Compound Poisson distributions for random dynamical systems

Lucas Amorim Vilas Boas
Joint PhD Program in Mathematics
Department of Mathematics, Faculty of Sciences of the University of Porto
2023

Supervisor
Jorge Miguel Milhazes Freitas, Associate Professor,
Department of Mathematics, Faculty of Sciences of the University of Porto

Co-supervisor
Sandro Vaienti, Full Professor,
Department of Mathematics, University of Toulon

Co-supervisor
Ana Cristina Moreira Freitas, Associate Professor,
Faculty of Economics of the University of Porto

1 2 9 0 UNIVERSIDADE D COIMBRA

UNIVERSITÉ DE TOULON

U.PORTO

FC FACULDADE DE CIÊNCIAS UNIVERSIDADE DO PORTO

PhD

UNIVERSITÉ DE TOULON

U.PORTO
FC FACULDADE DE CIÊNCIAS UNIVERSIDADE DO PORTO

COTUTELLE

ÉCOLE DOCTORALE 548 : Mer et Sciences

Laboratorie CPT – Centre de Physique Théorique

THÈSE présentée par :

Lucas AMORIM VILAS BOAS

soutenue le : **15 décembre 2023**

pour obtenir le grade de Docteur en Mathématique

Spécialité : Mathématique

TITRE de la thèse

Compound Poisson distributions for random dynamical systems

THÈSE dirigée par :

VAIENTI Sandro	Professeur, Université de Toulon
FREITAS Jorge Milhazes	Professeur, Universidade do Porto
FREITAS Ana Moreira	Professeur, Universidade do Porto

JURY :

ALVES José	Professeur, Universidade do Porto
SAUSSOL Benoît	Professeur, Université Aix-Marseille
SOUSA Ercília	Professeur, University of Coimbra
TODD Mike	Professeur, University of St. Andrews
VAIENTI Sandro	Professeur, Université de Toulon
VARANDAS Paulo	Professeur, Universidade do Porto
AIMINO Romain	Post-Doc, Universidade do Porto

Compound Poisson distributions for random dynamical systems

Lucas Amorim Vilas Boas

1 2 9 0 UNIVERSIDADE D COIMBRA

U. PORTO

UC|UP Joint PhD Program in Mathematics

Programa Interuniversitário de Doutoramento em Matemática

PhD Thesis | Tese de Doutoramento

January 2024

Acknowledgments

I would like to thank my supervisors Sandro Vaienti, Jorge Freitas and Ana Cristina Freitas for the adventurous research environment they've fostered while being always open to mathematical debates and providing constant support to me.

I also want to thank my collaborator Nicolai Haydn for introducing and welcoming me to the research agenda he shares with Sandro and for his huge availability for discussion.

I would like to express special gratitude to Paulo Varandas for his long-standing (non-) mathematical advisory and generous attitude, always open to guiding me through new ideas.

I want to thank Manuel Stadlbauer, Yushi Nakano, Cecilia González-Tokman, José Alves, Romain Aimino, and Victor Vargas for several useful discussions, as well as my colleagues Raquel Couto and Gustavo Pessil.

I also want to acknowledge the quality of the PhD joint program held between the University of Porto and the University of Coimbra in cotutelle with the University of Toulon, the scholarship provided by FCT (Fundação para a Ciência e a Tecnologia), reference number PD/BD/150458/2019, and the financial support provided by CMUP (Centro de Matemática da Universidade do Porto) and CNRS (Centre National de la Recherche Scientifique) to attend conferences. I want also to acknowledge the hospitality of the staff and the facilities provided in the Mathematics Department of the Faculty of Sciences of the University of Porto and the Centre de Physique Théorique in Marseille.

Finally, on the personal side, would like to thank my parents for supporting my decision to pursue a PhD in mathematics and especially my future wife, Brenda Poubel, for joining me in the adventure of moving to a new country and for the endless love I was given in this meantime.

Abstract

We obtain quenched limiting hitting distributions in the compound Poisson class for a certain family of random dynamical systems using a probabilistic block-approximation for the quenched hit-counting function up to annealed-Kac-normalized time. We consider general random targets with well-defined return statistics and systems with both quenched and annealed polynomial decay of correlations. The theory is made operational due to a result that allows certain hitting statistics to be recovered from the said return statistics, which are computable. Our examples include a class of random piecewise expanding one-dimensional systems, casting new light on the well-known deterministic dichotomy between periodic and aperiodic points, their usual extremal index formula $\mathrm{EI} = 1 - 1/JT^p(\zeta)$, and recovering the geometric case for general Bernoulli-driven systems, but distinct behavior otherwise. Future and on-going investigations aim to produce and accommodate examples of bonafide non-uniformly expanding random systems and targets approaching their neutral points.

Keywords: hitting statistics, compound Poisson distributions, random dynamical systems

Resumo

Obtemos distribuições limite do tempo de entrada *quenched* como sendo Poisson compostas para uma certa família de sistemas dinâmicos aleatórios usando uma aproximação probabilística em blocos para a função contagem de entradas *quenched* até o tempo normalizado segundo Kac-*annealed*. Consideramos alvos aleatórios gerais com estatísticas de retorno bem definidas e sistemas com decaimento polomial de correlações, tanto *quenched* quanto *annealed*. A teoria faz-se aplicável devido a um resultado que permite certas estatísticas de entrada serem recuperadas a partir das referidas estatísticas de retorno, que podem ser calculadas. Nossos exemplos incluem uma classe de sistemas unidimensionais expansores por partes, lançando nova luz na conhecida dicotomia determinística entre pontos periódicos e aperiódicos, suas fórmulas de indice extremal $\mathrm{EI} = 1 - 1/JT^p(\zeta)$ e recuperando o caso geométrico para sistemas dirigidos-por- Bernoulli gerais, bem como revelando comportamento distinto caso contrário. Investigações futuras e em andamento visam produzir e acomodar exemplos de sistemas aleatórios não uniformemente expansores com alvos se aproximando dos seus pontos neutros.

Palavras-chave: estatísticas de entrada, distribuições de Poisson compostas, sistemas dinâmicos aleatórios

Résumé

Nous obtenons des distributions d'entrée limites *quenched* dans la classe composée de Poisson pour une certaine famille de systèmes dynamiques aléatoires en utilisant une approximation probabiliste par bloc pour la fonction de comptage d'entrée quenched jusqu'au temps normalisé *annealed*-Kac. Nous considérons des cibles aléatoires générales avec des statistiques de retour bien définies et des systèmes avec une d´ecroissance polynomiale des corrélations à la fois quenched et *annealed*. La théorie est rendue opérationnelle grâce à un résultat qui permet de récupérer certaines statistiques d'entrée à partir desdites statistiques de retour, qui sont calculables. Nos exemples incluent une classe de systèmes unidimensionnels à expansion aléatoire par morceaux, jetant un nouvel éclairage sur la dichotomie déterministe bien connue entre les points périodiques et apériodiques, leur formule d'indice extrême habituelle $\mathrm{EI} = 1 - 1/JT^p(\zeta)$, et récupérer le cas géométrique pour les systèmes généraux pilotés par Bernoulli, mais comportement distinct dans le cas contraire. Les enquêtes futures et en cours visent à produire et à prendre en compte des exemples de véritables systèmes aléatoires à expansion non uniforme et de cibles s'approchant de leurs points neutres.

Mots clés: statistiques d'entrée, statistiques de Poisson composée, perturbations aléatoires de systèmes dynamiques

Contents

1 Introduction **1**

2 Prerequisites **5**

2.1 Poisson-type processes 5

2.1.1 Poisson distribution and process 5

2.1.2 Compound Poisson distribution and process 10

2.1.3 Poisson point process 11

2.1.4 Compound Poisson point process 14

2.1.5 Lifted compound Poisson point process 16

2.2 Extreme value theory and hitting time statistics (iid case) 17

2.2.1 Extreme value theory 17

2.2.2 Hitting statistics 24

3 State of the art: Compound Poisson distributions for dynamical systems **31**

3.1 Deterministic systems 31

3.1.1 Point processes 32

3.1.2 Spectral methods 40

3.1.3 Probabilistic approximation with Chen-Stein method 46

3.2 Random systems 48

3.2.1 Point processes 51

3.2.2 Spectral methods 53

3.2.3 Probabilistic approximation 56

4 Compound Poisson distributions for random dynamical systems: a probabilistic block-approximation approach **57**

4.1 Assumptions and main results 57

4.1.1 General setup 57

4.1.2 Working setup 59

4.1.3 Main results 65

4.2 Proof of theorem 4.1.3 66

4.3 An abstract approximation theorem 77

4.4 Borel-Cantelli type lemmata 81

4.5 Proof of theorem 4.1.4 92

4.5.1 Applying the abstract approximation theorem 92
4.5.2 Estimating the error $\mathcal{R}^1$ 93
4.5.3 Estimating the error $\tilde{\mathcal{R}}^1$ 94
4.5.4 Estimating the error $\mathcal{R}^2$ 96
4.5.5 Estimating the error $\mathcal{R}^3$ 99
4.5.6 Controlling the total error 99
4.5.7 Convergence of the leading term to the compound Poisson distribution 100

5 Application: random piecewise expanding one-dimensional systems 103

6 Future research 119

Bibliography 121

Notation

$\mathbb{N}_{\geqslant 0} = \{0, 1, 2, \dots\}$
$\mathbb{N}_{\geqslant 1} = \{1, 2, 3, \dots\}$
$\mathscr{B}_X$ is the Borel σ-algebra associated to a topological space X
$\mathcal{P}(X)$ is the set of Borelian probability measures on a topological space X
$\mathcal{P}_T(X)$ is the set of T-invariant of Borelian probability measures on a topological space X
d_M is the native metric of a metric space M
$B_r(x,r) \subset M$, $x \in M$, $r > 0$, M metric space, is the open ball around x with radius r w.r.t to d_M

Chapter 1

Introduction

The field of hitting (or entry) time statistics has a long history in the area of dynamical systems. Given a measurable system $T : M \to M$ preserving a probability μ and given a family of sets U_n shrinking to the so-called target set, one studies the statistics of how long it takes for points in M to hit U_n as compared to the normalized time $t/\mu(U_n)$. Then one goes after the asymptotics of these statistics as the sets shrink. Mathematically, one considers

$$\lim_{n\to\infty} \mu\left(r_{U_n} \geqslant \frac{t}{\mu(U_n)}\right), \text{ with}$$
$$r_U(x) = \min\{n \geqslant 1 : T^n x \in U\} \ (x \in M).$$

A related problem is that of return time statistics, which pursues the same type of statistical characterization of the time needed for future visits to occur, but conditioning on starting on U_n:

$$\lim_{n\to\infty} \mu\left(r_{U_n} \geqslant \frac{t}{\mu(U_n)} \,\middle|\, U_n\right).$$

Limiting hitting and entry times statistics are intimately connected. They harmonize in a specific way that indicates their coincidence occurs precisely when both are exponentially distributed (see, e.g., [51]).

Another related problem is that of recurrence times, which considers the statistics of the time needed for points in the state space to return to shrinking neighborhoods of themselves.

