

P1 *There exists* $\delta_0 > 0$ *such that* $\displaystyle\int_S N(g)^{\delta_0} \mu(dg) < +\infty$.

P2 *There exists no affine subspaces A of $\mathbb{R}^d$ such that $A \cap \mathcal{C}$ is non-empty and bounded and invariant under the action of all elements of the support of μ.*
P3 *There exists $n_0 \geq 1$ such that $\mu^{*n_0}(S^+) > 0$.*
P4 *The upper Lyapunov exponent γ_μ is equal to 0.*
P5 *There exists $\delta > 0$ such that $\mu\{g \in S : \forall x \in \mathbb{X}, \log|gx| \geq \delta\} > 0$.*

Condition P1 requires exponential moments of the quantity $\log N(g)$. Condition P2 roughly speaking requires that the dimension of the smallest closed semigroup which contains the support of μ can not be reduced. Condition P3 is called "Contraction property", it says that there is a positive probability that after a finite number n_0 of times, the product of n_0 matrices has positive coefficients, then those products are contracting. Condition P4 does not imply that the event $[\tau > n]$ occurs with positive probability when n tends to infinity, so P5 ensures this fact.

In this paper, we establish the asymptotic behaviour of $\mathbb{P}_{x,a}(\tau > n)$ by studying a $\widetilde{P}_+$-harmonic function V. More precisely, Proposition 1.1 concerns the existence of a $\widetilde{P}_+$-harmonic function and its properties whereas Theorem 1.2 is about the limit behaviour of the exit time τ.

Proposition 1.1. *Assume hypotheses P1-P5.*

(1) *For any $x \in \mathbb{X}$ and $a \geq 0$, the sequence $\left(\mathbb{E}_{x,a}[S_n; \tau > n]\right)_{n\geq 0}$ converges to the function $V(x,a) := a - \mathbb{E}_{x,a}M_\tau$, for $(M_n)_{n\geq 0}$ is a martingale defined in Proposition 2.6.*
(2) *For any $x \in \mathbb{X}$ the function $V(x,\cdot)$ is increasing on $\mathbb{R}^+$.*
(3) *There exist $c > 0$ and $A > 0$ such that for any $x \in \mathbb{X}$ and $a \geq 0$,*

$$\frac{1}{c} \vee (a - A) \leq V(x,a) \leq c(1+a).$$

(4) *For any $x \in \mathbb{X}$, the function $V(x,.)$ satisfies $\lim_{a\to+\infty} \dfrac{V(x,a)}{a} = 1$.*
(5) *The function V is $\widetilde{P}_+$-harmonic.*

The function V contains information of the part of the trajectory which stays in $\mathbb{R}^+$ as stated in Theorem 1.2.

Theorem 1.2. *Assume P1-P5. Then for any $x \in \mathbb{X}$ and $a \geq 0$,*

$$\mathbb{P}_{x,a}(\tau > n) \sim \frac{2V(x,a)}{\sigma\sqrt{2\pi n}} \text{ as } n \to +\infty,$$

where $\sigma^2 := \lim_{n\to+\infty} \dfrac{1}{n}\mathbb{E}_x[S_n^2]$ is the variance of the semi-markovian random walk $(S_n)_{n\geq 0}$. Moreover, there exists a constant c such that for any $x \in \mathbb{X}$, $a \geq 0$ and $n \geq 1$,

$$\sqrt{n}\mathbb{P}_{x,a}(\tau > n) \leq cV(x,a).$$

As a direct consequence, we prove that the sequence $(\frac{S_n}{\sigma\sqrt{n}})_{n\geq 1}$, conditioned to the event $\tau > n$, converges in distribution to the Rayleigh law as stated below.

Theorem 1.3. *Assume P1-P5. For any $x \in \mathbb{X}$, $a \geq 0$ and $t > 0$,*

$$\lim_{n\to+\infty} \mathbb{P}_{x,a}\left(\frac{S_n}{\sqrt{n}} \leq t \mid \tau > n\right) = 1 - \exp\left(-\frac{t^2}{2\sigma^2}\right).$$

In section 2, we approximate the chain $(S_n)_{n\geq 0}$ by a martigale and in section 3, we study the harmonic function V and state the proof of Proposition 1.1. We use the coupling argument to prove Theorem 1.2 and Theorem 1.3 in section 4. At last, in section 5 we check general conditions to apply an invariant principle stated in Theorem 2.1 in Grama et al. (2014).

Throughout this paper, we denote the absolute constants such as $C, c, c_1, c_2, \ldots$ and the constants depending on their indices such as $c_\varepsilon, c_p, \ldots$. Notice that they are not always the same when used in different formulas. The integer part of a real constant a is denoted by $[a]$.

2. Approximation of the chain $(S_n)_{n\geq 0}$

In this section, we discuss the spectral properties of P and then utilise them to approximate the chain $(S_n)_{n\geq 0}$. Throughout this section, we assume that conditions P1-P4 hold true.

2.1. *Spectral properties of the operators P and its Fourier transform.*

Following Hennion (1997), we endow $\mathbb{X}$ with a bounded distance d such that g acts on $\mathbb{X}$ as a contraction with respect to d for any $g \in S$. For any $x, y \in \mathbb{X}$, we write:

$$\mathsf{m}(x,y) = \min_{1\leq i\leq d} \left\{ \frac{x_i}{y_i} | y_i > 0 \right\}$$

and it is clear that $0 \leq \mathsf{m}(x,y) \leq 1$. For any $x, y \in \mathbb{X}$, let $d(x,y) := \varphi(\mathsf{m}(x,y)\,\mathsf{m}(y,x))$, where φ is the one-to-one function defined for any $s \in [0,1]$ by $\varphi(s) := \dfrac{1-s}{1+s}$. Setting $c(g) := \sup\{d(g\cdot x, g\cdot y), x, y \in \mathbb{X}\}$ for $g \in S$; the proposition below gives some more properties of d and $c(g)$.

Proposition 2.1. *Hennion (1997) The quantity d is a distance on $\mathbb{X}$ satisfying the following properties:*

(1) $\sup\{d(x,y) : x, y \in \mathbb{X}\} = 1$.
(2) $|x-y| \leq 2d(x,y)$ *for any* $x, y \in \mathbb{X}$.
(3) $c(g) \leq 1$ *for any* $g \in S$, *and* $c(g) < 1$ *if and only if* $g \in S^+$.
(4) $d(g\cdot x, g\cdot y) \leq c(g)\, d(x,y) \leq c(g)$ *for any and* $x, y \in \mathbb{X}$.
(5) $c(gh) \leq c(g)\, c(h)$ *for any* $g, h \in S$.

From now on, we consider a sequence $(g_n)_{n\geq 0}$ of i.i.d. S-valued random variables, we set $a_k := \rho(g_k, X_{k-1})$ for $k \geq 1$ and hence $S_n = a + \sum_{k=1}^n a_k$ for $n \geq 1$. In order to study the asymptotic behavior of the process $(S_n)_{n\geq 0}$, we need to consider the "Fourier transform" of the random variables a_k, under $\mathbb{P}_x, x \in \mathbb{X}$, similarly for classical random walks with independent increments on $\mathbb{R}$. Let P_t be the family of "Fourier operators" defined for any $t \in \mathbb{R}$, $x \in \mathbb{X}$ and any bounded Borel function $\varphi : \mathbb{X} \to \mathbb{C}$ by

$$P_t\varphi(x) := \int_S e^{it\rho(g,x)}\varphi(g\cdot x)\mu(dg) = \mathbb{E}_x\left[e^{ita_1}\varphi(X_1)\right] \tag{2.1}$$

and for any $n \geq 1$,

$$P_t^n\varphi(x) = \mathbb{E}[e^{it\log|L_n x|}\varphi(L_n\cdot x)] = \mathbb{E}_x[e^{itS_n}\varphi(X_n)]. \tag{2.2}$$

Moreover, we can imply that

$$\begin{aligned} P^m P_t^n \varphi(x) &= \mathbb{E}\left[e^{it\log|g_{m+n}\cdots g_{m+1}(L_m\cdot x)|}\varphi(L_{m+n}\cdot x)\right] \\ &= \mathbb{E}_x\left[e^{it(a_{m+1}+\cdots+a_{m+n})}\varphi(X_{n+m})\right] \end{aligned} \tag{2.3}$$

and when $\varphi = 1$, we obtain

$$\mathbb{E}_x\left[e^{itS_n}\right] = P_t^n 1(x) \quad \text{and} \quad \mathbb{E}_x\left[e^{it(a_{m+1}+\cdots+a_{m+n})}\right] = P^m P_t^n 1(x).$$

We consider the space $C(\mathbb{X})$ of continuous functions from $\mathbb{X}$ to $\mathbb{C}$ endowed with the norm of uniform convergence $|.|_\infty$. Let L be the subset of Lipschitz functions on $\mathbb{X}$ defined by

$$L := \{\varphi \in C(\mathbb{X}) : |\varphi|_L := |\varphi|_\infty + m(\varphi) < +\infty\},$$

where $m(\varphi) := \sup_{\substack{x,y\in\mathbb{X}\\ x\neq y}} \frac{|\varphi(x)-\varphi(y)|}{d(x,y)}$. The spaces $(C(\mathbb{X}), |\cdot|_\infty)$ and $(L, |\cdot|_L)$ are Banach spaces and the canonical injection from L into $C(\mathbb{X})$ is compact. The norm of a bounded operation A from L to L is denoted by $|A|_{L\to L} := \sup_{\varphi\in L} |A\varphi|_L$. We denote L' the topological dual of L endowed with the norm $|\cdot|_{L'}$ corresponding to $|\cdot|_L$; notice that any probability measure ν on $\mathbb{X}$ belongs to L'.
For further uses, we state here some helpful estimations.

Lemma 2.2. *For $g \in S$, $x, y, z \in \mathbb{X}$ such that $d(x, y) < 1$,*

$$|\rho(g, x)| \leq 2\log N(g), \tag{2.4}$$

and for any $t \in \mathbb{R}$,

$$|e^{it\rho(g,y)} - e^{it\rho(g,z)}| \leq \Big(4\min\{2|t|\log N(g), 1\} + 2C|t|\Big)d(y, z), \tag{2.5}$$

where $C = \sup\{\frac{1}{u}\log\frac{1}{1-u} : 0 < u \leq \frac{1}{2}\} < +\infty$.

Proof. For the first assertion, from (1.1), we can imply that $|\log|gx|| \leq \log N(g)$. For the second assertion, we refer to the proof the Theorem III.2 in Hennion and Hervé (2008).

□

The following Proposition is a combination of several results stated in Hennion and Hervé (2008); for the sake of completeness, we present below the main steps of its proof. Denote $\varepsilon(t) := \int_S \min\{2|t|\log N(g), 2\}\mu(dg)$. Notice that $\lim_{t\to 0}\varepsilon(t) = 0$.

Proposition 2.3. *Under hypotheses* **P1, P2, P3** *and* **P4**, *for any $t \in \mathbb{R}$, the operator P_t acts on L and there exists on $\mathbb{X}$ a unique P-invariant probability measure ν. Furthermore,*

(1) *If $\Pi : L \to L$ denotes the rank one operator defined by $\Pi(\varphi) = \nu(\varphi)1$ for any function $\varphi \in L$ and $R := P - \Pi$, the operator $R : L \to L$ satisfies*

$$\Pi R = R\Pi = 0,$$

and its spectral radius is less than 1. In other words, there exist constants $C > 0$ and $0 < \kappa < 1$ such that $|R^n|_{L\to L} \leq C\kappa^n$ for any $n \geq 1$.

(2) *There exist $\epsilon > 0$ and $0 \leq r_\epsilon < 1$ such that for any $t \in [-\epsilon, \epsilon]$, there exist a complex number λ_t close to 1 with modulus less than or equal to 1, a rank one operator Π_t and an operator R_t on L with spectral radius less than or equal to r_ϵ such that*

$$P_t = \lambda_t\Pi_t + R_t \quad \text{and} \quad \Pi_t R_t = R_t\Pi_t = 0.$$

Moreover, $C_P := \sup_{\substack{-\epsilon\le t\le\epsilon \\ n\ge 0}} |P_t^n|_{L\to L} < +\infty$.

(3) *For any* $p \geq 1$,

$$\sup_{n\geq 0}\sup_{x\in\mathbb{X}} \mathbb{E}_x|\rho(g_{n+1}, X_n)|^p < +\infty. \tag{2.6}$$

Proof. **(a)** We first check that P_t acts on $(L, |\cdot|_L)$ for any $t \in \mathbb{R}$. On one hand, $|P_t\varphi|_\infty \leq |\varphi|_\infty$ for any $\varphi \in L$. On the other hand, by (2.5) for any $x, y \in \mathbb{X}$ such that $x \neq y$,

$$\begin{aligned}
\frac{|P_t\varphi(x) - P_t\varphi(y)|}{d(x,y)} &\leq \int_S \left(\left| \frac{e^{it\rho(g,x)} - e^{it\rho(g,y)}}{d(x,y)} \right| |\varphi(g\cdot x)| + \left| \frac{\varphi(g\cdot x) - \varphi(g\cdot y)}{d(x,y)} \right| \right) \mu(dg) \\
&\leq |\varphi|_\infty(4\varepsilon(t) + 2C|t|) + \int_S \left(\frac{|\varphi(g\cdot x) - \varphi(g\cdot y)|}{d(g\cdot x, g\cdot y)} \frac{d(g\cdot x, g\cdot y)}{d(x,y)} \right) \mu(dg), \\
&\leq |\varphi|_\infty(4\varepsilon(t) + 2C|t|) + m(\varphi),
\end{aligned}$$

which implies $m(P_t\varphi) \leq |\varphi|_\infty(4\varepsilon(t) + 2C|t|) + m(\varphi) < +\infty$. Therefore $P_t\varphi \in L$.

(b) Let μ^{*n} be the distribution of the random variable L_n and set

$$c(\mu^{*n}) := \sup\left\{ \int_S \frac{d(g\cdot x, g\cdot y)}{d(x,y)} d\mu^{*n}(g) : x, y \in \mathbb{X}, x \neq y \right\}.$$

By hypothesis P3 and Proposition 2.1 (3), it holds $c(\cdot) < 1$. Furthermore, $c(\mu^{*(m+n)}) \leq c(\mu^{*m})c(\mu^{*n})$ for any $m, n > 0$; in other words, the sequence $(c(\mu^{*n}))_{n\geq 1}$ is sub-multiplicative. Thus, its limit $\kappa := \lim_{n\to+\infty} \left(c(\mu^{*n})\right)^{\frac{1}{n}}$ does exist and is strictly less than 1.

The existence and unicity of an invariant probability measure ν on $\mathbb{X}$ is a direct consequence of this contraction property, this is Theorem II-1 in Hennion and Hervé (2008).

(c) Now, let us achieve the proof of assertion (1) of the Proposition. Let Π be the rank one projection on L defined by $\Pi\varphi = \nu(\varphi)\mathbf{1}$ for any $\varphi \in L$. Let $R := P - \Pi$. By definition, we obtain $P\Pi = \Pi P = \Pi$ and $\Pi^2 = \Pi$ which implies $\Pi R = R\Pi = 0$ and $R^n = P^n - \Pi$ for any $n \geq 1$.

The same argument as in **(a)**, with $t = 0$ yields

$$m(P^n\varphi) \leq m(\varphi)c(\mu^{*n}).$$

Notice that $P^n(\varphi - \Pi\varphi)$ belongs to $Ker\Pi$ for any $\varphi \in L$ and $n \geq 0$ and that $m(\varphi) \leq |\varphi|_L \leq 3m(\varphi)$ for any $\varphi \in Ker\Pi$. Hence $|P^n(\varphi - \Pi\varphi)|_L \leq 3c(\mu^{*n})|\varphi|_L$ which yields

$$|R^n|_{L\to L} = |P^n - \Pi|_{L\to L} = |P^n(I - \Pi)|_{L\to L} \leq 3c(\mu^{*n}).$$

Therefore, the spectral radius of R is less than or equal to κ given above.

(d) By Hypothesis P1, the function $t \mapsto P_t$ is analytic near 0. The theory of perturbations (see Dunford and Schwarz (1958) Chapter VII, section 6) allows to extend the decomposition $P = \Pi + R$ to the operator P_t when t is close to 0. Indeed, for $\epsilon > 0$ small enough, there exists $r_\epsilon \in [0, 1[$ such that, for any $t \in [-\epsilon; \epsilon]$, the operator P_t may be decomposed as $P_t = \lambda_t\Pi_t + R_t$, where the spectral radius of R_t is less than or equal to r_ϵ and λ_t is the unique eigenvalue of P_t with modulus

greater than r_ϵ; furthermore, the eigenvalue λ_t is simple. In order to control P_t^n, we ask λ_t^n to be bounded. Notice that by Hypothesis P1, the function $t \mapsto P_t$ is analytic near 0. To prove that the sequence $(P_t^n)_t$ is bounded in L, it suffices to check $|\lambda_t| \leq 1$ for any $t \in [-\epsilon, \epsilon]$.
When $\varphi(x) = \mathbf{1}(x)$, equality (2.2) becomes

$$P_t^n \mathbf{1}(x) = \mathbb{E}\left[e^{it\rho(L_n,x)}\right] = \lambda_t^n \Pi_t \mathbf{1}(x) + R_t^n \mathbf{1}(x). \tag{2.7}$$

By the theory of perturbations, since $t \mapsto P_t$ is analytic around 0, the same holds for the map $t \mapsto \lambda_t$; thus, its local expansion near 0 is

$$\lambda_t = 1 + t\lambda_0' + \frac{t^2}{2}\lambda_0''[1 + o(1)]. \tag{2.8}$$

Taking the first derivative of (2.7) with respect to t, we may write for any $n \geq 0$,

$$\begin{aligned}
\mathbb{E}(i\rho(L_n,x)e^{it\rho(L_n,x)}) &= \frac{d}{dt}P_t^n \mathbf{1}(x) \\
&= \frac{d}{dt}\left(\lambda_t^n \Pi_t \mathbf{1}(x) + R_t^n \mathbf{1}(x)\right) \\
&= n\lambda_t^{n-1}\lambda_t' \Pi_t \mathbf{1}(x) + \lambda_t^n \Pi_t' \mathbf{1}(x) + (R_t^n)' \mathbf{1}(x)
\end{aligned}$$

Recall that $\lambda_0 = 1$, $\Pi_0 1(x) = 1$ and $|R^n|_{L\to L} \leq C r_\epsilon^n$; hence,

$$i\mathbb{E}(\rho(L_n,x)) = n\lambda_0' + \Pi_0' 1(x) - [(R_t^n)' 1(x)]_{t=0},$$

which readily imply $\lambda_0' = i \lim_{n\to+\infty} \frac{1}{n}\mathbb{E}[\rho(L_n,x)] = i\gamma_\mu = 0$.

Similarly, by taking the second derivative of (2.7), one may write

$$\begin{aligned}
-\mathbb{E}(\rho(L_n,x)^2 e^{it\rho(L_n,x)}) &= \frac{d^2}{dt^2}\left(\lambda_t^n \Pi_t \mathbf{1}(x) + R_t^n \mathbf{1}(x)\right) \\
&= n(n-1)\lambda_t^{n-2}(\lambda_t')^2 \Pi_t \mathbf{1}(x) + n\lambda_t^{n-1}\lambda_t'' \Pi_t \mathbf{1}(x) + \\
&\quad + 2n\lambda_t^{n-1}\lambda_t' \Pi_t' \mathbf{1}(x) + \lambda_t^n \Pi_t'' \mathbf{1}(x) + (R_t^n)'' \mathbf{1}(x).
\end{aligned}$$

This equality for $t = 0$ yields

$$-\mathbb{E}(\rho(L_n,x)^2) = n\lambda_0'' + \Pi_0'' \mathbf{1}(x) + [(R_t^n)'' \mathbf{1}(x)]_{t=0}$$

so that $\lambda_0'' = -\lim_{n\to+\infty} \frac{1}{n}\mathbb{E}[\rho(L_n,x)^2] = -\sigma^2$. Notice that $\lambda_0'' = -\sigma^2 < 0$ by Lemma 5.3 in Bougerol and Lacroix (1985).

Therefore, for t close to 0, expression (2.8) becomes

$$\lambda_t = 1 - \frac{\sigma^2}{2}t^2(1 + o(1)),$$

hence $|\lambda_t| \leq 1$ for t small enough.

(e) In particular, inequality (1.1) implies $|\rho(g,x)| \leq \log N(g)$ for any $x \in \mathbb{X}$. Therefore, for any $p \geq 1$, $x \in \mathbb{X}$ and $n \geq 1$, Hypothesis P1 yields

$$\mathbb{E}_x|\rho(g_{n+1}, X_n)|^p \leq \frac{p!}{\delta_0^p}\mathbb{E}_x e^{\delta_0|\rho(g_{n+1},X_n)|} \leq \frac{p!}{\delta_0^p}\mathbb{E}N(g)^{\delta_0} < +\infty.$$

□

2.2. *Martingale approximation of the chain* $(S_n)_{n\geq 0}$.
As announced in the abstract, we approximate the process $(S_n)_{n\geq 0}$ by a martingale $(M_n)_{n\geq 0}$. In order to construct the suitable martingale, we introduce the operator $\overline{P}$ and then find the solution of the Poisson equation as follows. First, it is neccessary to introduce some notation and basic properties. Let $g_0 = I$ and $X_{-1} := X_0$. The sequence $((g_n, X_{n-1}))_{n\geq 0}$ is a Markov chain on $S \times \mathbb{X}$, starting from (Id, x) and with transition operator $\overline{P}$ defined by: for any $(g, x) \in S \times \mathbb{X}$ and any bounded measurable function $\phi : S \times \mathbb{X} \to \mathbb{R}$,

$$\overline{P}\phi(g,x) := \int_{S\times\mathbb{X}} \phi(h,y)\overline{P}((g,x),dhdy) = \int_S \phi(h, g\cdot x)\mu(dh) \tag{2.9}$$

(in other words, the measure $\overline{P}((g,x),dhdy)$ on $S \times \mathbb{X}$ equals $\delta_{g\cdot x}(dy)\mu(dh)$).
Notice that by (2.4), under assumption **P1**, for any $g \in S$ and $x \in \mathbb{X}$, the function $h \mapsto \rho(h, g \cdot x)$ is μ-integrable, so that $\overline{P}\rho(g, x)$ is well defined.

Lemma 2.4. *The function* $\bar{\rho} : x \mapsto \int_S \rho(g,x)\mu(dg)$ *belongs to* L *and for any* $g \in S$, $x \in \mathbb{X}$ *and* $n \geq 0$,

$$\overline{P}^{n+1}\rho(g,x) = P^n\overline{\rho}(g\cdot x). \tag{2.10}$$

Proof. (1) For any $x \in \mathbb{X}$, definition of ρ and (2.4) yield

$$|\overline{\rho}(x)| \leq \int_S |\log|gx||\mu(dg) \leq \int_S 2\log N(g)\mu(dg) \leq c(\delta_0)\int_S 2N(g)^{\delta_0}\mu(dg) < +\infty,$$

where $c(\delta_0)$ is a constant depending on δ_0. Hence $|\overline{\rho}|_\infty < +\infty$. For any $x, y \in \mathbb{X}$ such that $d(x,y) > \frac{1}{2}$, we can see that

$$|\rho(g,x) - \rho(g,y)| \leq |\rho(g,x) - \rho(g,y)|2d(x,y) \leq 8\log N(g)d(x,y). \tag{2.11}$$

For any $x, y \in \mathbb{X}$ such that $d(x,y) \leq \frac{1}{2}$, applying Lemma III.1 in Hennion and Hervé (2008), we obtain

$$|\rho(g,x) - \rho(g,y)| \leq 2\log\frac{1}{1-d(x,y)} \leq 2Cd(x,y), \tag{2.12}$$

where C is given in Lemma 2.2. For any $x, y \in \mathbb{X}$, by (2.11) and (2.12) we obtain

$$\begin{aligned} |\overline{\rho}(x) - \overline{\rho}(y)| &\leq \int_S |\rho(g,x) - \rho(g,y)|\,\mu(dg) \\ &\leq \int_S [8\log N(g) + 2C]d(x,y)\mu(dg). \end{aligned}$$

Thus $m(\overline{\rho}) = \sup_{x,y\in\mathbb{X}, x\neq y} \dfrac{|\overline{\rho}(x) - \overline{\rho}(y)|}{d(x,y)} < +\infty$.

(2) From (2.9) and definition of ρ, it is obvious that

$$\overline{P}\rho(g,x) = \int_S \rho(h, g\cdot x)\mu(dh) = \overline{\rho}(g\cdot x),$$

which yields

$$\begin{aligned}\overline{P}^2\rho(g,x) &= \overline{P}(\overline{P}\rho)(g,x) = \int_{S\times\mathbb{X}} (\overline{P}\rho)(k,y)\overline{P}((g,x),dkdy)\\ &= \int_{S\times\mathbb{X}} \overline{\rho}(k\cdot y)\overline{P}((g,x),dkdy)\\ &= \int_S \overline{\rho}(k\cdot(g\cdot x))\mu(dk) = P\overline{\rho}(g\cdot x).\end{aligned}$$

By induction, we obtain $\overline{P}^{n+1}\rho(g,x) = P^n\overline{\rho}(g\cdot x)$ for any $n\geq 0$.

□

Formally, the solution $\theta : S\times\mathbb{X}\to\mathbb{R}$ of the equation $\theta - \overline{P}\theta = \rho$ is the function

$$\theta : (g,x)\mapsto \sum_{n=0}^{+\infty}\overline{P}^n\rho(g,x).$$

Notice that we do not have any spectral property for $\overline{P}$ and ρ does not belong to L. However, we still obtain the convergence of this series by taking into account the important relation (2.10), as shown in the following lemma.

Lemma 2.5. *The sum $\theta = \sum_{n=0}^{+\infty}\overline{P}^n\rho$ exists and satisfies the Poisson equation $\rho = \theta - \overline{P}\theta$. Moreover,*

$$|\overline{P}\theta|_\infty = \sup_{g\in S, x\in\mathbb{X}} |\theta(g,x) - \rho(g,x)| < +\infty; \tag{2.13}$$

and for any $p\geq 1$, it holds

$$\sup_{n\geq 0}\sup_{x\in\mathbb{X}} \mathbb{E}_x|\theta(g_{n+1},X_n)|^p < +\infty. \tag{2.14}$$

Proof. **(1)** Since P acts on $(L,|\cdot|_L)$ and $\overline{\rho}\in L$ from Lemma 2.4, we obtain $P\overline{\rho}\in L$. Thanks to definition of ρ, (1.3) and P4, it follows that

$$\nu(\overline{\rho}) = \int_{\mathbb{X}}\overline{\rho}(x)\nu(dx) = \int_S\int_{\mathbb{X}}\rho(g,x)\nu(dx)\mu(dg) = \gamma_\mu = 0.$$

Proposition 2.3 and the relation (2.10) yield for any $x\in\mathbb{X}$ and $n\geq 0$,

$$\overline{P}^{n+1}\rho(g,x) = P^n\overline{\rho}(g\cdot x) = \Pi\overline{\rho}(g\cdot x) + R^n\overline{\rho}(g\cdot x) = \nu(\overline{\rho}) + R^n\overline{\rho}(g\cdot x) = R^n\overline{\rho}(g\cdot x)$$

and there exist $C>0$ and $0<\kappa<1$ such that for any $x\in\mathbb{X}$ and $n\geq 0$,

$$|R^n\overline{\rho}(x)| \leq |R^n\overline{\rho}|_L \leq |R^n|_{L\to L}\,|\overline{\rho}|_L \leq C\kappa^n.$$

Hence for any $g\in S$ and $x\in\mathbb{X}$,

$$\left|\sum_{n=1}^{+\infty}\overline{P}^n\rho(g,x)\right| \leq \sum_{n=0}^{+\infty}|P^n\overline{\rho}(g\cdot x)| \leq C\sum_{n=0}^{+\infty}\kappa^n = \frac{C}{1-\kappa} < +\infty.$$

Therefore, the function $\theta = \sum_{n=0}^{+\infty} \overline{P}^n \rho$ exists and obviously satisfies the Poisson equation $\rho = \theta - \overline{P}\theta$. Finally, it is evident that

$$\sup_{g\in S, x\in\mathbb{X}} |\theta(g,x) - \rho(g,x)| = \sup_{g\in S, x\in\mathbb{X}} \left| \sum_{n=1}^{+\infty} \overline{P}^n \rho(g,x) \right| < +\infty.$$

(2) Indeed, from (2.6), (2.13) and Minkowski's inequality, the assertion arrives. □

Now we contruct a martingale to approximate the Markov walk $(S_n)_{n\geq 0}$. Hence, from the definition of S_n and the Poisson equation, by adding and removing the term $\overline{P}\theta(g_0, X_{-1})$, we obtain

$$\begin{aligned} S_n &= a + \rho(g_1, X_0) + \ldots + \rho(g_n, X_{n-1}) \\ &= a + \overline{P}\theta(g_0, X_{-1}) - \overline{P}\theta(g_n, X_{n-1}) + \sum_{k=0}^{n-1} \left[\theta(g_{k+1}, X_k) - \overline{P}\theta(g_k, X_{k-1})\right]. \end{aligned}$$

Let $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $\mathcal{F}_n := \sigma\{g_k : 0 \leq k \leq n\}$ for $n \geq 1$.

Proposition 2.6. *For any $n \geq 0$, $x \in \mathbb{X}$, $a \geq 0$ and $p > 2$, the sequence $(M_n)_{n\geq 0}$ defined by*

$$M_0 := S_0 \text{ and } M_n := M_0 + \sum_{k=0}^{n-1} \left[\theta(g_{k+1}, X_k) - \overline{P}\theta(g_k, X_{k-1})\right] \tag{2.15}$$

is a martingale in $L^p(\Omega, \mathbb{P}_{x,a}, (\mathcal{F}_n)_{n\geq 0})$ satisfying the properties:

$$\sup_{n\geq 0} |S_n - M_n| \leq 2|\overline{P}\theta|_\infty \quad \mathbb{P}_{x,a}\text{-a.s.} \tag{2.16}$$

$$\sup_{n\geq 1} n^{-\frac{p}{2}} \sup_{x\in\mathbb{X}} \mathbb{E}_{x,a}|M_n|^p < +\infty. \tag{2.17}$$

From now on, we set $A := 2|\overline{P}\theta|_\infty$.

Proof. By definition (2.15), martingale property arrives.

(1) From the construction of M_n and (2.13), we can see easily that

$$\sup_{n\geq 0} |S_n - M_n| = \sup_{n\geq 0} \left|\overline{P}\theta(g_0, X_{-1}) - \overline{P}\theta(g_n, X_{n-1})\right| \leq 2\left|\overline{P}\theta\right|_\infty < +\infty \quad \mathbb{P}_{x,a}\text{-a.s.}.$$

(2) Denote $\xi_k := \theta(g_{k+1}, X_k) - \overline{P}\theta(g_k, X_{k-1})$. Thus $M_n = M_0 + \sum_{k=0}^{n-1} \xi_k$. Using Burkholder's inequality, for any $p \geq 1$, there exists some positive constant c_p such that for $0 \leq k < n$,

$$(\mathbb{E}_{x,a}|M_n|^p)^{\frac{1}{p}} \leq c_p \left(\mathbb{E}_{x,a} \left| \sum_{k=0}^{n-1} \xi_k^2 \right|^{\frac{p}{2}} \right)^{\frac{1}{p}}.$$

Now, with $p > 2$, applying Hölder's inequality, we obtain

$$\left| \sum_{k=0}^{n-1} \xi_k^2 \right| \leq n^{1-\frac{2}{p}} \left(\sum_{k=0}^{n-1} |\xi_k|^p \right)^{\frac{2}{p}},$$

which implies

$$\mathbb{E}_{x,a}\left|\sum_{k=0}^{n-1}\xi_k^2\right|^{\frac{p}{2}} \leq n^{\frac{p}{2}-1}\mathbb{E}_{x,a}\sum_{k=0}^{n-1}|\xi_k|^p \leq n^{\frac{p}{2}}\sup_{0\leq k\leq n-1}\mathbb{E}_{x,a}|\xi_k|^p.$$

Since $(M_n)_n$ is a martingale, by using the convexity property, we can see that for any $k \geq 0$,

$$\left|\overline{P}\theta(g_k, X_{k-1})\right|^p = \left|\mathbb{E}_{x,a}\Big[|\theta(g_{k+1}, X_k)||\mathcal{F}_k\Big]\right|^p \leq \mathbb{E}_{x,a}\Big[|\theta(g_{k+1}, X_k)|^p|\mathcal{F}_k\Big],$$

which implies $\mathbb{E}_{x,a}\left|\overline{P}\theta(g_k, X_{k-1})\right|^p \leq \mathbb{E}_{x,a}\left|\theta(g_{k+1}, X_k)\right|^p$. Therefore, we obtain

$$\begin{aligned}
\Big(\mathbb{E}_{x,a}|M_n|^p\Big)^{\frac{1}{p}} &\leq c_p\left(n^{\frac{p}{2}}\sup_{0\leq k\leq n-1}\mathbb{E}_{x,a}|\xi_k|^p\right)^{\frac{1}{p}} \leq c_p n^{\frac{1}{2}}\sup_{0\leq k\leq n-1}\Big(\mathbb{E}_{x,a}|\xi_k|^p\Big)^{\frac{1}{p}} \\
&\leq c_p n^{\frac{1}{2}}\sup_{0\leq k\leq n-1}\Big[\Big(\mathbb{E}_{x,a}|\theta(g_{k+1}, X_k)|^p\Big)^{1/p} + \Big(\mathbb{E}_{x,a}|\overline{P}\theta(g_k, X_{k-1})|^p\Big)^{1/p}\Big] \\
&\leq 2c_p n^{\frac{1}{2}}\sup_{0\leq k\leq n-1}\Big(\mathbb{E}_{x,a}|\theta(g_{k+1}, X_k)|^p\Big)^{\frac{1}{p}}.
\end{aligned}$$

Consequently, we obtain $\mathbb{E}_{x,a}|M_n|^p \leq (2c_p)^p n^{\frac{p}{2}}\sup_{0\leq k\leq n-1}\mathbb{E}_{x,a}|\theta(g_{k+1}, X_k)|^p$ and the assertion arrives by using (2.14).

□

3. Proof of Proposition 1.1

In this section we construct explicitly a $\widetilde{P}_+$-harmonic function V and study its properties. We begin with the first time the martingale $(M_n)_{n\geq 0}$ (2.15) visit $]-\infty, 0]$, defined by

$$T = \min\{n \geq 1 : M_n \leq 0\}.$$

The equality $\gamma_\mu = 0$ yields $\liminf_{n\to+\infty} S_n = -\infty$ $\mathbb{P}_{x,a}$-almost surely for any $x \in \mathbb{X}$ and $a \geq 0$. Thus, by Lemma 2.6, $\liminf_{n\to+\infty} M_n = -\infty$ $\mathbb{P}_{x,a}$-almost surely so that $T < +\infty$ $\mathbb{P}_{x,a}$-a.s. for any $x \in \mathbb{X}$ and $a \geq 0$.

3.1. *On the properties of T and $(M_n)_n$.*

We need to control the first moment of the random variable $|M_{T\wedge n}|$ under $\mathbb{P}_x$; we consider the restriction of this variable to the event $[T \leq n]$ in Lemma 3.1 and control the remaining term in Lemma 3.4.

Lemma 3.1. *There exists $\varepsilon_0 > 0$ and $c > 0$ such that for any $\varepsilon \in (0, \varepsilon_0), n \geq 1, x \in \mathbb{X}$ and $a \geq n^{\frac{1}{2}-\varepsilon}$,*

$$\mathbb{E}_{x,a}\Big[|M_T|; T \leq n\Big] \leq c\frac{a}{n^\varepsilon}.$$

Proof. For any $\varepsilon > 0$, consider the event $A_n := \{\max_{0\leq k\leq n-1}|\xi_k| \leq n^{\frac{1}{2}-2\varepsilon}\}$, where $\xi_k = \theta(g_{k+1}, X_k) - \overline{P}\theta(g_k, X_{k-1})$; then

$$\mathbb{E}_{x,a}\Big[|M_T|; T \leq n\Big] = \mathbb{E}_{x,a}\Big[|M_T|; T \leq n, A_n\Big] + \mathbb{E}_{x,a}\Big[|M_T|; T \leq n, A_n^c\Big]. \quad (3.1)$$

On the event $[T \leq n] \cap A_n$, we have $|M_T| \leq |\xi_{T-1}| \leq n^{\frac{1}{2}-2\varepsilon}$. Hence for any $x \in \mathbb{X}$ and $a \geq n^{\frac{1}{2}-\varepsilon}$,

$$\mathbb{E}_{x,a}\Big[|M_T|; T \leq n, A_n\Big] \leq \mathbb{E}_{x,a}\Big[|\xi_{T-1}|; T \leq n, A_n\Big] \leq n^{\frac{1}{2}-2\varepsilon} \leq \frac{a}{n^{\varepsilon}}. \tag{3.2}$$

Let $M_n^* := \max\limits_{1\leq k\leq n} |M_k|$; since $|M_T| \leq M_n^*$ on the event $[T \leq n]$, it is clear that, for any $x \in \mathbb{X}$ and $a \geq 0$,

$$\begin{aligned}
\mathbb{E}_{x,a}\Big[|M_T|; T \leq n, A_n^c\Big] &\leq \mathbb{E}_x[M_n^*; A_n^c] \\
&\leq \mathbb{E}_{x,a}\Big[M_n^*; M_n^* > n^{\frac{1}{2}+2\varepsilon}, A_n^c\Big] + n^{\frac{1}{2}+2\varepsilon}\mathbb{P}_{x,a}(A_n^c) \\
&\leq \int_{n^{\frac{1}{2}+2\varepsilon}}^{+\infty} \mathbb{P}_{x,a}(M_n^* > t)dt + 2n^{\frac{1}{2}+2\varepsilon}\mathbb{P}_{x,a}(A_n^c). \qquad (3.3)
\end{aligned}$$

We bound the probability $\mathbb{P}_{x,a}(A_n^c)$ by using Markov's inequality, martingale definition and (2.14) as follows: for any $p \geq 1$,

$$\begin{aligned}
\mathbb{P}_{x,a}(A_n^c) &\leq \sum_{k=0}^{n-1} \mathbb{P}_{x,a}\left(|\xi_k| > n^{\frac{1}{2}-2\varepsilon}\right) \\
&\leq \frac{1}{n^{(\frac{1}{2}-2\varepsilon)p}} \sum_{k=0}^{n-1} \mathbb{E}_{x,a}|\xi_k|^p \\
&\leq \frac{2^p}{n^{(\frac{1}{2}-2\varepsilon)p}} \sum_{k=0}^{n-1} \mathbb{E}_{x,a}|\theta(g_{k+1}, X_k)|^p \\
&= \frac{c_p}{n^{\frac{p}{2}-1-2\varepsilon p}}.
\end{aligned}$$

For any $a \geq n^{\frac{1}{2}-\varepsilon}$, it follows that

$$n^{\frac{1}{2}+2\varepsilon}\mathbb{P}_{x,a}(A_n^c) \leq an^{3\varepsilon}\mathbb{P}_{x,a}(A_n^c) \leq \frac{c_p a}{n^{\frac{p}{2}-1-2\varepsilon p-3\varepsilon}}. \tag{3.4}$$

Now we control the integral in (3.3). Using Doob's maximal inequality for martingales and (2.17), we receive for any $p \geq 1$,

$$\mathbb{P}_x(M_n^* > t) \leq \frac{1}{t^p}\mathbb{E}_x\Big[|M_n|^p\Big] \leq c_p \frac{n^{\frac{p}{2}}}{t^p},$$

which implies for any $a \geq n^{\frac{1}{2}-\varepsilon}$,

$$\int_{n^{\frac{1}{2}+2\varepsilon}}^{+\infty} \mathbb{P}_x(M_n^* > t)dt \leq \frac{c_p}{p-1}\frac{n^{\frac{p}{2}}}{n^{(\frac{1}{2}+2\varepsilon)(p-1)}} \leq \frac{c_p}{p-1}\frac{a}{n^{2\varepsilon p-3\varepsilon}}. \tag{3.5}$$

Taking (3.3), (3.4) and (3.5) altogether, we obtain for some c_p',

$$\mathbb{E}_{x,a}\Big[|M_T|; T \leq n, A_n^c\Big] \leq c_p'\left(\frac{a}{n^{2\varepsilon p-3\varepsilon}} + \frac{a}{n^{\frac{p}{2}-1-2\varepsilon p-3\varepsilon}}\right). \tag{3.6}$$

Finally, from (3.1), (3.2) and (3.6), we obtain for any $a \geq n^{\frac{1}{2}-\varepsilon}$,

$$\mathbb{E}_{x,a}\Big[|M_T|; T \leq n\Big] \leq \frac{a}{n^{\varepsilon}} + c_p'\frac{a}{n^{\varepsilon}}\left(\frac{1}{n^{2\varepsilon p-4\varepsilon}} + \frac{1}{n^{\frac{p}{2}-1-2\varepsilon p-4\varepsilon}}\right).$$

Fix $p > 2$. Then there exist $c > 0$ and $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$ and $a \geq n^{\frac{1}{2}-\varepsilon}$,

$$\mathbb{E}_{x,a}\Big[|M_T|, T \leq n\Big] \leq c\frac{a}{n^{\varepsilon}}$$

which proves the lemma.

□

For fixed $\varepsilon > 0$ and $a \geq 0$, we consider the first time $\nu_{n,\varepsilon}$ when the process $(|M_k|)_{k\geq 1}$ exceeds $2n^{\frac{1}{2}-\varepsilon}$. Let $(B_t)_{t\geq 0}$ be a standard Brownian motion on the probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ and for any $a \geq 0$, let τ_a^{bm} be the first time the process $(a+\sigma B_t)_{t\geq 0}$ becomes non-positive. It is connected to Lemma 4.3 where $\mathbb{P}(\tau_a^{bm} > n)$ is controlled uniformly in a under condition $a \leq \theta_n\sqrt{n}$ with $\lim_{n\to+\infty}\theta_n = 0$ which we take into account here by setting

$$\nu_{n,\varepsilon} := \min\{k \geq 1 : |M_k| \geq 2n^{\frac{1}{2}-\varepsilon}\}.$$

Notice first that for any $\varepsilon > 0, x \in \mathbb{X}$ and $a \geq 0$ the sequence $(\nu_{n,\varepsilon})_{n\geq 1}$ tends to $+\infty$ a.s. on $(\Omega, \mathcal{B}(\Omega), \mathbb{P}_{x,a})$. The following lemma yields a more precise control of this property.