Hitting, return and recurrence times can be considered quantitative refinements to classical theorems such as Poincaré recurrence theorem and Kac's theorem.

The first contributions in this direction were by [31], [75] and [54], considering (higher order) return times of, respectively, the Gauss map, Markov chains and Axiom A systems. They find that limiting first return times are exponentially distributed when the target consists of a generic point. The number of returns to these shrinking sets then has Poisson distribution. In a nutshell, these are consequences of decorrelation and the fact that generic points are aperiodic.

Not much later, [27] addressed uniformly expanding maps of the interval and [49] paved the way for the study of hitting time statistics for general measurable systems (symbolic ones included) equipped with a generating partition an a probability with certain mixing features. Again these results are in the exponential distribution domain.

The literature in this field then grew rapidly, with generalizations in many directions: non-generic (periodic) singleton targets, general targets, random systems, non-uniformly hyperbolic systems, partially hyperbolic systems, zero-entropy systems and more.

Hitting and return statistics for general measurable systems with a partition and a sufficiently mixing equilibrium state are well reasonably well understood, see, e.g. [4] and [48] for general targets. These, however, use U_n as cylinder sets, which carry more structure than usual metric balls. When in metric spaces and working with certain systems (e.g., with non-uniformly hyperbolic), the problem can be more subtle when U_n's (or U_ρ's with $\rho \to 0$) are taken to be balls of shrinking radii (around some set).

A major break-through in the study of return time statistics for non-uniformly hyperbolic maps was [55], where a set of conditions implying exponential statistics for return times to cylinders or balls around generic points was given. Applications to interval maps with intermittent points were provided. Also, a useful result to deal with non-uniformly hyperbolic systems was [21], which showed that the return time statistics for a map coincide with those of the associated first return map. Chapter 3.1 covers some of these topics, in the perspective we discuss next.

When one considers targets made of a periodic (non generic) point, the picture changes, due to the local recurrence the dynamics produces around the target. This is going to produce clusters of visits and return times will not be exponentially distributed anymore: the new effect will create return time statistics with a point mass on the origin. The adequate approach is to consider hitting statistics up to normalized time. Instead of a Poisson process, a compound Poisson process then appears, which usually is of Polya-Aeppli type. This is sometimes referred to as 'dichotomy' in the literature. General targets can produce more intricate behavior but still in the compound Poisson class. Chapter 3.1 covers these topics.

The techniques applied to differentiable systems in order to address hitting statistics to balls around general targets are of varied nature. One relies on the relationship between hitting time statistics and extreme values (see e.g. [39]), leveraging on the previously available theory of stationary stochastic processes and the machinery of point processes (see e.g. [34]). Another one adapts the spectral method introduced by [60] (see e.g. [8] and [11]). Finally, techniques based on probabilistic approximation methods were also used (see e.g. [53]). Chapter 3.1 covers these techniques.

The theory was also brought to the realm of random dynamical systems. The picture is still pretty much incomplete, however. Let us now focus on the quenched situation, meaning that a statistical result is stated for almost all realizations of the noise. In the context, random subshifts were treated in [52] and [30], while random piecewise expanding maps of the interval were treated using point processes in [36] and using the spectral method in [8].

In the quenched approach to hitting statistics, time normalization can either occur in quenched manner (with division by a noise-dependent probability of the random shrinking set) or in an annealed manner (with division by an averaged probability of the random shrinking set). The second approach ([52], [36]) is appealing from an applied perspective, representing the fact that an experimentalist is not informed about the noise path (not even in its entire past) and deterministically decides how long to watch his experiment. But the first approach ([30] and [8]) is also theoretically appealing and works with milder hypotheses on the system producing the noise.

This thesis presents contributions to the theory of quenched hitting statistics of random dynamical systems based on a probabilistic block-approximation approach, generalizing the deterministic theory developed in [53] after the approach introduced in [25]. The main result, theorem 4.1.4, concludes that the limiting behavior is described by a compound Poisson distribution. Results are in collaboration with Nicolai Haydn and Sandro Vaienti.

The strategy goes like this. We use a probabilistic approximation (theorem 4.3.1) where the quenched hit-counting function up to annealed-Kac-normalized time is split into equally time-sized blocks which are mimicked by an independency of random variables distributed just like each of them. The said approximation goes for any given noise realization ω and ω-dependent leading terms and errors appear. Both of them are tamed by a lemmata which ultimately resorts to a Borel-Cantelli argument in order to show that they converge to the desired quantities almost surely (see section 4.4), thus establishing the quenched result we are after.

The aforementioned Borel-Cantelli argument produces, after convergence, some statistical quantities that reveal the hitting behavior of the dynamics to the target, the so-called λ_ℓ's (see equation (4.2)). It turns out that these essential quantities are not easily calculated or even shown to exist. To make our theory operational, we show that λ_ℓ's can be explicitly written in terms of the so-called α_ℓ's, another set of statistical quantities that reveal the return behavior of the dynamics to the target set (see equation (4.7)). Within the theory, the latter quantities are assumed to exist, while in the examples one tries to calculate them.

In view of the current state of the comparable literature for random dynamical systems, reviewed in chapter 3.2, important advantages can be identified in our approach.

First, it handles general random target sets, the only requirement being that they present well-defined return statistics α_ℓ's, which allow us to represent the intensity and multiplicity distribution of the limiting compound Poisson distributions very explicitly, given in terms of the dynamics and the target.

Second, our approach relies on polynomial decay of correlations, indicating its potential to cover non-uniformly expanding maps with parabolic fixed points and targets moving randomly so as to possibly approach the parabolic locus (provided that they do so slowly/unlikely enough). That is currently a work in progress. Here, the examples we will provide are certain one-dimensional piecewise expanding systems. Alternative techniques based on spectral theory and Lasota-Yorke inequalities by design will not cover polynomial decay of correlations (not directly, at least).

Third, our assumptions on the quasi-invariant family of measures do not consider their absolute continuity with respect to the Lebesgue measure. Regularity assumed is in a dimensional sense.

A drawback of our approach is that results are just along sufficiently fast shrinking neighborhoods of the target set. In other words, the results are about subsequential convergence (though many subsequences qualify) instead of plain convergence.

Let us note that some of the above topics are intimately connected: annealed decay of correlations, annealed normalization of time, the Borel-Cantelli argument and the need for (deterministic) subsequences.

The thesis is organized as follows. In chapter 2, prerequisites in probability theory are collected for the convenience of the reader. This chapter in mainly oriented to present the parallelism between hitting time statistics and extreme value theory (in its presentation with point processes) in the iid case. The reader can skip this chapter and consult it only upon need.

Chapter 3 reviews the literature addressing compound Poissonian extreme value and hitting statistics, both in deterministic and random cases, using techniques different from the aforementioned probabilistic block-approximation, namely: point processes, spectral methods and the Chen-Stein method. Section 3.1.1 relies on the notation introduced in section 2.2.2. We do not draw direct point-by-point comparisons with the approach we will develop, but we present enough of a review to enable the reader to undertake an informed appraisal of what we will present in chapter 4. The review of the deterministic theory in section 3.1 is not only to complete an overview of the literature, but also to prepare the ground for the random theory in section 3.2.

The specialist reader might skip directly to chapter 4 and 5, where the aforementioned contributions of this thesis are presented. We start the chapter redefining the basic objects of random dynamical systems previously introduced in section 3.2, so that chapter 4 can stand independently. To avoid a sizeable redundancy, however, we go directly to the random theory, without reviewing the deterministic counterpart preceding it, found in [53].

Chapter 2

Prerequisites

This chapter introduces the probabilistic foundations required by the following two chapters. It initially covers Poisson-type processes, aiming towards a comprehensive understanding of compound Poisson point processes. Finally, it covers extreme value theory and hitting time statistics in the iid case.

2.1 Poisson-type processes

In this section, we introduce many definitions related to Poisson-type processes.

We start with the less general version to convey intuition.

Section 2.1.1 considers Poisson processes on the positive real line. Definition 2.1.8 introduces Poisson processes with real intensities and its properties are discussed; whereas definition 2.1.10 generalizes the latter into Poisson processes with measure intensities having no atoms.

Section 2.1.2 considers compound Poisson processes, again on the positive real line. Definition 2.1.12 introduces compound Poisson processes with real intensities (and independent multiplicity kernels); whereas definition 2.1.13 generalizes the former into compound Poisson processes with measure intensities having no atoms (and independent multiplicity kernels).

Then we generalize considerably.

In section 2.1.3, definition 2.1.17 introduces Poisson point processes with general measure intensities on general spaces. This concept encompasses all of those in section 2.1.1 and it is claimed to be non-void.

In section 2.1.4, definition 2.1.21 introduces compound Poisson point processes with general measure intensities and general multiplicity kernels on general spaces.

2.1.1 Poisson distribution and process

Here we follow [47]. The basic pieces to be considered are the following distributions.

Definition 2.1.1. The **Poisson distribution with intensity** $\gamma \in \mathbb{R}_{>0}$ is the probability measure on $\mathbb{N}_{\geqslant 0}$, denoted Poi_γ, given by the following probability mass function

$$Poi_\gamma(\{n\}) = \frac{\gamma^n e^{-\gamma}}{n!} \; (n \geqslant 0).$$

Mean and variance of Poi_γ are both γ.

Definition 2.1.2. The **exponential distribution with intensity** $\gamma \in \mathbb{R}_{>0}$ is the probability measure on $\mathbb{R}_{\geqslant 0}$, denoted Exp_γ, given by the following probability density function (with respect to Lebesgue measure)

$$PDF_{Exp_\gamma}(x) = \begin{cases} \gamma e^{-\gamma x}, \text{ for } x \geqslant 0 \\ 0, \text{ otherwise} \end{cases}.$$

Mean and variance of Exp_γ are, respectively, $1/\gamma$ and $1/\gamma^2$.

Definition 2.1.3. The **binomial distribution with** $n \in \mathbb{N}_{\geqslant 0}$ **trials and success probability** $p \in (0,1)$ is the probability measure on $\mathbb{N}_{\geqslant 0}$, denoted $Bin_{n,p}$, given by the following probability mass function

$$Bin_{n,p}(\{k\}) = \binom{n}{k} p^k (1-p)^{n-k} \ (k \in \mathbb{N}_{\geqslant 0}).$$

This distribution gives the probability that, in an independent coin-tossing experiment with success probability p, among the first n trials, k successes occur. Its mean and variance are, respectively, np and $np(1-p)$.

Definition 2.1.4. The **geometric distribution with success probability** $p \in (0,1)$ is the probability measure on $\mathbb{N}_{\geqslant 1}$, denoted Geo_p, given by the following probability mass function

$$Geo_p(\{n\}) = (1-p)^{n-1} p \ (n \in \mathbb{N}_{\geqslant 1})$$

This distribution gives the probability that, in an independent coin-tossing experiment with success probability p, the first occurrence of success requires n trials. Its mean and variance are, respectively, $1/p$ and $(1-p)/p^2$.

Theorem 2.1.5 (Poisson's theorem for distributions)**.** *If $p_n = \gamma/n$ (so that the sequence of probabilities Bin_{n,p_n} all have mean γ), then*

$$Bin_{n,p_n}(\{k\}) \xrightarrow{n\to\infty} Poi_\gamma(\{k\}) \ (k \in \mathbb{N}_{\geqslant 0}).$$

The motivation for the previous distributions and their relationship will be given next.

Definition 2.1.6. A triplet of stochastic processes $\big((X_i)_{i\in\mathbb{N}_{\geqslant 1}}, (S_n)_{n\in\mathbb{N}_{\geqslant 1}}, (N_t)_{t\in\mathbb{R}_{\geqslant 0}}\big)$ on probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is said to comprise an **inter arrival times - arrival epochs - counting (IAC) process**, if separately they verify, respectively,

1.i. X_i is $\mathbb{R}_{>0}$-valued $(i \in \mathbb{N}_{\geqslant 1})$

2.i. S_n is $\mathbb{R}_{>0}$-valued $(n \in \mathbb{N}_{\geqslant 1})$

ii. $\exists \Omega_* \in \mathscr{F}, \mathbb{P}(\Omega_*) = 1, \forall \omega \in \Omega_* : n < n' \Rightarrow S_n(\omega) < S_{n'}(\omega)$

3.i. N_t is $\mathbb{N}_{\geqslant 0}$-valued $(t \in \mathbb{R}_{>0})$

ii. $\exists \Omega_* \in \mathscr{F}, \mathbb{P}(\Omega_*) = 1, \forall \omega \in \Omega_* :$

a. $N_0(\omega) = 0$,

b. $t \in \mathbb{R}_{\geqslant 0} \mapsto N_t(\omega) \in \mathbb{N}_{\geqslant 0}$ is left-continuous with finitely many discontinuities on bounded subsets of $\mathbb{R}_{\geqslant 0}$,

c. $t \leqslant t' \Rightarrow N_t(\omega) \leqslant N_{t'}(\omega)$,

d. On the discontinuities referred to in (b), the right-limits minus respective left-limits always equal 1,

and together they verify

a. $S_n(\omega) = \sum_{i=1}^{n} X_i(\omega)$ ($n \in \mathbb{N}_{\geqslant 1}, \omega \in \Omega$), i.e.,
$X_i(\omega) = S_i(\omega) - S_{i-1}(\omega)$ ($i \in \mathbb{N}_{\geqslant 1}$, with $S_0 :\equiv 0$)

b. $S_n(\omega) = \inf\{t \in \mathbb{R}_{>0} : N_t(\omega) = n\}$ ($n \geqslant \mathbb{N}_{\geqslant 1}, \omega \in \Omega$), i.e.,
$N_t(\omega) = \sup\{n \in \mathbb{N}_{\geqslant 1} : S_n(\omega) \leqslant t\}$ ($t \in \mathbb{R}_{>0}, \omega \in \Omega$).
In particular, $\{S_n \leqslant t\} = \{N_t \geqslant n\}$.

The last two conditions say that there are consistency relationships between the entries of the triplet in such a way that knowing any of them is enough to recover the remaining two. In particular, once the (joint) distribution of one of them is identified (or assigned), those of the remaining ones follow. If one wants to assign (joint) distributions to the entries of an IAC process, doing it directly to $(X_i)_{i \in \mathbb{N}_{\geqslant 1}}$ is clearly the easier path (e.g., this might be an iid sequence, while the other entries of the triplet can not).

To aid one's intuition, notice that IAC processes can model any phenomena of (non-overlapping) unitary-arrivals in continuous-time (or even discrete-time if X_i's are supported on $\mathbb{N}_{\geqslant 1}$), such as buses arriving at a stop, lightning bolts hitting planet Earth and calls reaching a call center (or even world cup trophies won by a country). So X_i measures how long one should wait for the i-th arrival after the $(i-1)$-th one has occurred, i.e., the i-th inter-arrival time; S_n measures when the n-th arrival occurs; and N_t counts how many arrivals have occurred in the interval $[0,t]$.

Of course, these different phenomena will be modeled by different IAC processes. A special one is the following.

Definition 2.1.7. A **Poisson IAC process with intensity** $\gamma \in \mathbb{R}_{>0}$ is an IAC process verifying one of the following equivalent condition:

1. $(X_i)_{i \in \mathbb{N}_{\geqslant 1}}$ is iid with $X_{i*}\mathbb{P} = Exp_\gamma$ ($i \in \mathbb{N}_{\geqslant 1}$).
2. $(S_n)_{n \in \mathbb{N}_{n \geqslant 1}}$ satisfies

$$PDF_{(S_1,\ldots,S_n)_*\mathbb{P}}(s_1,\ldots,s_n) = \begin{cases} \gamma^n e^{-\gamma s_n}, \text{ if } 0 \leqslant s_1 \leqslant \ldots \leqslant s_n \\ 0, \text{ otherwise} \end{cases} \quad (n \in \mathbb{N}_{n \geqslant 1}).$$

3. $(N_t)_{t \in \mathbb{R}_{t \geqslant 0}}$ has independent and stationary increments, and $N_{t*}\mathbb{P} = Poi_{\gamma t}$ ($t \in \mathbb{R}_{\geqslant 0}$).

Definition 2.1.8. A **Poisson process with intensity** $\gamma \in \mathbb{R}_{>0}$ is a stochastic process $(N_t)_{t \in \mathbb{R}_{>0}}$ verifying condition (3) of definition 2.1.6 and condition (3) of definition 2.1.7, i.e., it is the counting process of a Poisson IAC process with intensity $\gamma \in \mathbb{R}_{>0}$.

One of the main reasons why Poisson (IAC) processes are special is simply the memoryless property of exponentially distributed random variables, i.e., of their inter-arrival times.

A $\mathbb{R}_{>0}(\mathbb{N}_{\geqslant 1})$-valued random variable X on $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be memoryless when

$$\mathbb{P}(X > t + t' \mid X > t) = \mathbb{P}(X > t'), \ \forall t, t' \in \mathbb{R}_{\geqslant 0} \ (\forall t, t' \in \mathbb{N}_{\geqslant 0}),$$

or, equivalently,

$$\mathbb{P}(X > t + t') = \mathbb{P}(X > t')\mathbb{P}(X > t),\ \forall t, t' \in \mathbb{R}_{\geqslant 0}\ (\forall t, t' \in \mathbb{N}_{\geqslant 0}),$$

which is the case if, and only if,

$$\exists a \in \mathbb{R}_{>0} : \mathbb{P}(X > t) = a^t,\ \forall t \in \mathbb{R}_{\geqslant 0}\ (\forall t \in \mathbb{N}_{\geqslant 0}),$$

with, actually, $a \in (0,1)$, since when $t \to \infty$ the LHS evaluates to 0. In this case $a = e^{-\gamma}$ for some $\gamma \in \mathbb{R}_{>0}$ and the equivalence becomes

$$\exists \gamma \in \mathbb{R}_{>0} : \mathbb{P}(X > t) = e^{-\gamma t},\ \forall t \in \mathbb{R}_{\geqslant 0}\ (\forall t \in \mathbb{N}_{\geqslant 0}).$$

In the continuous-time case, differentiating on both sides leads to the following equivalent equation

$$PDF_{X_*\mathbb{P}}(t) = -\gamma e^{-\gamma t},\ \forall t \in \mathbb{R}_{\geqslant 0}.$$

It means that, in the continuous case, the memoryless random variables are exactly the exponentially distributed ones.

In the discrete-time case, the aforementioned condition, after manipulating some sums, is equivalent to

$$PMF_{X_*\mathbb{P}}(t) = (e^{-\gamma})^{t-1}(1 - e^{-\gamma}),\ \forall t \in \mathbb{N}_{\geqslant 1}$$

It means that, in the discrete case, the memoryless random variables are exactly the geometrically distributed ones.

This relationship is associated to the connection between Poisson processes and Bernoulli processes — as if the latter is the discrete version of the former. This will be discussed later in this section.

For now, we review other memoryless-related properties of Poisson processes. Notice that these properties should actually be used to prove the equivalence in definition 2.1.7.

The interpretation of the memoryless property should be clear: if a call center worker waited time t and no call arrived, then the probability that he will have to wait for an additional t' (at least) before a call arrives coincides with the probability that, from the very beginning, he would have waited for t' (at least) for a call a to arrive. In particular, the time t the worker spent waiting buys him no knowledge. He knows how much more he will have to wait just as much as he knew at the beginning of the experiment, or just as much as a newcomer coming to his table would know.

It is opportune to consolidate the interpretation of a Poisson distributed random variable N and a Poisson process $(N_t)_{t \in \mathbb{R}_{\geqslant 0}}$. A random variable N having Poisson distribution with intensity γ models the number of arrivals occurring in a unitary time interval in which the expected number of arrivals is γ and inter-arrival times are exponential (i.e. memoryless). A Poisson process $(N_t)_{t \in \mathbb{R}_{\geqslant 0}}$ with intensity γ models how the previous situation evolves in time, namely, the number of arrivals occurring in the time interval $[0,t]$ ($t \in \mathbb{R}_{\geqslant 0}$) when the expected number of arrivals in a unitary time interval is γ and inter-arrival times are exponential.

A consequence of the memoryless property is that if $(N_t)_{t\in\mathbb{R}_{\geqslant 0}}$ is a Poisson process with intensity γ, then, for any $t>0$, defining the random variable $Z = S_{\inf\{n\in\mathbb{N}_{\geqslant 1}:S_n>t\}} - t$, representative of the length of the interval from t until the first arrival after t, it holds that $PDF_{Z_*\mathbb{P}}(z) = 1-e^{-\gamma z}$ $(z\in\mathbb{R}_{>0})$, $Z \perp (S_i)_{i=1}^{\sup\{n\in\mathbb{N}_{\geqslant 1}:S_n\leqslant t\}}$ and $Z\perp (N_\tau)_{\tau\in(0,t]}$.

Moreover, conditioning on events $\{N_t = n\}$, it turns out that one has essentially no knowledge about when the associated arrivals of period $[0,t]$ happened, these being uniformly distributed on the region where they are supported:

$$\mathbb{P}(S_1\leqslant s \mid N_t = 1) = \begin{cases} s/t, \text{ if } 0<s\leqslant t \\ 0, \text{ otherwise}\end{cases}, \text{ i.e., } PDF_{S_{1*}\mathbb{P}|N_t}(s\mid 1) = \begin{cases} 1/t, \text{ if } 0<s\leqslant t \\ 0, \text{ otherwise}\end{cases}$$

and more generally

$$PDF_{(S_1,\ldots,S_n)_*\mathbb{P}|N_t}(s_1,\ldots,s_n\mid n) = \begin{cases} n!/t^n, \text{ if } 0<s_1<\ldots<s_n\leqslant t \\ 0, \text{ otherwise}\end{cases},$$

like the order statistics of n independent random variables that are uniformly distributed on $(0,t)$.

Now we reconsider discrete approximations to Poisson processes.