Lemma 3.2. *For any $\varepsilon \in (0, \frac{1}{2})$, there exists $c_\varepsilon > 0$ such that for any $x \in \mathbb{X}$, $a \geq 0$ and $n \geq 1$,*

$$\mathbb{P}_{x,a}(\nu_{n,\varepsilon} > n^{1-\varepsilon}) \leq \exp(-c_\varepsilon n^{\varepsilon}).$$

Proof. Let $m = [B^2 n^{1-2\varepsilon}]$ and $K = [n^{\varepsilon}/B^2]$ for some positive constant B. By (2.16), for n sufficiently great such that $A \leq n^{\frac{1}{2}-\varepsilon}$, we obtain for any $x \in \mathbb{X}$ and $a \geq 0$,

$$\begin{aligned}
\mathbb{P}_{x,a}(\nu_{n,\varepsilon} > n^{1-\varepsilon}) &\leq \mathbb{P}_{x,a}\left(\max_{1\leq k\leq n^{1-\varepsilon}} |M_k| \leq 2n^{\frac{1}{2}-\varepsilon}\right) \\
&\leq \mathbb{P}_{x,a}\left(\max_{1\leq k\leq K} |M_{km}| \leq 2n^{\frac{1}{2}-\varepsilon}\right) \\
&\leq \mathbb{P}_{x,a}\left(\max_{1\leq k\leq K} |S_{km}| \leq 3n^{\frac{1}{2}-\varepsilon}\right). \qquad (3.7)
\end{aligned}$$

Using Markov property, it follows that, for any $x \in \mathbb{X}$ and $a \geq 0$, from which by iterating K times, we obtain

$$\mathbb{P}_{x,a}\left(\max_{1\leq k\leq K} |S_{km}| \leq 3n^{\frac{1}{2}-\varepsilon}\right) \leq \left(\sup_{b\in\mathbb{R}, x\in\mathbb{X}} \mathbb{P}_{x,b}\left(|S_m| \leq 3n^{\frac{1}{2}-\varepsilon}\right)\right)^K. \qquad (3.8)$$

Denote $\mathbb{B}(b; r) = \{c : |b + c| \leq r\}$. Then for any $x \in \mathbb{X}$ and $b \in \mathbb{R}$

$$\mathbb{P}_{x,b}\left(|S_m| \leq 3n^{\frac{1}{2}-\varepsilon}\right) = \mathbb{P}_x\left(\frac{S_m}{\sqrt{m}} \in \mathbb{B}(b/\sqrt{m}; r_n)\right),$$

where $r_n = \frac{3n^{\frac{1}{2}-\varepsilon}}{\sqrt{m}}$. By the central limit theorem for $(S_n)_n$ (Theorem 5.1 property iii) Bougerol and Lacroix (1985)), we obtain as $m \to +\infty$,

$$\sup_{b\in\mathbb{R}, x\in\mathbb{X}} \left|\mathbb{P}_x\left(\frac{S_m}{\sqrt{m}} \in \mathbb{B}(b/\sqrt{m}; r_n)\right) - \int_{\mathbb{B}(b/\sqrt{m}; r_n)} \phi_{\sigma^2}(u)du\right| \to 0,$$

where $\phi_{\sigma^2}(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$ is the normal density of mean 0 and variance σ^2 on $\mathbb{R}$. Since $r_n \leq c_1 B^{-1}$ for some constant $c_1 > 0$, we obtain

$$\sup_{b\in\mathbb{R}}\int_{\mathbb{B}(b/\sqrt{m};r_n)}\phi_{\sigma^2}(u)du \le \int_{-r_n}^{r_n}\phi_{\sigma^2}(u)du \le \frac{2r_n}{\sigma\sqrt{2\pi}} \le \frac{2c_1}{B\sigma\sqrt{2\pi}}.$$

Choosing B and n great enough, for some $q_\varepsilon < 1$, we obtain

$$\sup_{b\in\mathbb{R},x\in\mathbb{X}}\mathbb{P}_{x,b}\left(|S_m| \le 3n^{\frac{1}{2}-\varepsilon}\right) \le \sup_{b\in\mathbb{R}}\int_{\mathbb{B}(b/\sqrt{m};r_n)}\phi_{\sigma^2}(u)du + o(1) \le q_\varepsilon.$$

Implementing this bound in (3.8) and using (3.7), it follows that

$$\sup_{a>0,x\in\mathbb{X}}\mathbb{P}_{x,a}(\nu_{n,\varepsilon} > n^{1-\varepsilon}) \le q_\varepsilon^K \le q_\varepsilon^{\frac{n^\varepsilon}{B^2}-1} \le ce^{-c_\varepsilon n^\varepsilon}$$

for some positive constants c and c_ε.

□

Lemma 3.3. *There exists $c > 0$ such that for any $\varepsilon \in (0, \frac{1}{2})$, $x \in \mathbb{X}$, $a \ge 0$ and $n \ge 1$,*

$$\sup_{1\le k\le n}\mathbb{E}_{x,a}[|M_k|;\nu_{n,\varepsilon} > n^{1-\varepsilon}] \le c(1+a)\exp(-c_\varepsilon n^\varepsilon)$$

for some positive constant c_ε which only depends on ε.

Proof. By Cauchy-Schwartz inequality, for any $x \in \mathbb{X}$, $a \ge 0$ and $1 \le k \le n$,

$$\mathbb{E}_{x,a}\left[|M_k|;\nu_{n,\varepsilon} > n^{1-\varepsilon}\right] \le \sqrt{\mathbb{E}_{x,a}|M_k|^2\mathbb{P}_{x,a}(\nu_{n,\varepsilon} > n^{1-\varepsilon})}.$$

By Minkowsky's inequality, (2.16) and the fact that $\frac{1}{n}\mathbb{E}_x|S_n|^2 \to \sigma^2$ as $n \to +\infty$, it yields

$$\sqrt{\mathbb{E}_{x,a}|M_k|^2} \le a + A + \sqrt{\mathbb{E}_{x,0}|S_k^2|} \le a + A + c\sqrt{k} \le a + A + c\sqrt{n}$$

for some $c > 0$ which does not depend on x. The claim follows by Lemma 3.2.

□

Lemma 3.4. *There exists $c > 0$ and $\varepsilon_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$, $x \in \mathbb{X}$, $a \ge 0$ and $n \ge 1$,*

$$\mathbb{E}_{x,a}[M_n;T > n] \le c(1+a). \tag{3.9}$$

and

$$\lim_{a\to+\infty}\frac{1}{a}\lim_{n\to+\infty}\mathbb{E}_{x,a}[M_n;T > n] = 1. \tag{3.10}$$

Proof. **(1)** On one hand, we claim

$$\mathbb{E}_{x,a}[M_n;T > n,\nu_{n,\varepsilon} \le n^{1-\varepsilon}] \le \left(1 + \frac{c'_\varepsilon}{n^\varepsilon}\right)\mathbb{E}_{x,a}\left[M_{[n^{1-\varepsilon}]};T > [n^{1-\varepsilon}]\right] \tag{3.11}$$

and delay the proof of (3.11) at the end of the first part. On the other hand, by Lemma 3.3, there exists $c > 0$ such that for any $\varepsilon \in (0, \frac{1}{2})$, $x \in \mathbb{X}$, $a \ge 0$ and $n \ge 1$,

$$\begin{aligned}\mathbb{E}_{x,a}[M_n;T > n,\nu_{n,\varepsilon} > n^{1-\varepsilon}] &\le \sup_{1\le k\le n}\mathbb{E}_{x,a}\left[|M_k|;\nu_{n,\varepsilon} > n^{1-\varepsilon}\right] \\ &\le c(1+a)\exp(-c_\varepsilon n^\varepsilon).\end{aligned} \tag{3.12}$$

Hence combining (3.11) and (3.12), we obtain for any $x \in \mathbb{X}$ and $a \ge 0$,

$$\mathbb{E}_{x,a}[M_n;T > n] \le \left(1 + \frac{c'_\varepsilon}{n^\varepsilon}\right)\mathbb{E}_{x,a}\left[M_{[n^{1-\varepsilon}]};T > [n^{1-\varepsilon}]\right] + c(1+a)\exp(-c_\varepsilon n^\varepsilon). \tag{3.13}$$

Let $k_j := \left[n^{(1-\varepsilon)^j}\right]$ for $j \geq 0$. Notice that $k_0 = n$ and $[k_j^{1-\varepsilon}] \leq k_{j+1}$ for any $j \geq 0$. Since the sequence $((M_n)\mathbf{1}_{[T>n]})_{n\geq 1}$ is a submartingale, by using the bound (3.13), it yields

$$\begin{aligned}\mathbb{E}_{x,a}[M_{k_1}; T > k_1] &\leq \left(1 + \frac{c'_\varepsilon}{{k_1}^\varepsilon}\right) \mathbb{E}_{x,a}\left[M_{[{k_1}^{1-\varepsilon}]}; T > [{k_1}^{1-\varepsilon}]\right] + c(1+a)\exp(-c_\varepsilon k_1^\varepsilon) \\ &\leq \left(1 + \frac{c'_\varepsilon}{{k_1}^\varepsilon}\right) \mathbb{E}_{x,a}[M_{k_2}; T > k_2] + c(1+a)\exp(-c_\varepsilon k_1^\varepsilon).\end{aligned}$$

Let n_0 be a constant and $m = m(n)$ such that $k_m = \left[n^{(1-\varepsilon)^m}\right] \leq n_0$. After m iterations, we obtain

$$\mathbb{E}_{x,a}[M_n; T > n] \leq A_m\Big(\mathbb{E}_{x,a}[M_{k_m}; T > k_m] + c(1+a)B_m\Big), \tag{3.14}$$

where

$$A_m = \prod_{j=1}^{m}\left(1 + \frac{c'_\varepsilon}{k_{j-1}^\varepsilon}\right) \leq \exp\Big(2^\varepsilon c'_\varepsilon \frac{n_0^{-\varepsilon}}{1 - n_0^{-\varepsilon^2}}\Big), \tag{3.15}$$

and

$$B_m = \sum_{j=1}^{m} \frac{\exp\left(-c_\varepsilon k_{j-1}^\varepsilon\right)}{\left(1 + \frac{c'_\varepsilon}{k_{j-1}^\varepsilon}\right)\ldots\left(1 + \frac{c'_\varepsilon}{k_m^\varepsilon}\right)} \leq c_1 \frac{n_0^{-\varepsilon}}{1 - n_0^{-\varepsilon^2}} \tag{3.16}$$

from Lemma 5.6 in Grama et al. (2014). By choosing n_0 sufficient great, the first assertion of the lemma follows from (3.14), (3.15) and (3.16) taking into account that

$$\mathbb{E}_{x,a}[M_{k_m}; T > k_m] \leq \mathbb{E}_{x,a}[M_{n_0}; T > n_0] \leq \mathbb{E}_{x,a}|M_{n_0}| \leq a + c.$$

Before proving (3.11), we can see that there exist $c > 0$ and $0 < \varepsilon_0 < \frac{1}{2}$ such that for any $\varepsilon \in (0, \varepsilon_0)$, $x \in \mathbb{X}$ and $b \geq n^{\frac{1}{2}-\varepsilon}$,

$$\mathbb{E}_{x,b}[M_n; T > n] \leq \left(1 + \frac{c}{n^\varepsilon}\right) b. \tag{3.17}$$

Indeed, since $(M_n, \mathcal{F}_n)_{n\geq 1}$ is a $\mathbb{P}_{x,b}$- martingale, we obtain

$$\mathbb{E}_{x,b}[M_n; T \leq n] = \mathbb{E}_{x,b}[M_T; T \leq n]$$

and thus

$$\begin{aligned}\mathbb{E}_{x,b}[M_n; T > n] &= \mathbb{E}_{x,b}[M_n] - \mathbb{E}_{x,a}[M_n; T \leq n] \\ &= b - \mathbb{E}_{x,b}[M_T; T \leq n] \\ &= b + \mathbb{E}_{x,b}[|M_T|; T \leq n].\end{aligned} \tag{3.18}$$

Hence (3.17) arrives by using Lemma 3.1. For (3.11), it is obvious that

$$\mathbb{E}_{x,a}\Big[M_n; T > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\Big] = \sum_{k=1}^{[n^{1-\varepsilon}]} \mathbb{E}_{x,a}\Big[M_n; T > n, \nu_{n,\varepsilon} = k\Big]. \tag{3.19}$$

Denote $U_m(x,a) := \mathbb{E}_{x,a}[M_m; T > m]$. For any $m \geq 1$, by the Markov property applied to $(X_n)_{n\geq 1}$, it follows that

$$\begin{aligned}\mathbb{E}_{x,a}\Big[M_n; T > n, \nu_{n,\varepsilon} = k\Big] &= \int \mathbb{E}_{y,b}[M_{n-k}; T > n-k] \\ &\qquad \mathbb{P}_{x,a}(X_k \in dy, M_k \in db; T > k, \nu_{n,\varepsilon} = k) \\ &= \mathbb{E}_{x,a}\Big[U_{n-k}(X_k, M_k); T > k, \nu_{n,\varepsilon} = k\Big]. \qquad (3.20)\end{aligned}$$

From the definition of $\nu_{n,\varepsilon}$, we can see that $[\nu_{n,\varepsilon} = k] \subset \Big[|M_k| \geq n^{\frac{1}{2}-\varepsilon}\Big]$, and by using (3.17), on the event $[T > k, \nu_{n,\varepsilon} = k]$ we have $U_{n-k}(X_k, M_k) \leq \Big(1 + \frac{c}{(n-k)^\varepsilon}\Big) M_k$. Therefore (3.20) becomes

$$\mathbb{E}_{x,a}[M_n; T > n, \nu_{n,\varepsilon} = k] \leq \left(1 + \frac{c}{(n-k)^\varepsilon}\right) \mathbb{E}_{x,a}[M_k; T > k, \nu_{n,\varepsilon} = k]. \qquad (3.21)$$

Combining (3.19) and (3.21), it follows that, for n sufficiently great,

$$\begin{aligned}\mathbb{E}_{x,a}[M_n; T > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}] &\leq \sum_{k=1}^{[n^{1-\varepsilon}]} \left(1 + \frac{c}{(n-k)^\varepsilon}\right) \mathbb{E}_{x,a}[M_k; T > k, \nu_{n,\varepsilon} = k] \\ &\leq \left(1 + \frac{c'_\varepsilon}{n^\varepsilon}\right) \sum_{k=1}^{[n^{1-\varepsilon}]} \mathbb{E}_{x,a}[M_k; T > k, \nu_{n,\varepsilon} = k],\end{aligned}$$

for some constant $c'_\varepsilon > 0$. Since $(M_n \mathbf{1}_{[T>n]})_{n\geq 1}$ is a submartingale, for any $x \in \mathbb{X}$, $a \geq 0$ and $1 \leq k \leq [n^{1-\varepsilon}]$,

$$\mathbb{E}_{x,a}[M_k; T > k, \nu_{n,\varepsilon} = k] \leq \mathbb{E}_{x,a}\Big[M_{[n^{1-\varepsilon}]}; T > [n^{1-\varepsilon}], \nu_{n,\varepsilon} = k\Big].$$

This implies

$$\begin{aligned}\mathbb{E}_{x,a}[M_n; T > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}] &\leq \left(1 + \frac{c'_\varepsilon}{n^\varepsilon}\right) \sum_{k=1}^{[n^{1-\varepsilon}]} \mathbb{E}_{x,a}\left[M_{[n^{1-\varepsilon}]}; T > [n^{1-\varepsilon}], \nu_{n,\varepsilon} = k\right] \\ &\leq \left(1 + \frac{c'_\varepsilon}{n^\varepsilon}\right) \mathbb{E}_{x,a}\left[M_{[n^{1-\varepsilon}]}; T > [n^{1-\varepsilon}]\right].\end{aligned}$$

(2) Let $\delta > 0$. From (3.15) and (3.16), by choosing n_0 sufficiently great, we obtain $A_m \leq 1 + \delta$ and $B_m \leq \delta$. Together with (3.14), since $(M_n \mathbf{1}_{[T>n]})_{n\geq 1}$ is a submartingale, we obtain for $k_m \leq n_0$,

$$\mathbb{E}_{x,a}[M_n; T > n] \leq (1+\delta)\Big(\mathbb{E}_{x,a}[M_{n_0}; T > n_0] + c(1+a)\delta\Big).$$

Moreover, the sequence $\mathbb{E}_{x,a}[M_n; T > n]$ is increasing, thus it converges $\mathbb{P}_{x,a}$-a.s. and

$$\lim_{n\to+\infty} \mathbb{E}_{x,a}[M_n; T > n] \leq (1+\delta)\Big(\mathbb{E}_{x,a}[M_{n_0}; T > n_0] + c(1+a)\delta\Big).$$

By using (3.18), we obtain

$$a \leq \lim_{n\to+\infty} \mathbb{E}_{x,a}[M_n; T > n] \leq (1+\delta)\Big(a + \mathbb{E}_x|M_{n_0}| + c(1+a)\delta\Big).$$

Hence the assertion follows since $\delta > 0$ is arbitrary.

□

3.2. *On the stopping time τ.*
We now state some useful properties of τ and S_τ. Denote $\tau_a := \min\{n \geq 1 : S_n \leq -a\}$ and $T_a := \min\{n \geq 1 : M_n \leq -a\}$ for any $a \geq 0$.

Lemma 3.5. *There exists $c > 0$ such that for any $x \in \mathbb{X}$, $a \geq 0$ and $n \geq 1$,*

$$\mathbb{E}_{x,a}[S_n, \tau > n] \leq c(1+a).$$

Proof. (2.16) yields $\mathbb{P}_x(\tau_a \leq T_{a+A}) = 1$ and $A + M_n \geq S_n > 0$ on the event $[\tau > n]$. By (3.9), it follows that

$$\begin{aligned} \mathbb{E}_{x,a}[S_n; \tau > n] &\leq \mathbb{E}_{x,a}[A + M_n; \tau > n] \\ &\leq \mathbb{E}_{x,a+A}[M_n; T > n] \\ &\leq c_1(1 + a + A) \leq c_2(1+a). \end{aligned}$$

□

Proposition 3.6. *There exists $c > 0$ such that for any $x \in \mathbb{X}$ and $a \geq 0$,*

$$\mathbb{E}_{x,a}|S_\tau| \leq c(1+a) < +\infty$$

and

$$\mathbb{E}_{x,a}|M_\tau| \leq c(1+a) < +\infty. \tag{3.22}$$

Proof. By (2.16), since $(M_n)_n$ is a martingale, we can see that

$$\begin{aligned} -\mathbb{E}_{x,a}[S_\tau; \tau \leq n] &\leq -\mathbb{E}_{x,a}[M_\tau; \tau \leq n] + A \\ &= \mathbb{E}_{x,a}[M_n; \tau > n] - \mathbb{E}_{x,a}[M_n] + A \\ &\leq \mathbb{E}_{x,a}[S_n; \tau > n] + 2A. \end{aligned}$$

Hence by Lemma 3.5, for any $x \in \mathbb{X}$ and $a \geq 0$,

$$\begin{aligned} \mathbb{E}_{x,a}\left[|S_\tau|; \tau \leq n\right] &\leq \mathbb{E}_{x,a}\left|S_{\tau \wedge n}\right| \\ &= \mathbb{E}_{x,a}\left[S_n; \tau > n\right] - \mathbb{E}_{x,a}\left[S_\tau; \tau \leq n\right] \\ &\leq 2\mathbb{E}_{x,a}[S_n; \tau > n] + 2A \\ &\leq c(1+a) + 2A. \end{aligned}$$

By Lebesgue's Monotone Convergence Theorem, it yields

$$\mathbb{E}_{x,a}|S_\tau| = \lim_{n \to +\infty} \mathbb{E}_{x,a}\left[|S_\tau|; \tau \leq n\right] \leq c(1+a) + 2A < +\infty.$$

By (2.16), the second assertion arrives.

□

3.3. *Proof of Proposition 1.1.*
By definition of τ_a, we can see that $\mathbb{E}_{x,a}M_\tau = a + \mathbb{E}_x M_{\tau_a}$ and $\mathbb{P}_{x,a}(\tau > n) = \mathbb{P}_x(\tau_a > n)$.

(1) By (3.22) and Lebesgue's Dominated Convergence Theorem, for any $x \in \mathbb{X}$ and $a \geq 0$,

$$\lim_{n \to +\infty} \mathbb{E}_{x,a}[M_\tau; \tau \leq n] = \mathbb{E}_{x,a}M_\tau = a - V(x,a),$$

where $V(x,a)$ is the quantity defined by: for $x \in \mathbb{X}$ and $a \in \mathbb{R}$,

$$V(x,a) := \begin{cases} -\mathbb{E}_x M_{\tau_a} & \text{if} \quad a \geq 0, \\ 0 & \text{if} \quad a < 0. \end{cases}$$

Since $(M_n, \mathcal{F}_n)_{n\geq 1}$ is a $\mathbb{P}_{x,a}$-martingale,

$$\mathbb{E}_{x,a}[M_n;\tau > n] \quad = \quad \mathbb{E}_{x,a}M_n - \mathbb{E}_{x,a}[M_n;\tau \leq n] = a - \mathbb{E}_{x,a}[M_\tau;\tau \leq n], \tag{3.23}$$

which implies

$$\lim_{n\to+\infty} \mathbb{E}_{x,a}[M_n;\tau > n] = V(x,a).$$

Since $|S_n - M_n| \leq A$ $\mathbb{P}_x$-a.s. and $\lim_{n\to+\infty} \mathbb{P}_{x,a}(\tau > n) = 0$, it follows that

$$\lim_{n\to+\infty} \mathbb{E}_{x,a}[S_n;\tau > n] = \lim_{n\to+\infty} \mathbb{E}_{x,a}[M_n;\tau > n] = V(x,a).$$

(2) The assertion arrives by taking into account that $0 \leq a \leq a'$ implies $\tau_a \leq \tau_{a'}$ and

$$\mathbb{E}_x[a + S_n;\tau_a > n] \leq \mathbb{E}_x[a' + S_n;\tau_{a'} > n].$$

(3) Lemma 3.5 and assertion 1 imply that $V(x,a) \leq c(1+a)$ for any $x \in \mathbb{X}$ and $a \geq 0$. Besides, (3.23) and (2.16) yield

$$\mathbb{E}_{x,a}[M_n;\tau > n] \geq a - \mathbb{E}_{x,a}[S_\tau;\tau \leq n] - A \geq a - A,$$

which implies

$$V(x,a) \geq a - A. \tag{3.24}$$

Now we prove $V(x,a) \geq 0$. Assertion 2 implies $V(x,0) \leq V(x,a)$ for any $x \in \mathbb{X}$ and $a \geq 0$. From P5, let $E_\delta := \{g \in S : \forall x \in \mathbb{X}, \log|gx| \geq \delta\}$ and choose a positive constant k such that $k\delta > 2A$. Hence, for any $g_1, \ldots, g_k \in E_\delta$ and any $x \in \mathbb{X}$, we obtain $\log|g_k \ldots g_1 x| \geq k\delta > 2A$. It yields

$$\begin{aligned}
V(x,0) &= \lim_{n\to+\infty} \mathbb{E}_x[S_n;\tau > n] \\
&\geq \liminf_{n\to+\infty} \int_{E_\delta} \ldots \int_{E_\delta} \mathbb{E}_{g_k\ldots g_1\cdot x, \log|g_k\ldots g_1 x|}[S_{n-k};\tau > n-k]\mu(dg_1)\ldots\mu(dg_k) \\
&\geq \liminf_{n\to+\infty} \int_{E_\delta} \ldots \int_{E_\delta} V(g_k\ldots g_1\cdot x, 2A)\mu(dg_1)\ldots\mu(dg_k) \\
&\geq A\Big(\mu(E_\delta)\Big)^k > 0,
\end{aligned}$$

where the last inequality comes from (3.24) by applying to $a = 2A$.

(4) Equation (3.24) yields $\lim_{a\to+\infty} \dfrac{V(x,a)}{a} \geq 1$. By (2.16), it yields that $\mathbb{P}_x(\tau_a < T_{A+a}) = 1$, which implies

$$\mathbb{E}_{x,a}[S_n;\tau > n] \leq \mathbb{E}_{x,a}[A+M_n;\tau > n] \leq \mathbb{E}_{x,a}[A+M_n;T_A > n] = \mathbb{E}_{x,a+A}[M_n;T > n].$$

From (3.10), we obtain $\lim_{n\to+\infty} \dfrac{V(x,a)}{a} \leq 1$.

(5) For any $x \in \mathbb{X}$, $a \geq 0$ and $n \geq 1$, we set $V_n(x,a) := \mathbb{E}_{x,a}[S_n;\tau > n]$. By assertion 1, we can see $\lim_{n\to+\infty} V_n(x,a) = V(x,a)$. By Markov property, we obtain

$$\begin{aligned}
V_{n+1}(x,a) &= \mathbb{E}_{x,a}\left[\mathbb{E}\left[S_1 + \sum_{k=1}^{n} \rho(g_{k+1}, X_k); S_1 > 0, \ldots, S_{n+1} > 0 | \mathcal{F}_1\right]\right] \\
&= \mathbb{E}_{x,a}\left[V_n(X_1, S_1);\tau > 1\right].
\end{aligned}$$

By Lemma 3.5, we obtain $\sup_{x\in\mathbb{X}, a\geq 0} V_n(x,a) \leq c(1+a)$ which implies $\mathbb{P}$-a.s.

$$V_n(X_1, S_1)\mathbf{1}_{[\tau>1]} \leq c(1+S_1)\mathbf{1}_{[\tau>1]}.$$

Lebesgue's Dominated Convergence Theorem and (1.2) yield

$$\begin{aligned} V(x,a) = \lim_{n\to+\infty} V_{n+1}(x,a) &= \lim_{n\to+\infty} \mathbb{E}_{x,a}[V_n(X_1,S_1);\tau>1] \\ &= \mathbb{E}_{x,a}[V(X_1,S_1);\tau>1] \\ &= \widetilde{P}_+ V(x,a). \end{aligned}$$

□

4. Coupling argument and proof of Theorems 1.2 and 1.3

Firstly, we check that the weak invariance principle with rate, Theorem 2.1 in Grama et al. (2014), may be applied. The hypotheses C1, C2 and C3 of this theorem are given in terms of Fourier transform of the partial sums of S_n; combining the expressions (2.1), (2.2), (2.3) and the properties of the Fourier operators $(P_t)_t$, we verify in the next section that these conditions are satisfied in our context. This leads to the following simpler but sufficient statement.

Theorem 4.1. *Assume P1-P4. There exist*

- $\varepsilon_0 > 0$, *and* $c_0 > 0$,
- *a probability space* $(\widetilde{\Omega}, B(\widetilde{\Omega}))$,
- *a family* $(\widetilde{\mathbb{P}}_x)_{x\in\mathbb{X}}$ *of probability measures on* $(\widetilde{\Omega}, B(\widetilde{\Omega}))$,
- *a sequence* $(\tilde{a}_k)_k$ *of real-valued random variables on* $(\widetilde{\Omega}, B(\widetilde{\Omega}))$ *such that* $\mathcal{L}\Big((\tilde{a}_k)_k/\widetilde{\mathbb{P}}_x\Big) = \mathcal{L}\Big((a_k)_k/\mathbb{P}_x\Big)$ *for any* $x \in \mathbb{X}$,
- *and a sequence* $(\widetilde{W}_i)_{i\geq 1}$ *of independent standard normal random variables on* $(\widetilde{\Omega}, \mathcal{B}(\widetilde{\Omega}))$

such that for any $x \in \mathbb{X}$,

$$\widetilde{\mathbb{P}_x}\left(\sup_{1\leq k\leq n}\left|\sum_{i=1}^{k}(\tilde{a}_i - \sigma\widetilde{W}_i)\right| > n^{\frac{1}{2}-\varepsilon_0}\right) \leq c_0 n^{-\varepsilon_0}. \tag{4.1}$$

In order to simplify the notations, we identify $(\widetilde{\Omega}, \mathcal{B}(\tilde{\Omega}))$ and $(\Omega, \mathcal{B}(\Omega))$. Notice that if (4.1) holds true for ε_0 then it also holds true true for any $\varepsilon \leq \varepsilon_0$.

Under hypotheses **P1**–**P4**, we may apply this Theorem to the sequence $(a_k)_{k\geq 0} = (\rho(g_k, X_{k-1}))_{k\geq 0}$. Thus, the process $(\log|L_n x|)_{n\geq 0}$ satisfies the following property: there exists $\varepsilon_0 > 0$ and $c_0 > 0$ such that for any $\varepsilon \in (0, \varepsilon_0]$ and $x \in \mathbb{X}$,

$$\begin{aligned} &\mathbb{P}\left(\sup_{0\leq t\leq 1} |\log|L_{[nt]}x| - \sigma B_{nt}| > n^{\frac{1}{2}-\varepsilon}\right) \\ &= \mathbb{P}_x\left(\sup_{0\leq t\leq 1} |S_{[nt]} - \sigma B_{nt}| > n^{\frac{1}{2}-\varepsilon}\right) \leq c_0 n^{-\varepsilon}, \end{aligned} \tag{4.2}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion on the probability space $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ and $\sigma > 0$ is the variance of the sequence $(S_n)_{n\geq 0}$. For any $a \geq 0$, let τ_a^{bm} be the first time the process $(a + \sigma B_t)_{t\geq 0}$ becomes non-positive:

$$\tau_a^{bm} = \inf\{t \geq 0 : a + \sigma B_t \leq 0\}.$$

The following lemma is due to Levy Levy (1937) (Theorem 42.I, pp. 194-195).

Lemma 4.2. (1) *For any $a \geq 0$ and $n \geq 1$,*

$$\mathbb{P}(\tau_a^{bm} > n) = \mathbb{P}\left(\sigma \inf_{0\leq u\leq n} B_u > -a\right) = \frac{2}{\sigma\sqrt{2\pi n}} \int_0^a \exp\left(-\frac{s^2}{2n\sigma^2}\right) ds.$$

(2) *For any a, b such that $0 \leq a < b < +\infty$ and $n \geq 1$,*

$$\mathbb{P}(\tau_a^{bm} > n, a + \sigma B_n \in [a, b]) = \frac{1}{\sigma\sqrt{2\pi n}} \int_a^b \left[\exp\left(-\frac{(s-a)^2}{2n\sigma^2}\right) - \exp\left(-\frac{(s+a)^2}{2n\sigma^2}\right)\right] ds.$$

From Lemma 4.2, we can obtain the next result.

Lemma 4.3. (1) *There exists a positive constant c such that for any $a \geq 0$ and $n \geq 1$,*

$$\mathbb{P}(\tau_a^{bm} > n) \leq c\frac{a}{\sigma\sqrt{n}}. \tag{4.3}$$

(2) *For any sequence of real numbers $(\alpha_n)_n$ such that $\alpha_n \to 0$ as $n \to +\infty$, there exists a positive constant c such that for any $a \in [0, \alpha_n\sqrt{n}]$,*

$$\left|\mathbb{P}(\tau_a^{bm} > n) - \frac{2a}{\sigma\sqrt{2\pi n}}\right| \leq c\frac{\alpha_n}{\sqrt{n}}a. \tag{4.4}$$

We use the coupling result described in Theorem 4.1 above to transfer the properties of the exit time τ_a^{bm} to the exit time τ_a for great a.

4.1. *Proof of Theorem 1.2.*
(1) Let $\varepsilon \in (0, \min\{\varepsilon_0; \frac{1}{2}\})$ and $(\theta_n)_{n\geq 1}$ be a sequence of positive numbers such that $\theta_n \to 0$ and $\theta_n n^{\varepsilon/4} \to +\infty$ as $n \to +\infty$. For any $x \in \mathbb{X}$ and $a \geq 0$, we have the decomposition

$$P_n(x, a) := \mathbb{P}_{x,a}(\tau > n) = \mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} > n^{1-\varepsilon}) + \mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}). \tag{4.5}$$

It is obvious that from Lemma 3.2, we obtain

$$\sup_{x\in\mathbb{X}, a\geq 0} \mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} > n^{1-\varepsilon}) \leq \sup_{x\in\mathbb{X}, a\geq 0} \mathbb{P}_{x,a}(\nu_{n,\varepsilon} > n^{1-\varepsilon}) \leq e^{-c_\varepsilon n^\varepsilon}. \tag{4.6}$$

For the second term, by Markov's property,

$$\begin{aligned}\mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}) &= \mathbb{E}_{x,a}\left[P_{n-\nu_n}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}); \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] \\ &= I_n(x, a) + J_n(x, a),\end{aligned} \tag{4.7}$$

where

$$I_n(x, a) := \mathbb{E}_{x,a}\left[P_{n-\nu_n}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}); S_{\nu_{n,\varepsilon}} \leq \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right],$$

and $J_n(x, a) := \mathbb{E}_{x,a}\left[P_{n-\nu_n}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}); S_{\nu_{n,\varepsilon}} > \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right]$.

Now we control the quantity $P_{n-\nu_n}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}})$ by using the following lemma. The proofs of the lemmas stated in this subsection are postponed to the next subsection.

Lemma 4.4. (1) *There exists $c > 0$ such that for any n sufficiently great, $x \in \mathbb{X}$ and $a \in [n^{\frac{1}{2}-\varepsilon}, \theta_n n^{\frac{1}{2}}]$,*

$$\left|\mathbb{P}_{x,a}(\tau > n) - \frac{2a}{\sigma\sqrt{2\pi n}}\right| \leq c\frac{a\theta_n}{\sqrt{n}}. \tag{4.8}$$

(2) *There exists* $c > 0$ *such that for any* $x \in \mathbb{X}$, $a \geq n^{\frac{1}{2}-\varepsilon}$ *and* $n \geq 1$,

$$\mathbb{P}_{x,a}(\tau > n) \leq c\frac{a}{\sqrt{n}}. \tag{4.9}$$

Notice that for any $x \in \mathbb{X}, a \geq 0$ and $0 \leq k \leq n^{1-\varepsilon}$,

$$P_n(x,a) \leq P_{n-k}(x,a) \leq P_{n-[n^{1-\varepsilon}]}(x,a). \tag{4.10}$$

By definition of $\nu_{n,\varepsilon}$ and (2.16), as long as $A \leq n^{\frac{1}{2}-\varepsilon}$, we have $\mathbb{P}_{x,a}$-a.s.

$$S_{\nu_{n,\varepsilon}} \geq M_{\nu_{n,\varepsilon}} - A \geq 2n^{\frac{1}{2}-\varepsilon} - A \geq n^{\frac{1}{2}-\varepsilon}. \tag{4.11}$$

Using (4.8) and (4.10), (4.11) with θ_n replaced by $\theta_n \left(\frac{n}{n-n^{1-\varepsilon}}\right)^{\frac{1}{2}}$, for n sufficiently great, on the event $\left[S_{\nu_{n,\varepsilon}} \leq \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right]$, we obtain $\mathbb{P}_{x,a}$-a.s.

$$P_{n-\nu_{n,\varepsilon}}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}) = \frac{2(1+o(1))S_{\nu_{n,\varepsilon}}}{\sigma\sqrt{2\pi n}}.$$

Let

$$I'_n(x,a) := \mathbb{E}_{x,a}\left[S_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right], \tag{4.12}$$

$$J'_n(x,a) := \mathbb{E}_{x,a}\left[S_{\nu_{n,\varepsilon}}; S_{\nu_{n,\varepsilon}} > \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right]. \tag{4.13}$$

Hence

$$\begin{aligned} I_n(x,a) &= \frac{2(1+o(1))}{\sigma\sqrt{2\pi n}}\mathbb{E}_{x,a}\left[S_{\nu_{n,\varepsilon}}; S_{\nu_{n,\varepsilon}} \leq \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] \\ &= \frac{2(1+o(1))}{\sigma\sqrt{2\pi n}}\left[I'_n(x,a) - J'_n(x,a)\right], \\ J_n(x,a) &= \frac{c(1+o(1))}{\sqrt{n}}J'_n(x,a). \end{aligned}$$

Therefore (4.5) becomes

$$\left|\mathbb{P}_{x,a}(\tau > n) - \frac{2(1+o(1))}{\sigma\sqrt{2\pi n}}I'_n(x,a)\right| \leq C\left(n^{-\frac{1}{2}}J'_n(x,a)\right) + C'\left(e^{-c_\varepsilon n^\varepsilon}\right).$$

The first assertion of Theorem 1.2 immediately follows by noticing that the term J'_n is negligible and $\mathbb{P}_{x,a}(\tau > n)$ is dominated by the term I'_n as shown in the lemma below.

Lemma 4.5.

$$\lim_{n\to+\infty} I'_n(x,a) = V(x,a) \quad \textit{and} \quad \lim_{n\to+\infty} n^{2\varepsilon}J'_n = 0,$$

where I'_n *and* J'_n *are defined in (4.12) and (4.13).*

(2) By using Proposition 1.1 (3), it suffices to prove $\sqrt{n}\mathbb{P}_{x,a}(\tau > n) \leq c(1+a)$ for n great enough. For n sufficiently great, using (4.9) and (4.11), we obtain $\mathbb{P}_{x,a}$-a.s.

$$P_{n-[n^{1-\varepsilon}]}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}) \leq c\frac{S_{\nu_{n,\varepsilon}}}{\sqrt{n}}.$$

Combined with (4.7), it yields

$$\mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}) \leq \frac{c}{\sqrt{n}}I'_n. \tag{4.14}$$

Since $\tau_a < T_{a+A}$ $\mathbb{P}$-a.s. and (3.9), it follows that

$$I'_n(x,a) \leq \mathbb{E}_{x,a+A}[M_{\nu_{n,\varepsilon}}; T > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}] \leq c(1+a+A).$$

Hence (4.14) becomes

$$\mathbb{P}_{x,a}(\tau > n, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}) \leq \frac{c}{\sqrt{n}}(1+a+A). \tag{4.15}$$

Combining (4.5), (4.6) and (4.15), we obtain for n great enough,

$$\mathbb{P}_{x,a}(\tau > n) \leq e^{-c_\varepsilon n^\varepsilon} + \frac{c}{\sqrt{n}}(1+a+A) \leq c'(1+a).$$

□

4.2. *Proof of Theorem 1.3.*
Let us decompose $\mathbb{P}_{x,a}(S_n \leq t\sqrt{n} | \tau > n)$ as follows:

$$\frac{\mathbb{P}_{x,a}(S_n \leq t\sqrt{n}, \tau > n)}{\mathbb{P}_{x,a}(\tau > n)} = D_{n,1} + D_{n,2} + D_{n,3}, \tag{4.16}$$

where

$$\begin{aligned}
D_{n,1} &:= \frac{\mathbb{P}_{x,a}(S_n \leq t\sqrt{n}, \tau > n, \nu_{n,\varepsilon} > n^{1-\varepsilon})}{\mathbb{P}_{x,a}(\tau > n)}, \\
D_{n,2} &:= \frac{\mathbb{P}_{x,a}(S_n \leq t\sqrt{n}, \tau > n, S_n > \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon})}{\mathbb{P}_{x,a}(\tau > n)}, \\
D_{n,3} &:= \frac{\mathbb{P}_{x,a}(S_n \leq t\sqrt{n}, \tau > n, S_n \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon})}{\mathbb{P}_{x,a}(\tau > n)}.
\end{aligned}$$

Lemma 3.2 and Theorem 1.2 imply

$$\lim_{n\to+\infty} D_{n,1} = 0. \tag{4.17}$$

Theorem 1.2 and Proposition 1.1 (3) imply

$$\begin{aligned}
D_{n,2} &\leq \frac{\mathbb{P}_{x,a}(\tau > n, S_n > \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon})}{\mathbb{P}_{x,a}(\tau > n)} \\
&= \frac{1}{\mathbb{P}_{x,a}(\tau > n)} \mathbb{E}_{x,a}\left[P_{n-\nu_{n,\varepsilon}}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}); \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} > \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] \\
&\leq c\frac{\mathbb{E}_{x,a}\left[1 + S_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} > \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right]}{\mathbb{P}_{x,a}(\tau > n)\sigma\sqrt{n - n^{1-\varepsilon}}} \\
&\leq c'\frac{\mathbb{E}_{x,a}\left[S_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} > \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] + \mathbb{P}_{x,a}(\tau > \nu_{n,\varepsilon})}{V(x,a)\sqrt{1 - n^{-\varepsilon}}}.
\end{aligned}$$

Since $\mathbb{P}_{x,a}(\tau < +\infty) = 1$ and $\mathbb{P}_{x,a}(\nu_{n,\varepsilon} < +\infty) = 0$, Lemma 4.5 yields

$$\lim_{n\to+\infty} D_{n,2} = 0. \tag{4.18}$$

Now we control $D_{n,3}$. Let $H_m(x,a) := \mathbb{P}_{x,a}(S_m \leq t\sqrt{n}, \tau > m)$. We claim the following lemma and postpone its proof at the end of this section.

Lemma 4.6. *Let $\varepsilon \in (0, \varepsilon_0), t > 0$ and $(\theta_n)_{n\geq 1}$ be a sequence such that $\theta_n \to 0$ and $\theta_n n^{\varepsilon/4} \to +\infty$ as $n \to +\infty$. Then for any $x \in \mathbb{X}$, $n^{1/2-\varepsilon} \leq a \leq \theta_n\sqrt{n}$ and $1 \leq k \leq n^{1-\varepsilon}$,*

$$\mathbb{P}_{x,a}\Big(S_{n-k} \leq t\sqrt{n}, \tau > n-k\Big) = \frac{2a}{\sigma^3\sqrt{2\pi n}} \int_0^t u \exp\Big(-\frac{u^2}{2\sigma^2}\Big) du(1+o(1)).$$

It is noticeable that on the event $[\tau > k, S_k \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} = k]$, the random variable $H_{n-k}(X_k, S_k)$ satisfies the hypotheses of Lemma 4.6. Hence

$$\begin{aligned}
& \mathbb{P}_{x,a}(S_n \leq t\sqrt{n}, \tau > n, S_n \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}) \\
= \; & \mathbb{E}_{x,a}\Big[H_{n-\nu_{n,\varepsilon}}(X_{\nu_{n,\varepsilon}}, S_{\nu_{n,\varepsilon}}); \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\Big] \\
= \; & \sum_{k=1}^{[n^{1-\varepsilon}]} \mathbb{E}_{x,a}\Big[H_{n-k}(X_k, S_k); \tau > k, S_k \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} = k\Big] \\
= \; & \frac{2(1+o(1))}{\sigma^3\sqrt{2\pi n}} \int_0^t u \exp\left(\frac{-u^2}{2\sigma^2}\right) du\, \mathbb{E}_{x,a}\Big[S_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\Big].
\end{aligned}$$

Lemma 4.5 yield as $n \to +\infty$,

$$\mathbb{E}_{x,a}\Big[S_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, S_{\nu_{n,\varepsilon}} \leq \theta_n\sqrt{n}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\Big] = V(x,a)(1+o(1)).$$

Therefore, Theorem 1.2 yields

$$\begin{aligned}
D_{n,3} &= \frac{2V(x,a)(1+o(1))}{\mathbb{P}_{x,a}(\tau > n)\sigma^3\sqrt{2\pi n}} \int_0^t u \exp\left(\frac{-u^2}{2\sigma^2}\right) du \\
&= \frac{1+o(1)}{\sigma^2} \int_0^t u \exp\left(\frac{-u^2}{2\sigma^2}\right) du. \qquad (4.19)
\end{aligned}$$

The assertion of the theorem arrives by combining (4.16), (4.17), (4.18) and (4.19). □

4.3. *Proof of Lemma 4.4.*
(1) Fix $\varepsilon > 0$ and let

$$A_{n,\varepsilon} := \left[\sup_{0\leq t\leq 1} |S_{[nt]} - \sigma B_{nt}| \leq n^{\frac{1}{2}-2\varepsilon}\right].$$

For any $x \in \mathbb{X}$, (4.2) implies $\mathbb{P}_x(A^c_{n,\varepsilon}) \leq c_0 n^{-2\varepsilon}$. Denote $a^\pm := a \pm n^{\frac{1}{2}-2\varepsilon}$ and notice that for $a \in [n^{\frac{1}{2}-\varepsilon}, \theta_n\sqrt{n}]$,

$$0 \leq a^\pm \leq 2\theta_n\sqrt{n}. \qquad (4.20)$$

Using (4.4) and (4.20), for any $x \in \mathbb{X}$ and $a \in [n^{\frac{1}{2}-\varepsilon}, \theta_n\sqrt{n}]$, we obtain

$$-\frac{ca^\pm\theta_n}{\sqrt{n}} \pm \frac{2n^{-2\varepsilon}}{\sigma\sqrt{2\pi}} \leq \mathbb{P}_x(\tau^{bm}_{a^\pm} > n) - \frac{2a}{\sigma\sqrt{2\pi n}} \leq \frac{ca^\pm\theta_n}{\sqrt{n}} \pm \frac{2n^{-2\varepsilon}}{\sigma\sqrt{2\pi}}. \qquad (4.21)$$

For any $a \geq n^{\frac{1}{2}-\varepsilon}$, we have $\big[\tau^{bm}_{a^-} > n\big] \cap A^c_{n,\varepsilon} \subset [\tau_a > n] \cap A^c_{n,\varepsilon} \subset \big[\tau^{bm}_{a^+} > n\big] \cap A^c_{n,\varepsilon}$, which yields

$$\mathbb{P}_x(\tau^{bm}_{a^-} > n) - \mathbb{P}_x(A^c_{n,\varepsilon}) \leq \mathbb{P}_x(\tau_a > n) \leq \mathbb{P}_x(\tau^{bm}_{a^+} > n) + \mathbb{P}_x(A^c_{n,\varepsilon})$$

for any $x \in \mathbb{X}$. It follows that

$$\begin{cases} \mathbb{P}_x(\tau_a > n) - \mathbb{P}_x(\tau_{a^+}^{bm} > n) \leq c_0 n^{-2\varepsilon}, \\ \mathbb{P}_x(\tau_{a^-}^{bm} > n) - \mathbb{P}_x(\tau_a > n) \leq c_0 n^{-2\varepsilon}. \end{cases} \tag{4.22}$$

The fact that $\theta_n n^{\varepsilon/4} \to +\infty$ yields for n great enough

$$\theta_n \frac{a}{\sqrt{n}} \geq \frac{n^{\frac{1}{2}-\varepsilon}}{n^{\varepsilon}\sqrt{n}} = n^{-2\varepsilon}. \tag{4.23}$$

From (4.21), (4.22) and (4.23), it follows that for any $a \in [n^{\frac{1}{2}-\varepsilon}, \theta_n\sqrt{n}]$,

$$\left| \mathbb{P}_x(\tau_a > n) - \frac{2a}{\sigma\sqrt{2\pi n}} \right| \leq c(1+\theta_n)n^{-2\varepsilon} + c_1 \frac{\theta_n a}{\sqrt{n}} \leq c_2 \frac{\theta_n a}{\sqrt{n}}.$$

(2) For n great enough, condition $a \geq n^{\frac{1}{2}-\varepsilon}$ implies $a^+ \leq 2a$. From (4.3) and (4.22) , since $n^{-2\varepsilon} \leq \frac{a}{\sqrt{n}}$, for any $x \in \mathbb{X}$,

$$\mathbb{P}_x(\tau_a > n) \leq c\frac{a}{\sigma\sqrt{n}} + c_0 n^{-2\varepsilon} \leq c_1 \frac{a}{\sqrt{n}}.$$

□

4.4. *Proof of Lemma 4.5.*
(1) We prove that $\lim_{n\to+\infty} \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] = V(x,a)$. Then, the assertion arrives by using (2.16) and taking into account that $\mathbb{P}_x(\tau_a < +\infty) = 1$ and $\mathbb{P}_x(\lim_{n\to+\infty} \nu_{n,\varepsilon} = +\infty) = 1$. For $x \in \mathbb{X}$ and $a \geq 0$, we obtain

$$\begin{aligned} \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] &= \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] \\ &= \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau > \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}]\right] \\ &\quad -\mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau > \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}], \nu_{n,\varepsilon} > n^{1-\varepsilon}\right]. \end{aligned}$$

By using Lemma 3.3,

$$\mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau > \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}], \nu_{n,\varepsilon} > n^{1-\varepsilon}\right] \leq c(1+a)e^{-c_\varepsilon n^{\varepsilon}}.$$

Using the facts that $(M_n)_{n\geq 0}$ is a martingale and $\mathbb{P}_x\left(\lim_{n\to+\infty} \nu_{n,\varepsilon} = +\infty\right) = 1$, we obtain

$$\begin{aligned} \lim_{n\to+\infty} \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}}; \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] &= \lim_{n\to+\infty} \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau > \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}]\right] \\ &= a - \lim_{n\to+\infty} \mathbb{E}_{x,a}\left[M_{\nu_{n,\varepsilon}\wedge[n^{1-\varepsilon}]}; \tau \leq \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}]\right] \\ &= a - \lim_{n\to+\infty} \mathbb{E}_{x,a}\left[M_\tau; \tau \leq \nu_{n,\varepsilon} \wedge [n^{1-\varepsilon}]\right] \\ &= a - \mathbb{E}_{x,a}[M_\tau] = V(x,a). \end{aligned}$$

(2) Let $b = a + A$. Remind that $M_n^* = \max_{1\leq k\leq n} |M_k|$. We obtain

$$\begin{aligned} \mathbb{E}_{x,a}\left[S_{\nu_{n,\varepsilon}}; S_{\nu_{n,\varepsilon}} > \theta_n n^{\frac{1}{2}}, \tau > \nu_{n,\varepsilon}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] &\leq \mathbb{E}_{x,b}\left[M_{\nu_{n,\varepsilon}}; M_{\nu_{n,\varepsilon}} > \theta_n n^{\frac{1}{2}}, \nu_{n,\varepsilon} \leq n^{1-\varepsilon}\right] \\ &\leq \mathbb{E}_{x,b}\left[M_{[n^{1-\varepsilon}]}^*; M_{[n^{1-\varepsilon}]}^* > \theta_n n^{\frac{1}{2}}\right]. \end{aligned}$$

Since $\theta_n n^{\varepsilon/4} \to +\infty$ as $n \to +\infty$, it suffices to prove that for any $\delta > 0$, $x \in \mathbb{X}$ and $b \in \mathbb{R}$,

$$\lim_{n\to+\infty} n^{2\varepsilon}\mathbb{E}_x\left[b + M_n^*; M_n^* > n^{\frac{1}{2}+\delta}\right] = 0.$$

Obviously, by (3.5),

$$\begin{aligned}\mathbb{E}_x\left[b + M_n^*; M_n^* > n^{\frac{1}{2}+\delta}\right] &\leq b\mathbb{P}_x\left(M_n^* > n^{\frac{1}{2}+\delta}\right) + \mathbb{E}_x\left[M_n^*; M_n^* > n^{\frac{1}{2}+\delta}\right] \\ &= \left(b + n^{\frac{1}{2}+\delta}\right)\mathbb{P}_x\left(M_n^* > n^{\frac{1}{2}+\delta}\right) + \int_{n^{\frac{1}{2}+\delta}}^{+\infty} \mathbb{P}_x(M_n^* > t)dt \\ &\leq c\left(b + n^{\frac{1}{2}+\delta}\right)n^{-p\delta} + cn^{-p\delta+\frac{1}{2}+\delta}.\end{aligned}$$

Since p can be taken arbitrarily great, it follows that $\lim_{n\to+\infty} n^{2\varepsilon} J_n' = 0$.

□

4.5. *Proof of Lemma 4.6.*
Recall that $a^{\pm} = a \pm n^{1/2-2\varepsilon}$ and denote $t^{\pm} = t \pm 2n^{-2\varepsilon}$. For any $1 \leq k \leq n^{1-\varepsilon}$,

$$\{\tau_{a^-}^{bm}\} \cap A_{n,\varepsilon} \subset \{\tau_a > n-k\} \cap A_{n,\varepsilon} \subset \{\tau_{a^+}^{bm}\} \cap A_{n,\varepsilon}$$

and

$$\{a^- + \sigma B_{n-k} \leq t^-\sqrt{n}\} \cap A_{n,\varepsilon} \subset \{a + S_{n-k} \leq t\sqrt{n}\} \cap A_{n,\varepsilon} \subset \{a^+ + \sigma B_{n-k} \leq t^+\sqrt{n}\} \cap A_{n,\varepsilon},$$

which imply

$$\begin{aligned}&\mathbb{P}_x(\tau_{a^-}^{bm} > n-k, a^- + \sigma B_{n-k} \leq t^-\sqrt{n}) - \mathbb{P}_x(A_{n,\varepsilon}^c) \\ &\qquad \leq \mathbb{P}_x(\tau_a > n-k, a + S_{n-k} \leq t\sqrt{n}) \leq \\ &\qquad\qquad \mathbb{P}_x(\tau_{a^+}^{bm} > n-k, a^+ + \sigma B_{n-k} \leq t^+\sqrt{n}) + \mathbb{P}_x(A_{n,\varepsilon}^c).\end{aligned} \tag{4.24}$$

Moreover, by Lemma 4.2, we obtain

$$\mathbb{P}_x\left(\tau_{a^+}^{bm} > n-k, a^+ + \sigma B_{n-k} \leq t^+\sqrt{n}\right) = \frac{2a}{\sigma^3\sqrt{2\pi n}}\int_0^t u\exp\left(-\frac{u^2}{2\sigma^2}\right)du(1+o(1)) \tag{4.25}$$

and similarly,

$$\mathbb{P}_x\left(\tau_{a^-}^{bm} > n-k, a^- + \sigma B_{n-k} \leq t^-\sqrt{n}\right) = \frac{2a}{\sigma^3\sqrt{2\pi n}}\int_0^t u\exp\left(-\frac{u^2}{2\sigma^2}\right)du(1+o(1)) \tag{4.26}$$

Therefore, from (4.24), (4.25), (4.26) and $\mathbb{P}_x(A_{n,\varepsilon}^c) \leq cn^{-2\varepsilon}$, it follows that

$$\mathbb{P}_x\left(\tau_a > n-k, a + S_{n-k} \leq t\sqrt{n}\right) = \frac{2a}{\sigma^3\sqrt{2\pi n}}\int_0^t u\exp\left(-\frac{u^2}{2\sigma^2}\right)du(1+o(1)).1$$

□

5. On conditions C1-C3 of Theorem 2.1 in Grama et al. (2014)

Let $k_{gap}, M_1, M_2 \in \mathbb{N}$ and $j_0 < \ldots < j_{M_1+M_2}$ be natural numbers. Denote $a_{k+J_m} = \sum_{l\in J_m} a_{k+l}$, where $J_m = [j_{m-1}, j_m), m = 1, \ldots, M_1+M_2$ and $k \geq 0$. Consider the vectors $\bar{a}_1 = (a_{J_1}, \ldots, a_{J_{M_1}})$ and $\bar{a}_2 = (a_{k_{gap}+J_{M_1+1}}, \ldots, a_{k_{gap}+J_{M_1+M_2}})$. Denote by $\phi_x(s,t) = \mathbb{E}e^{is\bar{a}_1+it\bar{a}_2}$, $\phi_{x,1}(s) = \mathbb{E}e^{is\bar{a}_1}$ and $\phi_{x,2}(s) = \mathbb{E}e^{it\bar{a}_2}$ the characteristic functions of $(\bar{a}_1, \bar{a}_2)$, $\bar{a}_1$ and $\bar{a}_2$, respectively. For the sake of brevity, we denote $\phi_1(s) = \phi_{x,1}(s), \phi_2(t) = \phi_{x,2}(t)$ and $\phi(s,t) = \phi_x(s,t)$.

We first check that conditions C1-C3 hold and then prove the needed lemmas.

5.1. *Statement and proofs of conditions C1-C3.*
C1: There exist positive constants $\varepsilon_0 \leq 1, \lambda_0, \lambda_1, \lambda_2$ such that for any $k_{gap} \in \mathbb{R}, M_1, M_2 \in \mathbb{Z}_+$, any sequence $j_0 < \ldots < j_{M_1+M_2}$ and any $s \in \mathbb{R}^{M_1}, t \in \mathbb{R}^{M_2}$ satisfying $|(s,t)|_\infty \leq \varepsilon_0$,

$$|\phi(s,t) - \phi_1(s)\phi_2(t)| \leq \lambda_0 \exp(-\lambda_1 k_{gap}) \left(1 + \max_{m=1,\ldots,M_1+M_2} card(J_m)\right)^{\lambda_2(M_1+M_2)}.$$

C2: There exists a positive constant δ such that $\sup_{n\geq 0} |a_n|_{L^{2+2\delta}} < +\infty$.
C3: There exist a positive constant C and a positive number σ such that for any $\gamma > 0$, any $x \in \mathbb{X}$ and any $n \geq 1$,

$$\sup_{m\geq 0} \left| n^{-1} Var_{\mathbb{P}_x} \left(\sum_{i=m}^{m+n-1} a_i \right) - \sigma^2 \right| \leq C n^{-1+\gamma}.$$

Proposition 5.1. *Condition 1 is satisfied under hypotheses P1-P5.*

Proof. First, we prove the following lemma.

Lemma 5.2. *There exist two positive constants C and κ such that $0 < \kappa < 1$ and*

$$|\phi(s,t) - \phi_1(s)\phi_2(t)| \leq C C_P^{M_1+M_2} \kappa^{k_{gap}},$$

where C_P is defined in Proposition 2.3.

Proof. In fact, the characteristic functions of the ramdom variables $\bar{a}_1, \bar{a}_2$ and $(\bar{a}_1, \bar{a}_2)$ can be written in terms of operator respectively as follows:

$$\begin{aligned}
\phi_1(s) &= \mathbb{E}_x[e^{is\bar{a}_1}] = P^{j_0-1} P_{s_1}^{|J_1|} \ldots P_{s_{M_1}}^{|J_{M_1}|} \mathbf{1}(x), \\
\phi_2(t) &= \mathbb{E}_x[e^{it\bar{a}_2}] = P^{k_{gap}+j_{M_1}-1} P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1}(x), \qquad (5.1) \\
\phi(s,t) &= \mathbb{E}_x[e^{is\bar{a}_1+it\bar{a}_2}] = P^{j_0-1} P_{s_1}^{|J_1|} \ldots P_{s_{M_1}}^{|J_{M_1}|} P^{k_{gap}} P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1}(x).
\end{aligned}$$

Now we decompose $\phi(s,t)$ into the sum of $\phi_\Pi(s,t)$ and $\phi_R(s,t)$ by using the spectral decomposition $P = \Pi + R$ in Proposition 2.3, where

$$\begin{aligned}
\phi_\Pi(s,t) &= P^{j_0-1} P_{s_1}^{|J_1|} \ldots P_{s_{M_1}}^{|J_{M_1}|} \Pi P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1}(x), \\
\phi_R(s,t) &= P^{j_0-1} P_{s_1}^{|J_1|} \ldots P_{s_{M_1}}^{|J_{M_1}|} R^{k_{gap}} P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1}(x).
\end{aligned}$$

Since $\Pi(\varphi) = \nu(\varphi)\mathbf{1}$ for any $\varphi \in L$ and P_t acts on L, we obtain

$$\phi_\Pi(s,t) = P^{j_0-1} P_{s_1}^{|J_1|} \ldots P_{s_{M_1}}^{|J_{M_1}|} \mathbf{1}(x) \nu\left(P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1} \right).$$

Then setting $\psi_2(t) = \nu(P_{t_1}^{|J_{M_1+1}|} \ldots P_{t_{M_2}}^{|J_{M_1+M_2}|} \mathbf{1})$ yields

$$\begin{aligned}
\phi(s,t) &= \phi_1(s)\psi_2(t) + \phi_R(s,t) \\
&= \phi_1(s)\phi_2(t) + \phi_1(s)[\psi_2(t) - \phi_2(t)] + \phi_R(s,t),
\end{aligned}$$

which implies

$$|\phi(s,t) - \phi_1(s)\phi_2(t)| \leq |\phi_1(s)||\psi_2(t) - \phi_2(t)| + |\phi_R(s,t)|. \qquad (5.2)$$

On the one hand, we can see that $|\phi_1(s)| = \left|\left(P^{j_0-1}P_{s_1}^{|J_1|}\dots P_{s_{M_1}}^{|J_{M_1}|}\mathbf{1}\right)(x)\right| \leq C_P^{1+M_1}$ and $|\phi_R(s,t)| \leq C_P^{1+M_1+M_2}C_R\kappa^{k_{gap}}$. On the other hand, since ν is P-invariant measure and $(\nu-\delta_x)(\mathbf{1})=0$, by using again the expression $P=\Pi+R$, we obtain

$$\begin{aligned}
|\psi_2(t)-\phi_2(t)| &= \left|(\nu-\delta_x)\left(P^{k_{gap}+j_{M_1}-1}P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&\leq \left|(\nu-\delta_x)\left(\Pi P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&\qquad + \left|(\nu-\delta_x)\left(R^{k_{gap}+j_{M_1}-1}P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&= \left|(\nu-\delta_x)(\mathbf{1})\nu\left(P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&\qquad + \left|(\nu-\delta_x)\left(R^{k_{gap}+j_{M_1}-1}P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&= \left|(\nu-\delta_x)\left(R^{k_{gap}+j_{M_1}-1}P_{t_1}^{|J_{M_1+1}|}\dots P_{t_{M_2}}^{|J_{M_1+M_2}|}\mathbf{1}\right)\right| \\
&\leq CC_P^{M_2}\kappa^{k_{gap}+j_{M_1}-1}. \qquad (5.3)
\end{aligned}$$

Therefore, (5.2) follows.

□

Second, let $\lambda_2 = \max\{1, \log_2 C_P\}$. Since $\max\limits_{m=1,\dots,M_1+M_2} card(J_m) \geq 1$, we obtain

$$C_P^{M_1+M_2} \leq 2^{\lambda_2(M_1+M_2)} \leq \left(1+\max_{m=1,\dots,M_1+M_2} card(J_m)\right)^{\lambda_2(M_1+M_2)},$$

which implies that

$$|\phi(s,t)-\phi_1(s)\phi_2(t)| \leq C\kappa^{k_{gap}}\left(1+\max_{m=1,\dots,M_1+M_2} card(J_m)\right)^{\lambda_2(M_1+M_2)}.$$

Finally, let $\lambda_0 = C$ and $\lambda_1 = -\log\kappa$. Then the assertion arrives.

□

Proposition 5.3. *Condition 2 is satisfied under hypotheses P1-P5.*

Proof. Condition P1 implies that there exists $\delta_0 > 0$ such that $\mathbb{E}[N(g)^{\delta_0}] < +\infty$ and since $\mathbb{E}[N(g)^{\delta_0}] = \mathbb{E}[\exp(\delta_0 \log N(g))] = \sum\limits_{k=0}^{+\infty}\frac{\delta_0^k}{k!}\mathbb{E}[(\log N(g))^k]$, we obtain $\mathbb{E}|a_n|^k \leq \mathbb{E}[(\log N(g))^k] < +\infty$ for any $n \geq 0$ and any $k \geq 0$.

□

Proposition 5.4. *Condition 3 is satisfied under hypotheses P1-P5. More precisely, there exists a positive constant σ such that for any $x \in \mathbb{X}$ and any $n \geq 1$,*

$$\sup_{m\geq 0}\left|Var_{\mathbb{P}_x}\left(\sum_{k=m}^{m+n-1}a_k\right) - n\sigma^2\right| < +\infty. \qquad (5.4)$$

Proof. For any integer $m, n \geq 0$, we denote $S_{m,n} = \sum_{k=m}^{m+n-1} a_k$, $V_x(X) = Var_{\mathbb{P}_x}(X) = \mathbb{E}_x(X^2) - (\mathbb{E}_x X)^2$ and $Cov_x(X,Y) = Cov_{\mathbb{P}_x}(X,Y)$. Then

$$V_x(S_{m,n}) = \sum_{k=m}^{m+n-1} V_x(a_k) + 2\sum_{k=m}^{m+n-1}\sum_{l=1}^{m+n-k-1} Cov_x(a_k, a_{k+l}) \qquad (5.5)$$

and (5.4) becomes $\sup_{m\geq 0}|V_x(S_{m,n})-n\sigma^2|<+\infty$. We claim two lemmas and postpone their proofs until the end of this section.

Lemma 5.5. *There exist $C>0$ and $0<\kappa<1$ such that for any $x\in\mathbb{X}$, any $k\geq 0$ and any $l\geq 0$,*

$$|Cov_x(a_k,a_{k+l})|\leq C\kappa^l. \tag{5.6}$$

Lemma 5.6. *There exist $C>0$, $0<\kappa<1$ and a sequence $(s_n)_{n\geq 0}$ of real numbers such that for any $x\in\mathbb{X}$, any $k\geq 0$ and any $l\geq 0$,*

$$|Cov_x(a_k,a_{k+l})-s_l|\leq C\kappa^k, \tag{5.7}$$

$$|s_l|\leq C\kappa^l. \tag{5.8}$$

For the first term of the right side of (5.5), by combining Lemma 5.5 and Lemma 5.6, we obtain

$$|Cov_x(a_k,a_{k+l})-s_l|\leq C\kappa^{\max\{k,l\}}. \tag{5.9}$$

Inequality (5.7) implies $|V_x(a_k)-s_0|\leq C\kappa^k$, which yields for any integer $m,n\geq 0$,

$$\left|\sum_{k=m}^{m+n-1}V_x(a_k)-ns_0\right|\leq\sum_{k=m}^{m+n-1}|V_x(a_k)-s_0|\leq C\sum_{k=m}^{m+n-1}\kappa^k\leq\frac{C}{1-\kappa}<+\infty. \tag{5.10}$$

For the second term of the right side of (5.5), we can see that

$$\begin{aligned}&\left|\sum_{k=m}^{m+n-1}\sum_{l=1}^{m+n-k-1}Cov_x(a_k,a_{k+l})-\sum_{k=m}^{m+n-1}\sum_{l=1}^{+\infty}s_l\right|\\ \leq&\sum_{k=m}^{m+n-1}\sum_{l=1}^{m+n-k-1}|Cov_x(a_k,a_{k+l})-s_l|+\sum_{k=m}^{m+n-1}\sum_{l=m+n-k}^{+\infty}|s_l|\\ =&\ \Sigma_1(x,m,n)+\Sigma_2(x,m,n).\end{aligned} \tag{5.11}$$

On the one hand, by (5.7) and (5.9), we can see that for any $x\in\mathbb{X}$, any $m\geq 0$ and any $n\geq 1$,

$$\begin{aligned}\Sigma_1(x,m,n)&\leq\sum_{k=0}^{+\infty}\sum_{l=1}^{k}C\kappa^k+\sum_{k=0}^{+\infty}\sum_{l=k+1}^{+\infty}C\kappa^l\\ &\leq\sum_{k=0}^{+\infty}Ck\kappa^k+\sum_{k=0}^{+\infty}C\frac{\kappa^{k+1}}{1-\kappa}<+\infty.\end{aligned} \tag{5.12}$$

Similarly, on the other hand, by (5.8) we obtain for any $x\in\mathbb{X}$, any $m\geq 0$ and any $n\geq 1$,

$$\Sigma_2(x,m,n)\leq\sum_{k=0}^{n-1}\sum_{l=n-k}^{+\infty}C\kappa^l\leq\frac{C}{(1-\kappa)^2}<+\infty. \tag{5.13}$$

Combining (5.5),(5.10),(5.11),(5.12) and (5.13), we obtain

$$\sup_{m\geq 0}\left|V_x(S_{m,n})-n\sum_{l=0}^{+\infty}s_l\right|<+\infty. \tag{5.14}$$

In fact, by using Lemma 2.1 in Le Page et al. (to appear), Theorem 5 in Hennion (1997) implies that the sequence $(\frac{S_n}{\sqrt{n}})_{n\geq 1}$ converges weakly to a normal law with

variance σ^2. Meanwhile, under hypothesis P2, Corollary 3 in Hennion (1997) implies that the sequence $(|R_n|)_{n\geq 1}$ is not tight and thus $\sigma^2 > 0$, see Hennion (1997) for the definition and basic properties. Therefore, we can see that $Var_x S_n \sim n\sigma^2$ with $\sigma^2 > 0$, which yields $\sum_{l=0}^{+\infty} s_l = \sigma^2$.

□

5.2. *Proof of Lemma 5.5.*

Let $g(x) = \begin{cases} x & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| > 2. \end{cases}$ such that g is C^∞ on $\mathbb{R}$ and $|g(x)| \leq |x|$ for any $x \in \mathbb{R}$. Then $g \in L^1(\mathbb{R}) \cap C_c^1(\mathbb{R})$. Therefore, the Fourier transform of g is $\hat{g}$ defined as follows:

$$\hat{g}(t) := \int_{\mathbb{R}} e^{-itx} g(x) dx,$$

and the Inverse Fourier Theorem yields

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{itx} \hat{g}(t) dt.$$

Let $g_T(x) := Tg(\frac{x}{T})$ for any $T > 0$. Then $|\hat{g}_T|_1 = T|\hat{g}|_1 < +\infty$. Let $h_T(x, y) = g_T(x)g_T(y)$. Then $\hat{h}_T(x, y) = \hat{g}_T(x)\hat{g}_T(y)$. Let V and V' be two i.i.d. random variables with mean 0, independent of a_l for any $l \geq 0$ whose characteristic functions have the support included in the interval $[-\varepsilon_0, \varepsilon_0]$ for ε_0 defined in C1. Assume that $\mathbb{E}_|V|^n < +\infty$ for any $n > 0$. Let $Z_k = a_k + V$ and $Z'_{k+l} = a_{k+l} + V'$ and denote by $\widetilde{\phi}_1(s), \widetilde{\phi}_2(t)$ and $\widetilde{\phi}(s, t)$ the characteristic functions of Z_k, Z'_{k+l} and (Z_k, Z'_{k+l}), respectively.

We use the same notations introduced at the beginning of this section by setting $\phi_1(s) = \mathbb{E}_x[e^{isa_k}], \phi_2(t) = \mathbb{E}_x[e^{ita_{k+l}}]$ and $\phi(s, t) = \mathbb{E}_x[e^{isa_k + ita_{k+l}}]$. We also denote φ the characteristic function of V, that yields

$$\begin{aligned} \widetilde{\phi}_1(s) &= \mathbb{E}_[e^{isZ_k}] = \mathbb{E}_[e^{isa_k}]\mathbb{E}_[e^{isV}] = \phi_1(s)\varphi(s), \\ \widetilde{\phi}_2(t) &= \mathbb{E}_[e^{itZ'_{k+l}}] = \mathbb{E}_[e^{ita_{k+l}}]\mathbb{E}_[e^{itV'}] = \phi_2(t)\varphi(t), \\ \widetilde{\phi}(s, t) &= \mathbb{E}_[e^{isZ_k + itZ'_{k+l}}] = \mathbb{E}_[e^{isa_k + ita_{k+l}}]\mathbb{E}_[e^{isV}]\mathbb{E}_[e^{itV'}] = \phi(s, t)\varphi(s)\varphi(t). \end{aligned} \tag{5.15}$$

Then we can see that $\widetilde{\phi}_1$ and $\widetilde{\phi}_2$ have the support in $[-\varepsilon_0, \varepsilon_0]$. We perturb a_k and a_{k+l} by adding the random variables V and V' with mean 0 and the support of their characteristic functions are on $[-\varepsilon_0, \varepsilon_0]$. We explicit the quantity $Cov_x(a_k, a_{k+l})$:

$$Cov_x(a_k, a_{k+l}) = \mathbb{E}_x[a_k, a_{k+l}] - \mathbb{E}_x a_k \mathbb{E}_x a_{k+l}. \tag{5.16}$$

On the one hand, we can see that

$$\begin{aligned} \mathbb{E}_x[a_k a_{k+l}] = \mathbb{E}_x[Z_k Z'_{k+l}] &= \mathbb{E}_x[h_T(Z_k; Z'_{k+l})] + \mathbb{E}_x[Z_k Z'_{k+l}] - \mathbb{E}_x[h_T(Z_k; Z'_{k+l})] \\ &= \frac{1}{(2\pi)^2} \mathbb{E}_x \int\int \hat{h}_T(s, t) e^{isZ_k + itZ'_{k+l}} ds dt + R_0 \\ &= \frac{1}{(2\pi)^2} \int\int \hat{h}_T(s, t) \mathbb{E}_x \left[e^{isZ_k + itZ'_{k+l}} \right] ds dt + R_0 \\ &= \frac{1}{(2\pi)^2} \int\int \hat{h}_T(s, t) \widetilde{\phi}(s, t) ds dt + R_0, \end{aligned} \tag{5.17}$$

where $R_0 = \mathbb{E}_x[Z_k Z'_{k+l}] - \mathbb{E}_x[h_T(Z_k; Z'_{k+l})]$. On the other hand, we obtain

$$\begin{aligned}\mathbb{E}_x a_k = \mathbb{E}_x Z_k &= \mathbb{E}_x g_T(Z_k) + \mathbb{E}_x Z_k - \mathbb{E}_x g_T(Z_k) \\ &= \frac{1}{2\pi}\int \hat{g}_T(s)\widetilde{\phi}_1(s)ds + R_1, \end{aligned} \tag{5.18}$$

where $R_1 = \mathbb{E}_x Z_k - \mathbb{E}_x g_T(Z_k)$ and

$$\begin{aligned}\mathbb{E}_x a_{k+l} = \mathbb{E}_x Z'_{k+l} &= \mathbb{E}_x g_T(Z'_{k+l}) + \mathbb{E}_x Z'_{k+l} - \mathbb{E}_x g_T(Z'_{k+l}) \\ &= \frac{1}{2\pi}\int \hat{g}_T(t)\widetilde{\phi}_2(t)dt + R_2, \end{aligned} \tag{5.19}$$

where $R_2 = \mathbb{E}_x Z'_{k+l} - \mathbb{E}_x g_T(Z'_{k+l})$. From (5.16), (5.17), (5.18) and (5.19), since $\hat{h}_T(s,t) = \hat{g}_T(s)\hat{g}_T(t)$, we obtain

$$\begin{aligned}Cov_x(a_k, a_{k+l}) &= \frac{1}{(2\pi)^2}\int\int \hat{h}_T(s,t)\widetilde{\phi}(s,t)dsdt + R_0 \\ &\quad - \left(\frac{1}{2\pi}\int \hat{g}_T(s)\widetilde{\phi}_1(s)ds + R_1\right)\left(\frac{1}{2\pi}\int \hat{g}_T(t)\widetilde{\phi}_2(t)dt + R_2\right) \\ &= \frac{1}{(2\pi)^2}\int\int \hat{h}_T(s,t)\left[\widetilde{\phi}(s,t) - \widetilde{\phi}_1(s)\widetilde{\phi}_2(t)\right]dsdt + R \end{aligned} \tag{5.20}$$

where $R = R_0 - R_1R_2 - R_1\frac{1}{2\pi}\int \hat{g}_T(t)\widetilde{\phi}_2(t)dt - R_2\frac{1}{2\pi}\int \hat{g}_T(s)\widetilde{\phi}_1(s)ds$. Since $\hat{g}_T \in L_1(\mathbb{R})$ and applying Lemma 5.2 for $j_0 = k, j_1 = k+1, j_2 = k+2, k_{gap} = l, M_1 = M_2 = 1$, we obtain

$$\begin{aligned}|Cov_x(a_k, a_{k+l})| &\le \frac{1}{(2\pi)^2}\int\int \left|\hat{h}_T(s,t)\right|\left|\widetilde{\phi}(s,t) - \widetilde{\phi}_1(s)\widetilde{\phi}_2(t)\right|dsdt + |R| \\ &\le \frac{1}{(2\pi)^2}\int\int \left|\hat{h}_T(s,t)\right|\left|\phi(s,t)\varphi(s)\varphi(t) - \phi_1(s)\phi_2(t)\varphi(s)\varphi(t)\right|dsdt + |R| \\ &\le \sup_{|s|,|t|\le\varepsilon_0}|\phi(s,t) - \phi_1(s)\phi_2(t)|\left(\int |\hat{g}_T(s)|\,ds\right)^2 + |R| \\ &\le CT^2\kappa^l + |R|. \end{aligned} \tag{5.21}$$

It remains to bound of $|R|$. On the one hand, we can see that

- $|R_1| = |\mathbb{E}_x[Z_k - g_T(Z_k)]| = \mathbb{E}_x\left|[Z_k - g_T(Z_k)]\mathbf{1}_{[|Z_k|>T]}\right| \le 2T^{-1}\mathbb{E}_x|Z_k|^2$,
- $\left|\frac{1}{2\pi}\int \hat{g}_T(s)\widetilde{\phi}_1(s)ds\right| = |\mathbb{E}_x g_T(Z_k)| \le \mathbb{E}_x|Z_k| \le \mathbb{E}_x|a_k| + \mathbb{E}_x|V| \le C$,
- $|R_2| = \left|\mathbb{E}_x[Z'_{k+l} - g_T(Z'_{k+l})]\right| \le 2T^{-1}\mathbb{E}_x|Z'_{k+l}|^2$,
- $\left|\frac{1}{2\pi}\int \hat{g}_T(t)\widetilde{\phi}_2(t)dt\right| = \left|\mathbb{E}_x g_T(Z'_{k+l})\right| \le \mathbb{E}_x|a_{l+k}| + \mathbb{E}_x|V'| \le C$.

On the other hand, similarly for $|R_0|$, we obtain

$$\begin{aligned}|R_0| &= \mathbb{E}_x\left[\left|Z_kZ'_{k+l} - h_T(Z_k, Z'_{k+l})\right|\left(\mathbf{1}_{[|Z_k|>T]} + \mathbf{1}_{[|Z_k|\le T]}\right)\left(\mathbf{1}_{[|Z'_{k+l}|>T]} + \mathbf{1}_{[|Z'_{k+l}|\le T]}\right)\right] \\ &\le \mathbb{E}_x\left[\left|Z_kZ'_{k+l} - h_T(Z_k, Z'_{k+l})\right|\left(\mathbf{1}_{[|Z_k|>T]} + \mathbf{1}_{[|Z'_{k+l}|>T]}\right)\right] \\ &\le 2\mathbb{E}_x\left|Z_kZ'_{k+l}\mathbf{1}_{[|Z_k|>T]}\right| + 2\mathbb{E}_x\left|Z_kZ'_{k+l}\mathbf{1}_{[|Z'_{k+l}|>T]}\right|. \end{aligned}$$

For any positive δ, let $q_\delta = \frac{\delta+1}{\delta}$, by Holder's inequality, we obtain

$$\mathbb{E}_x \left| Z_k Z'_{k+l} \mathbf{1}_{[|Z_k|>T]} \right| \leq \left(\mathbb{E}_x |Z_k|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} \left(\mathbb{E}_x |Z'_{k+l}|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} \mathbb{P}_x(|Z_k| > T)^{\frac{1}{q_\delta}}.$$

By Minkowski's inequality,

$$\left(\mathbb{E}_x |Z_k|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} \leq \left(\mathbb{E}_x |a_k|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} + \left(\mathbb{E}_x |V|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} < C,$$
$$\left(\mathbb{E}_x |Z'_{k+l}|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} \leq \left(\mathbb{E}_x |a_{l+k}|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} + \left(\mathbb{E}_x |V'|^{2+2\delta}\right)^{\frac{1}{2+2\delta}} < C.$$

By Markov's inequality,

$$\mathbb{P}_x(|Z_k| > T) \leq \frac{1}{T^{q_\delta}} \mathbb{E}_x |Z_k|^{q_\delta} \leq \frac{C}{T^{q_\delta}},$$
$$\mathbb{P}_x(|Z'_{k+l}| > T) \leq \frac{1}{T^{q_\delta}} \mathbb{E}_x |Z'_{k+l}|^{q_\delta} \leq \frac{C}{T^{q_\delta}}.$$

Hence $|R_0| \leq CT^{-1}$ for $T > 1$ and thus $|R| \leq CT^{-1}$.

Thus, (5.21) becomes $|Cov_x(a_k, a_{k+l})| \leq CT^2\kappa^l + CT^{-1}$. By choosing $T = \kappa^{-\alpha}$ with $\alpha > 0$, we obtain

$$|Cov_x(a_k, a_{k+l})| \leq C\kappa^{l-2\alpha} + C\kappa^\alpha \leq C' \max\{\kappa^{l-2\alpha}, \kappa^\alpha\}.$$

Now we choose $\alpha > 0$ such that $l - 2\alpha > 0$, for example, let $\alpha = \frac{l}{4}$, we obtain

$$|Cov_x(a_k, a_{k+l})| \leq C\kappa^{\frac{l}{4}}.$$

5.3. *Proof of Lemma 5.6.*
Inequality (5.8) follows by setting $k = l$ in (5.6) and (5.7). It suffices to prove (5.7). Recall the definition in (5.15) and let

$$\begin{aligned} \psi(s) &= \nu(P_s\mathbf{1})\varphi(s), \\ \psi(s,t;l) &= \nu(P_s P^{l-1} P_t \mathbf{1})\varphi(s)\varphi(t), \\ \widetilde{\psi}(s,t;l) &= \psi(s,t;l) - \psi(s)\psi(t), \\ \widetilde{\phi}_0(s,t) &= \widetilde{\phi}(s,t) - \widetilde{\phi}_1(s)\widetilde{\phi}_2(t), \\ s_{l,T} &= \frac{1}{(2\pi)^2} \int\int \hat{h}_T(s,t)\widetilde{\psi}(s,t;l)dsdt. \end{aligned} \tag{5.22}$$

Then (5.20) implies

$$|Cov_x(a_k, a_{k+l}) - s_{l,T}| \leq \left| \frac{1}{(2\pi)^2} \int\int \hat{h}_T(s,t)[\widetilde{\phi}_0(s,t) - \widetilde{\psi}(s,t;l)]dsdt \right| + |R|.$$

We claim that

$$\left| \frac{1}{(2\pi)^2} \int\int \hat{h}_T(s,t)[\widetilde{\phi}_0(s,t) - \widetilde{\psi}(s,t;l)]dsdt \right| \leq C\kappa^{k-1}T^2, \tag{5.23}$$

which implies

$$|Cov_x(a_k, a_{k+l}) - s_{l,T}| \leq C\kappa^{k-1}T^2 + CT^{-1}, \tag{5.24}$$

which yields for any $k, m \geq 1$,

$$|Cov_x(a_k, a_{k+l}) - Cov_x(a_m, a_{m+l})| \leq C\kappa^{\min\{k-1,m-1\}}T^2 + CT^{-1}. \tag{5.25}$$

By choosing $T = \kappa^{-\frac{1}{4}\min\{k-1,m-1\}}$, we obtain

$$|Cov_x(a_k, a_{k+l}) - Cov_x(a_m, a_{m+l})| \leq C\kappa^{\min\{\frac{k-1}{4}, \frac{m-1}{4}\}}. \tag{5.26}$$

Hence we can say that $(Cov_x(a_k, a_{k+l}))_l$ is a Cauchy sequence, thus it converges to some limit, denoted by $s_l(x)$. When $k \to +\infty$, (5.24) becomes

$$|s_l(x) - s_{l,T}| \leq CT^{-1}.$$

Now let $T = T(\ell) = \kappa^{-\ell}$, we obtain $|s_l(x) - s_{l,T(\ell)}| \leq C\kappa^{\ell}$. Let $\ell \to +\infty$, we can see that $s_{l,T(\ell)} \to s_l(x)$. Since $s_{l,T(\ell)}$ does not depend on x, so is $s_l(x)$, i.e. $s_l(x) = s_l$. Now let $m \to +\infty$ in (5.26), we obtain

$$|Cov_x(a_k, a_{k+l}) - s_l| \leq C\kappa^{\frac{k-1}{4}}.$$

Now we prove the claim (5.23). By definitions in (5.15) and (5.22), we obtain

$$\left|\widetilde{\phi}_0(s,t) - \widetilde{\psi}(s,t;l)\right| \leq \left|\widetilde{\phi}(s,t) - \psi(s,t;l)\right| + \left|\widetilde{\phi}_1(s)\widetilde{\phi}_2(t) - \psi(s)\psi(t)\right|. \tag{5.27}$$

On the one hand, we can see that

$$\begin{aligned}
&\left|\widetilde{\phi}(s,t) - \psi(s,t;k)\right| \\
= \;& \left|P^{k-1}P_sP^{l-1}P_t\mathbf{1}(x)\varphi(s)\varphi(t) - \nu(P_sP^{l-1}P_t\mathbf{1})\varphi(s)\varphi(t)\right| \\
= \;& \left|\Pi P_sP^{l-1}P_t\mathbf{1}(x) + R^{k-1}P_s\Pi P_t\mathbf{1}(x) + R^{k-1}P_sR^{l-1}P_t\mathbf{1}(x) - \nu(P_sP^{l-1}P_t\mathbf{1})\right| \\
= \;& \left|R^{k-1}P_s\mathbf{1}(x)\nu(P_t\mathbf{1}) + R^{k-1}P_sR^{l-1}P_t\mathbf{1}(x)\right| \leq C\kappa^{k-1}.
\end{aligned} \tag{5.28}$$

On the other hand,

$$\begin{aligned}
\left|\widetilde{\phi}_1(s)\widetilde{\phi}_2(t) - \psi(s)\psi(t)\right| &= \left|[\widetilde{\phi}_1(s) - \psi(s)]\widetilde{\phi}_2(t) + \psi(s)[\widetilde{\phi}_2(t) - \psi(t)]\right| \\
&\leq \left|\widetilde{\phi}_1(s) - \psi(s)\right| + \left|\widetilde{\phi}_2(t) - \psi(t)\right| \\
&\leq |\phi_1(s)\varphi(s) - \psi(s)| + |\phi_2(t)\varphi(t) - \psi(t)|,
\end{aligned}$$

where as long as $k \geq 2$,

$$\begin{aligned}
|\phi_1(s)\varphi(s) - \psi(s)| &= \left|\left[\Pi P_s\mathbf{1}(x) + R^{k-1}P_s\mathbf{1}(x)\right]\mathbb{E}_x[e^{isV}] - \nu(P_s\mathbf{1})\mathbb{E}_x[e^{isV}]\right| \\
&\leq \left|[\Pi P_s\mathbf{1}(x) - \nu(P_s\mathbf{1})] + R^{k-1}P_s\mathbf{1}(x)\right| \\
&= \left|R^{k-1}P_s\mathbf{1}(x)\right| \leq C\kappa^{k-1}.
\end{aligned}$$

Similarly, we obtain

$$\left|\widetilde{\phi}_1(s)\widetilde{\phi}_2(t) - \psi(s)\psi(t)\right| \leq C\kappa^{k-1}. \tag{5.29}$$

Therefore, (5.27), (5.28) and (5.29) imply $\left|\widetilde{\phi}_0(s,t) - \widetilde{\psi}(s,t;l)\right| \leq C\kappa^{k-1}$ which yields the assertion of the claim.