Definition 2.1.9. A **Bernoulli process with success probability** $p\in(0,1)$ is a $\{0,1\}$-valued stochastic process $(Y_j)_{j\in\mathbb{N}_{\geqslant 1}}$ on $(\Omega,\mathscr{F},\mathbb{P})$ which is iid and such that $\mathbb{P}(Y_j=1)=p$ and $\mathbb{P}(Y_j=0)=1-p$ $(j\in\mathbb{N}_{\geqslant 0})$.

We interpret that $Y_j = Y_j^p = 1$ indicates an arrival occurred at time j. This implies that the associated counting process $N_t^p = \sum_{j=1}^t Y_j^p$ $(t\in\mathbb{N}_{\geqslant 0})$ is such that $N^p_{t\,*}\mathbb{P} = Bin_{t,p}$ $(t\in\mathbb{N}_{\geqslant 0})$ and the associated inter-arrival times process $(X_i^p)_{i\in\mathbb{N}_{\geqslant 1}}$ is iid and such that $X^p_{i\,*}\mathbb{P} = Geo_p$ $(i\in\mathbb{N}_{\geqslant 1})$. Naturally, an IAC can be induced from them. We then want to consider associated rescaled-in-time and rescaled-in-success-probability IAC's (a k-indexed family of them): $N_t^{(k)} := \sum_{j=1}^{\lfloor t2^j\rfloor} Y_j^{\gamma 2^{-j}}$ $(k\in\mathbb{N}_{\geqslant 0}, t\in\mathbb{R}_{\geqslant 0})$. Under this rescaling, the arrival rate is kept constant, i.e., $\mathbb{E}_{\mathbb{P}}(N_1^{(k)}) = \mathrm{Mean}(Bin_{2^k,\gamma 2^{-k}}) = \gamma$ $(k\in\mathbb{N}_{\geqslant 0})$ and, in fact,

$$N_t^{(k)}{}_*\mathbb{P}(\{n\}) \xrightarrow{k\to\infty} \tilde{N}_{t*}\tilde{\mathbb{P}}(\{n\}),\ (n\in\mathbb{N}_{\geqslant 0}, t\in\mathbb{N}_{\geqslant 0}),$$

where $(\tilde{N}_t)_{t\in\mathbb{R}_{\geqslant 0}}$ in $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{\mathbb{P}})$ is a Poisson process with intensity γ. This result is known as Poisson's theorem for processes.

Finally, we consider Poisson processes with measure intensities (a.k.a., possibly inhomogeneous intensities) instead of real intensities (a.k.a. homogeneous).

Definition 2.1.10. A **Poisson process with intensity** $\gamma\in\mathcal{M}^{Rad}(\mathbb{R}_{\geqslant 0})$ (without atoms) is a stochastic process $(N_t)_{t\in\mathbb{R}_{\geqslant 0}}$ verifying condition (3) of definition 2.1.6, having independent (but not anymore stationary) increments and satisfying

$$(N_{t'}-N_t)_*\mathbb{P} = Poi_{\gamma((t,t'])}\ (t,t'\in\mathbb{R}_{\geqslant 0}, t<t').$$

Here is a subtlety. Despite the above counting process having independent increments, in general, its associated inter-arrival process is no longer independent (nor identically distributed, obviously).

Namely, conditioning on the value of the counting process at a certain time will not change the statistics of how larger it is going to be at a certain time afterward; however, it will change the statistics of the following inter-arrival time:

$$\mathbb{P}(X_2 > t|X_1) = \mathbb{P}(N_{X_1+t} - N_{X_1} = 0) = \mathrm{Poi}_{\gamma((X_1,X_1+t])}(\{0\}) = e^{-\gamma((X_1,X_1+t])}.$$

When in the above $\gamma = rLeb$, the integral on the RHS is always $r(t'-t)$, so increments are again stationary and $N_{t*}\mathbb{P} = Poi_{rt}$ $(t \in \mathbb{R}_{\geqslant 0})$. Therefore a Poisson process with measure intensity $rLeb$ is simply a Poisson process with real intensity r.

The above definition would still be meaningful for $\gamma \in \mathcal{M}^{Rad}(\mathbb{R}_{\geqslant 0})$ admitting atoms away from 0. However, this would break condition (3.i.d) of definition 2.1.6. For pedagogic reasons, we do not relax this condition now. In definition 2.1.17, it will be relaxed to handle $\gamma \in \mathcal{M}^{Rad}(\mathbb{R}_{\geqslant 0})$ in general.

We will note later that the definitions of this section are non-void — in the more general context of section 2.1.3, definition 2.1.17.

2.1.2 Compound Poisson distribution and process

A limitation of modeling using random variables with Poisson distributions (or Poisson processes) is that they can not accommodate batch arrivals. Picture the situation where a call center is taking orders for a burger shop. When an order arrives, how many burgers are demanded is of interest. The total demand in a unitary time interval will be modeled with a random variable having a compound Poisson distribution. A compound Poisson process will model how the previous situation evolves in time.

Definition 2.1.11. The **compound Poisson distribution with intensity** $\gamma \in \mathbb{R}_{>0}$ **and multiplicity distribution** $(\lambda_\ell)_{\ell \in \mathbb{N}_{\geqslant 1}} \in \mathcal{P}(\mathbb{N}_{\geqslant 1})$, $\sum_{\ell=1}^{\infty} \ell \lambda_\ell < \infty$, denoted $CPD_{\gamma,(\lambda_\ell)_\ell}$, is the distribution of a random variable $M : (\Omega, \mathscr{F}, \mathbb{P}) \to \mathbb{N}_{\geqslant 0}$ given by

$$M(\omega) = \sum_{j=1}^{N(\omega)} Q_j(\omega),$$

where N is a $\mathbb{N}_{\geqslant 0}$-valued random variable on $(\Omega, \mathscr{F}, \mathbb{P})$ having Poisson distribution with intensity γ and $(Q_j)_{j \in \mathbb{N}_{j\geqslant 1}}$ is a sequence of $\mathbb{N}_{\geqslant 1}$-valued random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ which are iid, independent of N and whose entries have distribution $\mathbb{P}(Q_j = \ell) = \lambda_\ell$ $(j, \ell \in \mathbb{N}_{\geqslant 1})$.

Its probability mass function is given indirectly by

$$CPD_{\gamma,(\lambda_\ell)_\ell}(\{k\}) = \sum_{n=1}^{k} \mathbb{P}(N = n)\mathbb{P}(R_n = k) = \sum_{n=1}^{k} \frac{\gamma^n e^{-\gamma}}{n!} \mathbb{P}(R_n = k) \ (k \in \mathbb{N}_{\geqslant 0}), \tag{2.1}$$

where $R_n = \sum_{j=1}^{n} Q_j$.

Mean and variance of $CPD_{\gamma,(\lambda_\ell)_\ell}$ are, respectively, $\gamma \sum_{\ell=1}^{\infty} \ell \lambda_\ell \in \mathbb{R}_{>0}$ and $\gamma \sum_{\ell=1}^{\infty} \ell^2 \lambda_\ell \in \mathbb{R}_{>0} \cup \{\infty\}$.

Definition 2.1.12. A **compound Poisson process with intensity** $\gamma \in \mathbb{R}_{>0}$ **and multiplicity distribution** $(\lambda_\ell)_{\ell \in \mathbb{N}_{\geqslant 1}} \in \mathcal{P}(\mathbb{N}_{\geqslant 1})$, $\sum_{\ell=1}^{\infty} \ell \lambda_\ell < \infty$, is an $\mathbb{N}_{\geqslant 0}$-valued stochastic process $(M_t)_{t \in \mathbb{R}_{\geqslant 0}}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ that:

i. $\exists \Omega_* \in \mathscr{F}, \mathbb{P}(\Omega_*) = 1, \forall \omega \in \Omega_*$:

a. $M_0(\omega) = 0$,

b. $t \in \mathbb{R}_{\geqslant 0} \mapsto M_t(\omega) \in \mathbb{N}_{\geqslant 0}$ is left-continuous with finitely many discontinuities on bounded subsets of $\mathbb{R}_{\geqslant 0}$,

c. $t \leqslant t' \Rightarrow M_t(\omega) \leqslant M_{t'}(\omega)$,

d. On the discontinuities referred to in (b), the right-limits minus respective left-limits are always in $\mathbb{N}_{\geqslant 1}$,

ii. has independent and stationary increments, and

iii. $M_{t*}\mathbb{P} = CPD_{\gamma t,(\lambda_\ell)_\ell}$ $(t \in \mathbb{R}_{\geqslant 0})$.

If $\lambda_1 = 1$ (and the remaining λ_ℓ's are 0), a compound Poisson process (distribution) with real intensity γ and multiplicity distribution $(\lambda_\ell)_\ell$ reduces to a Poisson process (distribution) with real intensity λ.

Finally, we consider compound Poisson processes with measure intensities instead of real intensities.

Definition 2.1.13. A **compound Poisson process with intensity** $\gamma \in \mathcal{M}^{Rad}(\mathbb{R}_{\geqslant 0})$ (without atoms) **and multiplicity distribution** $(\lambda_\ell)_{\ell \in \mathbb{N}_{\geqslant 1}} \in \mathcal{P}(\mathbb{N}_{\geqslant 1})$, $\sum_{\ell=1}^{\infty} \ell \lambda_\ell < \infty$, is an $\mathbb{N}_{\geqslant 0}$-valued stochastic process $(M_t)_{t \geqslant \mathbb{R}_{\geqslant 0}}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ that:

i. as definition 2.1.12,

ii. has independent increments, and

iii. $(M_{t'} - M_t)_* \mathbb{P} = CPD_{\gamma((t,t']),(\lambda_\ell)_\ell}$ $(t, t' \in \mathbb{R}_{\geqslant 0}, t < t')$.

Once again, when in the above $\gamma = rLeb$, the integral on the RHS is always $r(t' - t)$, so increments are again stationary and $N_{t*}\mathbb{P} = CPD_{rt,(\lambda_\ell)_\ell}$ $(t \in \mathbb{R}_{\geqslant 0})$. Therefore a compound Poisson process with measure intensity *rLeb* and multiplicity distribution $(\lambda_\ell)_\ell$ is simply a compound Poisson process with real intensity r and multiplicity distribution $(\lambda_\ell)_\ell$.

Moreover, if $\lambda_1 = 1$ (and the remaining λ_ℓ's are 0), a compound Poisson process (distribution) with measure intensity λ as above and multiplicity distribution $(\lambda_\ell)_\ell$ reduces to a Poisson process (distribution) with measure intensity λ.

One could possibly consider generalizations of compound Poisson processes in which not only the intensity γ, but also the multiplicity distribution $(\lambda_\ell)_\ell$, varies in a prescribed way along the positive real line (this, however, would break the independence between multiplicities and the number of arrivals) — e.g., back to the burger shop illustration, it might be that both propensity of call/order arrivals, but also their associated demand profile, varied with time. More about this in definition 2.1.21.

Also, as noted at the end of section 2.1.1, technical extensions concerning γ's atomicity are going to be handled later. See definition 2.1.21.