Acknowledgements. The author thanks the referee for many useful remarks and comments which improve this paper and the Vietnam Institute for Advanced Studies in Mathematics in Ha Noi for generous hospitality and financial support during the first semester 2017.

References

Y. Benoist and J-F. Quint. Central limit theorem for linear groups. Ann. Probab. **44** (2), 1308–1340 (2016).

E. Bolthausen. On a functional central limit theorem for random walks conditioned to stay positive. Ann. Probability **4** (3), 480–485 (1976).

Ph. Bougerol and J. Lacroix. Products of random matrices with applications to Schrödinger operators, volume 8 of Progress in Probability and Statistics. Birkhäuser Boston, Inc., Boston, MA (1985).

D. Denisov and V. Wachtel. Random walks in cones. Ann. Probab. **43** (3), 992–1044 (2015).

N. Dunford and J. Schwarz. Linear Operators 1. Interscience (1958).

H. Furstenberg and H. Kesten. Products of random matrices. Ann. Math. Statist. **31**, 457–469 (1960).

I. Grama, E. Le Page and M. Peigné. On the rate of convergence in the weak invariance principle for dependent random variables with applications to Markov chains. Colloq. Math. **134** (1), 1–55 (2014).

Y. Guivarc'h and A. Raugi. Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. Z. Wahrsch. Verw. Gebiete **69** (2), 187–242 (1985).

H. Hennion. Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes. Z. Wahrsch. Verw. Gebiete **67** (3), 265–278 (1984).

H. Hennion. Limit theorems for products of positive random matrices. Ann. Probab. **25** (4), 1545–1587 (1997).

H. Hennion and L. Hervé. Stable laws and products of positive random matrices. J. Theoret. Probab. **21** (4), 966–981 (2008).

D. L. Iglehart. Functional central limit theorems for random walks conditioned to stay positive. Ann. Probab. **2**, 608–619 (1974a).

D. L. Iglehart. Random walks with negative drift conditioned to stay positive. J. Appl. Probability **11**, 742–751 (1974b).

D. L. Iglehart. Conditioned limit theorems for random walks. Academic Press, New York (1975).

W. D. Kaigh. An invariance principle for random walk conditioned by a late return to zero. Ann. Probability **4** (1), 115–121 (1976).

E. Le Page. Théorèmes limites pour les produits de matrices aléatoires. In Probability measures on groups (Oberwolfach, 1981), volume 928 of Lecture Notes in Math., pages 258–303. Springer, Berlin-New York (1982).

E. Le Page, M. Peigné and C. Pham. The survival probability of a critical multi-type branching process in i.i.d. random environment. Ann. Probab. (to appear).

P. Levy. Théorie de l'addition des variables aléatoires. Gauthier-Villars (1937).

M. Shimura. A class of conditional limit theorems related to ruin problem. Ann. Probab. **11** (1), 40–45 (1983).

Chapitre 2

The Galton-Watson branching process

This chapter discuss single-type and multi-type branching processes in fixed environment and in random environment. The original branching process was considered by Galton and Watson in the 1870's while seeking a quantitative explanation for the disappearing family names phenomenon, even in a growing population. In general it describes the evolution of successive generations of a "population" under the following assumptions :

1. *the initial generation has one member (the ancestor) ;*
2. *the number of children of every member in every generation of the population is random and follows the same offspring distribution denoted p ;*
3. *the random variables that represent the number of children of each member in all generations of the population are independent.*

The behavior of this process, in particular the probability of the extinction of the population, is governed by the value m of the expectation of the offspring distribution p. When $m < 1$ (subcritical case), the process becomes extinct almost surely, the same property holds for $m = 1$ (critical case) ; however, when $m > 1$ (supercritical case), the process may survive indefinitely.

During the 20th century, multi-type Galton-Watson processes was introduced since in several cases, the individuals involved in a branching process are not all alike. Several examples are presented in [Mode, 1971] ; for instance, we may mention models coming from

1. Population Genetics : when considering inheritance of alleles, a multi-type branching process with types corresponding to the genotypes, can be used as a model.
2. Physics : cosmic-ray cascades involve both electrons and photons, with electrons producing photons and photons producing electrons. Such an example is modeled by a two-type branching process.

In this model, the population has p types with $p \geq 2$ and the offspring distribution is replaced by a family $\mathbf{p} = (p^i)_{1 \leq i \leq p}$ of probability measures on $\mathbb{N}^p$ and the corresponding

expectation becomes a matrix $M = (m_{i,j})_{1\leq i,j\leq p}$. A natural assumption is to assume that this matrix M is positive regular : all entries are finite and there exists a natural number $n \geq 1$ such that all entries of M^n are positive. Then the Perron-Frobenius Theorem ensures that M has a positive eigenvalue of multiplicity 1 (the Perron-Frobenius root of M), denoted by ρ and the modulus of all other eigenvalues of M is strictly less than ρ. The behavior of multi-type Galton-Watson processes are similar to the single-type ones and depending to the value of ρ that they are distinguished into three cases : when $\rho < 1$ (subcritical case), when $\rho = 1$ (critical case) and when $\rho > 1$ (supercritical case).

Branching processes in random environment (BPRE) were introduced in the 1960's (see for instance [Smith et Wilkinson, 1969]) to describe the development of populations whose evolution may be affected by environmental factors ; they became a central topic of research.

In the single-type case in random environment, the behavior of these processes is mainly determined by the 1-dimensional random walk generated by the logarithms of the expected population sizes denoted by m_k, for $k \geq 0$, of the respective generations. The theory of fluctuations of random walks on $\mathbb{R}$ with i.i.d. increments allows to classify BPRE in three classes - supercritical, critical or subcritical - of single-type BPRE, according to the fact that the associated random walk tends to $+\infty$, oscillates or tends to $-\infty$ (see for instance the fundamental papers [Athreya et Karlin, 1971b] and [Athreya et Karlin, 1971a]). In particular, when $\mathbb{E}(|\log m_k|) < +\infty$, the BPRE is supercritical (resp., critical or subcritical) when $\mathbb{E}(\log m_k) > 0$ (resp., $\mathbb{E}(\log m_k) = 0$ or $\mathbb{E}(\log m_k) < 0$).

There exist numerous statements concerning the asymptotic behavior of the probability of non-extinction up to time n, the distribution of the population size conditioned on survival up to moment n, large deviation type results (see for instance [Afanasyev, 1993], [Bansaye et Berestycki, 2008], [Dyakonova *et al.*, 2004]). In the critical case, the branching process is degenerate with probability one [Athreya et Karlin, 1971b] ; with the assumption that the generating functions of the BPRE are linear-fractional, in 1976, M. V. Kozlov [Kozlov, 1977] proved that the probability of non-extinction up to time n is equivalent to $c_1/\sqrt{n}$ as $n \to +\infty$, for some explicit constant $c_1 > 0$; in 2003, in the general case, J. Geiger and G. Kersting [Geiger *et al.*, 2003] strengthened this result. Let us recall that the probability of non-extinction up to time n is equivalent to $1/n$ when the offspring distribution is fixed, that is to say it does not vary randomly ; in other words, branching processes in random environment die more slowly. In the supercritical and subcritical cases, similar studies were done (see for instance [Athreya et Karlin, 1971b], [Athreya et Karlin, 1971a] and [Geiger *et al.*, 2003]), we do not go into detail about these cases since they are outside the scope of this thesis.

It is of interest to prove analogues of the statements above for the multi-type BPRE. The main difficulty which appears is that the role of the random walk associated to the BPRE in this case is played by the logarithms of the norm of some $\mathbb{R}^p$-valued Markov chain whose increments are governed by i.i.d. random $p \times p$-matrices M_k for $k \geq 0$; the coefficients of these matrices M_k are non-negative and correspond to the expected population sizes of the respective generations, according to the types of the particles and their direct parents. Therefore, the multi-type BPRE constitute a relevant application area to the studies of products of random matrices. Nevertheless, as in the single-type case, the

set of multi-type BPRE is divided into three classes : they are supercritical (resp., critical or subcritical) when the upper Lyapunov exponent of the associated random matrices is positive (resp., null or negative) [Kaplan, 1974].

Specification of the asymptotic behavior of the probability of non-extinction up to time n of the multi-type BPRE is important. As for single-type BPRE, the first case was studied when the generating functions defining the random environment are linear-fractional ; it yields to explicit formulas which are easier to tackle.

In [Dyakonova, 1999], E. E. Dyakonova obtained in the critical case an equivalence of the survival probability at time n, under the restrictive assumption that the mean matrices M_k have a concordant deterministic Perron-Frobenius eigenvector. It happens for instance when the M_k are upper-triangular matrices with strictly decreasing coefficients on the diagonal. In this case, the behavior of the BPRE is governed by the one-dimensional random walk whose increments are the logarithm of the spectral radius of the M_k ; this allows one to apply the same techniques as in the single-type case.

To tackle the general case when the action of the random matrices M_k is strongly irreducible, limit theorems on the fluctuations of the norm of products of random matrices are required, which were recently achieved in [Grama *et al.*, 2014] and [Grama *et al.*, 2017] with asymptotic results on the tails of certain hitting time distribution.

Let us briefly mention two main approaches to answer the main question of this chapter : quenched and annealed. Under quenched approach, we fix some $\omega \in \Omega$ and the characteristics of a BPRE are treated as random variables whose distributions correspond to the choice of ω ; in other words, there is only one layer of randomness and the environment is fixed with ω. The annealed approach studies the mean values of the mentioned characteristics and can be viewed as an averaged analysis of the more detailed quenched one. This is the context of the work presented here.

In this chapter, we introduce several classical results for single-type and multi-type Galton-Watson processes in fixed and in random environment. The principal result is about the multi-type process in random environment, located at the last section.

2.1 Single-type Galton-Watson process in fixed environment

2.1.1 Model description

Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space. We are interested in a population which evolves with time and denote by Z_n the number of individuals in the n^{th} generation of this population. At each generation n, each individual $i \in \{1, ..., Z_n\}$ has N_i^n children, where $\left(N_i^n, i \geqslant 1,\, n \geqslant 0\right)$ are i.i.d. random variables with distribution $\mathbb{P}\left(N_i^n = k\right)$, $k \geqslant 0$. The number of individuals at generation $n+1$ is therefore :

$$Z_{n+1} = N_1^n + N_2^n + ... + N_{Z_n}^n,$$

with the convention if $Z_n = 0$ then $Z_{n+1} = 0$. Let $\mathcal{F}_n$ be the σ-algebra generated by N_l^k for $k \leq n$ and $l \geq 0$. Since Z_{n+1} is a sum of i.i.d. random variables, its distribution depends only on the value of Z_n and not on past values. More precisely, $(Z_n)_{n\geq 0}$ is a time-homogeneous

Markov chain. Having defined the process, we want to know some of its properties : the probability distribution and moments of Z_n ; the fluctuations of Z_n as $n \to \infty$; the probability that the random sequence $Z_0, Z_1, Z_2, ..$ eventually goes to zero. The following basic result was discovered by Watson and has been rediscovered a number of times since.

What are the basic assumptions ?

1. the initial generation numbered 0 has one member (the ancestor),
2. the number of children of every member in every generation of the population is random and follows a *offspring distribution* denoted $\mathbf{p}$,
3. the $\mathbf{p}$-distributed random variables that represent the number of children of each member in all generations of the population are independent.

We denote $\mathbf{p} = (p(k))_{k\geq 0}$ the distribution of the random variables $N_{n,i}$.

We consider some trivial cases.

1. If $p(0) = 0$ then $\mathbb{P}(\text{extinction}) = 0$; furthermore, if p is degenerate, that is $p(1) = 1$, then $Z_n = 1$ $\mathbb{P}$-a.s. for any $n \geq 0$.
2. If $p(0) = 1$ then $Z_n = 0$ $\mathbb{P}$-a.s. for any $n \geq 1$ and $\mathbb{P}(\text{extinction}) = 1$.
3. If $0 < p(0) < 1$ and $p(0) + p(1) = 1$, then $\mathbb{P}(Z_n \neq 0, Z_{n+1} = 0) = p(1)^n p(0)$, which yields
$$\mathbb{P}(\text{extinction}) = \sum_{n=0}^{+\infty} \mathbb{P}(Z_n \neq 0, Z_{n+1} = 0) = \frac{p(0)}{1 - p(1)} = 1.$$

Therefore it is more interesting to omit these 3 particular cases and assume that p satisfies the following assumption

$$0 < p(0) < 1 \quad \text{and} \quad p(0) + p(1) < 1. \tag{2.1.1}$$

We also assume that p is *non-degenerate*, that is $p(k) \neq 1$ for any $k \geq 0$, and *aperiodic* that is $\gcd\{k \in \mathbb{N}^* \mid p(k) > 0\} = 1$.

Denote by $f(s)$ the probability generating function of the offspring distribution $\mathbf{p}$ defined by $f(s) = \sum_{k=0}^{+\infty} p(k)s^k$, and by $g_n(s)$, for $n \geq 0$, the probability generating function of the random variable Z_n. Obviously $g_0(s) = f(s)$ and $g_n(s) = \mathbb{E}[s^{Z_n}]$. Moreover, under condition (2.1.1), the function f is strictly convex.

2.1.2 Classical properties

Proposition 2.1.1 *The function g_n is the n-fold composition of f, that is, for any $n \geq 0$,*

1. $g_0(s) = s$,
2. $g_{n+1}(s) = \underbrace{f \circ f \circ \ldots \circ f}_{n \text{ times}}(s)$.

Proof. Easily we see that $g_0(s) = \mathbb{E}[s^{Z_0}] = \mathbb{E}[s^1]$. The second statement is obtained as

follows, using the fact that N_l^n are independent on $\mathcal{F}_{n-1}$,

$$\begin{aligned} g_{n+1}(s) = \mathbb{E}\left(s^{Z_{n+1}}\right) &= \sum_{i=0}^{+\infty} \mathbb{E}\left[s^{N_1^n+\ldots+N_{Z_n}^n} 1_{[Z_n=i]}\right] \\ &= \sum_{i=0}^{+\infty} \mathbb{E}\left[s^{N_1^n+\ldots+N_i^n} 1_{[Z_n=i]}\right] \\ &= \sum_{i=0}^{+\infty} \mathbb{E}\left[\mathbb{E}\left[s^{N_1^n+\ldots+N_i^n} 1_{[Z_n=i]}|\mathcal{F}_n\right]\right] \\ &= \sum_{i=0}^{+\infty} \left[\mathbb{E}\left(s^{N_1^1}\right)\right]^i \mathbb{P}\left(Z_n = i\right). \end{aligned}$$

Since $\mathbb{E}\left(s^N\right) = f(s)$, we see that

$$g_{n+1}(s) = \sum_{i=0}^{+\infty} \left[f(s)\right]^i \mathbb{P}\left(Z_n = i\right) = \mathbb{E}\left[\left(f(s)\right)^{Z_n}\right] = g_n\left(f(s)\right).$$

By induction, we obtain : $g_n = \underbrace{f \circ f \circ \ldots \circ f}_{n \text{ times}}$, which is the n^{th} iterate of function f.

□

Proposition 2.1.1 enables us to calculate the generating function of Z_n, and hence the probability distribution of Z_n by simply computing the iterates of f. From the point of view of probability theory, the main value of Proposition 2.1.1 is that it enables us to calculate the generating function of Z_n thanks to what we can obtain various asymptotic laws of behavior for Z_n when n is large.

Proposition 2.1.2 *Assume that* $\mathbb{E}((N_1^1)^2) < \infty$. *We set* $m := \mathbb{E}(N_1^1)$ *and* $\sigma^2 := Var(N_1^1)$. *Then, for any* $n \geq 1$,

1. $\mathbb{E}(Z_n) = m^n$,
2. $Var(Z_n) = \begin{cases} n\sigma^2 & when \quad m = 1, \\ \sigma^2\left(m^{n-1}\dfrac{m^n-1}{m-1}\right) & when \quad m \neq 1. \end{cases}$

2.1.3 Classification of the Galton-Watson process

Definition 2.1.3 *A Galton-Watson process with mean offspring number m is said to be* super-critical *if* $m > 1$, critical *if* $m = 1$, *or* subcritical *if* $m < 1$.

We are interested in the probability of extinction of the population namely the behavior of the non-decreasing sequence $(\mathbb{P}(Z_n = 0))_n$ whose limit e is called the extinction probability.

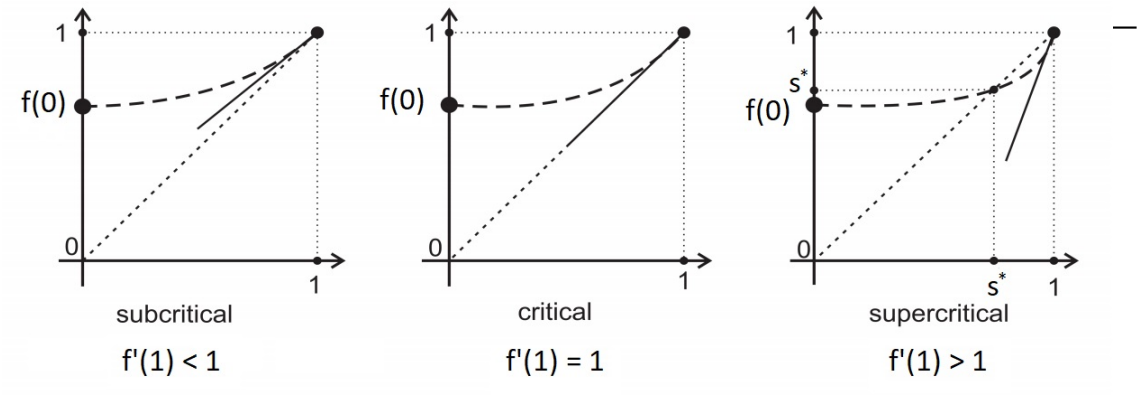

FIGURE 2.1 – Three cases

Proposition 2.1.4 *If $m \leq 1$, the extinction probability e is 1. If $m > 1$, the extinction probability is the unique non-negative solution less than 1 of the equation $f(s) = s$.*

Proof. One easily checks that f is strictly convex on $[0,1]$, so there are two cases to consider, as described in Figure 2.1 :

1. If the slope $f'(1) \leqslant 1$, the graph of $f(s)$ stands above the diagonal line and so, the equation $f(s) = s$ has no solution other than $s = 1$.
2. If $f'(1) > 1$ then the graph of $f(s)$ does intersect the diagonal line at some point other than $(1,1)$ so the solution is s^* with $s^* < 1$.

It is obvious that $(e_n)_n$ is a non-decreasing sequence bounded by 1. We also have $g_{n+1}(0) = e_{n+1} \to e$ and $g_n(0) = e_n \to e$. Since $g_{n+1}(0) = f(g_n(0))$, then $e = f(e)$.

For $n = 0$, we obtain $e_0 = \mathbb{P}(Z_0 = 0) = 0 \leqslant s^*$. Since f is an increasing function on $[0,1]$, one gets $e_1 = f(e_0) \leqslant f(s^*) = s^*$. By induction $e_n \leqslant s^*$ for any $n \geq 0$, and letting n tend to ∞, this yields $e \leqslant s^*$. This implies e is the smallest non-negative fixed-point of f. □

2.1.4 Speed of convergence

The first two precise results about the extinction of the GW process are due to [Kolmogorov, 1938] and the third one is due to [Kesten et Stigum, 1966].

Theorem 2.1.5 *Let $(Z_n)_{n\geq 0}$ be a GW process whose offspring distribution p satisfies (2.1.1) and has finite second moment.*

1. *In the subcritical case, the process $(Z_n)_{n\geq 0}$ becomes extinct $\mathbb{P}$-a.s. and there exists a positive constant c such that*
$$\mathbb{P}(Z_n \neq 0) \sim cm^n \quad as \quad n \to +\infty. \tag{2.1.2}$$
2. *In the critical case, the process $(Z_n)_{n\geq 0}$ becomes extinct $\mathbb{P}$-a.s. and*
$$\mathbb{P}(Z_n \neq 0) \sim \frac{2}{\sigma^2 n} \quad as \quad n \to +\infty. \tag{2.1.3}$$

3. *In the supercritical case, the martingale* $(Z_n/m^n)_{n\geq 0}$ *converges* $\mathbb{P}$*-a.s. to some random variable* W *which satisfies the two properties :*

$$\mathbb{E}(W) \leq 1 \qquad \textit{and} \qquad \mathbb{P}(W = 0) = e, \tag{2.1.4}$$

where $e \in]0,1[$ *is the probability of extinction of* $(Z_n)_{n\geq 0}$.

The statement concerning the supercritical case holds under a weaker moment assumption (namely when the N_i^n belong to $L^1 \ln^+ L^1$) and is due to Kesten and Stigum under this assumption; it is a bit out of the scope of this thesis, so we omit its proof. We refer to the book [Athreya et Ney, 1972] and readers can find a proof of another approach in [Abraham et Delmas, 2016], based on the notion of Galton-Watson trees.

Before giving the proof of the first two statements, let us say a few comments. In the subcritical and critical cases, the extinction probability at time n converge to 1, exponentially quickly in the subcritical case and only polynomially quickly (with speed $1/n$) in the critical one.

In the supercritical case, since the non-extinction event $\left(\bigcap_{n\geq 0}(Z_n \geq 1)\right)$ has probability $1 - e$, the events $(W > 0)$ and $\bigcap_{n\geq 0}(Z_n \geq 1)$ equal $\mathbb{P}$-almost surely. In other words, for $\mathbb{P}$-almost $\omega \in \Omega$,

1. either $Z_n(\omega) = 0$ for n large enough (this occurs with probability e);
2. or $(Z_n(w))_{n\geq 0}$ tends to $+\infty$ exponentially quickly, namely $Z_n(\omega) \sim m^n Z(w)$ as $n \to +\infty$, with $Z(\omega) > 0$ (this occurs with probability $1 - e$).

Proof. (1) Notice the two following points :

1. $\mathbb{P}(Z_n \geq 1) \leq \mathbb{E}[Z_n]$, where $\mathbb{E}[Z_n] = m^n$,
2.

$$\begin{aligned} \mathbb{P}(Z_{n+1} > 0) &= 1 - \mathbb{P}(Z_{n+1} = 0) \\ &= 1 - f(g_n(0)) \\ &= \frac{1 - f(g_n(0))}{1 - g_n(0)} \mathbb{P}(Z_n > 0). \end{aligned}$$

Since $f \in C^\infty[0,1]$ and f is continuous, there exists $s_n \in [g_n(0), 1]$ such that $f'(s_n) = \dfrac{1 - f(g_n(0))}{1 - g_n(0)}$. Moreover, $f'(s_n) < f'(1)$ by strict convexity of f, which yields

$$\mathbb{P}(Z_{n+1} > 0) \leq f'(1)\mathbb{P}(Z_n > 0) = m\mathbb{P}(Z_n > 0).$$

Hence, the sequence $(\mathbb{P}(Z_n > 0)/m^n)_n$ is decreasing, thus it converges to some non-negative limit c.

It remains to check that $c > 0$. By using the Cauchy-Schwartz inequality, we obtain $\mathbb{E}[Z_n] = \mathbb{E}[Z_n \mathbf{1}_{[Z_n>0]}] \leq \sqrt{\mathbb{E}[Z_n^2]\mathbb{P}(Z_n > 0)}$, together with Proposition 2.1.2, it follows that

$$\mathbb{P}(Z_n > 0) \geq \frac{(\mathbb{E}[Z_n])^2}{\mathbb{E}[Z_n^2]} = \frac{(1-m)m^{n+1}}{(1-m^n)\sigma^2 + (1-m)m^{n+1}},$$

which implies

$$\liminf_{n\to+\infty} \frac{\mathbb{P}(Z_n > 0)}{m^n} \geq \liminf_{n\to+\infty} \frac{(1-m)m}{(1-m^n)\sigma^2 + (1-m)m^{n+1}} = \frac{m(1-m)}{\sigma^2} > 0.$$

Hence $c \geq \dfrac{m(1-m)}{\sigma^2} > 0$.

(2) In critical case, we can see that $\sigma^2 = Var\xi = f''(1)$. Then by using Taylor expansion for $1-x$ and 1, we obtain $f(1-x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(a)}{n!}(-x)^n$, which yields

$$f(1-x) = 1 - x + \frac{\sigma^2}{2!}x^2(1+\epsilon(x)).$$

Let $v_n := \mathbb{P}(Z_1 > 0)$. By the two points in the first part of this proof, we see that $\lim_{n\to+\infty} v_n = 0$ and $v_{n+1} = 1 - f(1-v_n)$ for any $n \geq 0$. Moreover,

$$\begin{aligned}
\frac{1}{v_{n+1}} - \frac{1}{v_n} &= \frac{1}{1-f(1-v_n)} - \frac{1}{v_n} \\
&= \frac{1}{v_n}\frac{1}{1-\frac{\sigma^2}{2}v_n(1-\epsilon(v_n))} - \frac{1}{v_n} \\
&= \frac{1}{v_n}\left[1 + \frac{\sigma^2}{2}v_n(1+\epsilon(v_n))\right] - \frac{1}{v_n} \\
&= \frac{\sigma^2}{2}(1+\epsilon(v_n)) \to \frac{\sigma^2}{2},
\end{aligned}$$

which implies

$$\frac{1}{v_n} = \frac{1}{v_0} + \sum_{k=0}^{n-1}\left(\frac{1}{v_{k+1}} - \frac{1}{v_k}\right) \sim n\frac{\sigma^2}{2}. \tag{2.1.5}$$

Then $\mathbb{P}(Z_n > 0) \sim \frac{2}{n\sigma^2}$.

□

2.2 Multi-type Galton-Watson process in fixed environment

The materials in this section follows the style of [Harris, 1963]. We extend our problem to a little more difficult level : the Galton-Watson branching processes with many types of particles. The question considered is whether the asymptotic behavior of the process change compared to those of the single-type case.

In the reproduction of certain bacteria, the usual form may produce a mutant form that behaves differently. Another example is about mitochondrial DNA, it represents a mainstay of evolutionary biology. Normally, a normal mt DNA gives birth to two other normal mt DNA but someday, it changes. This time it separates into one normal and one abnormal which we may call mutant mt DNA. The noticeable feature is that only the normal mt DNA can do this, producing to either two normal, or one normal and one

mutant. Mutants only can give birth to mutants. This is just a very simple example. It can be a bit more complicated as follows.

We fix $p \geq 2$ and for each generation $n \geq 0$, we define Z_n as a p-dimensional vector $(Z_n(1), \dots Z_n(p))$ whose j^{th} entry gives the number of individuals of type j in the n^{th} generation of the branching process. We write $(Z_n)_{n\geq 0}$ for the sequence of the random variables Z_n for $n \geq 0$. Because all individuals reproduce independently and offspring distributions do not change between generations, the random vector Z_n depends only on the vector Z_{n-1}. In other words, the Galton-Watson process $(Z_n)_{n\geq 0}$ is a temporally homogeneous vector Markov chain.

Consider a p-dimensional vector $\tilde{\alpha} = (\alpha_1, \dots, \alpha_p) \in \mathbb{N}^p$; the vector $\tilde{\alpha}$ represents the offspring of each type created by a individual, with α_1 equalling the number of offspring of type $1, \dots,$ and α_p the number of offspring of type p. For every type $i = 1, 2, ..., p,$, let $\mathbf{p}^i(\tilde{\alpha})$ be the probability that an individual of type i has α_1 children of type $1, \dots, \alpha_p$ children of type p. The probability measure $\mathbf{p}^i$ is the offspring distribution for individuals of type i.

Let us assume that $Z_0 = \tilde{e}_i = (0, \dots, 1, \dots, 0)$ (all the entries equal 0 except the i^{th} one which equals 1). Then Z_1 has distribution $\mathbf{p}^i$ whose p-dimensional generating function is of the form

$$f^i(\mathbf{s}) = f^i(s_1, \dots, s_p) = \sum_{\tilde{\alpha}\in\mathbb{N}^p} p^i(\alpha_1, \dots, \alpha_p) s_1^{\alpha_1} \dots s_p^{\alpha_p}.$$

It is obvious that if $Z_n = (\alpha_1, \dots, \alpha_p)$ then Z_{n+1} is a sum of $\alpha_1 + \dots + \alpha_p$ vectors and if $Z_n = 0$ then $Z_{n+1} = 0$. The generating function of Z_n given $Z_0 = \tilde{e}_i$ is denoted f_n^i and we set $f_n(\mathbf{s}) = (f_n^1(\mathbf{s}), \dots, f_n^p(\mathbf{s}))$.

Theorem 2.2.1 *The generating function f_n^i are functional iterates as follows :*

$$\begin{aligned} f_{n+1}^i(\mathbf{s}) &= f^i(f_n^1(\mathbf{s}), \dots, f_n^p(\mathbf{s})), \quad n = 0, 1, \dots \\ f_0^i(\mathbf{s}) &= s_i, \qquad i = 1, \dots, p. \end{aligned}$$

In particular, $\mathbf{f}_{n+k}(\mathbf{s}) = \mathbf{f}_n[\mathbf{f}_k(\mathbf{s})]$ for any $n, k = 0, 1, \dots$

Firstly, we state several basic definitions. The multi-type process $(Z_n)_{n\geq 1}$ is irreducible if for every pair of types i and j in $\{1, \dots, p\}$, there exists some natural number n such that $\mathbb{P}[Z_n(j) \geq 1 | Z_0 = \tilde{e}_i] > 0$. If for some n the statement holds for all i and j, then the process is said to be positively regular.

From here to the end of this subsection, we assume that the distribution $\mathbf{p}$ has moment of order 2. Thus, the function $f = (f^1, \dots, f^p)$ is twice differentiable, let $M = (M(i, j))_{i,j}$ be the matrix of the first moments of the offspring generating function f,

$$M(i, j) = \mathbb{E}[Z_1(j) | Z_0 = \tilde{e}_i] = \frac{\partial f^i(1, \dots, 1)}{\partial s_j}, \quad i, j = 1, \dots, p.$$

Assume also that $n_{i,j}$ are finite and they are not all equal to 0. A straightforward generalization yields to the formula $\mathbb{E}[Z_{n+k} | Z_k] = Z_k M^n$, for any $n, k = 0, 1, 2, \dots$

A vector and a matrix are said to be *positive* if all their entries are positive.

A Galton-Watson process, or its mean matrix M, are said positively regular when there exists $k \geq 1$ such that M^k is positive; this condition is stricter than the process

being positively regular . The process is said to be *singular* if the generating functions $f^i(s_1, \ldots, s_p)$, for $1 \leq i \leq p$, are all linear in $s_1, \ldots, s_p$, with no constant terms; that is to say each individual has exactly one child.

The following classical result for positively regular matrices will be of interest in the sequel.

Theorem 2.2.2 *[Harris, 1963] (Perron-Frobenius) Let M be a non-negative matrix of order p such that M^k is positive for some positive integer k . Let ρ be its spectral radius. Then ρ is a simple eignevalue of M, greater in module than any other eigenvalues; it corresponds to positive right and left eigenvectors $\mu = (\mu_i)_{1\leq i\leq p}$ and $\nu = (\nu_i)_{1\leq i\leq p}$. Moreover, for $n \geq 1$,*

$$M^n = \rho^n M_1 + M_2^n,$$

where $M_1 = (\mu_i \nu_j)_{1\leq i,i\leq p}$ and $M_2 = M - \rho M_1$, with the normalisation $\sum_{i=1}^p \mu_i \nu_i = 1$. Hence $M_1 M_1 = M_1$, $M_1 M_2 = M_2 M_1 = 0$ and $\rho(M_2) < \rho(M) = \rho$.

As we may expect, the eigenvalue ρ plays a role similar to the expectation m in one type case; more precisely, it is the required indicator of criticality.

Let $\mathbf{e} = (e^1, \ldots, e^p)$ be the extinction vector probability, where e^i is the extinction probability given $Z_0 = \tilde{e}_i$:

$$e^k = \lim_{n\to+\infty} \mathbb{P}(Z_n = \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i).$$

If u and v are two vectors, then $u > v$ (or $u \geq v$) if $u - v$ has all its coefficients positive (or non-negative).

Theorem 2.2.3 *Suppose the process is positive regular and not singular. If $\rho \leq 1$, then $\mathbf{e} = \mathbf{1}$. If $\rho > 1$, then $\mathbf{0} \leq \mathbf{e} < \mathbf{1}$ and $\mathbf{e}$ satisfies the equation $\mathbf{f}(\mathbf{e}) = \mathbf{e}$. Moreover, if $\mathbf{e}_1$ is any vector in the unit cube other than $\mathbf{1}$, then $\lim_{n\to+\infty} \mathbf{f}_n(\mathbf{e}_1) = \mathbf{e}$. More precisely, the only solutions of the equation $\mathbf{f}(\mathbf{e}) = \mathbf{e}$ in the unit cube are $\mathbf{e}$ and $\mathbf{1}$.*

Remarks. 1. If the process is singular, then M is a Markov matrix with $\rho = 1$ and $\mathbb{E}[Z_n] = Z_0$ for any $n \geq 1$. It is obvious that the population survives forever.

2. For process that is not positive regular, problems arise similar to those encountered in the theory of Markov chains with decomposable matrices, that is matrices similar to a block-diagonal ones.

Here we state the results for the three regime with speed of convergence. We refer to Chapter 5 of the book [Athreya et Ney, 1972] for details and comments.

Theorem 2.2.4 *Let $(Z_n)_{n\geq 0}$ be a positively regular multi-type Galton-Watson process with moment of order 2, that is $\mathbb{E}(|Z_1|^2/Z_0 = \tilde{e}_i) < +\infty$ for any $1 \leq i \leq p$.*

1. *In the subcritical case ($\rho < 1$), the process $(Z_n)_{n\geq 0}$ becomes extinct $\mathbb{P}$-a.s. and for any $1 \leq i \leq p$, there exists a positive constant c_i such that*

$$\mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) \sim c_i \rho^n \quad as \quad n \to +\infty. \tag{2.2.1}$$

2. *In the critical case ($\rho = 1$), the process $(Z_n)_{n\geq 0}$ becomes extinct $\mathbb{P}$-a.s. and, for any $1 \leq i \leq p$, there exists a positive constant c_i such that*

$$\mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) \sim \frac{c_i}{n} \quad as \quad n \to +\infty. \tag{2.2.2}$$

3. *In the supercritical case ($\rho > 1$), the sequence $(Z_n/\rho^n)_{n\geq 0}$ converges $\mathbb{P}$-a.s. to $W\nu$ where W is a nonnegative random variable and ν is a left eigenvector of M associated with ρ; furthermore, $\mathbb{P}(W = 0|Z_0 = \tilde{e}_i) = e^i$ where $\mathbf{e} = (e^i)_{1\leq i\leq p}$ is the extinction probability of $(Z_n)_{n\geq 0}$.*

The constants c_i are given explicitly in [Athreya et Ney, 1972], in term of the Hessian matrices of the generating functions f_i and the eigenvectors μ and ν of the mean matrix M. The convergence in supercritical case holds under weaker assumption moment (moment of order $L^1 \ln^+ L^1$ are only required), as in the one type case.

2.3 Single-type Galton-Watson process in random environment

In the previous section, we present a model for the size over time of a population which evolves in a fixed environment. Indeed, the offspring distribution f is always the same : the members of the population produce and die according to the same law of chance. Furthermore, they do not interfere with one another. Unfortunately, natural processes of multiplication are often affected by many factors which introduce variations over time and also dependencies. In the nature, plant and animal populations depend on their surrounding environment which is constantly changing. Their survival depends on environmental conditions such as food, water availability and temperature. Under favorable environment, the number of offspring increases, otherwise it declines.

In this section, we take into account some variations of the offspring distributions over time; more precisely, a natural idea is to assume that the environment varies randomly and i.i.d., that is to replace the offspring distribution μ by a sequence of i.i.d. random probability meausre $(\mu_i)_{i\geq 0}$, or equivalently replace the generating function f by a sequence $(f_n)_{n\geq 0}$ of i.i.d. random generating functions. The random variable f_n is the generating function of the random offspring distribution of members of generation n.

One important remark is that there are two level of randomness that arise. Firstly, we choose a random family of distributions $(f_n)_{n\geq 0}$, then according to these distributions, the branching process $(Z_n)_n$ evolves stochastically, that is to say the particles at time n take f_n as their reproduction law. After, the number of offspring of each particle is again a random variable.

2.3.1 Model description

Our population starts with one ancestor. The particles at time n produce children with the offspring distribution f_n. Let Z_n denotes the size of the population at time n. The generating function of Z_n can be expressed by a product of generating functions :

$$\mathbb{E}\left(s_n^Z | Z_0, \ldots, Z_{n-1}, f_0, \ldots, f_{n-1}\right) = f_{n-1}(s)^{Z_{n-1}}, \tag{2.3.1}$$

which yields

$$\mathbb{E}[s^{Z_n}|f_0,\dots,f_{n-1}] = f_0(f_1(\dots f_{n-1}(s)\dots)).$$

Limiting to the interesting case, we assume

$$\mathbb{P}(\mu_0(0) < 1) = 1 \quad \text{and} \quad \mathbb{P}(\mu_0(0) + \mu_0(1) < 1) > 0. \tag{2.3.2}$$

2.3.2 Classification

Given the environment $\bar{f} = (f_n)_{n\geq 0}$, the "quenched" probability of survival q_n at time n is

$$q_n(\bar{f}) := \mathbb{P}(Z_n > 0|f_0,\dots,f_{n-1}) = 1 - f_0(f_1(\dots f_{n-1}(0)\dots)), \quad n \geq 1. \tag{2.3.3}$$

The "annealed probability" of extinction at time n is obtained averaging on the environment

$$\mathbb{P}(Z_n > 0) = \mathbb{E}(q_n(\bar{f})) = \mathbb{E}\Big(1 - f_0(f_1(\dots f_{n-1}(0)\dots))\Big). \tag{2.3.4}$$

Since the sequence $(q_n(\bar{f}))_n$ decreases $\mathbb{P}$-a.s., then it converges $\mathbb{P}$-a.s. to some limit $q(\bar{f})$ which is called the survival probability conditioned on the environment $\bar{f}$. The extinction of the Galton-Watson process is given by the following statement whose proof is detailed in [Athreya et Karlin, 1971b].

Theorem 2.3.1 *(1) Assume that* $\mathbb{E}\big(\log f_0'(1)\big)^+ < +\infty$. *If* $\mathbb{E}\Big(\log f_0'(1)\Big) \leq 0$ *then*

$$\mathbb{P}\Big(q(\bar{f}) = 0\Big) = 1.$$

In other words, conditioned on almost all environment $\bar{f}$, *the process* $(Z_n)_{n\geq 0}$ *becomes extinct a.s.*

(2) Assume $\mathbb{E}\big(\log f_0'(1)\big)^+ < +\infty$ *and* $\mathbb{E}(-\log(1 - f_0(0)) < +\infty$.[1] *If* $\mathbb{E}\Big(\log f_0'(1)\Big) > 0$ *then*

$$\mathbb{P}\Big(q(\bar{f}) > 0\Big) = 1.$$

In other words, conditioned on almost all random environment $\bar{f}$, *the process* $(Z_n)_{n\geq 0}$ *survives with some strictly positive probability.*

Consequently, we may classify the Galton-Watson process in random environment by the value of $\mathbb{E}\log f_0'(1)$ (when it exists, that is when $\mathbb{E}(\log f_0'(1))^+ < +\infty$). The Galton-Watson process is said to be :

(i) *subcritical* if $\mathbb{E}\Big(\log f_0'(1)\Big) < 0$;

(ii) *critical* if $\mathbb{E}\Big(\log f_0'(1)\Big) = 0$;

(iii) *supercritical* if $\mathbb{E}\Big(\log f_0'(1)\Big) > 0$ (and $\mathbb{E}(-\log 1 - f_0'(0)) < +\infty$).

1. The condition $\mathbb{E}(-\log(1-f_0(0)) < +\infty$ is not natural at the first sight but it appears in a crucial way in the proof in [Athreya et Karlin, 1971b] ; it is a way to control that the random probability measures corresponding to the functions $f_0, f_1, \dots$ are not close to the Dirac mass at 0, whose generating function equals 1 on $[0,1]$.

2.3.3 The particular case of linear-fractional generating functions

We consider the case when the offspring generating function are linear-fractional, that is when it is of the form

$$f(s) := 1 - \frac{\alpha}{1-\beta} + \frac{\alpha s}{1-\beta s}, \quad \forall z \in [0,1], \tag{2.3.5}$$

where $0 < \alpha, \beta < 1$ and $\frac{\alpha}{1-\beta} \leq 1$. In other words, the function f is the generating function of the probability measure $\mathbf{p}$ which is the convex combination $a\delta_0 + (1-a)\mathcal{G}_p$ of the Dirac mass δ_0 at 0 and the geometric distribution $\mathcal{G}_p$ with parameter $1-\beta$; obviously, $\mathbf{p}$ is a probability distribution if and only if $\frac{\alpha}{1-\beta} \leq 1$ (in this case, we have $\alpha = (1-a)p$ and $\beta = 1-p$).

In the context of random environment, we consider a sequence $\bar{f} = (f_n)_{n\geq 0}$ of i.i.d. random variables with values in the set of linear-fractional generating functions defined by :

$$\forall x \in [0,1[\quad f_n(x) := 1 - \frac{\alpha_n}{1-\beta_n} + \frac{\alpha_n x}{1-\beta_n x}.$$

Equivalently, $((\alpha_n, \beta_n))_{n\geq 0}$ is a sequence of i.i.d. random variables with values in the set $\{(\alpha,\beta) \mid 0 < \alpha, \beta < 1 \text{ and } \frac{\alpha}{1-\beta} \leq 1\}$. This class of generating offspring functions for Galton-Watson processes is extensively studied, in fixed and random environment; not only each function has linear-fractional form but also the iterates $f_1(f_2 \dots f_n(s)\dots)$.

Besides, the first and second moments of f can be found explicitly : $f'(1) = \frac{\alpha}{(1-\beta)^2}$ and $f''(1) = \frac{2\alpha\beta}{(1-\beta)^3}$. Then formula (2.3.5) may be rewritten as

$$\forall s \in [0,1] \qquad 1 - f(s) = \left(\frac{1}{f'(1)(1-s)} + \frac{f''(1)}{2f'(1)^2}) \right)^{-1}. \tag{2.3.6}$$

The work [Kozlov, 1977] by Kozlov is the first main contribution to the study of the speed of convergence to the extinction of critical Galton-Watson in i.i.d. random environment; he considered only linear-fractional offspring distributions. The starting point of his proof is the following formula for the superposition of linear-fractional generating functions, based on (2.3.6) : for any $0 \leq s \leq 1$, setting $\prod_0 = 1$ and $\prod_j := \prod_{i=0}^{j-1} f_i'(1)$ for $j \geq 1$, it follows that

$$1 - f_0(\cdots f_{n-1}(s)\cdots) = \left(\frac{1}{(1-s)\prod_n} + \sum_{j=0}^{n-1} \frac{1}{\prod_j} \frac{f_j''(1)}{2f_j'(1)^2} \right)^{-1}. \tag{2.3.7}$$

Thanks to this formula, Kozlov succeeded in translating the probability of survival into

the language of one dimensional random walks. Namely,

$$\begin{aligned} q_n(\bar{f}) &:= \mathbb{P}(Z_n > 0 \mid f_0, \cdots, f_{n-1}) \\ &= 1 - f_0(\cdots f_{n-1}(0)\cdots) \\ &= \left(e^{-S_n} + \sum_{j=0}^{n-1} \eta_{j+1} e^{-S_i} \right)^{-1}, \end{aligned} \tag{2.3.8}$$

where for any $j \geq 1$,

(i) $S_j := \ln(f_0 \circ \ldots \circ f_{j-1})'(1) = \sum_{i=1}^{j} X_i$ with $X_i = \ln f_{i-1}'(1)$ for $i \geq 1$;

(ii) $\eta_j = \frac{f_{j-1}''(1)}{2f_{j-1}'(1)^2}$.

A similar decomposition exists when the f_n are not linear-fractional; see [Geiger et Kersting, 2000]. It is well known that, under suitable moment conditions on the X_j and η_j and when the X_j are centered, the quantity $\sum_{j=0}^{n-1} \eta_{j+1} e^{-S_i}$ behaves as the minimum of the partial sums $S_0 = 0, \ldots S_n$, see [Grincevičjus, 1974]; thus, the study of $q_n(\bar{f})$ is connected to the fluctuation theory of the one-dimensional random walk $(S_n)_{n\geq 0}$.

2.3.4 Statements of results

In this section, we just state some main results about the speed to extinction of Galton-Watson processes in i.i.d. random environment. There exists a huge literature on the subject, in particular the ones with weaker moment assumptions (see [Afanasyev *et al.*, 2005] and references therein). We begin with a result on critical single-type branching process in random environment.

Theorem 2.3.2 *[Geiger et Kersting, 2000](Critical case) Suppose* $\mathbb{E}\log f_0'(1) = 0, 0 < \mathbb{E}(\log f'(1))^2 < +\infty$ *and* $\mathbb{E}\left(\frac{f''(1)}{f'(1)^2}(1 + \log_+ f'(1))\right) < +\infty$. *Then, for some* $0 < \beta < +\infty$,

$$\mathbb{P}(Z_n > 0) = \mathbb{E}[q_n(\bar{f})] \sim \frac{\beta}{\sqrt{n}} \quad as \quad n \to +\infty.$$

Let us emphasize that $(\mathbb{P}(Z_n > 0))_{n\geq 0}$ converges to 0 more slowly than in the fixed environment, where the speed of convergence to extinction was $1/n$, up to some constant. This can be explained by the averaging effect induced by the random environment.

In the subcritical case, there are three subcases to consider with different speeds of convergence to extinction; this is an important difference compared to the fixed environment case. We refer for instance to [Geiger *et al.*, 2003] among many other articles existing on this subject.

Theorem 2.3.3 *(Subcritical case) [Geiger* et al., *2003] Let* $\mathbb{E}(\log f_0'(1)) < 0$.

1. **Strongly subcritical case.**
Assume that
$$\mathbb{E}(f_0'(1)\log f_0'(1)) < 0 \quad \textit{and} \quad \mathbb{E}(Z_1 \log^+ Z_1) < +\infty,$$
then there exists $0 < \beta_1 \leq 1$ *such that*
$$\mathbb{P}(Z_n > 0) \sim \beta_1 \Big(\mathbb{E}(f_0'(1)\Big)^n \quad \textit{as} \quad n \to +\infty.$$
Moreover,
$$\lim_{n\to+\infty} \mathbb{P}(Z_n = k | Z_n > 0) = q_1(k), \quad k \geq 1,$$
where
$$\sum_{k\geq 1}^{+\infty} q_1(k) = 1 \quad \textit{and} \quad \sum_{k\geq 1}^{+\infty} k q_1(k) < +\infty.$$

2. **Intermediate subcritical case.**
Assume that
$$\mathbb{E}(f_0'(1)\log f_0'(1)) = 0,$$
$$\mathbb{E}(f_0'(1)\log^2 f_0'(1)) < +\infty \quad \textit{and} \quad \mathbb{E}((1+\log^- f_0'(1))f_0''(1)) < +\infty,$$
then there exists $0 < \beta_2 < +\infty$ *such that*
$$\mathbb{P}(Z_n > 0) \sim \frac{\beta_2}{\sqrt{n}} \Big(\mathbb{E}(f_0'(1)\Big)^n \quad \textit{as} \quad n \to +\infty.$$
Moreover,
$$\lim_{n\to+\infty} \mathbb{P}(Z_n = k | Z_n > 0) = q_2(k), \quad k \geq 1,$$
where $\displaystyle\sum_{k\geq 1}^{+\infty} q_2(k) = 1.$

3. **Weakly subcritical case.**
Assume
$$0 < \mathbb{E}(f_0'(1)\log f_0'(1)) < +\infty$$
and the distribution of $\log f_0'(1)$ *is not supported by any non-centered lattice, and that the following integrability conditions are satisfied,*
$$\mathbb{E}\left(\frac{f_0''(1)}{f_0'(1)^{1-\epsilon}}\right) < +\infty, \quad \textit{and} \quad \mathbb{E}\left(\frac{f_0''(1)}{f_0'(1)^{2-\epsilon}}\right) < +\infty.$$
Then there exists $0 < \beta_3 < +\infty$ *and* $\gamma \in]0,1[$ *such that*
$$\mathbb{P}(Z_n > 0) \sim \frac{\beta_3}{n^{3/2}}\gamma^n \quad \textit{as} \quad n \to +\infty. \tag{2.3.9}$$
Moreover,
$$\lim_{n\to+\infty} \mathbb{P}(Z_n = k | Z_n > 0) = q_3(k), \quad k \geq 1,$$
where $\displaystyle\sum_{k\geq 1}^{+\infty} q_3(k) = 1.$

Let us just explain briefly how these three subcases appear ; we do not get into the details. Firstly, the function $\phi : t \mapsto \mathbb{E}(e^{t \log f_0'(1))})$ is well defined, convex on $[0,1]$ and it satisfies

$$\phi(0) = 1, \quad \phi(1) = \mathbb{E} f_0'(1), \quad \phi'(0) = \mathbb{E}\ \log f_0'(1) \quad \text{and} \quad \phi'(1) = \mathbb{E}\ f_0'(1) \log f_0'(1).$$

Consequently,

1. in the strongly subcritical case, the function ϕ strictly decreases on $[0,1]$, it reaches its minimum at 1 and $\phi'(1) < 0$; in particular $\phi(1) = \mathbb{E} f_0'(1) < 1$, and by convexity $\mathbb{E} \log f_0'(1) < 0$.
2. in the intermediate subcritical case, the function ϕ reaches its minimum at 1 and $\phi'(1) = 0$; again $\phi(1) = \mathbb{E} f_0'(1) < 1$, and by convexity $\mathbb{E} \log f_0'(1) < 0$.
3. in the weakly subcritical case, the function ϕ reaches its minimum at some $t_0 \in]0,1[$ and $\phi'(t_0) = 0$; the constant γ in (2.3.9) equals $\phi(t_0)$ and belongs to $]0,1[$. Notice that $\phi(1) = \mathbb{E} f_0'(1)$ may be greater than 1 in this case.

The proof of Theorem 2.3.3 is based on a classical argument in the theory of random walk : using the function ϕ, one changes the distribution of the increments of the random walk $(\log \Pi_n)_{n\geq 1}$. In the first sub-case, the new random walk has a negative first moment, which corresponds to the rate of convergence to 0 of the process $(Z_n)_{n\geq 1}$; in the intermediate one, the new random walk is centered, that's why the quantity $\frac{1}{\sqrt{n}}$ appears ; at last, in the weakly case, the change of measure yields also to a centered random walk and the coefficient $1/n^{3/2}$ comes from classical results on fluctuations of such random walks.

Eventually, in the supercritical case, a version of Kesten-Stigum's theorem [Kesten et Stigum, 1966] completes the picture of analogues in random environment. Many works were done since the appearence of the couple of seminal papers by K. B. Athreya and S. Karlin [Athreya et Karlin, 1971b], [Athreya et Karlin, 1971a] ; let us state their version in random environment of Kesten-Stigum's theorem , base on the fact that $(Z_N/\Pi_n)_{n\geq 0}$ is a martingale.

Theorem 2.3.4 *(supercritical case) [Athreya et Karlin, 1971a] Assume that* $\mathbb{E}(\log f_0'(1)) > 0$*; then the process* $(Z_n/\Pi_n)_{n\geq 0}$ *is a non-negative martingale with respect to the filtration* $(\sigma(Z_0, \ldots, Z_n, f_0, f_1 \ldots, f_n))_{n\geq 0}$ *and hence* $\lim_{n\to+\infty} Z_n/\Pi_n = W$ *exists* $\mathbb{P}$*-a.s.*

Suppose in addition that $\mathbb{E}\left(\dfrac{f_0''(1)}{f_0'(1)}\right) < +\infty$. [2]

Then $\mathbb{E}[W|\bar{f}] = 1$ *and* $\mathbb{P}(W \neq 0|\bar{f}) = q(\bar{f})$ $\mathbb{P}$*-a.s.*

In others words, for almost all environment $\bar{f}$, the process $(Z_n)_{n\geq 0}$, given this environment $\bar{f}$, becomes extinct with probability $1 - q(\bar{f})$ or behave as $e^{S_n} W \to +\infty$ a.s. with probability $q(\bar{f})$.

2. Notice that weaker moment assumptions of type $L^1 \log L^1$ are required, as for the Kesten-Stigum's theorem in fixed environment.

2.4 Multi-type Galton-Watson process in random environment

E.E. Dyakonova studied multi-type Galton-Watson process in random environment under some restrictive assumption : the mean matrices of the random environment have a common eigenvector. For example, those matrices are upper triangular with strictly decreasing entries on the diagonal. In this case, the behavior of the BPRE is governed by the one-dimensional random walk $(S_n)_{n\geq 0}$ with increments $\log \rho(M_k)$. The strategy of her proof is based on the same techniques as in single-type case and the study of the extinction probability $q_n(\bar{f})$ is connected to the fluctuation theory of the one-dimensional random walk $(S_n)_{n\geq 0}$.

In the general case, that is when the action of the random matrices is strongly irreducible, the one-dimensional random walk $(S_n)_{n\geq 0}$ mentioned above is replaced by the process $(\log |\tilde{x} M_0 \cdots M_{n-1}|)_{n\geq 0}$, for some row vector $\tilde{x}$ with positive entries. To control the fluctuations of this process, we apply the result presented in the previous section. This is the main "new" ingredient to study the asymptotic behavior of $q_n(\bar{f})$ in the multitype case. Nevertheless, we have to restrict our study to the case when the matrices M_k belong to some proper subset $S^+(B)$ of the semi-group S^+ of $p \times p$ matrices with positive entries; namely,

$$S^+(B) = \{M = (M(i,j))_{1\leq i,j\leq p} \mid \frac{1}{B} \leq \frac{M(i,j)}{M(k,l)} \leq B \text{ for all } 1 \leq i,j,k,l \leq p\},$$

where B is a positive constant. Let us explain how this restriction is neccessarry and present the main argument we need to achieve the proof; this corresponds to Lemma 3.1 in the article presented at the end of this section.

Firstly, in [Geiger et Kersting, 2000], J. Geiger and G. Kersting prove that for any real constant C,

$$\sum_{n\geq 0} \mathbb{E}\left[e^{-S_n}; S_1 > C, \ldots, S_n > C\right] < +\infty \tag{2.4.1}$$

holds, where $S_n = Y_1 + \ldots + Y_n$ for i.i.d. random variables Y_k on $\mathbb{R}$ for $k \geq 1$. This is an important step in their work and they prove that it can be reduced to the case when $C = 0$.

In this case, inequality (2.4.1) is a direct consequence of the so-called "duality property" which states that the random vectors $(Y_1, \ldots, Y_n)$ and $(Y_n, \ldots, Y_1)$ have the same distribution; consequently, the same property holds for the vectors $(S_1, S_2, \ldots, S_n)$ and $(Y_n, Y_n + Y_{n-1}, \ldots, Y_n + \cdots + Y_1)$. Hence,

$$\begin{aligned}
&\mathbb{E}\left[e^{-S_n}; S_1 > 0, \ldots, S_n > 0\right] \\
&\quad = \mathbb{E}\left[e^{-S_n}; Y_1 > 0, Y_1 + Y_2 > 0, \ldots, Y_1 + \cdots + Y_n > 0\right] \\
&\quad = \mathbb{E}\left[e^{-S_n}; Y_n > 0, Y_n + Y_{n-1} > 0, \ldots, Y_n + \cdots + Y_1 > 0\right] \\
&\quad = \mathbb{E}\left[e^{-S_n}; S_n > S_{n-1}, S_n > S_{n-2}, \ldots, S_n > 0\right] \\
&\quad = \mathbb{E}\left[e^{-S_n}; n \text{ is a strict ladder epoch of the random walk } (S_k)_{k\geq 0}\right].
\end{aligned}$$

By a straightforward argument, one easily deduces that (2.4.1) holds when $C = 0$.

After, let us consider the multi-type case in random environment when S_n is replaced by $S_n(\tilde{x}) = \log|\tilde{x}R_n|$, where $R_n = M_0 \dots M_{n-1}$ for any $n \geq 0$ and $\tilde{x}$ is a row vector with positive entries with L^1-norm equal to 1 (the set of such $\tilde{x}$ is denoted by $\mathbb{X}$ in the following article). Following the same strategy, we want to overestimate the quantity

$$\mathbb{E}\Big[|\tilde{x}R_n|^{-1}; |\tilde{x}R_1| > C, \dots, |\tilde{x}R_n| > C\Big].$$

The "duality property" $\mathcal{L}(M_0, \dots, M_{n-1}) = \mathcal{L}(M_{n-1}, \dots, M_0)$ yields that two following events are equal in distribution

$$[|\tilde{x}R_k| > C] \quad \stackrel{\text{dist}}{=} \quad [|\tilde{x}M_{n-1} \dots M_{n-k}| > C].$$

Indeed, firstly, notice that the "right product" R_k is replaced by the left product $M_{n-1} \dots M_{n-k}$; secondly, the quantity $|\tilde{x}M_{n-1} \dots M_{n-k}|$ cannot be expressed in terms of $\tilde{x}L_n$ and $\tilde{x}L_{n-k}$ as in the case of classical random walks, where $L_n = M_{n-1} \dots M_0$. The difficuty comes from the starting vector $\tilde{x}$, so in order to overcome this problem, we try to "forget" x by temporarily replace the quantity $|\tilde{x}R_n|$ by $|R_n|$ (and similarly $|L_n x|$ by $|L_n|$, where x denotes the column vector corresponding to $\tilde{x}$), which is possible when the matrices M_n belong to $S^+(B)$, since in this case the quantities $|\tilde{x}R_n|$ and $|R_n|$ are "comparable" in the following sense (see Lemma 1.3.2) : there exists a positive constant c greater than 1 depending on B such that, for any $\tilde{x} \in \mathbb{X}$ and any $n \geq 1$,

$$\frac{1}{c}|R_n| \leq |\tilde{x}R_n| \leq c|R_n| \quad \text{(similarly} \quad \frac{1}{c}|L_n| \leq |L_n x| \leq c|L_n|)$$

(one write $|R_n| \overset{c}{\asymp} |\tilde{x}R_n|$ and $|L_n| \overset{c}{\asymp} |L_n x|$ for short). As a direct consequence, changing if necessary the value of c, it also holds $|L_n| \overset{c}{\asymp} |M_{n-1} \dots M_{n-k}||L_{n-k}|$ for any $0 \leq k \leq n$, by using Lemma 1.3.2. Hence, in distribution

$$[|\tilde{x}R_k| > C] \subset \left[|R_k| > \frac{C}{c}\right] \quad \stackrel{\text{dist}}{=} \quad [|M_{n-1} \dots M_{n-k}| > \frac{C}{c}]$$

with

$$\begin{aligned}\left[|M_{n-1} \dots M_{n-k}| > \frac{C}{c}\right] &\subset \left[|L_n| > \frac{C}{c^2}|L_{n-k}|\right] \\ &\subset \left[|L_n x| > \frac{C}{c^4}|L_{n-k}x|\right].\end{aligned}$$

More generally, it holds

$$\begin{aligned}&\mathbb{E}\Big[|\tilde{x}R_n|^{-1}; |\tilde{x}R_1| > C, \dots, |\tilde{x}R_n| > C\Big] \\ &\qquad \leq c\mathbb{E}\Big[|L_n x|^{-1}; |L_n x| > \frac{C}{c^4}|L_{n-1}x|, \dots, |L_n x| > \frac{C}{c^4}|x|\Big].\end{aligned}$$

If $C \geq c^4$, it holds

$$\mathbb{E}\Big[|\tilde{x}R_n|^{-1}; |\tilde{x}R_1| > C, \dots, |\tilde{x}R_n| > C\Big]$$

$$\leq c\mathbb{E}\Big[|L_n x|^{-1}; n \text{ is a strict ladder epoch of the random process } (|L_k x|)_{k\geq 0}\Big]$$

so that $\sum_{n\geq 0}\mathbb{E}\Big[|\tilde{x}R_n|^{-1}; |\tilde{x}R_1| > C, \ldots, |\tilde{x}R_n| > C\Big] < +\infty$. The proof is a bit more delicate for the other values of C, the details are given in the following article.

To summarize this introduction, the study of the probability of extinction of multi-type Galton-Watson processes we present here follows the strategy developed by J. Geiger and G. Kersting and needs two main ingredients : the fluctuation theory for product of random matrices and a "key lemma" (see Lemma 1.3.2) which allows to compare the norm of some product of matrices with the norm of vectors obtained by their action on $\mathbb{X}$.

The Annals of Probability
2018 Vol. 46, No. 5, 2946–2972

THE SURVIVAL PROBABILITY OF A CRITICAL MULTI-TYPE BRANCHING PROCESS IN I.I.D. RANDOM ENVIRONMENT

BY E. LE PAGE*, M. PEIGNÉ † AND C. PHAM †

*Université de Bretagne-Sud * and Université Fr. Rabelais Tours†*

We study the asymptotic behavior of the probability of non-extinction of a critical multi-type Galton-Watson process in i.i.d. random environments by using limits theorems for products of positive random matrices. Under suitable assumptions, the survival probability is proportional to $1/\sqrt{n}$

1. Introduction. Branching processes in random environments (BPREs) were introduced in the 1960s (see for instance [22]) to describe the development of populations whose evolution may be affected by environmental factors; they have been a central topic of research.

In the single-type case, the behaviour of these processes is mainly determined by the 1-dimensional random walk generated by the logarithms of the expected population sizes $m_k, k \geq 0$, of the respective generations. The theory of fluctuations of random walks on $\mathbb{R}$ with i.i.d. increments allows one to classify BPREs in three classes - supercritical, critical or subcritical - of single-type BPREs, according to the fact that the associated random walk tends to $+\infty$, oscillates or tends to $-\infty$ (see for instance the fundamental papers [2] and [3]). In particular, when $\mathbb{E}(|\log m_k|) < +\infty$, the BPRE is supercritical (resp., critical or subcritical) when $\mathbb{E}(\log m_k) > 0$ (resp., $\mathbb{E}(\log m_k) = 0$ or $\mathbb{E}(\log m_k) < 0$). There exist numerous statements concerning the asymptotic behaviour of the probability of non-extinction up to time n, the distribution of the population size conditionally to survival up to moment n, large deviation type results (see for instance [1], [4], [7]). In the critical case, the branching process is degenerate with probability one [2]; M. V. Kozlov [19] (for BPREs with linearly fractional generating functions) and J. Geiger and G. Kersting [10] in the general case strengthened this result and proved that the probability of non-extinction up to time n is equivalent to $c_1/\sqrt{n}$ as $n \to +\infty$, for some explicit constant $c_1 > 0$ (see also [15]). Let us recall that this probability of non-extinction up to time n is equivalent to $1/n$ when the offspring distribution is fixed, that is it does not vary randomly; in other words, branching processes in random environment die more slowly. In the supercritical and subcritical cases, similar studies have been done (see for instance [2], [3], [11] and [16]), we do not go into detail about these cases since they are outside the scope of this paper.

It is of interest to prove analogues of the above statements for the multi-type BPREs

MSC 2010 subject classifications: Primary 60J80; secondary 60F17, 60K37
Keywords and phrases: multi-type branching process, survival probability, random environment, product of random matrices, critical case.

$(Z_n)_{n\geq 0}$. The main difficulty which appears is that the role of the random walk associated to the BPRE in this case is played by the logarithm of the norm of some $\mathbb{R}^p$-valued Markov chain whose increments are governed by i.i.d. random $p \times p$-matrices $M_k, k \geq 0$; the coefficients of these matrices M_k are non negative and correspond to the expected population sizes of the respective generations, according to the types of the particles and their direct parents. Products of random matrices have been the object of many investigations and many limit theorems do exist in this context: for instance, the law of large numbers, the central limit theorem, the large deviations principle (see [5], [20] and references therein). Unfortunately, the theory of their fluctuations remains a field in which practically no research has been done and the multi-type BPREs constitute a relevant application area for this study. Nevertheless, as in the single-type case, the set of multi-type BPREs may be divided into three classes: they are supercritical (resp., critical or subcritical) when the upper Lyapunov exponent of the associated random matrices is positive, null or negative [18].

A problem of particular importance is to specify the asymptotic behaviour of their probability of non-extinction up to time n. As for single-type BPREs, the case has been first studied when the generating functions defining the random environment are linear fractional; it yields to explicit formulas which are easier to tackle.

In [6], E. E. Dyakonova obtained in the critical case an equivalence of the survival probability at time n, under the restrictive assumption that the mean matrices M_k have a concordant deterministic Perron-Frobenius eigenvector. It happens for instance when the M_k are upper-triangular matrices with strictly decreasing coefficients on the diagonal. In this case, the behaviour of the BPRE is governed by the 1-dimensional random walk whose increments are the logarithm of the spectral radius of the M_k; this allows one to apply the same techniques as in the single-type case.

To tackle the "general" case when the action of the random matrices M_k is strongly irreducible (see hypothesis H2 below), limit theorems on the fluctuations of the norm of products of random matrices were required; they have been recently achieved in [12] and [13] with asymptotic results on the tails of certain hitting time distribution.

In this paper, we investigate the asymptotic behaviour of the probability of non-extinction up to time n of critical multi-type BPREs and obtain an optimal result in the case of linear fractional generating functions. To formulate our main results, we first introduce some standard notations and definitions.

We fix an integer $p \geq 2$ and denote $\mathbb{R}^p$ (resp., $\mathbb{N}^p$) the set of p-dimensional column vectors with real (resp., non-negative integers) coordinates; for any column vector $x = (x_i)_{1\leq i\leq p} \in \mathbb{R}^p$, we denote $\tilde{x}$ the row vector $\tilde{x} := (x_1, \ldots, x_p)$. Let $\mathbf{1}$ (resp., $\mathbf{0}$) be the column vector of $\mathbb{R}^p$ whose all coordinates equal 1 (resp., 0). We fix a basis $\{e_i, 1 \leq i \leq p\}$ in $\mathbb{R}^p$ and denote $|.|$ the corresponding L^1 norm. We also consider the general linear semi -group S^+ of $p \times p$ matrices with positive coefficients. We endow S^+ with the L^1-norm denoted also by $|.|$.

The multi-type Galton-Watson process is a temporally homogeneous vector Markov process $(Z_n)_{n\geq 0}$ whose states are column vectors in $\mathbb{N}^p$. We always assume that Z_0 is

non-random. For any $1 \leq i \leq p$, the i-th component $Z_n(i)$ of Z_n may be interpreted as the number of objects of type i in the n-th generation.

We consider a measurable function $\xi \mapsto f_\xi$ from $\mathbb{R}$ to the set of multi-variate probability generating functions $f_\xi = (f_\xi^{(i)})_{1\leq i\leq p}$ defined by

$$f_\xi^{(i)}(s) = \sum_{\alpha\in\mathbb{N}^p} p_\xi^{(i)}(\alpha)s^\alpha,$$

for any $s = (s_i)_{1\leq i\leq p} \in [0,1]^p$, where

(i) $\alpha = (\alpha_i)_i \in \mathbb{N}^p$ and $s^\alpha = s_1^{\alpha_1}\dots s_p^{\alpha_p}$;

(ii) $p_\xi^{(i)}(\alpha) = p_\xi^{(i)}(\alpha_1,\dots,\alpha_p)$ is the probability that an object of type i in environment ξ has α_1 children of type $1,\dots,\alpha_p$ children of type p.

Let $(\xi_n)_{n\geq 0}$ be a sequence of real valued i.i.d. random variables defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. The Galton-Watson process with p types of particles in a random environment $(\xi_n)_{n\geq 0}$ describes the evolution of a particle population $Z_n = (Z_n(i))_{1\leq i\leq p}$ for $n \geq 0$.

We assume that for any $n \geq 0, i = 1,\dots,p$ and $\xi \in \mathbb{R}$, if $\xi_n = \xi$ then each of the $Z_n(i)$ particles of type i existing at time n produces offspring in accordance with the p-dimensional generating function $f_\xi^{(i)}$ independently of the reproduction of other particles of all types.

If $Z_0 = \tilde{e}_i$ then Z_1 has the generating function

$$f_{\xi_0}^{(i)}(s) = \sum_{\alpha\in\mathbb{N}^p}^{+\infty} p_{\xi_0}^{(i)}(\alpha)s^\alpha.$$

In general, if $Z_n = (\alpha_i)_{1\leq i\leq p}$, then Z_{n+1} is the sum of $\alpha_1 + \ldots + \alpha_p$ independent random vectors, where α_i particles of type i have the generating function $f_{\xi_n}^{(i)}$ for $i = 1,\dots,p$. It is obvious that if $Z_n = \tilde{\mathbf{0}}$, then $Z_{n+1} = \tilde{\mathbf{0}}$.

Denote $f_n = f_{\xi_n}$. By the above descriptions, (written in equation (2.1) in [18]) for any $s = (s_i)_{1\leq i\leq p}, 0 \leq s_i \leq 1$,

$$\mathbb{E}\left(s^{Z_n}|Z_0,\dots,Z_{n-1},f_0,\dots,f_{n-1}\right) = f_{n-1}(s)^{Z_{n-1}}$$

which yields (Lemma 2.1 in [18])

$$\begin{aligned}\mathbb{E}\left(s^{Z_n}|f_0^{(i)},f_1,\dots,f_{n-1}\right) &:= \mathbb{E}\left(s^{Z_n}|Z_0=\tilde{e}_i,f_0,\dots,f_{n-1}\right)\\ &= f_0^{(i)}(f_1(\dots f_{n-1}(s)\dots)).\end{aligned}$$

In particular, the probability of non-extinction $q_n^{(i)}$ at generation n given the environment $f_0^{(i)}, f_1, \dots f_{n-1}$ is

$$\begin{aligned}q_n^{(i)} &:= \mathbb{P}(Z_n \neq \tilde{\mathbf{0}}|f_0^{(i)},f_1,\dots,f_{n-1})\\ &= 1 - f_0^{(i)}(f_1(\dots f_{n-1}(\tilde{\mathbf{0}})\dots)) = \tilde{e}_i(\mathbf{1} - f_0(f_1(\dots f_{n-1}(\tilde{\mathbf{0}})\dots))),\end{aligned} \tag{1.1}$$

so that

$$\mathbb{E}[q_n^{(i)}] = \mathbb{E}[\mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | f_0^{(i)}, f_1, \ldots, f_{n-1})] = \mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i).$$

As in the classical one-type case, the asymptotic behaviour of the quantity above is controlled by the mean of the offspring distributions. We assume that the offspring distributions have finite first and second moments; the generating functions $f_\xi^{(i)}, \xi \in \mathbb{R}, 1 \leq i \leq p$, are thus C^2-functions on $[0,1]^p$ and we introduce:

(i) the random mean matrices $M_n = (M_n(i,j))_{1 \leq i,j \leq p} = \left(\dfrac{\partial f_n^{(i)}(\mathbf{1})}{\partial s_j} \right)_{i,j}$ taken from the vector-valued random generating function $f_n(s)$ at $s = \mathbf{1}$, namely

$$M_n = \begin{pmatrix} \dfrac{\partial f_n^{(1)}(\mathbf{1})}{\partial s_1} & \cdots & \dfrac{\partial f_n^{(1)}(\mathbf{1})}{\partial s_p} \\ \vdots & & \\ \dfrac{\partial f_n^{(p)}(\mathbf{1})}{\partial s_1} & \cdots & \dfrac{\partial f_n^{(p)}(\mathbf{1})}{\partial s_p} \end{pmatrix}.$$

(ii) the random Hessian matrices $B_n^{(i)} = (B_n^{(i)}(k,l))_{1 \leq k,l \leq p} = \left(\frac{\partial^2 f_n^{(i)}}{\partial s_k \partial s_l}(\mathbf{1}) \right)_{k,l}$, $1 \leq i \leq p$, taken from the real-valued random generating function $f_{\xi_n}^{(i)}(s)$ at $s = \mathbf{1}$.

For any $1 \leq i \leq p$, the random variables M_n and $B_n^{(i)}$ are i.i.d. in n. The common law of the M_n is denoted by μ.

Let R_n and L_n denote the right and the left product of random matrices $M_k, k \geq 0$, respectively, $R_n = M_0 M_1 \ldots M_{n-1}$ and $L_n = M_{n-1} \ldots M_1 M_0$.

By [9], if $\mathbb{E}(\max(0, \log |M_0|)) < +\infty$, then the sequence $\left(\dfrac{1}{n} \log |R_n| \right)_{n \geq 1}$ converges $\mathbb{P}$-almost surely to some constant limit $\pi_\mu := \lim\limits_{n \to +\infty} \dfrac{1}{n} \mathbb{E}[\log |R_n|]$. Furthermore, by [18], if there exists a constant $A > 0$ such that $\dfrac{1}{A} \leq M_n(i,j) \leq A$ and $0 \leq B_n^{(i)}(k,l) \leq A$ $\mathbb{P}$-almost surely for any $1 \leq i,j,k,l \leq p$, then the process $(Z_n)_{n \geq 0}$ becomes extinct $\mathbb{P}$-almost surely if and only if $\pi_\mu \leq 0$.

In the present work, we focus our attention on the so-called **critical case**, that is $\pi_\mu = 0$, and specify the speed of extinction of the Galton-Watson process.

We define the cone $\mathcal{C}$ and the space $\mathbb{X}$ respectively as follows:

$$\mathcal{C} := \{\tilde{x} = (x_1, \ldots, x_p) \in \mathbb{R}^p \mid \forall i = 1, \ldots, p, \ x_i \geq 0\},$$

and

$$\mathbb{X} := \{\tilde{x} \in \mathcal{C} \mid |\tilde{x}| = 1\}.$$

In the sequel, we consider:

- the right and the left linear actions of S^+ on $\mathcal{C}$ defined by:

$$(\tilde{x}, g) \mapsto \tilde{x}g \quad \text{and} \quad (\tilde{x}, g) \mapsto gx$$

for any $\tilde{x} \in \mathcal{C}$ and $g \in S^+$,

- the right and the left projective actions of S^+ on $\mathbb{X}$ defined by:

$$(\tilde{x}, g) \mapsto \tilde{x} \cdot g = \frac{\tilde{x}g}{|\tilde{x}g|} \quad \text{and} \quad (\tilde{x}, g) \mapsto g \cdot x = \frac{gx}{|gx|}$$

for any $\tilde{x} \in \mathbb{X}$ and $g \in S^+$.

For any $g = (g(i,j))_{1 \le i,j \le d} \in S^+$, let $v(g) := \min_{1 \le j \le d} \Big(\sum_{i=1}^{d} g(i,j) \Big)$. Then, for any $x \in \mathcal{C}$,

$$0 < v(g)\ |x| \le |gx| \le |g|\ |x|. \tag{1.2}$$

We set $N(g) := \max\left(\frac{1}{v(g)}, |g|\right)$. We also introduce some proper subset of S^+ which is of interest in the sequel: for any constant $B \ge 1$, let $S^+(B)$ denote the set of $p \times p$ matrices $g = (g(i,j))_{1 \le i,j \le p}$ with positive coefficients such that for any $1 \le i, j, k, l \le p$,

$$\frac{1}{B} \le \frac{g(i,j)}{g(k,l)} \le B. \tag{1.3}$$

On the product space $\mathbb{X} \times S^+$ we define the function ρ by setting $\rho(\tilde{x}, g) := \log |\widetilde{x}g|$ for $(\tilde{x}, g) \in \mathbb{X} \times S^+$. This function satisfies the cocycle property, namely for any $g, h \in S^+$ and $\tilde{x} \in \mathbb{X}$,

$$\rho(\tilde{x}, gh) = \rho(\widetilde{x} \cdot g, h) + \rho(\tilde{x}, g). \tag{1.4}$$

Under hypothesis H3 introduced below, there exists a unique μ-invariant measure ν on $\mathbb{X}$ such that for any continuous function φ on $\mathbb{X}$,

$$(\mu * \nu)(\varphi) = \int_{S^+} \int_{\mathbb{X}} \varphi(\tilde{x} \cdot g) \nu(d\tilde{x}) \mu(dg) = \int_{\mathbb{X}} \varphi(\tilde{x}) \nu(d\tilde{x}) = \nu(\varphi).$$

Moreover, the upper Lyapunov exponent π_μ defined above coincides with the quantity $\int_{\mathbb{X} \times S^+} \rho(\tilde{x}, g) \mu(dg) \nu(d\tilde{x})$ and is finite [5].

In the sequel, we first focus our attention on "linear-fractional multi-dimensional generating functions" $f = f_{\gamma,M}$ of the form:

$$f(s) = f_{\gamma,M}(s) = \mathbf{1} - \frac{1}{1 + \tilde{\gamma}(\mathbf{1} - s)} M(\mathbf{1} - s), \tag{1.5}$$

for any $s \in [0,1]^p$, where $\tilde{\gamma} = (\gamma, \ldots, \gamma) \in \mathbb{R}^p$ with $\gamma > 0$ and $M \in S^+$. For such a function $f = f_{\gamma,M}$, we set $\gamma = \gamma(f)$ and $M = M(f)$ and notice that $M(f)$ equals the mean matrix $\left(\frac{\partial f^{(i)}}{\partial s_j}(\mathbf{1})\right)_{1 \le i,j \le p}$.

Here, we specify hypotheses concerning the distribution μ of the mean matrices $M_n = M(f_n), n \geq 1$.

H1. *There exists* $\epsilon_0 > 0$ *such that* $\int_{S^+} N(g)^{\epsilon_0} \mu(dg) < \infty$.

H2. *(Strong irreducibility). There exists no affine subspaces* $\mathcal{A}$ *of* $\mathbb{R}^d$ *such that* $\mathcal{A} \cap \mathcal{C}$ *is non-empty, bounded and invariant under the action of all elements of the support of* μ.

H3. *There exists* $B \geq 1$ *such that* $\mu(S^+(B)) = 1$.

H4. *The upper Lyapunov exponent* π_μ *of* μ *is equal to 0.*

H5. *There exists* $\delta > 0$ *such that* $\mu(E_\delta) > 0$, *where*

$$E_\delta := \{g \in S^+ \mid \forall \tilde{x} \in \mathbb{X}, \ \ \log|\tilde{x}g| \geq \delta\}.$$

If the variables f_n are linear fractional generating functions, then we introduce an additional hypothesis.

H6. *There exists* $B' \geq 1$ *such that* $\dfrac{1}{B'} \leq \gamma(f_n) \leq B' \quad \mathbb{P} - a.s.$

Here comes the main result of this paper.

THEOREM 1.1. *Assume that the random variables* f_n *are linear fractional generating functions and that hypotheses* H1-H6 *hold. Then, for any* $i \in \{1, \ldots, p\}$, *there exists a real number* $\beta_i \in (0, +\infty)$ *such that*

$$\lim_{n \to +\infty} \sqrt{n} \mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) = \beta_i. \tag{1.6}$$

For general generating functions, we obtain the following weaker result.

THEOREM 1.2. *Assume that the random variables* f_n *are* C^2*-functions on* $[0,1]^p$ *such that*

1. *there exists* $A > 0$ *such that for any* $i, k, l \in \{1, \ldots, p\}$,

$$\frac{\partial^2 f_n^{(i)}}{\partial s_k \partial s_l}(\mathbf{1}) \leq A \frac{\partial f_n^{(i)}}{\partial s_k}(\mathbf{1}) \quad \mathbb{P} - \text{a.s.}$$

2. *the distribution* μ *of the matrices* $M_n = \left(\frac{\partial f_n^{(i)}}{\partial s_j}(\mathbf{1})\right)_{1 \leq i,j \leq p}$ *satisfies hypotheses* H1-H5.

Then, there exist real constants $0 < c_1 < c_2 < +\infty$ *such that for any* $i \in \{1, \ldots, p\}$, *and* $n \geq 1$,

$$\frac{c_1}{\sqrt{n}} \leq \mathbb{P}(Z_n \neq \tilde{\mathbf{0}} | Z_0 = \tilde{e}_i) \leq \frac{c_2}{\sqrt{n}}. \tag{1.7}$$

In particular, under weaker assumptions than in [18], this theorem states that the process $(Z_n)_{n \geq 0}$ becomes extinct $\mathbb{P}$-a.s. in the critical case. Notice that by a recent work

[8], equivalence (1.6) holds true even when the f_n are not assumed to be linear fractional generating functions. The authors in [8] apply the proof in the present paper with two new ingredients: first, an extension to the multitype case of Geiger & Kersting's decomposition of the extinction probability ([10], formula (2.2)) and second, a variation of Lemma 3.1 below, taking into account the residual term which appears in this expression of the extinction probability.

Notation. *Let $c > 0$ and ϕ, ψ be two functions of some variable x; we shall write $\phi \overset{c}{\preceq} \psi$ (or simply $\phi \preceq \psi$) when $\phi(x) \leq c\psi(x)$ for any value of x. The notation $\phi \overset{c}{\asymp} \psi$ (or simply $\phi \asymp \psi$) means $\phi \overset{c}{\preceq} \psi \overset{c}{\preceq} \phi$.*

2. Preliminary results.

2.1. *Product of matrices with non-negative coefficients.* We describe in this section some properties of the set S^+. We first endow $\mathbb{X}$ with a distance d which is a variant of the Hilbert metric; it is bounded on $\mathbb{X}$ and any element $g \in S^+$ acts on $(\mathbb{X}, d)$ as a contraction. We summarize here its construction and its major properties.

For any $\tilde{x} = (x_i)_{1 \leq i \leq p}, \tilde{y} = (y_i)_{1 \leq i \leq p} \in \mathbb{X}$, we write

$$\mathsf{m}(\tilde{x}, \tilde{y}) = \min\left\{ \frac{x_i}{y_i} \middle| i = 1, \ldots, p \text{ such that } y_i > 0 \right\}$$

and we set

$$d(\tilde{x}, \tilde{y}) := \varphi\Big(\mathsf{m}(\tilde{x}, \tilde{y})\mathsf{m}(\tilde{y}, \tilde{x})\Big),$$

where φ is the one-to-one function on $[0, 1]$ defined by $\varphi(s) := \dfrac{1-s}{1+s}$. For $g \in S^+$, we set

$$c(g) := \sup\{d(\tilde{x} \cdot g, \tilde{y} \cdot g) \mid \tilde{x}, \tilde{y} \in \mathbb{X}\}.$$

We present some crucial properties of d.

Proposition 2.1. *([17]). The function d is a distance on $\mathbb{X}$ which satisfies the following properties:*

1. $\sup\{d(\tilde{x}, \tilde{y}) \mid \tilde{x}, \tilde{y} \in \mathbb{X}\} = 1$.
2. *For any $g = (g(i, j))_{i,j} \in S^+$*

$$c(g) = \max_{i,j,k,l \in \{1,\ldots,p\}} \frac{|g(i,j)g(k,l) - g(i,l)g(k,j)|}{g(i,j)g(k,l) + g(i,l)g(k,j)}.$$

   In particular, there exists $\kappa_B \in [0, 1)$ such that $c(g) \leq \kappa_B < 1$ for any $g \in S^+(B)$.
3. *$d(\tilde{x} \cdot g, \tilde{y} \cdot g) \leq c(g)d(\tilde{x}, \tilde{y}) \leq c(g)$ for any $\tilde{x}, \tilde{y} \in \mathbb{X}$ and $g \in S^+(B)$.*
4. *$c(gg') \leq c(g)c(g')$ for any $g, g' \in S^+(B)$.*

The following lemma follows from [9] Lemma 2, and its proof is presented in Section 4. It is the key argument in the sequel to control the asymptotic behaviour of the norm of products of matrices of $S^+(B)$. Let $T_{S^+(B)}$ be the semi-group generated by the set $S^+(B)$.

LEMMA 2.1. *For any $g \in T_{S^+(B)}$ and $1 \leq i, j, k, l \leq p$,*

$$g(i,j) \overset{B^2}{\asymp} g(k,l). \tag{2.1}$$

In particular, there exist $c > 1$ such that for any $g, h \in T_{S^+(B)}$ and $\tilde{x}, \tilde{y} \in \mathbb{X}$

1. $|gx| \overset{c}{\asymp} |g|$ *and* $|\tilde{y}g| \overset{c}{\asymp} |g|$,
2. $|\tilde{y}gx| \overset{c}{\asymp} |g|$,
3. $|gh| \overset{c}{\asymp} |g||h|$.