2.1.3 Poisson point process

In this section, we take Poisson processes to a more abstract level in which intensity of random arrivals over time will be understood as intensity of points (or, more precisely, unitary point-masses) occurring randomly over space. When the space is the positive real line, we are back to the section of Poisson processes. Overall, [76] and [66] are the main references in this section.

Let E be a locally compact complete separable metric space (thus Hausdorff with a countable basis) with its Borel σ-algebra $\mathscr{B}_E$.

Let $\mathcal{M}^{Rad}(E)$ be given the vague topology, i.e., the smallest topology making continuous the following family of evaluation maps $\{T_f : m \in \mathcal{M}^{Rad}(E) \mapsto m(f) \in [0,\infty] \mid f \in C_c^+(E)\}$. This is a complete separable metric space (see [79] section 2.3.3). The associated convergence is denoted $m_n \xrightarrow{v} m$.

Let $\mathcal{M}^{Rad}(E)$ be the σ-algebra

$$\mathscr{M}^{Rad}(E) = \sigma(\{T_F : m \in \mathcal{M}^{Rad}(E) \mapsto m(F) \in \mathbb{R} \mid F \in \mathscr{B}_E\}).$$

It coincides with the Borel σ-algebra associated to the vague topology, $\mathscr{B}_{\mathcal{M}^{Rad}(E)}$ (see [76] pg. 141).

Define (see [76] sec. 3.1, [79] def. 2.3.2 and [66] def. 2.4, def. 2.1, prop 6.2, prop. 6.3, cor. 6.5)

$$\begin{aligned}
&\mathcal{M}_p^{Rad}(E) \\
&\quad = \left\{m = \textstyle\sum_{i=1}^{\kappa} \delta_{x_i} : (x_i)_{i=1}^{\kappa} \subset E, \kappa \in \mathbb{N}_{\geqslant 0} \cup \{\infty\}, m(K) < \infty, \forall K \subset E \text{ compact}\right\} \\
&\quad = \left\{m = \textstyle\sum_{i=1}^{\kappa} \delta_{x_i} : (x_i)_{i=1}^{\kappa} \subset E, \kappa \in \mathbb{N}_{\geqslant 0} \cup \{\infty\}, \text{ with no accumulation points}\right\} \\
&\quad \subset \mathcal{M}^{Rad}(E),
\end{aligned}$$

where the sequences $(x_i)_{i=1}^{\kappa}$ above might have repetitions. This set is closed and hence measurable (see [76] pg. 140 and prop. 3.14). Therefore $\mathcal{M}_p^{Rad}(E)$, as a subset, inherits the structure from $\mathcal{M}^{Rad}(E)$. The inherited topology is its vague topology and makes it a complete separable metric space (see [76] prop. 3.17). The inherited σ-algebra concides with the analogous definition of $\mathscr{M}_p^{Rad}(E)$ and $\mathscr{B}_{\mathcal{M}_p^{Rad}(E)}$.

Define

$$\begin{aligned}
\mathcal{M}_{sp}^{Rad}(E) &= \left\{m \in \mathcal{M}_p^{Rad}(E) : m(\{x\}) \leqslant 1, \forall x \in E\right\} \\
&= \left\{m = \textstyle\sum_{i=1}^{\infty} \delta_{x_i} \in \mathcal{M}_p^{Rad}(E) : (x_i)_{i \in \mathbb{N}_{\geqslant 1}} \subset E \text{ disjoint}\right\},
\end{aligned}$$

as a subset of $\mathcal{M}_p^{Rad}(E)$. This subset will not be qualified topologically, but it is important to note that it is a measurable subset of $\mathcal{M}_p^{Rad}(E)$ (see [66] prop. 6.7).

Definition 2.1.14. A **random measure** is a random variable

$$Z : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}^{Rad}(E), \mathscr{M}^{Rad}(E)).$$

From now on we consider only the following special case of random measures.

Definition 2.1.15. A **point process** is a random variable

$$Z : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p^{Rad}(E), \mathscr{M}_p^{Rad}(E)).$$

Such a function is measurable if, and only if, $Z(\cdot)(F) : \Omega \to \mathbb{N}_{\geqslant 0} \cup \{\infty\}$ is measurable $\forall F \in \mathscr{B}_E$ (see [76] prop. 3.1).

Definition 2.1.16. A **simple point process** is a point process

$$Z : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p^{Rad}(E), \mathscr{M}_p^{Rad}(E))$$

so that $\mathbb{P}(Z \in \mathcal{M}_{sp}^{Rad}(E)) = 1$.

Two simple point processes N and N' on $(\Omega, \mathscr{F}, \mathbb{P})$ have a common distribution if the probabilities that their realizations avoid a set are coincident, for every set in a suitable family: if $\mathbb{P}(\{\omega : N(\omega)(F) = 0\}) = \mathbb{P}(\{\omega : N'(\omega)(F) = 0\})$ $\forall F \subset E, F \in \mathscr{R}$ (a ring generating $\mathscr{B}_E$), then $N_* \mathbb{P} = N'_* \mathbb{P}$ (see [76] prop. 3.23 and [58] thm. 3.3). From here on, $\mathbb{P}(\{\omega : N(\omega)(F) = k\})$ will be denoted simply $\mathbb{P}(N(F) = k)$.

A sequence of point processes $(Z_n)_{n \in \mathbb{N}_{\geqslant 1}}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ is said to converge in distribution to another point process Z_0 on $(\Omega, \mathscr{F}, \mathbb{P})$, denoted $Z_n \xrightarrow{d} Z_0$, if $Z_{n*}\mathbb{P} \xrightarrow{v} Z_{0*}\mathbb{P}$.

Convergence in distribution of point processes can be subtle to verify. Beyond portmanteau-type theorems, one can consider Kallenberg's criteria, which are sufficient conditions for a sequence of point processes to converge simple point processes with enough regularity (see [58] thm. 4.3 or [34] thm. 2.1 for a neat presentation).

Now we start considering Poisson point processes.

Definition 2.1.17. A **Poisson point process with intensity** $\gamma \in \mathcal{M}^{Rad}(E)$ (with or without atoms) is a point process

$$N : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p^{Rad}(E), \mathscr{M}_p^{Rad}(E))$$

satisfying:

i. $\forall k \in \mathbb{N}_{\geqslant 1}, \forall F_1, \ldots, F_k \in \mathscr{B}_E$ mutually disjoint:

$$\big(\omega \in \Omega \mapsto N(\omega)(F_i) \in \mathbb{N}_{\geqslant 0} \cup \{\infty\}\big)_{i=1}^{k} \text{ forms an independency}$$

ii. $\forall F \in \mathscr{B}_E, \forall n \in \mathbb{N}_{\geqslant 0}$:

$$\mathbb{P}(N(F) = n) = \begin{cases} \frac{\gamma(F)^n e^{-\gamma(F)}}{n!}, & \text{if } \gamma(F) < \infty \\ 0 & \text{, if } \gamma(F) = \infty \end{cases} .$$

It follows from the previous definition that $\gamma(F) = \infty \Rightarrow \mathbb{P}(N(F) = \infty) = 1$.

It can be shown that definition 2.1.17 is nonempty and that any two Poisson point processes with the same intensity will have the same distributions (see [76] prop. 3.6(i) and [66] prop. 3.2 and section 3.2).

Also, it can be shown that Poisson point processes are simple if, and only if, the associated measure intensities have no atoms (see [66] prop 6.9). Among them, we call homogeneous those having Lebesgue measure or volume measure (in case E is given a Riemannian manifold structure) intensities.

When $E = [0, \infty)$, there is a natural identification between I) simple point processes and II) counting processes introduced in section 2.1.1, definition 2.1.6: points in a realization of a simple point process are in correspondence with discontinuity points in a realization of a counting process. Likewise, I) Poisson point processes with measure intensities having no atoms are identified with II)

Poisson (counting) processes with measure intensities having no atoms, introduced in section 2.1.1, definition 2.1.10.

In our applications, E will be either $\mathbb{R}^d$, for some $d \in \mathbb{N}_{\geqslant 1}$, or some tangent space or bundle. These will be equipped, respectively, with Lebesgue and volume measures (induced from a Riemannian structure), which might be taken as intensities.

2.1.4 Compound Poisson point process

We noticed that, when $E = [0, \infty)$, there is a natural identification between I) Poisson point processes with measure intensities having no atoms and II) Poisson (counting) processes with measure intensities having no atoms. The latter could be adapted to handle batch arrivals, and so can the former. Batches will be reinterpreted in terms of (integer) weights independently assigned to points, or even abstract quantities independently attached to them.

The first step in this abstract construction is the following.

Definition 2.1.18. Consider a mark space $(H, \mathscr{B}_H)$ which is a locally compact complete separable metric space with its Borel σ-algebra $\mathscr{B}_H$. A **H-marked point process** over a point process

$$Z : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p^{Rad}(E), \mathscr{M}_p^{Rad}(E))$$

is an enlarged point process

$$X : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p^{Rad}(E \times H), \mathscr{M}_p^{Rad}(E \times H))$$

so that $\pi_1 \circ X = Z$.

Definition 2.1.19. (See [81] def. 2.10) Consider an H-marked point process X over Z as in definition 2.1.18. It is called an **H-marked point process with kernel** $K(e, dh)$, namely

$$\begin{array}{rccccl} K: & E \times \mathscr{B}_H & \to & [0,1] & \text{satisfying} & K(e, \cdot) \in \mathcal{P}^{Rad}(H), \forall e \in E \\ & (e, G) & \mapsto & K(e, G) & & K(\cdot, G) \text{ measurable}, \forall G \in \mathscr{B}_H \end{array}$$

if — after having fixed a representation $Z(\omega) = \sum_{n=1}^{\kappa(\omega)} \delta_{Z_n(\omega)}$, with random variables $\kappa : \Omega \to \mathbb{N}_{\geqslant 0} \cup \{\infty\}$ and $Z_n : \Omega \to E$ $(n \in \mathbb{N}_{\geqslant 1})$ — X has the structure

$$X(\omega) = \sum_{n=1}^{\kappa(\omega)} \delta_{(Z_n(\omega), W_n(\omega))}$$

for some sequence of random variables $W_n : \Omega \to H$ $(n \in \mathbb{N}_{\geqslant 1})$ satisfying that $\forall n \in \mathbb{N}_{\geqslant 1}, \forall G_1, \ldots, G_n \in \mathscr{B}_H$:

$$\mathbb{P}\Big(W_1 \in G_1, \ldots, W_n \in G_n \;\Big|\; \kappa \geqslant n, Z = (Z_1, \ldots, Z_n, \ldots)\Big)(e_1, \ldots, e_n, \ldots) = \prod_{i=1}^{n} K(e_i, G_i).$$

The above definition uses the representation of Z into Z_n's in an innocuous way because if a different representation is chosen there is no distributional implication, due to the fact that the probabilities of marks are independent after conditioning on the entire realization of the base point process. Similarly, the specific choice of W_n's, as long as the distributional equation is satisfied, is also innocuous in the same sense.

A special subcase of the previous definition is that in which the marks are completely independent of Z, i.e., $K(e,\cdot) \overset{\forall e}{=} \lambda(\cdot)$ — the latter being the space-independent distribution of marks.

Definition 2.1.20. An H-marked point process X with kernel $K(e,dh)$ (as in definition 2.1.19) over an Poisson point process Z with measure intensity $\gamma(de)$ (as in definition 2.1.17) is called an **H-marked Poisson point process with intensity $\gamma(de)$ and kernel $K(e,dh)$**.

For example, in the above definition, we can set $E = [0,\infty)$ and $H = \mathbb{R}^d$ to define a random walk in random times through $\mathbb{R}^d$, where particles start from the origin and move at the instants drawn in $E = [0,\infty)$ by the associated translation vector drawn in $\mathbb{R}^d$.

It can be shown that an H-marked Poisson point process $X : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p(E\times H), \mathscr{M}_p(E\times H))$ with intensity $\gamma(de)$ and kernel $K(e,dh)$ is actually a Poisson point process $X : (\Omega, \mathscr{F}, \mathbb{P}) \to (\mathcal{M}_p(E\times H), \mathscr{M}_p(E\times H))$ with intensity $K(e,dh)\gamma(de)$ (see [81] thm. 2.12).

Definition 2.1.21. Consider X, an $\mathbb{N}_{\geqslant 1}$-marked Poisson point process with intensity $\gamma(de)$ and kernel $K(e,dh)$, as in definition 2.1.20.

After representing it as in definition 2.1.19,

$$X(\omega) = \sum_{i=1}^{\kappa(\omega)} \delta_{(Z_n(\omega),W_n(\omega))},$$

it induces the following (non-simple) point process

$$M(\omega) = \sum_{j=1}^{\kappa(\omega)} W_n(\omega)\delta_{Z_n(\omega)}$$

which is referred to as a **compound Poisson point process with intensity $\gamma(de)$ and multiplicity kernel $K(e,dh)$**.

Changes in representation are again distributionally innocuous. Here the notation is slightly misleading, because, since the mark-space is $\mathbb{N}_{\geqslant 1}$, then $K(e,dh) \in \mathcal{P}(\mathbb{N}_{\geqslant 1}), \forall e \in E$. When the multiplicity kernel is space-independent, we have simply $K(e,dh) \overset{\forall e}{=} (\lambda_\ell)_{\ell\in\mathbb{N}_{\geqslant 1}} \in \mathcal{P}(\mathbb{N}_{\geqslant 1})$, for some $(\lambda_\ell)_{\ell\in\mathbb{N}_{\geqslant 1}} \in \mathcal{P}(\mathbb{N}_{\geqslant 1})$. In this case, we write just suppress the word "kernel" from the terminology.

The same stacking idea used in definition 2.1.21 could be applied to $H = \mathbb{R}_{\geqslant 0}$, but the random variable M introduced therein would not anymore be a point process.

Once more, when $E = [0,\infty)$, there is a natural identification between
I) compound Poisson point processes with measure intensity γ having no atoms and multiplicity $(\lambda_\ell)_\ell$ (as in above), and
II) compound Poisson processes with measure intensity γ having no atoms and multiplicity $(\lambda_\ell)_\ell$ (as in definition 2.1.13).

Moreover, a compound Poisson point process with measure intensity γ and multiplicity for $(\lambda_\ell)_\ell$, where $\lambda_1 = 1$, reduces to a Poisson point process with measure intensity γ.

2.1.5 Lifted compound Poisson point process

Definition 2.1.22. Consider $(E,\mathscr{E})$ and $(H,\mathscr{H})$ two locally compact complete separable metric spaces with their Borel σ-algebra. An **H-lifted compound Poisson process on the state space E relative to** $\{G_\alpha \in \mathscr{B}_H$, bounded $:\ \alpha \in \mathcal{A}\}$ **with intensity** $\alpha \in \mathcal{A} \mapsto \gamma_\alpha \in \mathcal{M}^{Rad}(E)$ **and multiplicity kernel** $\alpha \in \mathcal{A} \mapsto K_\alpha(e,dh)$ is a point process

$$\begin{array}{cccc} X: & (\Omega,\mathscr{F},\mathbb{P}) & \to & (\mathcal{M}_p^{Rad}(E\times H),\mathscr{M}_p^{Rad}(E\times H)) \\ & \omega & \mapsto & \sum_{n=1}^{\kappa(\omega)} \delta_{(A_n(\omega),B_n(\omega))} \end{array}$$

so that, for every $\alpha \in \mathcal{A}$, denoting

$$\begin{array}{ccc} X_\alpha := X|_{E\times G_\alpha} : (\Omega,\mathscr{F},\mathbb{P}) \to & (\mathcal{M}_p^{Rad}(E\times H),\mathscr{M}_p^{Rad}(E\times H)) \\ \omega & \mapsto \sum_{n=1}^{\kappa(\omega)} \delta_{(A_n(\omega),B_n(\omega))} \mathbb{1}_{E\times G_\alpha}(A_n(\omega),B_n(\omega)) \end{array},$$

it holds that $\pi_1 \circ X_\alpha$ is a compound Poisson point process with intensity $\gamma_\alpha(de)$ and multiplicity kernel $K_\alpha(e,dh)$.

Remark 1. Do not confuse the above with the stacking procedure of definition 2.1.21 which used an $\mathbb{N}_{\geqslant 1}$-marked Poisson point process to assign weights to the points of an underlying Poisson point process and thereby create a compound Poisson point process. The said $\mathbb{N}_{\geqslant 1}$-marked process and its induced compound Poisson point process have essentially the same information (exactly, if the Poisson point process is simple). The X_α's that were specified in definition 2.1.22 have more information than the compound Poisson point processes they project to, as if every mass-unit in such compound Poisson point process is assigned a value in H.

To interpret the object of the last definition, consider that a seismograph can only detect earthquakes with Richter magnitudes larger than 8. They might come in batches. We want to register not only how many come in each batch, but also how severe is each of these. The last attribute is in the H space. If we collapse this information, we recover the usual compound Poisson point process, say in the state space $E = [0,\infty)$, representing time. What if the seismograph can be tuned to detect smaller earthquakes? This is exactly changing G. For example, we could consider $\{G_\alpha = [\alpha,\infty) : \alpha \in \mathbb{R}_{>0}\}$, and G would first be G_8 and then be G_7. As α diminishes, (after projecting) one can see how the statistics given by the associated α-compound Poisson point processes change, i.e. how intensities and multiplicity kernels change. In our example, we expect intensities to increase as α decreases. Notice also that the H space can be richer, recording more information than simply the Ritcher magnitude of an earthquake.

Notice that in a compound Poisson point process, the unit masses that are stacked over a point (representing multiplicity) are indistinguishable and thus are not even ordered. If however, we could somehow distinguish them as to produce an order, we could see such a process as lifted into the space $H = \mathbb{N}_{\geqslant 1}$ space, where the new coordinate registers precisely the order. In general, all the information available to distinguish masses can be registered in a new lifted coordinate, i.e., in principle, one could enrich the lifting more and more. On the other hand, any H-lift of the original process can be used to distinguish and order its unit masses.

2.2 Extreme value theory and hitting time statistics (iid case)

2.2.1 Extreme value theory

Let $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ be a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega,\mathscr{F},\mathbb{P})$. Denote by $F_{X_0}(x) = \mathbb{P}(X_0 \leqslant x)$ the cumulative distribution function (cdf) of X_0. Also, denote $M_n = \max\{X_0,\dots,X_{n-1}\}$ $(n \geqslant 1)$, another $\mathbb{R}$-valued stochastic process on $(\Omega,\mathscr{F},\mathbb{P})$.

Extreme value theory wants to study the statistics of large realizations of the process, which can be understood from many perspectives. Let's start with the classical one and then look at some relevant equivalent perspectives.

We say that a cdf $G : \mathbb{R} \to \mathbb{R}$ is non-degenerate if its associated measure $\mathbb{G} \in \mathcal{M}^{Rad}(\mathbb{R})$ is not the Dirac measure of a point.

Definition 2.2.1. Let G be a non-degenerate cdf.

The (possibly empty) **maximum domain of attraction of** G is the set $MDA(G)$ comprised of iid $\mathbb{R}$-valued stochastic processes $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ on $(\Omega,\mathscr{F},\mathbb{P})$ for which there exists $(a_n)_{n\in\mathbb{N}_{\geqslant 0}} \subset \mathbb{R}_{>0}, (b_n)_{n\in\mathbb{N}_{\geqslant 0}} \subset \mathbb{R}$:

$$\begin{gathered}\left(\tfrac{M_n-b_n}{a_n}\right)_* \mathbb{P} \xrightarrow[n\to\infty]{v} \mathbb{G}, \text{ or, equivalently,}\\ \forall x \in cont(G) :\ \mathbb{P}(M_n \leqslant a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x).\end{gathered} \tag{2.2}$$

Denote $MDA = \{G \text{ non-degenerate cdf } : \ MDA(G) \neq \varnothing\}$.

Remark 2. If $\big((X_i)_{i\in\mathbb{N}_{\geqslant 0}}, (a_n)_{n\in\mathbb{N}_{\geqslant 0}}, (b_n)_{n\in\mathbb{N}_{\geqslant 0}}, G\big)$ satisfy equation 2.2, there can still be another $\big((X_i)_{i\in\mathbb{N}_{\geqslant 0}}, (\tilde{a}_n)_{n\in\mathbb{N}_{\geqslant 0}}, (\tilde{b}_n)_{n\in\mathbb{N}_{\geqslant 0}}, \tilde{G}\big)$ satisfying it, but they will relate like (see [79] lem. 1.2.7)

$$\frac{\tilde{a}_n}{a_n} \underset{n\to\infty}{\to} A \in \mathbb{R}_{>0},\ \frac{\tilde{b}_n - b_n}{a_n} \underset{n\to\infty}{\to} B \in \mathbb{R},\ \tilde{G}(x) = G(Ax+B). \tag{2.3}$$

Definition 2.2.2. The set of max-stable distributions is

$$MS = \left\{ G \text{ non-degenerate cdf} \;\middle|\; \begin{array}{l}\exists \alpha : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{>0}, \beta : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0} \text{ so that}\\ G(\alpha(t)x+\beta(t))^t = G(x), \forall t \in \mathbb{R}_{>0}, x \in \mathbb{R}\end{array}\right\}.$$

Consider $G \in MS$. If $(X_i)_{i\mathbb{N}_{\geqslant 0}}$ is an $\mathbb{R}$-valued stochastic process which is independent and whose entries are distributed like G, then M_n is distributed like G^n $(n \in \mathbb{N}_{\geqslant 0})$ and the condition above says that $(X_i)_{i\in\mathbb{N}_{\geqslant 0}} \in MDA(G)$ using $a_n := \alpha(n)$ and $b_n := \beta(n)$ $(n \in \mathbb{N}_{\geqslant 0})$.

Next is the main theorem characterizing the objects referred to in this section (see [80] thm. 1 and [79] prop 1.3.2, thm. 1.3.4 and rmk. 1.3.6).