2.2. *Conditioned product of random matrices.* Recall that $(M_n)_{n\geq 0}$ is a sequence of i.i.d. matrices whose law μ satisfies hypotheses H1-H5 and $R_n = M_0 \dots M_{n-1}$ and $L_n = M_{n-1} \dots M_0$ for $n \geq 1$. Consider the homogenous Markov chain $(X_n)_{n\geq 0}$ on $\mathbb{X}$, with initial value $X_0 = \tilde{x} \in \mathbb{X}$, defined by

$$X_n = \tilde{x} \cdot R_n, \quad n \geq 1.$$

Its transition probability P is given by: for any $\tilde{x} \in \mathbb{X}$ and any bounded Borel function $\varphi : \mathbb{X} \to \mathbb{R}$,

$$P\varphi(\tilde{x}) := \int_{S^+} \varphi(\tilde{x} \cdot g)\mu(dg).$$

The chain $(X_n)_{n\geq 0}$ has been the object of many studies, in particular there exists on $\mathbb{X}$ a unique P-invariant probability measure ν.

Indeed, by Proposition 2.1, for any $\tilde{x}, \tilde{y} \in \mathbb{X}$ and **any sequence $(g_n)_{n\geq 1}$ of elements** in $S^+(B)$, it follows that

$$d(\tilde{x} \cdot g_n \dots g_1, \tilde{y} \cdot g_n \dots g_1) \leq \kappa_B^n. \tag{2.2}$$

This yields $\sup_{k\geq 0} d(\tilde{x} \cdot g_{n+k} \cdots g_n \cdots g_1, \tilde{x} \cdot g_n \dots g_1) \to 0$ as $n \to +\infty$; the sequence $(\tilde{x} \cdot g_1 \cdots g_n)_{n\geq 0}$ thus converges in $\mathbb{X}$.

In particular, under hypothesis H3, the sequence $(\tilde{x} \cdot L_n)_{n\geq 0}$ converges $\mathbb{P}$-a.s. towards some $\mathbb{X}$-valued random variable X_∞. It follows that the Markov chain $(\tilde{x} \cdot R_n)_{n\geq 0}$ converges in distribution to the law ν of X_∞, which is the unique P-invariant probability measure on $\mathbb{X}$. Property (2.2) allows one to prove that the restriction of P to some suitable space of continuous functions from $\mathbb{X}$ to $\mathbb{C}$ is quasi-compact, which is a crucial ingredient to study the asymptotic behavior of $(\tilde{x} \cdot R_n)_{n\geq 0}$ ([5], [14], [17]).

In the sequel, we are interested in the left linear action $\tilde{x} \mapsto \widetilde{x}R_n$ of the right products R_n, for any $\tilde{x} \in \mathbb{X}$. Since it holds

$$\widetilde{x}R_n = e^{\log |\widetilde{x}R_n|} \ \widetilde{x} \cdot R_n,$$

we consider the random process $(S_n)_{n\geq 0}$ defined by: for any $\tilde{x} \in \mathbb{X}, a \in \mathbb{R}$ and $n \geq 1$,

$$S_0 = S_0(\tilde{x}, a) := a \quad \text{and} \quad S_n = S_n(\tilde{x}, a) := a + \log|\tilde{x}R_n|.$$

Iterating the cocycle property (1.4), the basic decomposition of $S_n(\tilde{x}, a)$ arrives:

$$S_n(\tilde{x}, a) = a + \log|\tilde{x}R_n| = a + \sum_{k=0}^{n-1} \rho(X_k, M_k). \tag{2.3}$$

Let us emphasize that for any $a \in \mathbb{R}$ the sequence $(X_n, S_n)_{n\geq 0}$ is a Markov chain on $\mathbb{X}\times\mathbb{R}$ whose transition probability $\widetilde{P}$ is defined by: for any $(\tilde{x}, a) \in \mathbb{X} \times \mathbb{R}$ and any bounded Borel function $\psi : \mathbb{X} \times \mathbb{R} \to \mathbb{C}$,

$$\widetilde{P}\psi(\tilde{x}, a) = \int_{S^+} \psi(\tilde{x} \cdot g, a + \rho(\tilde{x}, g))\mu(dg).$$

We set $\mathbb{R}_*^+ := \mathbb{R}^+\backslash\{0\}$ and denote $\widetilde{P}_+$ the restriction of $\widetilde{P}$ to $\mathbb{X}\times\mathbb{R}^+$ defined by: for $a \geq 0$ and any $\tilde{x} \in \mathbb{X}$,

$$\widetilde{P}_+((\tilde{x}, a), \cdot) = 1_{\mathbb{X}\times\mathbb{R}^+}(\cdot)\widetilde{P}((\tilde{x}, a), \cdot).$$

Fix $a \geq 0$ and denote by τ the first time greater than 1 the random process $(S_n)_{n\geq 0}$ becomes nonpositive:

$$\tau := \min\{n \geq 1 \mid S_n \leq 0\}. \tag{2.4}$$

For any $\tilde{x} \in \mathbb{X}$ and $a \geq 0$, let us denote $\mathbb{P}_{\tilde{x},a}$ the probability measure on $(\Omega, \mathcal{F}, \mathbb{P})$ conditioned to the event $[X_0 = \tilde{x}, S_0 = a]$ and $\mathbb{E}_{\tilde{x},a}$ the corresponding expectation; we omit the index a when $a = 0$ and denote $\mathbb{P}_{\tilde{x}}$ the corresponding probability. For any $n \geq 0$, let $m_n = m_n(\tilde{x}, a)$ be the minimum of the values $S_1(\tilde{x}, a), \ldots, S_n(\tilde{x}, a)$ and set

$$\mathbf{m}_n(\tilde{x}, a) := \mathbb{P}_{\tilde{x},a}[m_n > 0] = \mathbb{P}_{\tilde{x},a}(\tau > n).$$

We present a general result concerning the behavior of the tail distribution of the random variable τ.

The asymptotic behaviour of the probability $\mathbb{P}_{\tilde{x},a}(\tau > n)$ is established in [13] when the matrices M_n are invertible and in [21] when the M_n has non negative entries, under several conditions P1-P5; the first step is to establish the existence of a $\widetilde{P}_+$-harmonic function h on $\mathbb{X} \times \mathbb{R}^+$. Our hypotheses H1, H2, H4 and H5 are exactly P1, P2, P4 and P5 in [21] and H3 is obviously stronger than P3. Proposition 2.2 concerns the existence of the function h and its properties.

PROPOSITION 2.2. *([21]). Under hypotheses* H1-H5*, for any $x \in \mathbb{X}$ and $a \geq 0$, the sequence $\Big(\mathbb{E}_{x,a}[S_n; \tau > n]\Big)_{n\geq 0}$ converges to a function $h(\tilde{x}, a)$ which satisfies the following properties:*

1. *For any $x \in \mathbb{X}$, the function $h(\tilde{x}, a)$ is increasing on $\mathbb{R}^+$.*
2. *There exists $c > 0$ and $A > 0$ such that for any $x \in \mathbb{X}$ and $a \geq 0$,*

$$\frac{1}{c} \vee (a - A) \leq h(\tilde{x}, a) \leq c(1 + a). \tag{2.5}$$

3. *For any $x \in \mathbb{X}$, the function $h(\tilde{x}, a)$ satisfies $\lim_{a \to +\infty} \frac{h(\tilde{x}, a)}{a} = 1$.*
4. *The function h is $\widetilde{P}_+$-harmonic.*

This statement yields to the following theorem about the limit behaviour of $\mathbb{P}_{\tilde{x},a}(\tau > n)$ as $n \to +\infty$; the relation $u_n \sim v_n$ defines $\lim_{n \to +\infty} \frac{u_n}{v_n} = 1$.

THEOREM 2.1. *([21]). Under hypotheses* H1-H5*, for any $\tilde{x} \in \mathbb{X}$ and $a \geq 0$,*

$$\mathbb{P}_{\tilde{x},a}(\tau > n) \sim \frac{2}{\sigma\sqrt{2\pi n}} h(\tilde{x}, a) \quad \text{as } n \to \infty, \tag{2.6}$$

where $\sigma^2 := \lim_{n \to +\infty} \frac{1}{n} \mathbb{E}_x[S_n^2]$ is the variance of the semi-markovian random walk $(S_n)_{n \geq 0}$. Moreover, there exists a constant $c > 0$ such that for any $\tilde{x} \in \mathbb{X}$, $a \geq 0$ and $n \geq 0$,

$$\sqrt{n}\mathbb{P}_{\tilde{x},a}(\tau > n) \leq c\ h(\tilde{x}, a). \tag{2.7}$$

Remark. The fact that $\sigma^2 > 0$ is a direct consequence of hypotheses H2 and H5 (which implies in particular that the semi-group generated by the support of μ is unbounded); see [5], chap 6, Lemmas 5.2 and 5.3 and section 8 for more details.

3. Proof of Theorem 1.1.

3.1. *Expression of nonextinction probability.* For any $0 \leq k < n$, set $R_{k,n} := M_k \dots M_{n-1}$ and $R_{k,n} := I$ otherwise. Let $Y_{k,n} := R_{k,n} \cdot \mathbf{1}$. It is known (see Lemma 1 in [6]) that

$$(q_n^{(i)})^{-1} = \frac{1 + \tilde{\gamma}_0 M_1 \dots M_{n-1}\mathbf{1} + \tilde{\gamma}_1 M_2 \dots M_{n-1}\mathbf{1} + \dots + \tilde{\gamma}_{n-1}\mathbf{1}}{\tilde{e}_i R_n \mathbf{1}}. \tag{3.1}$$

Since $\tilde{e}_i R_k R_{k,n}\mathbf{1} = \tilde{e}_i M_0 \dots M_{n-1}\mathbf{1}$ for any $1 \leq k \leq n$, we may rewrite (3.1) as

$$\left(q_n^{(i)}\right)^{-1} = \frac{1}{\tilde{e}_i R_n \mathbf{1}} + \sum_{k=0}^{n-1} \frac{\tilde{\gamma}_k Y_{k+1,n}}{\tilde{e}_i R_k Y_{k+1,n}} = \frac{1}{\tilde{e}_i R_n \mathbf{1}} + \sum_{k=0}^{n-1} \frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,n}}. \tag{3.2}$$

Notice that, for any fixed $k \geq 0$, the sequence $(Y_{k,n})_{n \geq 0}$ converges to some limit $Y_{k,\infty}$ ([9], [17]). Indeed, the matrices $M_l, l \geq 1$, all belong to $S^+(B)$, so do their transposes and we may write as in (2.2)

$$\sup_{m \geq 1} d(\widetilde{Y}_{k,n}, \widetilde{Y}_{k,n+m}) \leq \kappa_B^{n-k} \overset{n \to +\infty}{\longrightarrow} 0.$$

The sequence of row vectors $(\widetilde{Y}_{k,n})_{n\geq 0}$ is thus a Cauchy sequence in $\mathbb{X}$, it converges and so does $(Y_{k,n})_{n\geq 0}$; we denote $Y_{k,\infty}$ the limit of this last sequence. Using this key fact, we prove in Lemma 3.2 that the sequence $\left(q_n^{(i)}\right)_{n\geq 1}$ converges to some positive limit

$$q_\infty^{(i)} = \Big(\sum_{k=0}^{+\infty} \frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,\infty}}\Big)^{-1} \tag{3.3}$$

in the L^1-norm with respect to another probability measure $\widehat{\mathbb{P}}_{\tilde{x},a}$ on $(\Omega, \mathcal{F})$, corresponding to the Doob transform of $\mathbb{P}$ associated with the function h given by Theorem 2.1; the construction of $\widehat{\mathbb{P}}_{\tilde{x},a}$ is described in the following subsection.

3.2. *Construction of a new probability measure $\widehat{\mathbb{P}}_{\tilde{x},a}$ conditioned to the environment.* Since the function h is $\widetilde{P}_+$-harmonic on $\mathbb{X}\times\mathbb{R}^+$, it gives rise to a Markov kernel $\widetilde{P}_+^h$ on $\mathbb{X}\times\mathbb{R}^+$ defined by

$$\widetilde{P}_+^h\phi = \frac{1}{h}\widetilde{P}_+(h\phi)$$

for any bounded measurable function ϕ on $\mathbb{X}\times\mathbb{R}^+$. The kernels $\widetilde{P}_+$ and $\widetilde{P}_+^h$ are related to the stopping times τ by the following identity: for any $\tilde{x}\in\mathbb{X}$, $a\geq 0$ and $n\geq 1$,

$$\begin{aligned}(\widetilde{P}_+^h)^n\phi(\tilde{x},a) &= \frac{1}{h(\tilde{x},a)}\widetilde{P}_+^n(h\phi)(\tilde{x},a)\\ &= \frac{1}{h(\tilde{x},a)}\mathbb{E}_{\tilde{x},a}\left(h\phi(X_n,S_n);\tau>n\right)\\ &= \frac{1}{h(\tilde{x},a)}\mathbb{E}_{\tilde{x},a}\left(h\phi(X_n,S_n);m_n>0\right).\end{aligned}$$

This new Markov chain with kernel $\widetilde{P}_+^h$ allows one to change the measure on the canonical path space $((\mathbb{X}\times\mathbb{R})^{\otimes\mathbb{N}},\sigma(X_n,S_n:n\geq 0),\theta)$ of the Markov chain $(X_n,S_n)_{n\geq 0}$ [(1)] from $\mathbb{P}$ to the measure $\widehat{\mathbb{P}}_{\tilde{x},a}$ characterized by the property that

$$\widehat{\mathbb{E}}_{\tilde{x},a}[\varphi(X_0,S_0,\ldots,X_k,S_k)] = \frac{1}{h(\tilde{x},a)}\mathbb{E}_{\tilde{x},a}[\varphi(X_0,\ldots,S_k)h(X_k,S_k);\ m_k>0] \tag{3.4}$$

for any positive Borel function φ on $(\mathbb{X}\times\mathbb{R})^{k+1}$.

[1] θ denotes the shift operator on $(\mathbb{X}\times\mathbb{R})^{\otimes\mathbb{N}}$ defined by $\theta\Big((x_k,s_k)_{k\geq 0}\Big) = (x_{k+1},s_{k+1})_{k\geq 0}$ for any $(x_k,s_k)_{k\geq 0}$ in $(\mathbb{X}\times\mathbb{R})^{\otimes\mathbb{N}}$

For any $0 \leq k \leq n$,

$$\begin{aligned}
& \mathbb{E}_{\tilde{x},a}[\varphi(X_0, S_0, \ldots, X_k, S_k); m_n > 0] \\
= \; & \mathbb{E}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k); S_1 > 0, \ldots, S_n > 0] \\
= \; & \mathbb{E}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k); a + \rho(X_0, M_0) > 0, \\
& \qquad \ldots, a + \sum_{i=0}^{k-1} \rho(X_i, M_i) + \sum_{i=k}^{n-1} \rho(X_i, M_i) > 0] \\
= \; & \mathbb{E}_{\tilde{x},a}[\mathbb{E}[\varphi(X_0, \ldots, S_k); S_1 > 0, \ldots, S_k > 0, \\
& \qquad S_k + S_1 \circ \theta^k > 0, \ldots, S_k + S_{n-k} \circ \theta^k > 0 | M_0, \ldots, M_{k-1}]] \\
= \; & \mathbb{E}_{\tilde{x},a}\Big[\varphi(X_0, \ldots, S_k) \\
& \qquad \times \mathbb{E}[S_k + S_1 \circ \theta^k > 0, \ldots, S_k + S_{n-k} \circ \theta^k > 0 | M_0, \ldots, M_{k-1}]; m_k > 0\Big] \\
= \; & \mathbb{E}_{\tilde{x},a}\Big[\varphi(X_0, \ldots, S_k) \\
& \qquad \times \mathbb{P}_{X_k,S_k}(S_1 \circ \theta^k > 0, \ldots, S_{n-k} \circ \theta^k > 0); m_k > 0\Big] \\
= \; & \mathbb{E}_{\tilde{x},a}\Big[\varphi(X_0, \ldots, S_k)\mathbb{P}_{X_k,S_k}(\tau > n-k); m_k > 0\Big].
\end{aligned}$$

Hence,

$$\begin{aligned}
& \mathbb{E}_{\tilde{x},a}\left[\varphi(X_0, \ldots, S_k); m_n > 0\right] \\
& \qquad = \mathbb{E}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k)\mathbf{m}_{n-k}(X_k, S_k); m_k > 0]. \qquad (3.5)
\end{aligned}$$

Moreover, in view of Theorem 2.1, the dominated convergence theorem and (3.5), we obtain for any bounded function φ with compact support,

$$\begin{aligned}
& \lim_{n\to+\infty} \mathbb{E}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k) | m_n > 0] \\
= \; & \frac{1}{h(\tilde{x}, a)} \mathbb{E}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k) h(X_k, S_k); m_k > 0] \\
= \; & \widehat{\mathbb{E}}_{\tilde{x},a}[\varphi(X_0, \ldots, S_k)], \qquad (3.6)
\end{aligned}$$

which clarifies the interpretation of $\widehat{\mathbb{P}}_{\tilde{x},a}$.

We formalize in three steps the construction of a new probability measure, denoted again $\widehat{\mathbb{P}}_{\tilde{x},a}$, for each $\tilde{x} \in \mathbb{X}$ and $a \geq 0$, but defined this time on the bigger σ-algebra $\sigma(f_n, Z_n : n \geq 0)$. Retaining the notations from the previous parts, the measure $\widehat{\mathbb{P}}_{\tilde{x},a}$ is characterized by properties (3.4), (3.7) and (3.8).

Step 1. The marginal distribution of $\widehat{\mathbb{P}}_{\tilde{x},a}$ on $\sigma(X_n, S_n : n \geq 0)$ is $\widehat{\mathbb{P}}_{\tilde{x},a}$ characterized by the property (3.4).

Step 2. For any $n \geq 0$, the conditional distribution of $(f_0, \cdots, f_n)$ under $\widehat{\mathbb{P}}_{\tilde{x},a}$ given $X_0 = \tilde{x}_0 = \tilde{x}, \ldots, X_n = \tilde{x}_n, S_0 = s_0 = a, \ldots, S_n = s_n$ is given by:

$$\begin{aligned}
& \widehat{\mathbb{P}}_{\tilde{x},a}(f_k \in A_k, 0 \leq k \leq n | X_i = \tilde{x}_i, S_i = s_i, 0 \leq i \leq n) \\
& \qquad = \mathbb{P}(f_k \in A_k, 0 \leq k \leq n | X_i = \tilde{x}_i, S_i(\tilde{x}) = s_i, 0 \leq i \leq n), \qquad (3.7)
\end{aligned}$$

for any measurable sets $A_0, \dots, A_n$ in $\mathcal{H}_B$ and almost all $(\tilde{x}_i)_{0\le i\le n}$ and $(s_i)_{0\le i\le n}$ with respect to the law of $(X_0, \dots, X_n, S_0, \dots, S_n)$ under $\mathbb{P}$ (and also under $\widehat{\mathbb{P}}_{\tilde{x},a}$ since, by formula (3.4), the probability measure $\widehat{\mathbb{P}}_{\tilde{x},a}$ is absolutely continuous with respect to $\mathbb{P}$ on each σ-algebra $\sigma(X_0, \dots, S_n)$).

Step 3. The conditional distribution of $(Z_n)_{n\ge 0}$ under $\widehat{\mathbb{P}}_{\tilde{x},a}$ given $f_0^{(i)}, f_1, \dots$ is the same as under $\mathbb{P}$, namely

$$\begin{aligned}&\widehat{\mathbb{E}}_{\tilde{x},a}\left(s^{Z_n}|Z_0,\dots,Z_{n-1},f_0^{(i)},f_1,\dots,f_{n-1}\right)\\ &\qquad = f_{n-1}(s)^{Z_{n-1}} = \mathbb{E}\left(s^{Z_n}|Z_0,\dots,Z_{n-1},f_0^{(i)},f_1,\dots,f_{n-1}\right).\end{aligned} \tag{3.8}$$

3.3. *Proof of Theorem 1.1.* We follow J. Geiger and G. Kersting's approach. We fix $1 \le i \le p$. For any $\tilde{x} \in \mathbb{X}$ and $a \ge 0$ let us denote $\mathbb{P}^{(i)}_{\tilde{x},a}$ the probability measure on $(\Omega, \mathcal{F}, \mathbb{P})$ conditioned to the event $[X_0 = \tilde{x}, S_0 = a, Z_0 = \tilde{e}_i]$ and $\mathbb{E}^{(i)}_{\tilde{x},a}$ the corresponding expectation.

First, notice that the quantity $\mathbb{P}(Z_n \ne \tilde{\mathbf{0}} \mid Z_0 = \tilde{e}_i)$ equals $\mathbb{E}_{\tilde{x},a}(q_n^{(i)})$ for any $(x,a) \in \mathbb{X}\times\mathbb{R}^+$; thus we fix $(\tilde{x},a) \in \mathbb{X}\times\mathbb{R}^+$, $\rho > 1$ and $0 \le k \le n$ and decompose $\mathbb{P}(Z_n \ne \tilde{\mathbf{0}} \mid Z_0 = \tilde{e}_i)$ as

$$\mathbb{P}^{(i)}(Z_n \ne \tilde{\mathbf{0}}) = A_n + B_n + C_n - D_n, \tag{3.9}$$

where

- $A_n = A_n(\tilde{x},a) = \mathbb{P}^{(i)}_{\tilde{x},a}(Z_n \ne \tilde{\mathbf{0}}, m_n \le 0)$;
- $B_n = B_n(\tilde{x},a,\rho) = \mathbb{P}^{(i)}_{\tilde{x},a}(Z_n \ne \tilde{\mathbf{0}}, m_n > 0) - \mathbb{P}_x(Z_n \ne \tilde{\mathbf{0}}, m_{\rho n} > 0)$;
- $C_n = C_n(\tilde{x},a,\rho,k) = \mathbb{P}^{(i)}_{\tilde{x},a}(Z_k \ne \tilde{\mathbf{0}}, m_{\rho n} > 0)$;
- $D_n = D_n(\tilde{x},a,\rho,k) = \mathbb{P}^{(i)}_{\tilde{x},a}(Z_k \ne \tilde{\mathbf{0}}, Z_n = \tilde{\mathbf{0}}, m_{\rho n} > 0)$.

We decompose the proof in 5 steps and then conclude, letting first $k \to +\infty$, then $\rho \to 1$ and at last $a \to +\infty$.

1. The quantity $A = A(\tilde{x},a) := \limsup_{n\to+\infty} \sqrt{n} A_n(\tilde{x},a)$ tends to 0 when $a \to +\infty$.
2. The quantity $B = B(\tilde{x},a,\rho) := \limsup_{n\to+\infty} \sqrt{n} B_n(\tilde{x},a,\rho)$ tends to 0 when $\rho \to +1$.
3. As $n \to +\infty$, the sequence $\left(\sqrt{n} C_n(\tilde{x},a,\rho,k)\right)_{n\ge 0}$ converges to
$$C = C(\tilde{x},a,\rho,k) := c_1 \frac{h(\tilde{x},a)}{\sqrt{\rho}} \widehat{\mathbb{P}}_x(Z_k \ne \tilde{\mathbf{0}}).$$
4. For any $\rho > 1$ the quantity $D = D(\tilde{x},a,\rho,k) := \limsup_{n\to+\infty} \sqrt{n} D_n(\tilde{x},a,\rho,k)$ tends to 0 when $k \to +\infty$.
5. The sequence $(\widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \ne 0))_{k\ge 0}$ converges to some limit $v^{(i)}(\tilde{x},a) > 0$.

Step 1. We may write

$$\begin{aligned} A_n(\tilde{x},a) &= \mathbb{E}_{\tilde{x},a}\left(\mathbb{E}\left(Z_n \neq \tilde{\mathbf{0}} | f_0^{(i)}, f_1, \ldots, f_{n-1}\right); m_n \leq 0\right) \\ &= \mathbb{E}_{\tilde{x},a}\left(q_n^{(i)}; m_n \leq 0\right). \end{aligned} \tag{3.10}$$

To control the quantity $q_n^{(i)}$, we use the expression (3.2); Lemma 2.1 yields

$$\begin{aligned} \frac{1}{q_n^{(i)}} &= \frac{1}{\tilde{e}_i R_n \mathbf{1}} + \sum_{k=0}^{n-1} \frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,n}} \\ &\geq \max_{0\leq k\leq n-1}\left\{\frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,n}}\right\} \asymp \max_{0\leq k\leq n-1}\left\{\frac{1}{|\tilde{x}R_k|}\right\} \\ &\geq \frac{1}{\exp\left\{\min\limits_{0\leq k\leq n-1}(a+\ln|\tilde{x}R_k|)\right\}} \end{aligned}$$

so that $q_n^{(i)} \preceq \exp(m_n(\tilde{x},a))$ and by applying Theorem 2.1 equation (2.7), we obtain

$$\begin{aligned} A_n(\tilde{x},a) &= \mathbb{P}_{\tilde{x}}^{(i)}(Z_n \neq \tilde{\mathbf{0}}, m_n \leq -a) \\ &\preceq \mathbb{E}_{\tilde{x}}[\exp(m_n); m_n \leq -a] \\ &\leq \sum_{k=a}^{+\infty} e^{-k}\mathbb{P}_{\tilde{x}}(-k < m_n \leq -k+1) \\ &\leq \sum_{k=a}^{+\infty} e^{-k}\mathbb{P}_{\tilde{x},k}(\tau > n) \\ &\preceq \frac{1}{\sqrt{n}}\sum_{k=a}^{+\infty}(k+1)e^{-k}, \end{aligned} \tag{3.11}$$

where the last inequality comes from (2.5) and (2.7). Notice that the sum $\sum\limits_{k=a}^{+\infty}(k+1)e^{-k}$ becomes arbitrarily small for sufficiently great a. Consequently $A(\tilde{x},a) = \limsup\limits_{n\to+\infty}\sqrt{n}A_n(\tilde{x},a) \longrightarrow 0$ as $x \to +\infty$.

Step 2. Theorem 2.1 yields

$$\begin{aligned} 0 \leq B_n(\tilde{x},a,\rho) &= \mathbb{P}_{\tilde{x},a}^{(i)}(Z_n \neq 0, m_n > 0) - \mathbb{P}_{\tilde{x},a}^{(i)}(Z_n \neq 0, m_{\rho n} > 0) \\ &\sim c_1 \frac{h(\tilde{x},a)}{\sqrt{n}}\left(1-\frac{1}{\sqrt{\rho}}\right) \quad \text{as} \quad n \to +\infty \end{aligned}$$

with $c_1 = \dfrac{2}{\sigma\sqrt{2\pi}}$. Hence

$$B(\tilde{x},a,\rho) := \limsup_{n\to+\infty}\sqrt{n}B_n(\tilde{x},a,\rho) \leq c_1 h(\tilde{x},a)\left(1-\frac{1}{\sqrt{\rho}}\right) \overset{\rho\to 1}{\longrightarrow} 0.$$

Step 3. Fix $0 \leq k \leq n$. Using (3.5) and the fact that $1_{[m_k>0]}$ and $\mathbf{m}_{\rho n-k}(X_k, S_k)$ are measurable with respect to the σ-algebra $\sigma(f_0^{(i)}, f_1, \ldots, f_{k-1})$, we may write

$$\begin{aligned}
C_n(\tilde{x}, a, \rho, k) &= \mathbb{P}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}, m_{\rho n} > 0) \\
&= \mathbb{P}^{(i)}_{\tilde{x},a}\left(Z_k \neq \tilde{\mathbf{0}}, m_k > 0, \mathbf{m}_{\rho n-k}(X_k, S_k)\right) \\
&= \mathbb{E}^{(i)}_{\tilde{x},a}\left(\mathbb{E}(1_{Z_k \neq \tilde{\mathbf{0}}} 1_{[m_k>0]} \mathbf{m}_{\rho n-k}(X_k, S_k) \mid f_0^{(i)}, \ldots, f_{k-1})\right) \\
&= \mathbb{E}^{(i)}_{\tilde{x},a}\left(\mathbb{P}(Z_k \neq \tilde{\mathbf{0}} | f_0^{(i)} \ldots, f_{k-1}) 1_{[m_k>0]} \mathbf{m}_{\rho n-k}(X_k, S_k)\right) \\
&= \mathbb{E}_{\tilde{x},a}\left(q_k^{(i)}, m_{\rho n} > 0\right) \\
&= \mathbb{E}_{\tilde{x},a}\left(q_k^{(i)} \mid m_{\rho n} > 0\right) \mathbb{P}_{\tilde{x},a}(m_{\rho n} > 0).
\end{aligned}$$

By (3.6), since k is fixed

$$\lim_{n\to+\infty} \mathbb{E}_{\tilde{x},a}\left(q_k^{(i)} | m_{\rho n} > 0\right) = \widehat{\mathbb{E}}_{\tilde{x},a}\left(q_k^{(i)}\right) = \widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}).$$

Hence, by Theorem 2.1

$$\lim_{n\to+\infty} \sqrt{n} C_n(\tilde{x}, a, \rho, k) = c_1 \frac{h(\tilde{x}, a)}{\sqrt{\rho}} \widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}).$$

Step 4. As in the previous step, we may write

$$\begin{aligned}
D_n(\tilde{x}, a, \rho, k) &= \mathbb{P}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}, Z_n = \tilde{\mathbf{0}}, m_{\rho n} > 0) \\
&= \mathbb{P}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}, m_{\rho n} > 0) - \mathbb{P}^{(i)}_{\tilde{x},a}(Z_n \neq \tilde{\mathbf{0}}, m_{\rho n} > 0) \\
&= \mathbb{E}_{\tilde{x},a}\left(q_k^{(i)}, m_{\rho n} > 0\right) - \mathbb{E}_{\tilde{x},a}\left(q_n^{(i)}, m_{\rho n} > 0\right) \\
&= \mathbb{E}_{\tilde{x},a}\left(\left(q_k^{(i)} - q_n^{(i)}\right) \mathbf{m}_{(\rho-1)n}(X_n, S_n); m_n > 0\right) \\
&\preceq \frac{1}{\sqrt{(\rho-1)n}} \frac{1}{h(\tilde{x}, a)} \mathbb{E}_{\tilde{x},a}\left((q_k^{(i)} - q_n^{(i)}) h(X_n, S_n); m_n > 0\right),
\end{aligned}$$

where the last inequality comes from (2.7). Since $1_{[m_n>0]}$ and $h(X_n, S_n)$ are $\sigma(S_0, \ldots, S_n)$-measurable, we observe that

$$\begin{aligned}
\sqrt{n} D_n(\tilde{x}, a, \rho, k) &\preceq \frac{1}{\sqrt{\rho-1}} \widehat{\mathbb{E}}_{\tilde{x},a}\left(q_k^{(i)} - q_n^{(i)}\right) \\
&= \frac{1}{\sqrt{\rho-1}} \left(\widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}) - \widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_n \neq \tilde{\mathbf{0}})\right).
\end{aligned}$$

Hence

$$D(\tilde{x}, a, \rho, k) = \limsup_{n\to+\infty} \sqrt{n} D_n(\tilde{x}, a, \rho, k) \preceq \frac{1}{\sqrt{\rho-1}} \left(\widehat{\mathbb{P}}_x(Z_k \neq \tilde{\mathbf{0}}) - v^{(i)}(\tilde{x}, a)\right)$$

so that $D(\tilde{x}, a, \rho, k) \longrightarrow 0$ as $k \to +\infty$.

Step 5. We first state the following lemmas whose proofs follow in the next section.

LEMMA 3.1. *For any $\tilde{x} \in \mathbb{X}$ and $a \geq 0$,*

$$\widehat{\mathbb{E}}_{\tilde{x},a} \sum_{n=0}^{+\infty} e^{-S_n} < +\infty.$$

This allows one to identify the limit of the sequence $(\widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}))_{k\geq 0}$.

LEMMA 3.2. *For any $\tilde{x} \in \mathbb{X}$ and $a \geq 0$, the sequence $\left(q_n^{(i)}\right)_{n\geq 1}$ converges to $q_\infty^{(i)}$ in $L^1(\widehat{\mathbb{P}}_{\tilde{x},a})$; in particular*

$$\lim_{k\to+\infty} \widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_k \neq \tilde{\mathbf{0}}) = \widehat{\mathbb{E}}_{\tilde{x},a} q_\infty^{(i)}, \tag{3.12}$$

where $q_\infty^{(i)}$ is given by the expression (3.3).

From Lemma 3.2, it is obvious that $\widehat{\mathbb{E}}_{\tilde{x},a} q_\infty^{(i)} \leq 1$. On the other hand, the expression of $q_\infty^{(i)}$ combined with Lemma 3.1 and the fact that $|\tilde{x}R_n| \asymp \tilde{e}_i R_n Y_{n+1,\infty}$ (see Lemma 2.1 property 2) yields $\widehat{\mathbb{E}}_{\tilde{x},a} q_\infty^{(i)} > 0$. In other words, for any $a \geq 0$

$$0 < v^{(i)}(\tilde{x}, a) := \widehat{\mathbb{E}}_{\tilde{x},a} q_\infty^{(i)} < +\infty.$$

This achieves the proof of Step 5.

Let us complete the proof of Theorem 1.1. Letting $n \to +\infty$ in expression (3.9) yields, for any $\tilde{x} \in \mathbb{X}, a \geq 0, \rho > 1$ and $k \geq 1$,

$$C - D \leq \liminf_{n\to+\infty} \sqrt{n}\mathbb{P}^{(i)}(Z_n \neq \tilde{\mathbf{0}}) \leq \limsup_{n\to+\infty} \sqrt{n}\mathbb{P}^{(i)}(Z_n \neq \tilde{\mathbf{0}}) \leq A + B + C - D.$$

Letting first $k \to +\infty$ then $\rho \to 1$ yields

$$\begin{aligned} c_1 h(\tilde{x}, a) v^{(i)}(\tilde{x}, a) &\leq \liminf_{n\to+\infty} \sqrt{n}\mathbb{P}(Z_n \neq \tilde{\mathbf{0}}) \\ &\leq \limsup_{n\to+\infty} \sqrt{n}\mathbb{P}(Z_n \neq \tilde{\mathbf{0}}) \\ &\leq c_1 h(\tilde{x}, a) v^{(i)}(\tilde{x}, a) + A(\tilde{x}, a). \end{aligned}$$

Since $v^{(i)}(\tilde{x}, a) > 0, h(\tilde{x}, a) > 0$ and $A(\tilde{x}, a) < +\infty$, we obtain in particular

$$0 < \liminf_{n\to+\infty} \sqrt{n}\mathbb{P}(Z_n \neq \tilde{\mathbf{0}}) \leq \limsup_{n\to+\infty} \sqrt{n}\mathbb{P}(Z_n \neq \tilde{\mathbf{0}}) < +\infty.$$

Finally, letting $a \longrightarrow +\infty$, we conclude that both limits

$$\lim_{n\to+\infty} \sqrt{n}\mathbb{P}^{(i)}(Z_n \neq \tilde{\mathbf{0}}) \quad \text{and} \quad \lim_{a\to+\infty} c_1 h(\tilde{x}, a) v^{(i)}(\tilde{x}, a)$$

exist, coincide and belong to $]0, +\infty[$.

3.4. *Proof of Theorem 1.2.* First, for any $n \geq 1$ and $s = (s_1, \ldots, s_p)$, we denote $F_n(s) = f_{\xi_0}(f_{\xi_1}(\ldots(f_{\xi_{n-1}}(s))\ldots))$. By definition of $q_n^{(i)}$, we have for any $0 \leq k < n$,

$$q_n^{(i)} = \tilde{e}_i(F_k(\mathbf{1}) - F_k(z)),$$

where $z = z(k,n) = f_{\xi_k}(\ldots(f_{\xi_{n-1}}(0))\ldots)$; the Mean Value Theorem yields

$$q_n^{(i)} = \tilde{e}_i(F_k(\mathbf{1}) - F_k(z)) \leq \tilde{e}_i M_0 \ldots M_{k-1}\mathbf{1} \asymp |\tilde{x}R_k| = \exp(S_k(\tilde{x}, 0))$$

so that $q_n^{(i)} \preceq \exp(m_n(\tilde{x}, 0))$ and $\mathbb{E}[q_n^{(i)}] \preceq \mathbb{E}[e^{m_n(\tilde{x},0)}] = \mathbb{E}_{\tilde{x}}[e^{m_n}]$.

Using the same trick like in (3.11), we can deduce that there exists a constant c_2 such that

$$\mathbb{E}_{\tilde{x}}[e^{m_n}] = \mathbb{E}_{\tilde{x}}[e^{m_n}; m_n \leq 0] \sim \frac{c_2}{\sqrt{n}},$$

and thus the upper estimate in equation (1.7) arrives.

To obtain the lower estimate in (1.7), for any $\mathbb{R}$-valued multi-dimensional generating function $f(s)$, $s = (s_1, \ldots, s_p)^T$, we obtain (see for instance formulas (64) and (65) in [23])

$$f(s) \leq 1 - \left(\sum_{i=1}^{p} \frac{\partial f}{\partial s_i}(\mathbf{1})(1 - s_i)\right) \left(1 + \frac{\displaystyle\sum_{i,j=1}^{p} \frac{\partial^2 f}{\partial s_i \partial s_j}(\mathbf{1})(1 - s_j)(1 - s_i)}{\displaystyle\sum_{l=1}^{p} \frac{\partial f}{\partial s_l}(\mathbf{1})(1 - s_l)}\right)^{-1}. \tag{3.13}$$

We set $g_{\xi_n}(s) = \mathbf{1} - \dfrac{M_{\xi_n}(\mathbf{1} - s)}{1 + \widetilde{\gamma}_{\xi_n}(\mathbf{1} - s)}$, where M_{ξ_n} is the mean matrix of $f_{\xi_n}(s)$ and $\widetilde{\gamma}_{\xi_n} = (A, \ldots, A)$; thus, hypothesis H6 holds for g_{ξ_n} with $B' = A$. Applying inequality (3.13) with $f = f_{\xi_n}^{(i)}$, we may write

$$f_{\xi_n}^{(i)}(s) \leq g_{\xi_n}^{(i)}(s), \quad i = 1, \ldots, p,$$

which yields

$$\mathbb{E}[\mathbf{1} - g_{\xi_0}(g_{\xi_1}(\ldots(g_{\xi_n - 1}(0))\ldots))] \leq \mathbb{E}[\mathbf{1} - f_0(f_1(\ldots(f_{n-1}(0))\ldots))]. \tag{3.14}$$

The lower estimate in equation (1.7) appears by applying Theorem 1.1 to the left side of equation (3.14). Therefore, the assertion of the Theorem 1.2 arrives.

4. Proof of lemmas. We first give some hints for the proof of Lemma 2.1 and describe later the proofs of Lemmas 3.1 and 3.2.

4.1. *Proof of Lemma 2.1.* First, using (2.1), we obtain

$$|g| = \sum_{i,j=1}^{p} g(i,j) \overset{p^2B^2}{\asymp} g(k,l). \tag{4.1}$$

Further properties can be easily deduced from (4.1). Indeed, the assertions we need are obvious by noticing that

$$|gx| = \sum_{i,j=1}^{p} g(i,j)x_j \overset{p^3B^2}{\asymp} |g|,$$

$$\tilde{y}gx = \sum_{i,j=1}^{p} y_i g(i,j)x_j \overset{p^2B^2}{\asymp} |g|,$$

$$|gh| = \sum_{i,j,k=1}^{p} g(i,j)h(j,k) \overset{p^7B^4}{\asymp} |g||h|.$$

4.2. *Proof of Lemma 3.1.* Before going into the proof, we first claim that in the critical case, for any $\delta > 0$ and c given from Lemma 2.1, there exists an integer $\kappa = \kappa(\delta, c) \geq 1$ such that

$$\mu^{*\kappa}(E_\delta) := \mu^{*\kappa}\{g \mid \forall \tilde{x} \in \mathbb{X},\ \log|\tilde{x}g| \geq \delta\} > 0. \tag{4.2}$$

Indeed, let $\tau' := \inf\{n \geq 1 \mid \log|R_n| \geq \log c + \delta\}$. The random variable τ' is a stopping time with respect to the natural filtration $(\sigma(M_0, \ldots, M_k))_{k\geq 0}$; furthermore $\mathbb{P}(\tau' < +\infty) = 1$ since the Lyapunov exponent π_μ equals 0, which yields

$$\limsup_{n\to+\infty} \log|R_n| = \limsup_{n\to+\infty} \log|\tilde{x} \cdot R_n| = +\infty$$

(this last property is a direct consequence of the ergodic theorem applied here to the Birkhof sums $\log|\tilde{x} \cdot R_n|, n \geq 0$ [5]).

Therefore, for any $\delta > 0$ and c given from Lemma 2.1, there exists $\kappa \geq 1$ such that $\mathbb{P}(\tau' = \kappa) > 0$. Moreover, we also have

$$\begin{aligned} \mathbb{P}(\log|R_\kappa| \geq \log c + \delta) &\geq \mathbb{P}(\log|R_\kappa| \geq \log c + \delta, \tau' = \kappa) \\ &= \mathbb{P}(\log|R_{\tau'}| \geq \log c + \delta, \tau' = \kappa) \\ &= \mathbb{P}(\tau' = \kappa) > 0. \end{aligned}$$

Since for any $\tilde{x} \in \mathbb{X}$, $g \in S^+(B)$, $|gx| \geq \frac{|g|}{c}$, it follows that

$$\{g \mid \log|g| \geq \log c + \delta\} \subset \{g \mid \forall \tilde{x} \in \mathbb{X}, \log|\tilde{x}g| \geq \delta\}.$$

Thus,

$$\begin{aligned} 0 < \mathbb{P}(\log|R_\kappa| \geq \log c + \delta) &= \mu^{*\kappa}\{g \mid \log|g| \geq \log c + \delta\} \\ &\leq \mu^{*\kappa}\{g \mid \forall \tilde{x} \in \mathbb{X},\ \log|\tilde{x}g| \geq \delta\}, \end{aligned}$$

which is the assertion of the claim (4.2).

Let us go into the proof of Lemma 3.1. For any $\tilde{x} \in \mathbb{X}$, $a \geq 0$ and $\lambda \in (0,1)$, there exists some constant $C(\lambda) > 0$ such that $(t+1)e^{-t} \leq C(\lambda)e^{-\lambda t}$ for any $t > 0$. From now on, we fix such a $\lambda \in (0,1)$ and, for technical reasons which appear in Step 1 below, assume that $0 < \lambda \leq \epsilon_0$ where ϵ_0 is the constant of hypothesis H1. Hence

$$\begin{aligned}
\widehat{\mathbb{E}}_{\tilde{x},a}\Big[\sum_{n=0}^{+\infty} e^{-S_n}\Big] &\leq 1 + \frac{1}{h(\tilde{x},a)} \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-S_n} h(X_n, S_n); S_1 > 0, \ldots, S_n > 0\right) \\
&\leq 1 + \frac{c}{h(\tilde{x},a)} \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-S_n}(1+S_n); S_1 > 0, \ldots, S_n > 0\right) \\
&\leq 1 + \frac{c\, C(\lambda)}{h(\tilde{x},a)} \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-\lambda S_n}; S_1 > 0, \ldots, S_n > 0\right),
\end{aligned}$$

where c is introduced in inequality (2.5).

We define a function Φ for any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$ as follow:

$$\Phi(\tilde{x}, a) := \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-\lambda S_n}; S_1 > 0, \ldots, S_n > 0\right)$$

Notice that for any $\tilde{x} \in \mathbb{X}$, the function $\Phi(\tilde{x}, .)$ increases on $\mathbb{R}$. To prove Lemma 3.1, it is sufficient to check that for any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$

$$\Phi(\tilde{x}, a) < +\infty. \tag{4.3}$$

Let us explain briefly the strategy of our proof, inspired by Lemma 3.2 in [10]. J. Geiger and G. Kersting first use in a crucial way the so-called "duality principle" (that is the fact that for a classical random walk $(S_n)_{n\geq 1}$ on $\mathbb{R}$ the vectors $(S_1, S_2, \cdots, S_n)$ and $(S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n)$ have the same distribution) and they prove in this context of random walks with i.i.d. increments that the quantity $\sum_{n=1}^{+\infty} \mathbb{E}_a[e^{-\lambda S_n}; S_1 > 0, \ldots, S_n > 0]$ is finite when $a = 0$. Second, they extend this property for any $a \geq 0$.

In the multi-dimensional case, it is more complicated to apply the duality principle and we can only prove at the beginning that the quantity $\Phi(\tilde{x}, a_0)$ is finite for some $a_0 < 0$ (without any control on the value of a_0). To extend this property to $\Phi(\tilde{x}, a)$ for $a > a_0$, we use crucially hypothesis H5 ; we refer to Step 2 below, especially inequality (4.4), and emphasize that the argument holds only when $-a_0 \leq \delta$). To avoid this difficulty, we modify the function Φ by introducing the functions Φ_κ associated with the κth power of convolution $\mu^{*\kappa}$ of μ. For any $\tilde{x} \in \mathbb{X}, a \in \mathbb{R}$, let

$$\Phi_\kappa(\tilde{x}, a) := \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-\lambda S_{n\kappa}}; S_\kappa > 0, \ldots, S_{n\kappa} > 0\right).$$

The relation is that $\Phi(\tilde{x},a) \preceq \Phi_\kappa(\tilde{x},a)$ for any $\tilde{x} \in \mathbb{X}, a \in \mathbb{R}$, proved in Step 1 below. Then, by using the duality principle, we bound from above $\Phi_\kappa(\tilde{x},a)$ by a new quantity $\Psi_\kappa(\tilde{x})$ defined below for any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$. Finally, we prove $\Psi_\kappa(\tilde{x}) < +\infty$ by using the ascending ladder epochs associated to the Markov walk $(L_n \cdot x, \log|L_n x|)_{n\geq 0}$ and the Elementary Renewal Theorem.

We set $L_0 = 0$ and denote $L_n := M_{n-1} \dots M_0$ the left product of the matrices $M_0, \dots, M_n$ when $n \geq 1$. Set, for any $\tilde{x} \in \mathbb{X}$,

$$\Psi_\kappa(\tilde{x}) := \sum_{n=1}^{+\infty} \mathbb{E}\Big[|L_{n\kappa}x|^{-\lambda}; |L_{n\kappa}x| > |L_{(n-1)\kappa}x|, \dots, |L_{n\kappa}x| > 1\Big].$$

Property (4.3) is a direct consequence of the four following steps :

1. For any $\kappa \geq 1$, there exists a constant $C(\kappa) > 0$ such that, for any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$,
$$\Phi(\tilde{x},a) \leq C(\kappa)(1 + \Phi_\kappa(\tilde{x},a)).$$
2. If there exist some $\kappa \geq 1, \tilde{x}_0 \in \mathbb{X}$ and $a_0 < 0$ such that $0 < \Phi_\kappa(\tilde{x}_0, a_0) < +\infty$, then
$$\forall \tilde{x} \in \mathbb{X}, \forall a \in \mathbb{R} \qquad \Phi_\kappa(\tilde{x},a) < +\infty.$$
3. There exist $C_1 > 0$ and $a_1 < 0$ such that for any $\kappa \geq 1$, $\tilde{x} \in \mathbb{X}$ and $a < a_1$
$$\Phi_\kappa(\tilde{x},a) \overset{C_1}{\preceq} \Psi_\kappa(\tilde{x}).$$
4. For any $\kappa \geq 1$ and $\tilde{x} \in \mathbb{X}$
$$\Psi_\kappa(\tilde{x}) < +\infty.$$

With these steps at hand, we may prove (4.3) as follows. First, for any $a_0 \leq a_1 < 0$, we choose some δ_0 such that $\delta_0 > -a_0 > 0$. For each δ_0, there exists $\kappa_0 \geq 1$ such that $\mathbb{P}(\log|\tilde{x}R_{\kappa_0}| \geq \delta_0) > 0$ (see (4.2) above). Since $\delta_0 > -a_0$, we have $\mathbb{P}_{\tilde{x},a}(S_{\kappa_0} > 0) > 0$, which implies $\Phi_{\kappa_o}(\tilde{x}_0, a_0) > 0$. On the other hand, since $a_0 \leq a_1$, step 3 and step 4 yield $\Phi_{\kappa_o}(\tilde{x}_0, a_0) < +\infty$. Therefore, we can apply step 2 and it yields $\Phi_\kappa(\tilde{x},a) < +\infty$ for any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$. Finally, thanks to Step 1, (4.3) arrives.

Step 1. It is easy to see that

$$\begin{aligned}\Phi(\tilde{x},a) &\leq \sum_{r=1}^{\kappa-1} \mathbb{E}_{\tilde{x},a}[e^{-\lambda S_r}, S_r > 0] + \sum_{n=1}^{+\infty}\sum_{r=0}^{\kappa-1} \mathbb{E}_{\tilde{x},a}[e^{-\lambda S_{n\kappa+r}}; S_\kappa > 0, \dots, S_{n\kappa} > 0] \\ &\leq \kappa + \sum_{n=1}^{+\infty} \mathbb{E}_{\tilde{x},a}\left(e^{-\lambda S_{n\kappa}}; S_\kappa > 0, \dots, S_{n\kappa} > 0\right) \times \sum_{r=0}^{\kappa-1} \sup_{\tilde{y}\in\mathbb{X}} \mathbb{E}_{\tilde{y},0}\left(e^{-\lambda S_r}\right).\end{aligned}$$

By Hypothesis H1 and Lemma 2.1, for $0 \leq r < \kappa$,

$$0 \leq \mathbb{E}_{\tilde{y},0}[e^{-\lambda S_r}] = \mathbb{E}_{\tilde{y},0}[|L_r y|^{-\lambda}] \leq c^\lambda \mathbb{E}[|L_r|^{-\lambda}] \leq c^{(r+1)\lambda}\mathbb{E}[|M_1|^{-\lambda}] < +\infty.$$

Hence $0 \le \Phi(\tilde{x}, a) \le C(\kappa)(1 + \Phi_\kappa(\tilde{x}, a))$ with $C(\kappa) = \kappa\left(1 + c^{\kappa\lambda}\mathbb{E}[|M_1|^{-\lambda}]\right) < +\infty$.

Step 2. The inequality $\Phi_\kappa(\tilde{x}_0, a_0) > 0$ implies that $\mathbb{P}(\log|\tilde{x}_0 R_\kappa| > -a_0) > 0$; we thus fix $\delta > -a_0 > 0$ and $\kappa \ge 1$ such that $\mu^{*\kappa}(E_\delta) > 0$. Since $a_0 < 0$, this property may hold only when κ is large enough; this happens for instance when the support of μ is bounded. To simplify the notations, we assume that $-a_0 < \delta$ where δ is given by H5. We set $\kappa = 1$ and write

$$\begin{aligned}
\Phi(\tilde{x}_0, a_0) &= \sum_{n=1}^{+\infty} \mathbb{E}\Big[|\tilde{x}_0 R_n|^{-\lambda}; |\tilde{x}_0 R_1| > e^{-a_0}, \dots, |\tilde{x}_0 R_n| > e^{-a_0}\Big] \\
&\ge \int_{\{g \in S^+(B): |\tilde{x}_0 g| \ge e^{-a_0}\}} \sum_{n=2}^{+\infty} \mathbb{E}\Big[|\tilde{x}_0 g R_{1,n}|^{-\lambda}; |\tilde{x}_0 g| > e^{-a_0}, \\
&\qquad\qquad \dots, |\tilde{x}_0 g R_{1,n}| > e^{-a_0}\Big]\mu(dg) \\
&\ge \int_{E_\delta} \sum_{n=2}^{+\infty} \mathbb{E}\Big[|\tilde{x}_0 g R_{1,n}|^{-\lambda}; |\tilde{x}_0 g| \ge e^{\delta} > e^{-a_0}, |\tilde{x}_0 g R_{1,2}| > e^{-a_0}, \\
&\qquad\qquad \dots, |\tilde{x}_0 g R_{1,n}| > e^{-a_0}\Big]\mu(dg)
\end{aligned}$$

$$\begin{aligned}
&= \int_{E_\delta} |\tilde{x}_0 g|^{-\lambda} \sum_{m=1}^{+\infty} \mathbb{E}\Big[|(\tilde{x}_0 \cdot g) R_m|^{-\lambda}; |(\tilde{x}_0 \cdot g) R_1| > e^{-a_0 - \log|\tilde{x}_0 g|}, \\
&\qquad\qquad \dots, |(\tilde{x}_0 \cdot g) R_k| > e^{-a_0 - \log|\tilde{x}_0 g|}\Big]\mu(dg) \\
&= \int_{E_\delta} |\tilde{x}_0 g|^{-\lambda} \Phi(\tilde{x}_0 \cdot g, a_0 + \log|\tilde{x}_0 g|)\mu(dg),
\end{aligned}$$

so that

$$\Phi(\tilde{x}_0, a_0) \ge \int_{E_\delta} |\tilde{x}_0 g|^{-\lambda} \Phi(\tilde{x}_0 \cdot g, a_0 + \delta)\mu(dg). \tag{4.4}$$

Consequently, if $\Phi(\tilde{x}_0, a_0) < +\infty$ then $\Phi(\tilde{x}_0 \cdot g, a_0 + \delta) < +\infty$ for μ-almost all $g \in E_\delta$ and by iterating this argument, there thus exists a sequence $(g_k)_{k \ge 1}$ of elements of E_δ such that

$$\forall k \ge 1, \quad \Phi(\tilde{x}_0 \cdot g_1 \cdots g_k, a_0 + k\delta) < +\infty.$$

By Lemma 2.1, for any $\tilde{x}, \tilde{y} \in \mathbb{X}$ and $a \in \mathbb{R}$

$$\begin{aligned}
\Phi(\tilde{x}, a - \log c) &\le c^\lambda \sum_{n=1}^{+\infty} \mathbb{E}[|R_n|^{-\lambda}; |R_1| > e^{-a}, \dots, |R_n| > e^{-a}] \\
&\le c^{2\lambda} \Phi(\tilde{y}, a + \log c);
\end{aligned}$$

it follows that, by choosing k sufficiently great such that $a_0 + k\delta > a + 2\log c$, we have

$$\Phi(\tilde{x}, a) \leq \Phi(\tilde{x}_0 \cdot g_1 \cdots g_k, a + 2\log c) \leq \Phi(\tilde{x}_0 \cdot g_1 \cdots g_k, a_0 + k\delta) < +\infty.$$

Step 3. For any $0 \leq k < n$, denote $L_{n,k} := M_{n-1} \dots M_k$ and $L_{n,k} = I$ otherwise. Let $c > 1$ be the constant given by Lemma 2.1. For any $\tilde{x} \in \mathbb{X}$ and $a \in \mathbb{R}$, by using Lemma 2.1, we may write

$$\begin{aligned}\Phi_\kappa(\tilde{x}, a) &= \sum_{n=1}^{+\infty} \mathbb{E}\left(|\tilde{x}R_{n\kappa}|^{-\lambda}; |\tilde{x}R_\kappa| > e^{-a}, \dots, |\tilde{x}R_{n\kappa}| > e^{-a}\right) \\ &\leq c^\lambda \sum_{n=1}^{+\infty} \mathbb{E}\left(|R_{n\kappa}|^{-\lambda}; |R_\kappa| > \frac{e^{-a}}{c}, \dots, |R_{n\kappa}| > \frac{e^{-a}}{c}\right)\end{aligned}$$

so that, by duality principle and Lemma 2.1,

$$\begin{aligned}\Phi_\kappa(\tilde{x}, a) &\leq c^\lambda \sum_{n=1}^{+\infty} \mathbb{E}\left(|L_{n\kappa}|^{-\lambda}; |L_{n\kappa,(n-1)\kappa}| > \frac{e^{-a}}{c}, \dots, |L_{n\kappa}| > \frac{e^{-a}}{c}\right) \\ &= c^\lambda \sum_{n=1}^{+\infty} \mathbb{E}[|L_{n\kappa}|^{-\lambda}; |L_{n\kappa,(n-1)\kappa}| \times |L_{(n-1)\kappa}| > |L_{(n-1)\kappa}|\frac{e^{-a}}{c}, \\ &\qquad\qquad \dots, |L_{n\kappa}| > \frac{e^{-a}}{c}] \\ &\leq c^\lambda \sum_{n=1}^{+\infty} \mathbb{E}\left(|L_{n\kappa}|^{-\lambda}; |L_{n\kappa}| > |L_{(n-1)\kappa}|\frac{e^{-a}}{c^2}, \dots, |L_{n\kappa}| > \frac{e^{-a}}{c^2}\right) \\ &\leq c^{2\lambda} \sum_{n=1}^{+\infty} \mathbb{E}\left(|L_{n\kappa}x|^{-\lambda}; |L_{n\kappa}x| > |L_{(n-1)\kappa}x|\frac{e^{-a}}{c^4}, \dots, |L_{n\kappa}x| > \frac{e^{-a}}{c^4}\right)\end{aligned}$$

Consequently, setting $a_1 := -4\log c$ and using the fact that the map $a \mapsto \Phi_\kappa(\tilde{x}, a)$ is non decreasing for any $a \in \mathbb{R}$, one may write $\Phi_\kappa(\tilde{x}, a) \leq \Psi(\tilde{x})$ as long as $a < a_1$.

Step 4. To simplify the notations, we assume here $\kappa = 1$; the proof is the same when $\kappa \geq 2$. For any $\tilde{x} \in \mathbb{X}$ and $n \geq 0$, set $X'_n := L_n \cdot x$ and $S'_n := \log|L_n x|$; the random process $(X'_n, S'_n)_{n\geq 0}$ is a Markov walk on $\mathbb{X} \times \mathbb{R}$ starting from $(x, 0)$ and whose transitions are governed by the ones of the Markov chain $(X'_n)_{n\geq 0}$ on $\mathbb{X}$. To study the quantity $\Psi(\tilde{x})$, we follow the strategy developed in the case of one dimensional random walks on $\mathbb{R}$ with independent increments and we thus introduce the sequence $(\eta_j)_{j\geq 0}$ of ladder epochs of $(S'_n)_{n\geq 0}$ defined by

$$\eta_1 = 0,\ \eta_{j+1} = \eta_{j+1}(x) := \min\left\{n > \eta_j : \log|L_n x| > \log|L_{\eta_j} x|\right\},\ j \geq 0.$$

For any $\tilde{x} \in \mathbb{X}$, one may write

$$\begin{aligned} \Psi(\tilde{x}) &= \sum_{n=1}^{+\infty} \mathbb{E}\left(|L_n x|^{-\lambda}; \exists j \geq 1 \mid n = \eta_j\right) \\ &= \sum_{j=1}^{+\infty} \mathbb{E}\left(|L_{\eta_j} x|^{-\lambda}\right). \end{aligned} \tag{4.5}$$

Let Q' denote the transition kernel of the Markov walk $(X'_n, S'_n)_{n\geq 0}$ and $G_{Q'} := \sum_{n=0}^{+\infty} Q'^n$ its Green kernel. The sub-process $(X'_{\eta_j}, S'_{\eta_j})_{j\leq 0}$ is also a Markov chain, its transition kernel Q'_η is given by: for any bounded Borel function $\phi : \mathbb{X} \times \mathbb{R} \to \mathbb{C}$ and for any $x \in \mathbb{X}$, $a \in \mathbb{R}$,

$$\begin{aligned} Q'_\eta \phi(x,a) &= \mathbb{E}\left(\phi(X'_{\eta_1}, a + S'_{\eta_1}) | X'_0 = x\right) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}\left(\phi(L_n \cdot x, a + \log|L_n x|); \eta_1 = n\right) \\ &= \sum_{n=1}^{+\infty} \mathbb{E}[\phi(L_n \cdot x, a + \log|L_n x|); \\ &\qquad\qquad |L_1 x| \leq 1, \cdots, |L_{n-1} x| \leq 1, |L_n x| > 1]. \end{aligned}$$

Let G'_η denote the Green kernel associated with the process $(X'_{\eta_j}, S'_{\eta_j})_{j\geq 0}$; by (4.5)

$$\begin{aligned} 1 + \Psi(\tilde{x}) &= \sum_{j=0}^{+\infty} \mathbb{E}\left(|L_{\eta_j} x|^{-\lambda}\right) \\ &= \sum_{j=0}^{+\infty} \int_{\mathbb{X}} \int_{\mathbb{R}} e^{-\lambda a} (Q'_\eta)^j ((x,0), dyda) \\ &= \int_{\mathbb{X}} \int_{\mathbb{R}} e^{-\lambda a} G'_\eta ((x,0), dyda). \end{aligned}$$

The Markov walk $(X'_n, S'_n)_{n\geq 0}$ has been studied by many authors (see for instance [9], [14] or [13]). All the works are based on the fact that the transition kernel of the chain $(X'_n)_{n\geq 0}$ has some "nice" spectral properties, namely its restriction to the space of Lipschitz functions on $\mathbb{X}$ is quasi-compact. In particular, it allows these authors to prove that the classical renewal theorem remains valid for this Markov walk on $\mathbb{X} \times \mathbb{R}$ as long as it is not centered, that is $\pi_\mu = \lim_{n\to+\infty} \frac{1}{n} \mathbb{E}[\log|L_n|] \neq 0$; in this case one may prove in particular that, for any $\tilde{x} \in \mathbb{X}$, the quantity $G_{Q'}((x,0), \mathbb{X} \times [0,a])$ is equivalent to $\frac{a}{\pi}$ as $a \to +\infty$ [14]. For the behavior as $a \to +\infty$ of $G'_\eta((x,0), \mathbb{X} \times [0,a])$, the situation

is way different. On one hand, it is easier since for any $j \geq 1$ the random variables S'_{η_j} are strictly positive, one might thus expect a similar result; on the other hand, the control of the spectrum of the transition kernel Q'_η remains unfortunately unknown in this circumstance, in particular the transition kernel Q'_η does not even act on the space of continuous functions on $\mathbb{X}$. Nevertheless, we have the following weak result with the postponed proof at the end of this subsection.

FACT 4.1. *There exists $C > 0$ such that for any $\tilde{x} \in \mathbb{X}$ and $a \geq 0$*

$$G'_\eta((x,0), \mathbb{X} \times [0,a]) = \sum_{j=0}^{+\infty} \mathbb{P}(\log |L_{\eta_j} x| \leq a) \leq Ca.$$

It follows that

$$\begin{aligned} 1 + \Psi(\tilde{x}) &= \int_{\mathbb{X}} \int_{\mathbb{R}_*^+} e^{-\lambda a} G'_\eta((x,0), dyda) \\ &\leq e^\lambda \sum_{a=1}^{+\infty} e^{-\lambda a} G'_\eta((x,0), \mathbb{X} \times [a-1, a]) \\ &\leq e^\lambda \sum_{a=1}^{+\infty} e^{-\lambda a} G'_\eta((x,0), \mathbb{X} \times [0, a]) \\ &\leq Ce^\lambda \sum_{a=1}^{+\infty} a e^{-\lambda a} < +\infty. \end{aligned}$$

To complete the proof of Step 4, it remains to prove Fact 4.1. First, by definition of E_δ, for any $j \geq 0$ and $\tilde{x} \in \mathbb{X}$, we may write $S'_{\eta_{j+1}} - S'_{\eta_j} \geq \delta 1_{E_\delta}(M_{\eta_j})$; setting $\varepsilon_j := 1_{E_\delta}(M_{\eta_j})$, this yields $S'_{\eta_j} \geq \delta(\varepsilon_0 + \ldots + \varepsilon_{j-1})$ so that

$$\begin{aligned} G'_\eta((x,0), \mathbb{X} \times [0,a]) &= \sum_{j=0}^{+\infty} \mathbb{P}(X'_{\eta_j} \in \mathbb{X}, S'_{\eta_j} \in [0,a] | X'_0 = x) \\ &\leq \sum_{j=0}^{+\infty} \mathbb{E}\left(1_{[0,a]}(S'_{\eta_j}) | X'_0 = x\right) \\ &\leq \sum_{j=0}^{+\infty} \mathbb{E}\left(1_{[0,a]}(\delta(\varepsilon_0 + \cdots + \varepsilon_{j-1}))\right). \end{aligned}$$

To conclude, we use the fact that $(\varepsilon_i)_{i \geq 0}$ is a sequence of i.i.d. random variables; the Elementary Renewal Theorem for the Bernoulli random walk process $[(\varepsilon_0 + \ldots + \varepsilon_{j-1})]_{j \geq 0}$ implies

$$G'_\eta((x,0), X \times [0,a]) \leq \mathbb{E}\left(\sum_{j=1}^{+\infty} 1_{[0,a]}(\delta(\varepsilon_0 + \ldots + \varepsilon_{j-1}))\right) \preceq a.$$

4.3. *Proof of Lemma 3.2.* We claim that

$$\lim_{n\to+\infty} \widehat{\mathbb{E}}_{\tilde{x},a}\Big|\frac{1}{q_n^{(i)}} - \frac{1}{q_\infty^{(i)}}\Big| = 0. \tag{4.6}$$

By definition, the quantities $q_n^{(i)}$ are always less than or equal to 1. Therefore, (4.6) implies that the same property holds $\widehat{\mathbb{P}}_{\tilde{x},a}$-almost surely for $q_\infty^{(i)}$. Hence,

$$|q_n^{(i)} - q_\infty^{(i)}| = q_n^{(i)} q_\infty^{(i)} \Big|\frac{1}{q_n^{(i)}} - \frac{1}{q_\infty^{(i)}}\Big| \le \Big|\frac{1}{q_n^{(i)}} - \frac{1}{q_\infty^{(i)}}\Big|.$$

Using (4.6) again, we find that $\lim_{n\to+\infty} \widehat{\mathbb{E}}_{\tilde{x},a}|q_n^{(i)} - q_\infty^{(i)}| = 0$. In particular,

$$\lim_{n\to+\infty} \widehat{\mathbb{P}}^{(i)}_{\tilde{x},a}(Z_n \neq \tilde{\mathbf{0}}) = \lim_{n\to+\infty} \widehat{\mathbb{E}}_{\tilde{x},a} q_n^{(i)} = \widehat{\mathbb{E}}_{\tilde{x},a} q_\infty^{(i)},$$

which is the assertion of (3.12).

Finally, it remains to verify (4.6). From (3.2) and (3.3),

$$\begin{aligned}\Big|\frac{1}{q_n^{(i)}} - \frac{1}{q_\infty^{(i)}}\Big| \;\le\; & \Big|\frac{1}{\tilde{e}_i R_n \mathbf{1}}\Big| \\ & + \sum_{k=0}^{n-1} \Big|\frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,n}} - \frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,\infty}}\Big| + \sum_{k=n}^{+\infty} \Big|\frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,\infty}}\Big|\end{aligned}$$

so that

$$\widehat{\mathbb{E}}_{\tilde{x},a}\Big|\frac{1}{q_n^{(i)}} - \frac{1}{q_\infty^{(i)}}\Big| \le I_n + J_n + K_n$$

with

$$\begin{aligned} I_n &= \widehat{\mathbb{E}}_{\tilde{x},a}\Big|\frac{1}{\tilde{e}_i R_n \mathbf{1}}\Big|, \\ J_n &= \sum_{k=0}^{n-1} \widehat{\mathbb{E}}_{\tilde{x},a}\Big|\gamma_k \frac{\tilde{e}_i R_k (Y_{k+1,n} - Y_{k+1,\infty})}{(\tilde{e}_i R_k Y_{k+1,n})(\tilde{e}_i R_k Y_{k+1,\infty})}\Big|, \\ K_n &= \sum_{k=n}^{+\infty} \widehat{\mathbb{E}}_{\tilde{x},a}\Big|\frac{\gamma_k}{\tilde{e}_i R_k Y_{k+1,\infty}}\Big|. \end{aligned}$$

By using Lemma 2.1, it is obvious that Lemma 3.1 implies

$$\sum_{k=0}^{+\infty} \widehat{\mathbb{E}}_{\tilde{x},a}\left(\frac{1}{|R_k|}\right) < +\infty. \tag{4.7}$$

Besides, as a direct consequence of Lemma 2.1

$$I_n \preceq \widehat{\mathbb{E}}_{\tilde{x},a}\left(\frac{1}{|R_n|}\right), \tag{4.8}$$

$$\left|\gamma_k \frac{\tilde{e}_i R_k (Y_{k+1,n} - Y_{k+1,\infty})}{(\tilde{e}_i R_k Y_{k+1,n})(\tilde{e}_i R_k Y_{k+1,\infty})}\right| \preceq \frac{|Y_{k+1,n} - Y_{k+1,\infty}|}{|R_k|} \preceq \frac{1}{|R_k|}, \tag{4.9}$$

$$K_n \preceq \sum_{k=n}^{+\infty} \widehat{\mathbb{E}}_{\tilde{x},a}\left(\frac{1}{|R_k|}\right). \tag{4.10}$$

Hence

- (4.7) and (4.8) implies $I_n \to 0$ as $n \to \infty$.
- since $Y_{k+1,n}(\omega) \to Y_{k+1,\infty}(\omega)$ for any $\omega \in \Omega$ (see Section 3.1), this convergence holds in particular $\hat{\mathbb{P}}_{\tilde{x},a}$-almost surely; combining (4.7), (4.9) and the the dominated convergence Theorem, we obtain $J_n \to 0$ as $n \to +\infty$.
- by (4.7) and (4.10), the term C_l can be made arbitrarily small by choosing l sufficiently great.

Acknowledgments. We wish to thank the referees for many useful remarks and comments; their suggestions really improve the presentation of this manuscript. M. Peigné and C. Pham thank the Vietnam Institute for Advanced Studies in Mathematics (VIASM) in Ha Noi for generous hospitality and financial support during the first semester 2017; the present article has been finished during this period.

References.

[1] V.I. Afanasyev (1993) Limit theorem for the critical branching process in random environment. Discrete Math. Appl. **5**, no. **1**, 45–58. MR1221669

[2] K. Athreya, S. Karlin (1971) On branching process with random environment, I: extinction probabilities. Ann. Math. Statist. **42**, no. **5**, 1499–1520. MR0298780

[3] K. Athreya, S. Karlin (1971) On branching process with random environment, II: Limit theorem. Ann. Math. Statist. **42**, no. **6**, 1843–1858. MR0298781

[4] V. Bansaye, J. Berestycki (2009) Large deviations for Branching Processes in Random Environment. Markov Process. Related Fields, 15 (2009), 493–524. MR2598126

[5] Ph. Bougerol, J. Lacroix (1985) *Products of Random Matrices with Applications to Schrödinger Operators*. Birkhäuser. MR0886674

[6] E. E. Dyakonova (1999) Asymptotics behaviour of the probability of non-extinction for a multi-type branching process in a random environment. Discrete Math. Appl. **9**, no. **2**, 119–136. MR1702916

[7] E. E. Dyakonova, J. Geiger, V.A. Vatutin (2004) On the survival probability and a functional limit theorem for branching processes in random environment. Markov Process. related Fields, 10:2 , 289–306. MR2082575

[8] E. E. Dyakonova, V.A. Vatutin (2018) Multitype branching processes in random environment: survival probability for the critical case. Theory Probab. Appl., **62**, no. **4**, 506–521

[9] H. Furstenberg, H. Kesten (1960) Products of random matrices. Ann. Math. Statist. **31**, 457–469. MR0121828

[10] J. Geiger, G. Kersting (2002) The survival probability of a critical branching process in random environment. Theory. Probab. Appl., **45**, no. **3**, 518–526. MR1967796

[11] J. Geiger, G. Kersting, V. A. Vatutin (2003) Limit theorems for subcritical branching processes in random environment. Ann. de l'I.H.P. Probab. et Statist., **39**, no. **4**, 593–620. MR1983172

[12] I. Grama, E. Le Page, M. Peigné (2014) On the rate of convergence in the weak invariance principle for dependent random variables with applications to Markov chains. *Colloquium Mathematicum,* **134**, 1–55. MR3164936

[13] I. Grama, E. Le Page, M. Peigné (2016) Conditional limit theorems for products of random matrices. Probab. Theory Relat. Field . doi:10.1007/s00440-016-0719-z

[14] Y. Guivarc'h, J. Hardy (1988) Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov. Ann. Inst. H. Poincaré Probab. Statist. **24**, no. **1**, 73–98. MR0937957

[15] Y. Guivarc'h, E. Le Page, Q. Liu (2003) Normalisation d'un processus de branchement critique dans un environnement aléatoire. C. R. Acad. Sci. Paris Sér. I Math. 337, **9**, 603–608. MR2017734

[16] Y. Guivarc'h, Q. Liu (2001) Asymptotic properties of branching processes in a random environment. C. R. Acad. Sci. Paris Sér. I Math. 332, **4**, 339–344. MR1821473

[17] H. Hennion (1997) Limit theorems for products of positive random matrices, Ann. Probab. **25**, no. **4**, 1545–1587. MR1487428

[18] N. Kaplan (1974) Some results about multidimensional branching processes with random environments, Ann. Probab., no. **2**, 441-455. MR0378127

[19] M. V. Kozlov (1976) On the asymptotic bahaviour of the probability of non-extinction for critical branching processes in a random environment. Theor. of. Probab. and its Appl., Vol. **X**XI, no. **4**, 791-804. MR0428492

[20] E. Le Page (1982) Théorèmes limites pour les produits de matrices aléatoires, Springer Lecture Notes **928**, 258–303. MR0669072

[21] T.D.C. Pham (2017) Conditioned limit theorems for products of positive random matrices. ALEA, Lat. Am. J. Probab. Math. Stat. 15, 67–100 (2018)

[22] W.L. Smith, W. Wilkinson (1969) On branching processes in random environment. Adv. Math. Statist., 40:3, 814–827. MR0246380

[23] A. M. Zubkov (1994) Inequalities for the distribution of the numbers of simultaneous events. Survey of Appl. and Industry math. Ser. Probab. and Stat. **1**, 638–666.

Université de Bretagne-Sud
LMBA UMR CNRS 6205
Vannes, France.

E-mail: emile.le-page@univ-ubs.fr

Université Fr. Rabelais Tours
LMPT UMR CNRS 7350
Tours, France.

E-mail: Marc.Peigne@lmpt.univ-tours.fr
E-mail: Da-Cam.Pham@lmpt.univ-tours.fr

Chapitre 3

Simulation of Galton-Watson processes

The interest of this chapter is to simulate :

1. critical single-type process in fixed environment,

2. critical multi-type (two-type and three-type) process in fixed environment,

3. critical single-type process in random environment,

4. critical multi-type (two-type and three-type) process in random environment,

The code simulating the processes are in Python and C languages and accesible on the webpage of the author.

3.1 How to choose critical distributions ?

3.1.1 For single-type process in fixed environment

Let $p = [p(0), \ldots, p(d)]$ be the offspring distribution of the process, where d is the maximum number of children. Criticality requires that $\sum_{i=0}^{d} kp(k) = 1$. We first choose $p(k)$ for all $1 \leq k \leq d$, then it remains to find $p(0)$ such that it satisfies the probability property $\sum_{i=0}^{d} p(k) = 1$.

For instance, let $\mu = [0.5, 0.4, 0.1]$ be an offspring distribution, where 0.5 is the probability to have 0 child, 0.4 is the probability to have 1 child and 0.1 probability to have 2 children. The mean number of children is the value of the first derivative of the offspring generating function at 1, which is

$$f'(1) = 0 \times 0.5 + 1 \times 0.4 + 2 \times 0.1 = 0.6.$$

Therefore, criticality can be obtained for example in the distribution $[0.3, 0.4, 0.3]$.

3.1.2 For multi-type process in fixed environment

For multi-type-processes, we have to work with the mean matrices. If d is the number of types in the process, then $d \times d$ is the dimension of the mean matrix. In fixed environment, the criteria is that the Perron root of the mean matrix M equals 1. Notice that what we need is the offspring distribution in order to apply in Python code. If we start with a mean matrix with its Perron root 1, then this matrix satisfies the criteria to be in critical case, but to trace back to the possible offspring distributions is not an easy problem because we may end up with a systems of many variables which is not easy to solve. Therefore, we try to find a way that answer the following question : if we adjust the mean matrix, how does the corresponding offspring distribution change ?

Let us consider a two-type-process in the given environment $\mu = (\mu_1, \mu_2)$, where μ_i is a probability measure on $\mathbb{N}^2$ that gives the offspring distribution of parent typed i, for $i = 1, 2$. We assume that $\mu_i(1,0) = \mu_i(0,1) = 0$ for any $i = 1, 2$. Therefore, it is easy to see that the corresponding mean matrix M has the Perron root $\rho(M)$ greater than 1. The mean matrix M here should be checked if it is positively regular. After, we modify μ to μ' by creating a combination of μ with some other Dirac mass at $(0,0), (0,1)$ and $(1,0)$ as follows :

$$\mu'_i = a_i\delta_{(0,0)} + b_i\delta_{(1,0)} + c_i\delta_{(0,1)} + d_i\mu_i, \tag{3.1.1}$$

where $a_i, b_i, c_i, d_i \in]0,1[$ and $a_i + b_i + c_i + d_i = 1$, for $i = 1, 2$. From the expression (3.1.1) above, we can calculate the coefficients $m'_{i,j}$ of the mean matrix M' corresponding to μ'. The strategy is : from a super-critical mean matrix with a given distribution, we construct another new distribution depending on the old one such that the new mean matrix satisfies the critical criteria. Notice that if $M' = \frac{1}{\rho(M)}M$ then $\rho(M') = 1$. Solving the equation $M' = \frac{1}{\rho(M)}M$, we obtain

$$\begin{cases} b_1 + d_1 m_{1,1} = \frac{1}{\rho} m_{1,1} \\ c_1 + d_1 m_{1,2} = \frac{1}{\rho} m_{1,2} \\ b_2 + d_2 m_{2,1} = \frac{1}{\rho} m_{2,1} \\ c_2 + d_2 m_{2,2} = \frac{1}{\rho} m_{2,2}, \end{cases}$$

which implies

$$\begin{cases} b_1 = (\frac{1}{\rho} - d_1) m_{1,1} \\ c_1 = (\frac{1}{\rho} - d_1) m_{1,2} \\ b_2 = (\frac{1}{\rho} - d_2) m_{2,1} \\ c_2 = (\frac{1}{\rho} - d_2) m_{2,2}. \end{cases}$$

Then, we can choose d_1 and d_2 such that $0 < d_1, d_2 < \frac{1}{\rho}$ and find the other coefficients b_1, c_1, b_2, c_2. Since $a_i + b_i + c_i + d_i = 1$, we can find finally a_i for $i = 1, 2$. Notice that we can always choose d_i close enough to $\frac{1}{\rho}$ so that $b_i, c_i \in]0,1[$. With all the found coefficients, we can construct explicitly the new measure μ'.

Example 1 : Two-type process in fixed environment.

Let μ_1 and μ_2 given in Tables 3.3 be two offspring distributions, where x is the number of children of type 1 and y is the number of children of type 2. Then, the mean matrix M is found as

	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.25
$x=1$	0	0.25	0.5

TABLE 3.1 – μ_1

	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0.1	0.1
$x=2$	0.1	0.2	0.4

TABLE 3.2 – μ_2

TABLE 3.3 – Two-type process in fixed environment

$$M = \begin{pmatrix} 0.75 & 1.75 \\ 1.6 & 1.5 \end{pmatrix},$$

and its Perron root $\rho(M) = \dfrac{45 + \sqrt{4705}}{40}$. It is obvious that M is positively regular. Solving the equation $M' = \frac{1}{\rho(M)} M$ by choosing $d_1 = 0.3$ and $d_2 = 0.2$, we obtain

$$M' \simeq \begin{pmatrix} 0.2641 & 0.6162 \\ 0.5634 & 0.5282 \end{pmatrix},$$

which yields the new measure described in Table 3.6.

	$y=0$	$y=1$	$y=2$
$x=0$	0.5697	0.0912	0.075
$x=1$	0.0391	0.075	0.15

TABLE 3.4 – μ_1'

	$y=0$	$y=1$	$y=2$
$x=0$	0.3284	0.2282	0.02
$x=1$	0.2434	0.02	0.02
$x=2$	0.02	0.04	0.08

TABLE 3.5 – μ_2'

TABLE 3.6 – Two-type process in fixed environment

Example 2 : Three-type process in fixed environment.

Let μ_1, μ_2 and μ_3 given in Tables 3.10 be three offspring distributions, where x is the number of children of type 1, y is the number of children of type 2 and z is the number of children of type 3.

Then, the mean matrix M is found as

$$M = \begin{pmatrix} 1.7 & 0.6 & 0.7 \\ 0.6 & 1.8 & 0.6 \\ 0.5 & 0.7 & 1.7 \end{pmatrix},$$

with the Perron root $\rho(M) \sim 2.966$ and we round off $\frac{1}{\rho(M)} \sim 0.33715$. It is obvious that M is positively regular. Notice that the new measure is of the following combination.

$$\begin{aligned} \mu_i' \;=\; & a_0^i \delta_{(0,0,0)} + a_1^i \delta_{(1,0,0)} + a_2^i \delta_{(0,1,0)} + a_3^i \delta_{(0,0,1)} \\ & + a_4^i \delta_{(1,1,0)} + a_5^i \delta_{(1,0,1)} + a_6^i \delta_{(0,1,1)} + a_7^i \mu_i. \end{aligned} \tag{3.1.2}$$

$z=0$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0
$x=2$	0.2	0.1

$z=1$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.3
$x=2$	0.2	0.2

TABLE 3.7 – μ_1

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.3
$x=1$	0	0	0.1

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0.2	0.3

TABLE 3.8 – μ_2

$z=0$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0

$z=1$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.3

$z=2$	$y=0$	$y=1$
$x=0$	0.2	0.3
$x=1$	0.1	0.1

TABLE 3.9 – μ_3

TABLE 3.10 – Three-type process in fixed environment

Solving the equation $M' = \frac{1}{\rho(M)}M$, taking into account (3.1.2) and by choosing $a_7^1 = 0.02, a_7^2 = 0.01$ and $a_7^3 = 0.15$, we obtain the new measure described in Table 3.14.

$z=0$	$y=0$	$y=1$
$x=0$	0.36855	0.00029
$x=1$	0.269155	0.12
$x=2$	0.004	0.002

$z=1$	$y=0$	$y=1$
$x=0$	0.002005	0.07
$x=1$	0.15	0.006
$x=2$	0.004	0.004

TABLE 3.11 – μ_1'

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0.23855	0.50887	0.003
$x=1$	0.00629	0.04	0.001

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0.00629	0.04	0.001
$x=1$	0.15	0.002	0.003

TABLE 3.12 – μ_2'

$z=0$	$y=0$	$y=1$
$x=0$	0.457267	0.021005
$x=1$	0.023573	0.03

$z=1$	$y=0$	$y=1$
$x=0$	0.198155	0.08
$x=1$	0.04	0.045

$z=2$	$y=0$	$y=1$
$x=0$	0.03	0.045
$x=1$	0.015	0.015

TABLE 3.13 – μ_3'

TABLE 3.14 – Three-type process in fixed environment

3.1.3 For single-type process in random environment

Let $P^i = (p^i(k))_{1\leq k}$ be offspring distributions, for $1 \leq i \leq N$, where N is the number of offspring distributions in the environment and α_i the probability to choose the distribution

P^i. Let $m_i = \sum_{k=0}^{+\infty} kp^i(k)$. Then the criteria is $\mathbb{E}[\log f'(\mathbf{1})] = 0$, that is $\sum_{i=1}^{N} \alpha_i \log m_i = 0$. In this case, we have to choose P^i and also α_i.

For instance, given two offspring distributions μ_0 and μ_1, the criteria is the value $\mathbb{E}(\log f_0'(1)) = 0$. We introduce two ways to obtain the criticality : to choose the offspring distributions with the mean number of children equal 1 or to choose the distribution $\alpha = (\alpha_0, \alpha_1)$ such that

$$\mathbb{E}(\log f_0'(1)) = \alpha_0 \log m_0 + \alpha_1 \log m_1 = 0$$

More precisely, for the first way, we choose $\mu_0 = [0.25, 0.5, 0.25]$ and $\mu_1 = [0.4, 0.2, 0.4]$ with the mean values corresponding $m_0 = 1$ and $m_1 = 1$. Therefore, the criteria $\mathbb{E}(\log f_0'(1)) = \alpha_0 \log m_0 + \alpha_1 \log m_1 = 0$, no matter the value of α. For the second way, we choose $\mu_0 = [0.5, 0.3, 0.2]$ and $\mu_1 = [0.2, 0.3, 0.5]$ with $m_0 = 0.7$ and $m_1 = 1.3$. Then we solve a system of equations to find α :

$$\begin{cases} \alpha_0 \log m_0 + \alpha_1 \log m_1 = 0, \\ \alpha_0 + \alpha_1 = 1, \\ \alpha_0 \times \alpha_1 > 0. \end{cases}$$

The solution is $\alpha = \left(\dfrac{\log 1.3}{\log 1.3 - \log 0.7}, \dfrac{-\log 0.7}{\log 1.3 - \log 0.7} \right)$. For more samples of distribution, we do similarly.

Example 3 : Single-type process in random environment with the offspring distribution described as follows : environment distribution : $\alpha = \left(\dfrac{\log 1.3}{\log 1.3 - \log 0.7}, \dfrac{-\log 0.7}{\log 1.3 - \log 0.7} \right)$ and offspring distribution : $[0.5, 0.3, 0.2]$ and $[0.2, 0.3, 0.5]$.

3.1.4 For multi-type process in random environment

In fact the probability to fall into a sequence of matrices such that with respect to some distribution, the corresponding Lyapunov exponent is different from 0 is high. Therefore, we first calculate the Lyapunov exponent of those matrices, denoted by γ. Then using the technique presented in the part of multi-type in fixed environment with $\exp(\gamma)$ instead of the Perron root $\rho(M)$, we can obtain a new set of probability measures. The mean matrices corresponding to those new measures do have the Lyapunov's exponent 0.