Theorem 2.2.3. *[Fisher–Tippett–Gnedenko] The sets MDA and MS coincide. Moreover, if* $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ *a sequence of iid* $\mathbb{R}$*-valued random variables on* $(\Omega,\mathscr{F},\mathbb{P})$*,* $(a_n)_{n\in\mathbb{N}_{n\geqslant 0}} \subset \mathbb{R}_{>0}$*,* $(b_n)_{n\in\mathbb{N}_{n\geqslant 0}} \subset \mathbb{R}_{\geqslant 0}$ *and G non-degenerate cdf satisfies*

$$\forall x \in cont(G) :\ \mathbb{P}(M_n \leqslant a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x),$$

then G is a generalized extreme value (GEV) cdf

$$G(x) = G_{\xi,\sigma,\mu}(x) := \begin{cases} e^{-\left[0 \vee (1+\xi \frac{x-\mu}{\sigma})\right]^{-1/\xi}} &, \textit{if } \xi \neq 0 \\ e^{-e^{-\frac{x-\mu}{\sigma}}} &, \textit{if } \xi = 0 \end{cases} \qquad (x \in \mathbb{R}),$$

for some $\xi, \mu \in \mathbb{R}$, $\sigma \in \mathbb{R}_{>0}$, *with*

i) [light/exponential tail case] $\xi = 0$ *if, and only if, the associated GEV cdf is the cdf of the Gumbel distribution*

$$\Lambda_{\sigma,\mu}(x) = e^{-e^{-\frac{x-\mu}{\sigma}}} \quad (x \in \mathbb{R})$$

ii) [heavy/polynomial tail case] $\xi > 0$ *if, and only if, the associated GEV cdf is the cdf of the Fréchet distribution*

$$\Phi_{1/\xi,\sigma,\mu}(x) = \begin{cases} e^{-\left(\frac{x-\mu}{\sigma}\right)^{-1/\xi}} &, \textit{if } x > 0 \\ 0 &, \textit{if } x \leqslant 0 \end{cases} \qquad (x \in \mathbb{R})$$

iii) [upper bounded tail case] $\xi < 0$ *if, and only if, the associated GEV cdf is the cdf of the Weibull distribution*

$$\Psi_{1/\xi,\sigma,\mu}(x) = \begin{cases} e^{-\left(-\frac{x-\mu}{\sigma}\right)^{-1/\xi}} &, \textit{if } x < 0 \\ 1 &, \textit{if } x \geqslant 0 \end{cases} \qquad (x \in \mathbb{R}).$$

In particular, the set MDA = MS is exactly the space of GEV cdfs (or, specifically Gumbel, Fréchet and Weibull cdfs), which is closed under the modification introduced in equation 2.3: $G_{\xi,\sigma,\mu}(Ax+B) = G_{\xi,\sigma/A,(B-\mu)/A}(x)$. *Using this type of procedure one can always reduce the limit to standard type, i.e., to* $G_{\xi,1,0}$.

The mean and variance of $G_{\xi,\sigma,\mu}$ are

$$Mean_{G_{\xi,\sigma,\mu}} = \begin{cases} \mu + \sigma \cdot (\Gamma(1-\xi) - 1)/\xi, & \text{if } \xi \neq 0, \xi < 1 \\ \mu + \sigma \cdot \gamma &, \text{if } \xi = 0 \\ \infty &, \text{if } \xi \geqslant 1 \end{cases} \qquad \text{and}$$

$$Var_{G_{\xi,\sigma,\mu}} = \begin{cases} \sigma^2 \cdot (\Gamma(1-2\xi) - \Gamma(1-\xi)^2)/\xi^2, & \text{if } \xi \neq 0, \xi < 1/2 \\ \sigma^2 \cdot \pi^2/6 &, \text{if } \xi = 0 \\ \infty &, \text{if } \xi \geqslant 1/2 \end{cases} \qquad ,$$

where γ is the Euler–Mascheroni constant.

In the previous theorem, we can pack $a_n x + b_n$ into a single number u_n. But we might also want to consider similar asymptotics when u_n has not such a linear structure on x, or is even a standalone sequence, independent of x. In this direction, the following proposition is useful.

Proposition 2.2.4. *Let* $\tau \in \mathbb{R}_{>0}$ *and* $(u_n)_{n \in \mathbb{N}_{\geqslant 0}} \subset \mathbb{R}$. *Then*

$$n\mathbb{P}(X_0 > u_n) = n(1 - F_{X_0}(u_n)) \underset{n\to\infty}{\longrightarrow} \tau \Leftrightarrow \mathbb{P}(M_n \leqslant u_n) \underset{n\to\infty}{\longrightarrow} e^{-\tau}. \qquad (2.4)$$

The proof of the implication ($\Rightarrow$) in the previous proposition is simple:

$$\mathbb{P}(M_n \leqslant u_n) = (1 - \mathbb{P}(X_0 > u_n))^n \sim \left(1 - \frac{\tau}{n}\right)^n \to e^{-\tau}.$$

It is often possible to adjust some $(u_n)_{n\in\mathbb{N}_{\geqslant 0}}$ to each τ, defining some

$$\begin{array}{rccc} u: & \mathbb{N}_{\geqslant 0} \times \mathbb{R}_{>0} & \to & \mathbb{R} \\ & (n,\tau) & \mapsto & u_n(\tau) \end{array}$$

which guarantees the balancing condition on the LHS of equation 2.4 and is continuous, strictly decreasing in τ, strictly increasing in n to $\operatorname{esssup} X_0$ (for any τ). For example, when $X_{0*}\mathbb{P}$ has no atoms, F_{X_0} has range $[0,1]$ and admits the generalized inverse $F_{X_0}^{-1*}$, allowing one to define

$$u_n(\tau) := F_{X_0}^{-1*}(1 - \tau/n), \text{ for } n \geqslant \tau;\ u_n(\tau) := 0, \text{ for } n < \tau.$$

We can interpret the balancing condition on the LHS of equation 2.4 as saying that $(u_n(\tau))_{n\in\mathbb{N}_{\geqslant 0}}$ grows in a pace so that the expected number of times that $(X_i)_{i=0}^{n-1}$ exceeds $u_n(\tau)$ approaches τ.

The RHS of equation 2.4 and the conclusions of theorem 2.2.3 seem related. But notice that the former is not included in the latter: $\mathbb{P}(M_n \leqslant u_n(\tau)) \to e^{-\tau}$, but $\tau \in \mathbb{R}_{>0}$, $u_n(\tau)$ is (generally) not of the form $a_n\tau + b_n$ and $e^{-\tau}$ is not in $DMA = MS$.

To see one instance of this connection, the following proposition is worthwhile.

Proposition 2.2.5. *Let $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ be a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ whose entries have standard normal distribution.*

Then $\mathbb{P}(M_n \leqslant a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G_{0,1,0}(x) = e^{-e^{-x}}$ $(x \in \mathbb{R})$, where

$$a_n = (2\ln n)^{-1/2} \text{ and } b_n = (2\ln n)^{1/2} - 1/2(2\ln n)^{-1/2}(\ln\ln n + \ln 4\pi).$$

This conclusion is definitely in the realm of theorem 2.2.3. The proof develops some calculations which we omit, but its structure involves noticing that

$$\frac{1 - F_{\mathcal{N}(0,1)}(u)}{PDF_{\mathcal{N}(0,1)}(u)/u} \underset{u\to\infty}{\longrightarrow} Q \in \mathbb{R}_{>0},$$

and setting $\tau := e^{-x}$ with the accompanying $u_n(\tau)$ defined implicitly as (notice that $F_{\mathcal{N}(0,1)}$ is invertible)

$$1 - F_{\mathcal{N}(0,1)}(u_n(\tau)) = \frac{\tau}{n} = \frac{e^{-x}}{n},$$

which implies that $n(1 - F_{\mathcal{N}(0,1)}(u_n)) \overset{\forall n}{=} \tau$, so, by design, the LHS of equation 2.4 in proposition 2.2.4 is verified. Moreover, the limit we started with implies that

$$u_n(\tau) = a_n \cdot -\ln\tau + b_n + o(a_n) = a_n x + b_n + o(a_n). \tag{2.5}$$

Therefore one can apply proposition 2.2.4 to conclude, as desired, that

$$\begin{aligned} \mathbb{P}(M_n \leqslant u_n(\tau)) &\underset{n\to\infty}{\longrightarrow} e^{-\tau} = e^{-e^{-x}} \\ = \mathbb{P}(M_n \leqslant a_n x + b_n + o(a_n)) &\underset{n\to\infty}{\longrightarrow} \lim_n \mathbb{P}(M_n \leqslant a_n x + b_n). \end{aligned}$$

On the other side of this connection, we have no indication that $(a_n x + b_n)_n$ in theorem 2.2.3 has to satisfy the balancing condition on the LHS of equation 2.4. The following proposition elucidates what is going on (see [79] lem. 1.4.2).

Proposition 2.2.6. *Let $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ be a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega,\mathscr{F},\mathbb{P})$, $(a_n)_{n\in\mathbb{N}_{n\geqslant 0}} \subset \mathbb{R}_{>0}$, $(b_n)_{n\in\mathbb{N}_{n\geqslant 0}} \subset \mathbb{R}_{\geqslant 0}$ and G non-degenerate cdf. Then*

$$\begin{aligned} &\forall x \in cont(G) : \mathbb{P}(M_n \leqslant a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x) \text{ (i.e., equation 2.2)} \\ &\Leftrightarrow \forall x \in cont(G) :\ n(1 - \mathbb{P}(X_0 > a_n x + b_n)) \underset{n\to\infty}{\longrightarrow} -\ln G(x). \end{aligned}$$

Applying proposition 2.2.6 to the standard Gumbel case, the limit on the RHS of the conclusion is $-\ln G(x) = -\ln(e^{-e^{-x}}) = e^{-x}$. If we denote $u_n(x) := a_n x + b_n$, the limit rewrites as $n(1 - F_{X_0}(u_n(x))) \to e^{-x}$. Thus, the sequence $(a_n x + b_n)_n$ in theorem 2.2.3 really does not satisfy the balancing condition on the LHS of equation 2.4. Also, if we denote $\tau := e^{-x}$, the limit rewrites as $n(1 - F_{X_0}(a_n \cdot -\ln\tau + b_n)) \to \tau$, which provides us a normalizing sequence $u_n(\tau) := a_n \cdot -\ln\tau + b_n$, in the sense of the LHS of equation 2.4. Similar observations can be made in the Fréchet and Weibull cases.