More precisely, given μ and ν the two offspring probability measures, we can calculate the mean matrices M_μ and M_ν. Let $\alpha = (\alpha_1, \alpha_2)$ be the distribution of the environment, that is with probability α_1 we choose μ and with probability α_2 we choose ν as the offspring distribution. The Lyapunov's exponent is $\gamma = \lim\limits_{n \to +\infty} \dfrac{1}{n} \mathbb{E}|M_1 \dots M_n|$, where M_i are i.i.d. random matrices chosen from the set $\{M_\mu, M_\nu\}$ with distribution α. By using the technique of combination with Dirac mass applying to $\exp(\gamma)$, we can construct a new set of probability measures $\{\mu', \nu'\}$ with the corresponding mean matrices M'_μ and M'_ν. Since $M'_\mu = \frac{1}{\exp(\gamma)} M_\mu$ and $M'_\nu = \frac{1}{\exp(\gamma)} M_\nu$, the new Lyapunov's exponent γ' can be explicited as

follows

$$\begin{aligned}\gamma' &= \lim_{n\to+\infty}\frac{1}{n}\mathbb{E}|M_1'\dots M_n'| \\ &= \lim_{n\to+\infty}\frac{1}{n}\log|e^{-\gamma}M_n\dots e^{-\gamma}M_1| \\ &= \lim_{n\to+\infty}\frac{1}{n}\log e^{-n\gamma}|M_n\dots M_1| \\ &= -\gamma+\lim_{n\to+\infty}\frac{1}{n}\log|M_n\dots M_1| \\ &= 0. \end{aligned} \tag{3.1.3}$$

Notice that to obtain criticality, we have to deal with the Perron root and the limit for Lyapunov's exponent, which are not found absolutely precisely, thus rounding off numbers is neccessary. Therefore, criticality considered in python code is relative but quite fine. Furthermore, the mean matrices should satisfy the strong irreducibility, which we go into more detail in the next subsection.

Example 4 : Two-type process in random environment.

Environment distribution : [0.4, 0.6].

Offspring distribution : Tables 3.19.

	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.25
$x=1$	0	0.25	0.5

Table 3.15 – μ_1

	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0.1	0.1
$x=2$	0.1	0.2	0.4

Table 3.16 – μ_2

	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.4
$x=2$	0.4	0.2

Table 3.17 – ν_1

	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.2
$x=1$	0	0.2	0.1
$x=2$	0.2	0.2	0.1

Table 3.18 – ν_2

Table 3.19 – Two-type process in random environment

The Lyapunov's exponent is round off to $\exp(-\gamma) = 0.3995$. New measures are described in Table 3.24.

Example 5 : Three-type process in random environment.

Environment distribution : [0.4, 0.6].

Offspring distribution : Table 3.31. The Lyapunov's exponent is round off to $\exp(-\gamma) = 0.31956$. New measures are described in Table 3.38.

	$y=0$	$y=1$	$y=2$
$x=0$	0.45125	0.174125	0.075
$x=1$	0.074625	0.075	0.15

Table 3.20 – μ'_1

	$y=0$	$y=1$	$y=2$
$x=0$	0.18155	0.29925	0.02
$x=1$	0.3192	0.02	0.02
$x=2$	0.02	0.04	0.08

Table 3.21 – μ'_2

	$y=0$	$y=1$
$x=0$	0.4211	0.0897
$x=1$	0.2392	0.1
$x=2$	0.1	0.05

Table 3.22 – ν'_1

	$y=0$	$y=1$	$y=2$
$x=0$	0.22625	0.2994	0.03
$x=1$	0.32435	0.03	0.015
$x=2$	0.03	0.03	0.015

Table 3.23 – ν'_2

Table 3.24 – Two-type process in random environment

3.2 Matrices satisfy the strong irreducibility hypothesis

3.2.1 For 2×2 matrices

Step 1. Two positive matrices A and B with distinct eigenvectors create a semi-group $T_{(A,B)}$ which is strong irreducible, that is there is no finite union of proper subspaces U of $T_{(A,B)}$ such that for any $g \in T_{(A,B)} : g(U) \subset U$. Suppose that there exists a finite union U such that $g(U) \subset U$ for any $g \in U$. Denote by x_A, x'_A and x_B, x'_B the distinct eigenvectors of the matrices A and B. We consider the projective space $P(\mathbb{R}^2)$. Notice that for any $x \in P(\mathbb{R}^2)$ such that $x \neq x_A$, we obtain the convergence in direction $A^n \cdot x \to x_A$, where x_A is the dominant eigenvector of A.

Case 1 : when $\{x'_A, x'_B\} \not\subset U$, for any $x \in U$, we see that $A^n x \to x_A$ (! with finite union because $A^n x \neq x_A$).

Case 2 : when $\{x'_A\} \subset U$ and when $\{x'_B\} \not\subset U$, for any $x \in U$, we see that $A^n B x \to x_A$ (! with finite union).

Case 3 : when $\{x'_A, x'_B\} \subset U$, for any $x \in U$, there exists k such that

$$A^k x \in \{x'_A \mathbb{R}, N_\epsilon(x_A)\} \text{ and } x'_B \notin \{x'_A \mathbb{R}, N_\epsilon(x_A)\},$$

which yields $B^n A^k x \to x_B$ (! with finite union).

Step 2. If a semi-group is strong irreducible, then so are all the bigger semi-groups that contain it. We prove for instant a set of three matrices. For any positive matrix C, we consider the semi-group $T_{(A,B,C)}$ generated by the set of matrices $\{A, B, C\}$. Therefore $T_{(A,B)} \subset T_{(A,B,C)}$. Since $T_{(A,B)}$ is strong irreducible, so is $T_{(A,B,C)}$. Indeed, we can verify this property by using contradiction method. Suppose that $T_{(A,B,C)}$ does not satisfy the strong irreducible property, that is there exists some finite union U such that for any $g \in T_{(A,B,C)} : g(U) \subset U$, which yields for any $g \in T_{(A,B)}, g(U) \subset U$, which is contradiction because $T_{(A,B)}$ is strongly irreducible.

$z=0$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0
$x=2$	0.2	0.1

$z=1$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.3
$x=2$	0.2	0.2

TABLE 3.25 – μ_1

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.3
$x=1$	0	0	0.1

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0.2	0.3

TABLE 3.26 – μ_2

$z=0$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0

$z=1$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.3

$z=2$	$y=0$	$y=1$
$x=0$	0.2	0.3
$x=1$	0.1	0.1

TABLE 3.27 – μ_3

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0	0.3

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.2
$x=1$	0	0.2	0.2

TABLE 3.28 – ν_1

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.1
$x=1$	0	0	0.05
$x=2$	0.05	0.15	0.1

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0	0	0.05
$x=1$	0	0.1	0.1
$x=2$	0.1	0.1	0.1

TABLE 3.29 – ν_2

$z=0$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0
$x=2$	0.15	0.3

$z=1$	$y=0$	$y=1$
$x=0$	0	0
$x=1$	0	0.1
$x=2$	0.05	0.4

TABLE 3.30 – ν_3

TABLE 3.31 – Three-type process in random environment

3.2.2 For 3×3 matrices

The same idea is applied for 3×3 matrices like in case of 2×2 matrices. For instance, we consider two 3×3 matrices A and B, where A has three different eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$ with u_1, u_2, u_3 the three associate eigenvectors, and B has three different eigenvalues $\gamma_1 > \gamma_2 > \gamma_3$ with v_1, v_2, v_3 the three associate eigenvectors, such that $\lambda_i \neq \lambda_j$, $\gamma_i \neq \gamma_j$, $\mathbb{R}u_i \neq \mathbb{R}v_j$ and $\mathbb{R}(u_i \wedge u_j) \neq \mathbb{R}(v_k \wedge v_l)$, for $i \neq j, k \neq l$ and $i, j, k, l = 1, 2, 3$. In other words, there does not exist any directions, planes or combination of directions and planes that stay invariant under the action of products of matrices composed by A and B.

$z=0$	$y=0$	$y=1$
$x=0$	0.44132	0.041736
$x=1$	0.218252	0.03
$x=2$	0.03	0.015

$z=1$	$y=0$	$y=1$
$x=0$	0.048692	0.03
$x=1$	0.04	0.045
$x=2$	0.03	0.03

TABLE 3.32 – μ_1'

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0.33132	0.355208	0.03
$x=1$	0.071736	0.01	0.01

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0.051736	0.03	0.01
$x=1$	0.05	0.02	0.03

TABLE 3.33 – μ_2'

$z=0$	$y=0$	$y=1$
$x=0$	0.533276	0.023692
$x=1$	0.02978	0.01

$z=1$	$y=0$	$y=1$
$x=0$	0.133252	0.05
$x=1$	0.02	0.06

$z=2$	$y=0$	$y=1$
$x=0$	0.04	0.06
$x=1$	0.02	0.02

TABLE 3.34 – μ_3'

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0.519364	0.165208	0.02
$x=1$	0.013692	0.03	0.06

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0.011736	0.02	0.04
$x=1$	0.04	0.04	0.04

TABLE 3.35 – ν_1'

$z=0$	$y=0$	$y=1$	$y=2$
$x=0$	0.511974	0.048906	0.015
$x=1$	0.045862	0.15	0.0075
$x=2$	0.0075	0.0225	0.015

$z=1$	$y=0$	$y=1$	$y=2$
$x=0$	0.013258	0.03	0.0075
$x=1$	0.05	0.015	0.015
$x=2$	0.015	0.015	0.015

TABLE 3.36 – ν_2'

$z=0$	$y=0$	$y=1$
$x=0$	0.38643	0.025648
$x=1$	0.0267164	0.1
$x=2$	0.015	0.03

$z=1$	$y=0$	$y=1$
$x=0$	0.020758	0.05
$x=1$	0.05	0.01
$x=2$	0.005	0.04

TABLE 3.37 – ν_3'

TABLE 3.38 – Three-type process in random environment

3.3 Simulation with tree graphs and stairstep graphs

Notice that all the simulations presented in this section use the offspring distributions and environment distributions which are introduced in the following subsection.

3.3.1 Tree graphs

The vertical axe of the following graphs corresponds to the number of generation. Those trees are downward oriented and they represent some possible samples of random

trees generated with the following distributions.

• Single-type process in fixed environment :

Offspring distribution : $[0.3, 0.4, 0.3]$. Tree graphs are presented in Figure 3.1.

• Single-type process in random environment :

Environment distribution : $[\frac{\log 1.3}{\log 1.3 - \log 0.7}, \frac{-\log 0.7}{\log 1.3 - \log 0.7}]$.

Offspring distribution : $[0.5, 0.3, 0.2]$ and $[0.2, 0.3, 0.5]$. Tree graphs are presented in Figure 3.2.

• Two-type process in fixed environment :

Offspring distribution : Table 3.6. Tree graphs are presented in Figure 3.3.

• Two-type process in random environment :

Environment distribution : $[0.4, 0.6]$.

Offspring distribution : Table 3.24. Tree graphs are presented in Figure 3.4.

• Three-type process in fixed environment :

Offspring distribution : Table 3.14. Tree graphs are presented in Figure 3.5.

• Three-type process in random environment :

Environment distribution : $[0.4, 0.6]$.

Offspring distribution : Table 3.38. Tree graphs are presented in Figure 3.6.

3.3.2 Stairstep graphs

The Figure 3.13 describes 100 realizations of critical Galton-Watson branching processes in fixed and in random environment over 150 generations, with the according offspring distributions presented in subsection above. The horizontal axe is for the number of generations and the verical axe is for the size of the population.

Comments.

As it is supposed to be, the process dies more slowly in random environment rather than in fixed environment. Notice that there is no absolute comparison between the cases according to the distributions or to the number of types. However, intuitive prediction says that with more types of individual, the population dies more slowly and it survives better in random environment than in fixed environment.

The graphs are plotted in order to visualize the extinction probability. Following the graphs, it is clear that most of the trajectories end at 0 quickly after around 40 generations and only a few populations survive. This fact matches the probability of extinction theoretically. Moreover, we cannot say anything about the effect of the number of types on the extinction probability at the moment. To obtain a more persuasive conclusion, we need a more precise approximation, which can be done by optimizing the codes that simulates the process. However, this problem is out of the scope of this thesis. Still one question naturally arises from observing the graphs : what is the first moment that one of the types gets extinction. This can be the first step to the answer of the question about the effect of the number of types on the extinction probability.

3.3.3 Simulation of rate of convergence

In this subsection, we simulate critical Galton-Watson processes given the same offspring distributions as in previous subsection, with the maximum number of generations $1e2, 1e3, 1e4, 1e5$ and $1e6$. Each experiment is repeated $1e6$ times and Table 3.43 gives us the mean frequencies obtained. Notice that the codes used in simulating those processes are in C and by the help of the central computer in Insitut Denis Poisson, we could not calculate precisely the frequency and had to truncate the size of the population if it exceeds $1e6$ in all cases except single-type in fixed environment case, and consider that this population survives. This truncation undoubtedly perturbs the rate of convergence showed in Figures 3.14 and 3.15.

In order to understand Figures 3.14 and 3.15, we briefly provide some ordinary comments. The frequencies obtained play the role of survival propbability at time n, where n is the maximum number of generations in each experiment. Let q_n and q'_n be the survival probability in fixed and in random environment of a critical single-type Galton-Watson process. Then, it is well known that $q_n \sim \frac{c_1}{n}$ and $q'_n \sim \frac{c_2}{\sqrt{n}}$. We take the logarithm of base 10 and obtain $\log(q_n) \sim \log c_1 - \log n$ and $\log(q'_n) \sim \log c_2 - \frac{1}{2}\log n$. After, we plot the graphs of these two functions with the variable $\log n$ for x-axis and $\log(frequency(n))$ for y-axis. With the graphs obtained, they should show the difference of slopes between two graphs (slope -1 in fixed case and $-\frac{1}{2}$ in random case). However, because of technical contraints, we have to truncate the size of populations who exceed some certain number ($1e6$ in this context) and this truncation creates un effect, that is it increases the frequency of survival. However, this approach opened a door for further work after.

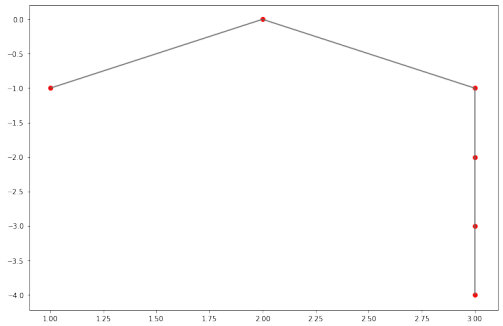

FIGURE 3.1 – Single-type process in fixed environment

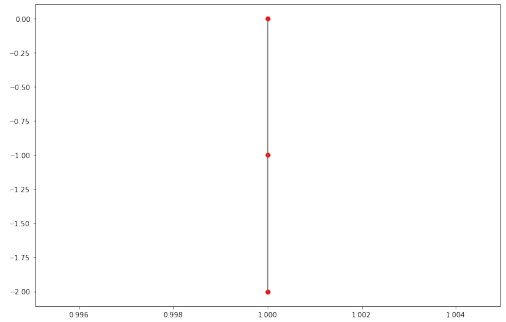

FIGURE 3.2 – Single-type process in random environment

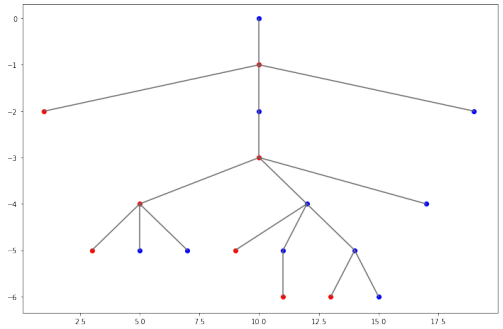

FIGURE 3.3 – Two-type process in fixed environment

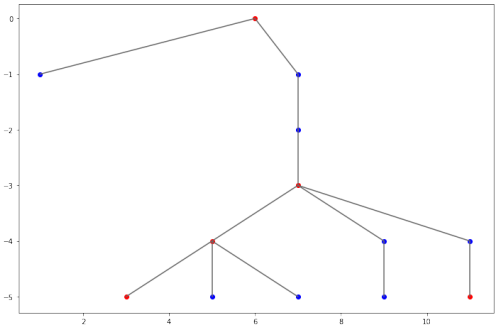

FIGURE 3.4 – Two-type process in random environment

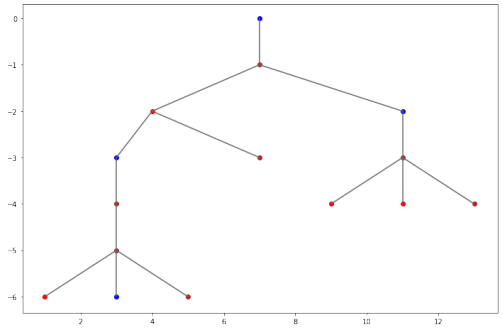

FIGURE 3.5 – Three-type process in fixed environment

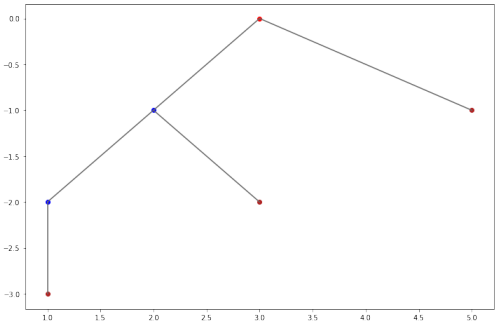

FIGURE 3.6 – Three-type process in random environment

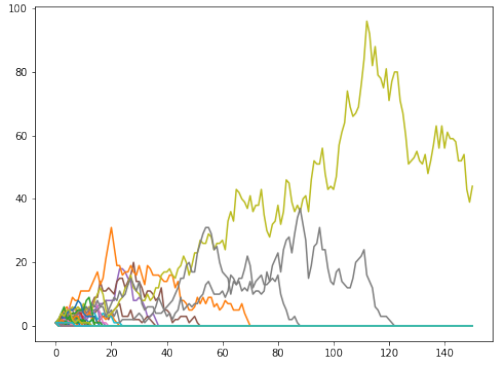

FIGURE 3.7 – Single-type process in fixed environment

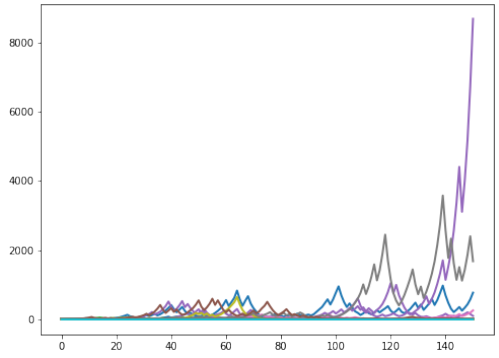

FIGURE 3.8 – Single-type process in random environment

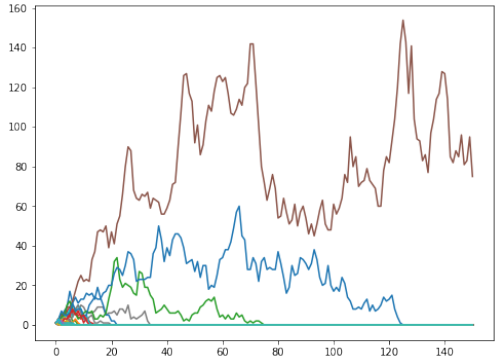

FIGURE 3.9 – Two-type process in fixed environment

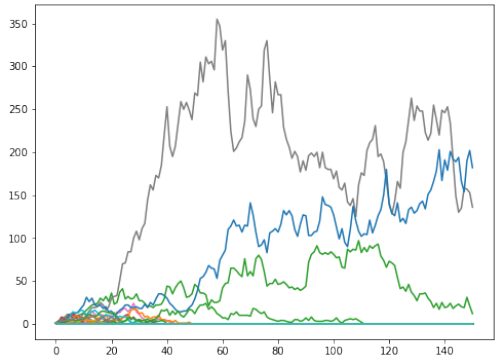

FIGURE 3.10 – Two-type process in random environment

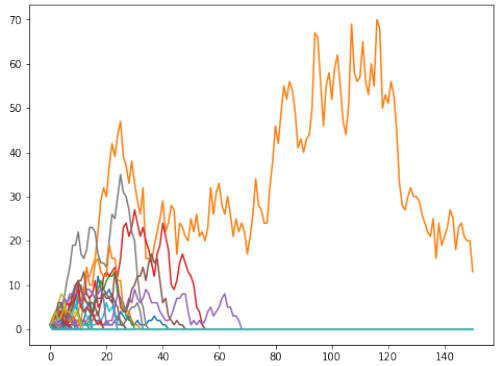

FIGURE 3.11 – Three-type process in fixed environment

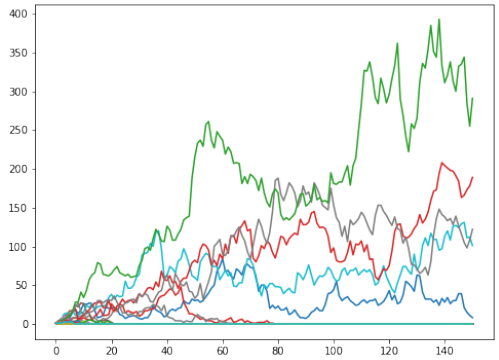

FIGURE 3.12 – Three-type process in random environment

FIGURE 3.13 – 100 possible realizations of populations over 150 generations

n	log(n, 10)	frequency(n)	log(frequency(n), 10)
1e2	2	0,969048	-0,01365471
1e3	3	0,996397167	-0,001567516
1e4	4	0,999693667	-0,000133059
1e5	5	0,999876333	-5,37111e-05
1e6	6	0,999968	-1,38976e-05

TABLE 3.39 – Single-type in fixed environment

n	log(n, 10)	frequency(n)	log(frequency(n), 10)
1e2	2	0,956372833	-0,019372769
1e3	3	0,983792	-0,007096713
1e4	4	0,983901333	-0,007048451
1e5	5	0,983890333	-0,007053306
1e6	6	0,984010333	-0,007000341

TABLE 3.40 – Single-type in random environment

n	log(n, 10)	frequency(n)	log(frequency(n), 10)
1e2	2	0,975969167	-0,010563903
1e3	3	0,997426	-0,001119315
1e4	4	0,999567333	-0,000187945
1e5	5	0,999850667	-6,48595e-05
1e6	6	0,999828333	-7,45603e-05

TABLE 3.41 – Multi-type in fixed environment

n	log(n, 10)	frequency(n)	log(frequency(n), 10)
1e2	2	0,985623	-0,006289171
1e3	3	0,998218167	-0,000774531
1e4	4	0,999616833	-0,000166439
1e5	5	0,999612833	-0,000168177
1e6	6	0,999616833	-0,000166439

TABLE 3.42 – Multi-type in random environment

TABLE 3.43 – Frequency

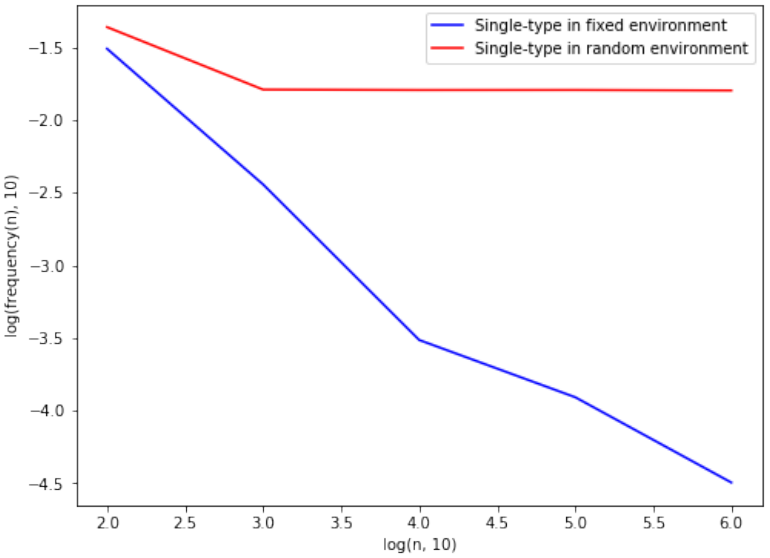

FIGURE 3.14 – Rate of convergence in single-type case

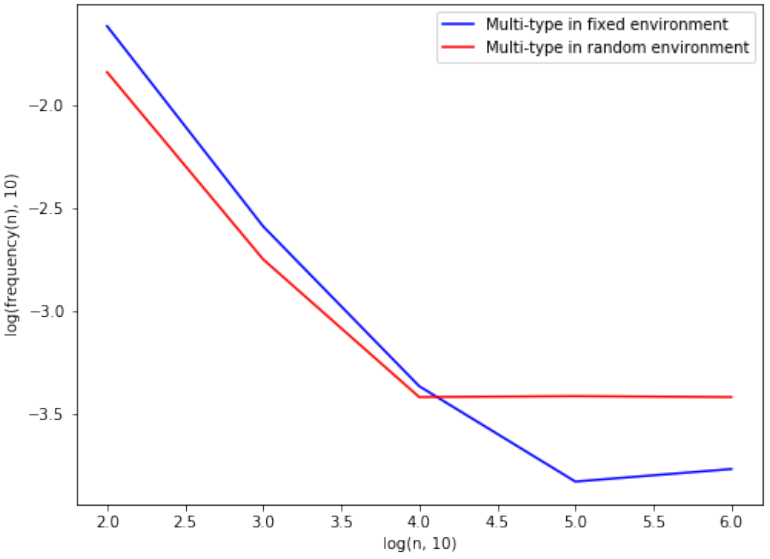

FIGURE 3.15 – Rate of convergence in multi-type case

Conclusion

In this chapter we revise principal results obtained and presented in section 2.6 and 3.4; we also discuss some questions for future research.

Conditioned limit theorems for products of positive random matrices

The fluctuations of a random sequence $S_n = a + \log|g_n \dots g_1 x|$ are studied, where g_j are i.i.d. random $d \times d$ matrices with non-negative entries. The contribution of the paper lies in the description of some aspects of the fluctuations of S_n : the behavior of the probability that the path $S_1, \dots, S_n$ is positive, and the distribution of S_n conditioned on $S_1, \dots, S_n$ to stay positive.

If τ denotes the first time the random walk $(S_n)_n$ becomes non-positive, the asymptotic expansion for $\mathbb{P}_{x,a}(\tau > n)$ is of order $\frac{c}{\sqrt{n}}$; the constant c is explicitly presented, it depends on the $\tilde{P}_+$-harmonic function $V(x, a)$.

Furthermore, it is shown that $S_n/\sqrt{n}$ conditioned on $[\tau > n]$ has asymptotically Rayleigh distribution.

The proofs are based on a martingale approximation, on error estimates in the central limit theorem for dependent random variables, and on a "coupling" of $\frac{S_{[nt]}}{\sigma\sqrt{n}}$ with Brownian motion.

The survival probability of a critical multi-type branching process in i.i.d. random environment

The paper investigates the asymptotic behavior of the quantity $\mathbb{P}(Z_n > 0)$ for the multi-type branching process in i.i.d. random environment, under reasonable assumptions. First, the relation $\mathbb{P}(Z_n > 0) \sim c/\sqrt{n}$ for the special case of random linear-fractional offspring distributions. Second, for a bigger class of random offspring distributions, it follows that $\mathbb{P}(Z_n > 0) \asymp 1/\sqrt{n}$. The proofs use techniques for univariate BPRE and methods from the theory of products of i.i.d. random matrices (result obtained from the article published in ALEA).

Some questions for future research

There is a huge literature on one type or multi-type Galton Watson processes and many questions remain open; in this section, we propose to describe some of them, closely related to the main results of this thesis.

On a local limit theorem for product of random matrices whose norms are conditioned on staying > 1

Once the behavior at infinity of the quantity $\mathbb{P}_{x,a}(\tau > n)$ is known, a natural question is to focus on the one of the probabilities $\mathbb{P}(\tilde{x}R_n \in K \text{ compact}, \tau_a > n)$ and to prove that it behaves as $\frac{c}{n^{3/2}}$. Results in this direction was obtained recently by R. Lauvergnat [Lauvergnat, 2017] who studied in his PhD thesis fluctuations of a class of Markov walks (or random walks with Markovian dependent increments). The question is still open for product of random matrices, works are in progress with E. Le Page. Such a result for product of positive random matrices would be of interest to describe the probability of extinction of sub-critical multi-type Galton-Watson processes in random environment, as in the one type case presented in Chapter 3 of this thesis.

On a central limit theorem for Galton-Watson processes conditioned on non-extinction at time n

It is known that **critical one type Galton-Watson processes** $(Z_n)_{n\geq 0}$ in random environment, conditioned on the event $[Z_n > 0]$, do exhibit "super critical behavior". Let us briefly explain this.

First, supercritical branching processes (whether classical or in random environment) obey the growth law $Z_n/\mu^n \to W$ a.s., where W is some typically non-degenerate random variable (see Kesten-Stigum's statement in Chapter 3 of this thesis). Conditioned on the event $[Z_n > 0]$, this kind of behavior can no longer be formulated as a statement on a.s. convergence, since the conditional probability measures depend on n.

Instead, in random environment, it holds some "central limit theorem" for the size of the population at time n, with a random normalization depending on the environment up to time n. The main ingredient is the fact that the process $(Z_n/f_0'(1)\dots f_{n-1}'(1))_{n\geq 1}$ is a martingale, where $f_0, f_1 \dots$ is the i.i.d. environment of the branching process. More precisely,

$$\mathcal{L}\left(\frac{Z_n}{f_0'(1)\dots f_{n-1}'(1)}/Z_n > 0\right) \longrightarrow \mathcal{L}(W)$$

for some positive and finite random variable W (see [Afanasyev *et al.*, 2005] for the details and interesting comments).

Thus, it is natural to ask whether or not such a result does exist for multi-type Galton-Watson processes $(Z_n)_{n\geq 0}$ in random environment. It is natural to replace the quantity $f_0'(1)\dots f_{n-1}'(1)$ by the product of the random mean matrices of the offspring distributions,

unfortunately the sequence

$$\left(\frac{Z_n^{(i)}\mathbf{1}}{e^{S_n(\tilde{x})}}\right)_{n\geq 1} = \left(\frac{Z_n^{(i)}\mathbf{1}}{\tilde{x}R_n\mathbf{1}}\right)_{n\geq 1}$$

is no more a martingale. Nevertheless, we may expect for a weak convergence of this sequence, which should allow us to conclude as in [Afanasyev *et al.*, 2005]; the delicate argument related to martingale needs to be settled, using ideas from [Cohn, 1989] and [Dolgopyat *et al.*, 2017].

On the affine recursion in $(\mathbb{R}^+)^p, p \geq 1$

We consider the random difference equation $X_{n+1} = A_{n+1}X_n + B_{n+1}$, where (A_n, B_n) are i.i.d. $(S^+(B) \times \mathbb{R}^{d+})$-valued. A direct computation yields for any $n \geq 1$,

$$X_n = A_n \dots A_1 X_0 + \sum_{k=1}^{n} A_n \dots A_{k+1} B_k.$$

A major improvement in critical case for one dimension was done in [Babillot *et al.*, 1997], when $\mathbb{E}(\log A_n) = 0$. Under general hypotheses, they proved the existence and unicity of an infinite Radon measure λ which is invariant for the process $(X_n)_n$ on $\mathbb{R}$. The unicity of λ base on the local contractivity of the semi-group of the random affine maps $x \mapsto A_n \dots A_1 x + \sum_{k=1}^{n} A_n \dots A_{k+1} B_k$. The demonstration used fine results on random walsk on affine groups of the real line and especially the fact that these random walks are transient since the affine group is unimodular.

In higher dimension, several results exist when the products $A_n \dots A_1$ are strongly contractive, namely when the Lyapunov's exponent of the product of matrices A_n is strictly less than 0. Nevertheless, the same question for critical case remains open. Under quite strong hypotheses on the matrices A_n, we prove the existence and unicity of an infinite Radon measure, using some results of products of matrices with non-negative entries presented in [Le Page *et al.*, 2018], [Pham, 2018] and adapting the strategy developped in [Brofferio, 2003].

Bibliographie

[Abraham et Delmas, 2016] Abraham, R. et Delmas, J. P. (2016). An introduction on galton-watson trees and their local limits. *ESAIM : Proceedings and Surveys*, 7. ESAIM : Proceedings and Surveys.

[Afanasyev, 1993] Afanasyev, V. I. (1993). A limit theorem for a critical branching process in a random environment. *Diskret. Mat.*, 5(1):45–58.

[Afanasyev *et al.*, 2005] Afanasyev, V. I., Geiger, J., Kersting, G. et Vatutin, V. A. (2005). Criticality for branching processes in random environment. *Ann. Probab.*, 33(2):645–673.

[Athreya et Karlin, 1971a] Athreya, K. B. et Karlin, S. (1971a). Branching processes with random environments. II. Limit theorems. *Ann. Math. Statist.*, 42:1843–1858.

[Athreya et Karlin, 1971b] Athreya, K. B. et Karlin, S. (1971b). On branching processes with random environments. I. Extinction probabilities. *Ann. Math. Statist.*, 42:1499–1520.

[Athreya et Ney, 1972] Athreya, K. B. et Ney, P. E. (1972). *Branching processes*. Springer-Verlag, New York-Heidelberg. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[Babillot *et al.*, 1997] Babillot, M., Bougerol, P. et Elie, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Ann. Probab.*, 25(1):478–493.

[Bansaye et Berestycki, 2008] Bansaye, V. et Berestycki, J. (2008). Large deviations for branching processes in random environment. *arXiv preprint arXiv :0810.4991*.

[Benoist et Quint, 2016] Benoist, Y. et Quint, J.-F. (2016). Central limit theorem for linear groups. *Ann. Probab.*, 44(2):1308–1340.

[Billingsley, 1968] Billingsley, P. (1968). *Convergence of probability measures*. John Wiley & Sons, Inc., New York-London-Sydney.

[Bolthausen, 1976] Bolthausen, E. (1976). On a functional central limit theorem for random walks conditioned to stay positive. *Ann. Probability*, 4(3):480–485.

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

[Brofferio, 2003] Brofferio, S. (2003). How a centred random walk on the affine group goes to infinity. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(3):371–384.

[Chow et Teicher, 2012] Chow, Y. S. et Teicher, H. (2012). *Probability theory : independence, interchangeability, martingales*. Springer Science & Business Media.

[Church, 1971] Church, J. D. (1971). On infinite composition products of probability generating functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 19:243–256.

[Cohn, 1989] Cohn, H. (1989). On the growth of the multitype supercritical branching process in a random environment. *Ann. Probab.*, 17(3):1118–1123.

[Dyakonova, 1999] Dyakonova, E. E. (1999). The asymptotics of the probability of nonextinction of a multidimensional branching process in a random environment. *Diskret. Mat.*, 11(1):113–128.

[Denisov et Wachtel, 2015] Denisov, D. et Wachtel, V. (2015). Random walks in cones. *Ann. Probab.*, 43(3):992–1044.

[Dolgopyat *et al.*, 2017] Dolgopyat, D., Hebbar, P., Koralov, L. et Perlman, M. (2017). Multi-type branching processes with time-dependent branching rates.

[Dyakonova *et al.*, 2004] Dyakonova, E. E., Geiger, J. et Vatutin, V. A. (2004). On the survival probability and a functional limit theorem for branching processes in random environment. *Markov Process. Related Fields*, 10(2):289–306.

[Feller, 1968] Feller, W. (1968). *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons, Inc., New York-London-Sydney.

[Furstenberg et Kesten, 1960] Furstenberg, H. et Kesten, H. (1960). Products of random matrices. *Ann. Math. Statist.*, 31:457–469.

[Geiger et Kersting, 2000] Geiger, J. et Kersting, G. (2000). The survival probability of a critical branching process in random environment. *Teor. Veroyatnost. i Primenen.*, 45(3):607–615.

[Geiger *et al.*, 2003] Geiger, J., Kersting, G. et Vatutin, V. A. (2003). Limit theorems for subcritical branching processes in random environment. *Ann. Inst. H. Poincaré Probab. Statist.*, 39(4):593–620.

[Grama *et al.*, 2014] Grama, I., Le Page, E. et Peigné, M. (2014). On the rate of convergence in the weak invariance principle for dependent random variables with applications to Markov chains. *Colloq. Math.*, 134(1):1–55.

[Grama *et al.*, 2017] Grama, I., Le Page, E. et Peigné, M. (2017). Conditioned limit theorems for products of random matrices. *Probab. Theory Related Fields*, 168(3-4):601–639.

[Grincevičjus, 1974] Grincevičjus, A. K. (1974). A central limit theorem for the group of linear transformations of the line. *Dokl. Akad. Nauk SSSR*, 219:23–26.

[Guivarc'h et Raugi, 1985] Guivarc'h, Y. et Raugi, A. (1985). Frontière de Furstenberg, propriétés de contraction et théorèmes de convergence. *Z. Wahrsch. Verw. Gebiete*, 69(2):187–242.

[Harris, 1963] Harris, T. E. (1963). *The Theory of branching processes*.

[Hennion, 1984] Hennion, H. (1984). Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes. *Z. Wahrsch. Verw. Gebiete*, 67(3):265–278.

[Hennion, 1997] Hennion, H. (1997). Limit theorems for products of positive random matrices. *Ann. Probab.*, 25(4):1545–1587.

[Hennion et Hervé, 2008] Hennion, H. et Hervé, L. (2008). Stable laws and products of positive random matrices. *J. Theoret. Probab.*, 21(4):966–981.

[Iglehart, 1974a] Iglehart, D. L. (1974a). Functional central limit theorems for random walks conditioned to stay positive. *Ann. Probability*, 2:608–619.

[Iglehart, 1974b] Iglehart, D. L. (1974b). Random walks with negative drift conditioned to stay positive. *J. Appl. Probability*, 11:742–751.

[Iglehart, 1975] Iglehart, D. L. (1975). Conditioned limit theorems for random walks. pages 167–194.

[Jagers, 2005] Jagers, P. (2005). Branching processes. *Encyclopedia of Biostatistics*, 1.

[Kaigh, 1976] Kaigh, W. D. (1976). An invariance principle for random walk conditioned by a late return to zero. *Ann. Probability*, 4(1):115–121.

[Kaplan, 1974] Kaplan, N. (1974). Some results about multidimensional branching processes with random environments. *Ann. Probability*, 2:441–455.

[Kersting, 2017] Kersting, G. (2017). A unifying approach to branching processes in varying environments. *arXiv preprint arXiv :1703.01960*.

[Kesten et Stigum, 1966] Kesten, H. et Stigum, B. P. (1966). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1211–1223.

[Kolmogorov, 1938] Kolmogorov, A. N. (1938). Zur lösung einer biologishen aufgabe. *Iz. NII. Mathem. Mekh. Tomskogo Univ. 2*, pages 1–6.

[Kozlov, 1977] Kozlov, M. V. (1977). On the asymptotic behavior of the probability of non-extinction for critical branching processes in a random environment. *Theory of Probability & Its Applications*, 21(4):791–804.

[Lauvergnat, 2017] Lauvergnat, R. (2017). *Théorèmes limites pour des marches aléatoires markoviennes conditionnées à rester positives*. Thèse de doctorat, Lorient.

[Le Page, 1982] Le Page, E. (1982). Théorèmes limites pour les produits de matrices aléatoires. *In Probability measures on groups (Oberwolfach, 1981)*, volume 928 de *Lecture Notes in Math.*, pages 258–303. Springer, Berlin-New York.

[Le Page et Peigné, 1997] Le Page, E. et Peigné, M. (1997). A local limit theorem on the semi-direct product of $\mathbf{R}^{*+}$ and $\mathbf{R}^d$. *Ann. Inst. H. Poincaré Probab. Statist.*, 33(2):223–252.

[Le Page *et al.*, 2018] Le Page, E., Peigné, M. et Pham, C. (2018). The survival probability of a critical multi-type branching process in i.i.d. random environment. *Ann. Probab.*, 46(5):2946–2972.

[Lindvall, 1992] Lindvall, T. (1992). *Lectures on the coupling method*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York. A Wiley-Interscience Publication.

[Mode, 1971] Mode, C. J. (1971). *Multitype branching processes. Theory and applications*. Modern Analytic and Computational Methods in Science and Mathematics, No. 34. American Elsevier Publishing Co., Inc., New York.

[Paley et Wiener, 1987] Paley, R. E. A. C. et Wiener, N. (1987). *Fourier transforms in the complex domain*, volume 19 de *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI. Reprint of the 1934 original.

[Pham, 2018] Pham, T. D. C. (2018). Conditioned limit theorems for products of positive random matrices. *ALEA Lat. Am. J. Probab. Math. Stat.*, 15(1):67–100.

[Smith et Wilkinson, 1969] Smith, W. L. et Wilkinson, W. E. (1969). On branching processes in random environments. *Ann. Math. Statist.*, 40:814–827.

[Spitzer, 1976] Spitzer, F. (1976). *Principles of random walk*. Springer-Verlag, New York-Heidelberg, second édition. Graduate Texts in Mathematics, Vol. 34.

[Vatutin et Dyakonova, 2017] Vatutin, V. A. et Dyakonova, E. E. (2017). Multitype branching processes in a random environment : nonextinction probability in the critical case. *Teor. Veroyatn. Primen.*, 62(4):634–653.

Résumé :

La théorie des processus de branchement multi-type en environnement i.i.d. est considérablement moins développée que dans le cas univarié, et les questions fondamentales ne sont pas résolues en totalité à ce jour. Les réponses exigent une compréhension profonde du comportement des produits des matrices i.i.d. à coefficients positifs.

Sous des hypothèses assez générales et lorsque les fonctions génératrices de probabilité des lois de reproduction sont "linéaire fractionnaires", nous montrons que la probabilité de survie à l'instant n du processus de branchement multi-type en environnement aléatoire est proportionnelle à $\frac{1}{\sqrt{n}}$ lorsque $n \to \infty$.

La démonstration de ce résultat suit l'approche développée pour étudier les processus de branchement uni-variés en environnement aléatoire i. i. d. Il utilise de façon cruciale des résultats récents portant sur les fluctuations des normes de produits de matrices aléatoires i.i.d.

Mots clés :

Multi-type branching process, Survival probability, Random environment, Critical case, Exit time, Markov chains, Product of random matrices.

Abstract :

The theory of multi-type branching process in i.i.d. environment is considerably less developed than for the univariate case, and fundamental questions are up to date unsolved. Answers demand a solid understanding of the behavior of products of i.i.d. matrices with non-negative entries.

Under mild assumptions, when the probability generating functions of the reproduction laws are fractional-linear, the survival probability of the multi-type branching process in random environment up to moment n is proportional to $\frac{1}{\sqrt{n}}$ as $n \to \infty$.

Techniques for univariate branching process in random environment and methods from the theory of products of i.i.d. random matrices are required.

Keywords :

Multi-type branching process, Survival probability, Random environment, Critical case, Exit time, Markov chains, Product of random matrices.