Summarizing: The discussion from proposition 2.2.5 until here was to shed light on how the condition

$$\mathbb{P}(M_n \leqslant a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G(x), \forall x \in \mathrm{cont}(G)$$

in theorem 2.2.3 relate to

$$n\mathbb{P}(X_0 > u_n(\tau)) \underset{n\to\infty}{\longrightarrow} e^{-\tau} \Leftrightarrow \mathbb{P}(M_n \leqslant u_n(\tau)) \underset{n\to\infty}{\longrightarrow} e^{-\tau}, \tau \in \mathbb{R}_{>0}$$

in proposition 2.2.4. In a pedestrian way: given sequences a_n, b_n and G verifying the first condition, one arranges the parametric sequence $u_n(\tau) := a_n \cdot G^{-1}(e^{-\tau}) + b_n$ ($\tau \in \mathbb{R}_{>0}$) which verifies the second condition. The converse is nonexistent, in the sense that the second condition has less information (no way to recover G from it).

From another angle, the first condition might be thought to be finer in the sense of giving a richer asymptotic description, whereas the second might be thought to be finer in the sense of giving an asymptotic description based on less structure.

The best one can do to complete the picture impaired by the 'nonexistent converse' is to consider that proposition 2.2.6 is an equivalence and that it has a counterpart in terms of the alternative balancing condition, as in 2.2.4. More clearly and conveniently, these equivalences are presented in parallelism, respectively, in theorem 2.2.7 (items 1 and 2) and theorem 2.2.8 (items 1 and 2). The reader is invited to read the two items of these two theorems right away.

The next theorem, theorem 2.2.7, describes the phenomena introduced in theorem 2.2.3 also in terms of tails and point processes (see [80] thm. 2, [79] thm. 1.4.1, [22] thm. 2.4).

Theorem 2.2.7 (Pickands, Balkema, de Haan, Resnick)**.** *Let $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ be a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega,\mathscr{F},\mathbb{P})$. Let $\xi\in\mathbb{R}$. Let $x^*=\sup\operatorname{supp}(X_{0*}\mathbb{P})$ and $x_*=\inf\operatorname{supp}(X_{0*}\mathbb{P})$. Then items 1, 2, 2' and 3 below are equivalent.*

1. $\exists(a_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{>0}$, $(b_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{\geqslant 0}$, $\forall x\in\mathbb{R}\cap cont(G_{\xi,1,0})$:

$$\mathbb{P}(M_n\leqslant a_nx+b_n)\underset{n\to\infty}{\longrightarrow}G_{\xi,1,0}(x).$$

2. $\exists(a_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{>0}$, $(b_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{\geqslant 0}$, $\forall x\in\mathbb{R}\cap cont(G_{\xi,1,0})$:

$$n\mathbb{P}(X_0>a_nx+b_n)\underset{n\to\infty}{\longrightarrow}-\ln G_{\xi,1,0}(x).$$

2'. $\exists\sigma:(0,\infty)\to(0,\infty)$, $\forall x\in\mathbb{R}\cap cont(G_{\xi,1,0})$:

$$\mathbb{P}\left(\frac{X_0-t}{\sigma(t)}>x\;\middle|\;X_0>t\right)=\frac{1-F_{X_0}(t+\sigma(t)x)}{1-F_{X_0}(t)}\underset{t\nearrow x^*}{\longrightarrow}-\ln G_{\xi,1,0}(x),$$

or, equivalently,

$$\lim_{t\nearrow x^*}\sup_{0<x<x^*-t}\left|\mathbb{P}(X-t\leqslant x\mid X_0>t)-\left[1-\left(-\ln G_{\xi,\sigma(t),0}(x)\right)\right]\right|=0.$$

In 2', x^* is finite if, and only if, $\xi<0$. In 1 and 2', the normalizing objects harmonize as $\sigma(t)=\alpha\left(\frac{1}{1-F_{X_0}(t)}\right)$ (recall that $a:\mathbb{N}_{\geqslant 0}\to\mathbb{R}_{>0}$ lifts to $\alpha:\mathbb{R}_{\geqslant 0}\to\mathbb{R}_{>0}$, see definition 2.2.2 and theorem 2.2.3).

3. $\exists(a_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{>0}$, $(b_n)_{n\in\mathbb{N}_{n\geqslant 0}}\subset\mathbb{R}_{\geqslant 0}$:

$$\tilde{N}_n^{(2)}:=\sum_{i=0}^{n-1}\delta_{\left(\frac{i}{n},\frac{X_i-b_n}{a_n}\right)}\xrightarrow[n\to\infty]{d}\tilde{N}^{(2)},$$

where $\tilde{N}^{(2)}$ is a Poisson point process on $[0,1)\times(x_*,x^*)$ with intensity measure γ given by $\gamma((t_1,t_2]\times(x,x^*))=(t_2-t_1)\cdot-\ln G_{\xi,1,0}(x)$.

In item 2 and the first part of item 2' of the previous theorem, tail mass is being evaluated, so the limit is the tail mass of a distribution, known as generalized Pareto distribution (GPD), whose cdf is

$$P_{\xi,\sigma,\mu}(x)=1-(-\ln G_{\xi,\sigma,\mu}(x))=\begin{cases}1-\left[0\vee(1+\xi\frac{x-\mu}{\sigma})\right]^{-1/\xi}, & \text{if } \xi\neq 0\\ 1-e^{-\frac{x-\mu}{\sigma}} & \text{, if } \xi=0\end{cases}.$$

This family covers the following subfamilies:

i) when $\xi=0$, $P_{0,\sigma,0}$ is the cdf of an exponential distribution with intensity $1/\sigma$,

ii) when $\xi>0$, $P_{\xi,\sigma,\sigma/\xi}$ is the cdf of a Pareto distribution with scale parameter σ/ξ and shape parameter $1/\xi$,

iii) when $\xi=-1$, $P_{-1,\sigma,0}$ is the cdf of a uniform distribution on $[0,\sigma]$.

The third characterization in theorem 2.2.7 interests us the most, so we look at a proof sketch.

To consider (3 $\Rightarrow$ 1), notice simply that

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leqslant x\right) = \mathbb{P}\Big(\tilde{N}_n^{(2)}\left([0,1] \times (x,x^*)\right) = 0\Big)$$

$$\to \mathbb{P}\Big(\tilde{N}^{(2)}\left([0,1] \times (x,x^*)\right) = 0\Big) = e^{-\gamma([0,1]\times(x,x^*))} = e^{\ln G_{\xi,1,0}(x)} = G_{\xi,1,0}(x),$$

as desired.

Now we consider (1 $\Rightarrow$ 3) (see [26] sec. 7.3.1). Let $A = [0,1] \times (x,x^*)$. Then the probability that each of the points $\left(\frac{i}{n}, \frac{X_i - b_n}{a_n}\right)$ $(i = 0, \ldots, n-1)$ is in A is worked out as follows:

$$\mathbb{P}(M_n \leqslant a_n x + b_n) \to G_{\xi,1,0}(x) \Leftrightarrow \mathbb{P}(X_0 \leqslant a_n x + b_n)^n \to G_{\xi,1,0}(x)$$

$$\Leftrightarrow n \ln \mathbb{P}(X_0 \leqslant a_n x + b_n) \to \ln G_{\xi,1,0}(x) \Leftrightarrow n \ln\left(1 - \mathbb{P}(X_0 > a_n x + b_n)\right) \xrightarrow{n} \ln G_{\xi,1,0}(x),$$

whereas,

$$\ln\left(1 - \mathbb{P}(X_0 > a_n x + b_n)\right) = -\mathbb{P}(X_0 > a_n x + b_n) + o_{n\to\infty}(\mathbb{P}(X_0 > a_n x + b_n)),$$

so

$$n\mathbb{P}(X_0 > a_n x + b_n) - n o_{n\to\infty}(\mathbb{P}(X_0 > a_n x + b_n)) \xrightarrow{n} -\ln G_{\xi,1,0}(x)$$

$$\Rightarrow n\mathbb{P}(X_0 > a_n x + b_n)\left[1 - \frac{o_{n\to\infty}(\mathbb{P}(X_0 > a_n x + b_n))}{\mathbb{P}(X_0 > a_n x + b_n)}\right] \xrightarrow{n} -\ln G_{\xi,1,0}(x)$$

$$\Rightarrow n\mathbb{P}(X_0 > a_n x + b_n) \xrightarrow{n} -\ln G_{\xi,1,0}(x).$$

On the other hand, since the X_i's are independent,

$$(\tilde{N}_n^{(2)}(A))_*\mathbb{P} = Bin\left(n, \mathbb{P}\left(\frac{X_0 - b_n}{a_n} > x\right)\right),$$

whose mean is $n\mathbb{P}\left(\frac{X_0 - b_n}{a_n} > x\right) \xrightarrow{n} -\ln G_{\xi,1,0}(x)$, so, using theorem 2.1.5, one finds that

$$(\tilde{N}_n^{(2)}(A))_*\mathbb{P} \xrightarrow{v} Poi_{-\ln G_{\xi,1,0}(x)}.$$

Finally, because the first component of $\sum_{i=0}^{n-1} \delta_{\left(\frac{i}{n}, \frac{X_i - b_n}{a_n}\right)}$ is evenly spread in the line (and not random at all), we conclude that, when $A = (t_1, t_2] \times (x, x^*)$ with $t_1, t_2 \in [0,1], t_1 < t_2$:

$$(\tilde{N}_n^{(2)}(A))_*\mathbb{P} \xrightarrow{v} Poi_{(t_2 - t_1)\,\cdot\,-\ln G_{\xi,1,0}(x)},$$

fulfilling, as desired, condition (ii) in definition 2.1.17. The remaining conditions are left for the reader.

On an intuitive level, the point process $\tilde{N}_n^{(2)}$ spreads homogeneously over the unit-time-interval and, independently, stacks on top of it, over the (n-normalized) severity axis, the associated (n-normalized) realizations of the process. As n grows, a given (n-normalized) realization, say $i = 7$, stacks-over closer to $t = 0$. But, at the same time, the larger the n, the more realizations (i.e., the more i's) get to be concentrated on top of the unit-time-interval: how many of them will be stacking-over close to a certain time t is just as many as there are close to time $t=0$. In the limit, we get a point process $\tilde{N}^{(2)}$ in the time $\times$ (normalized) severity space that realizes countably many points which, vaguely speaking, seem to have independent coordinates, being them uniformly distributed in time, whereas, in the severity space, iid according to some distribution (the same on each time fiber). To put this vague description into perspective, we discuss more precisely the behavior of the limit, say, when $\xi \geqslant 0$ and $x_* = -\infty$.

If the limit point process $\tilde{N}^{(2)}$ is projected into the time coordinate, we do not even get a random measure in the sense of definition 2.1.14, because almost surely there are infinitely many points in the compact set $[0,1]$, since realizations in the product space themselves already have infinitely many points. Notice that the mass randomly assigned to a set of the type $(t_1,t_2]\times(x_*,x^*)$ in expectation equals $(t_2-t_1)\cdot -\ln G_{\xi,1,0}(x_*) = \infty$. However, if only the x-final-tail of the limit point process is projected into the time coordinate, then we get a Poisson point process $\tilde{N}^{(1,x)}$ with intensity $r(x)Leb_{[0,1]}$, where $r(x) = -\ln G_{\xi,1,0}(x)$. So, according to definition 2.1.22, $\tilde{N}^{(2)}$ is an (x_*,x^*)-lifted (vacuously compound) Poisson point process on the state space $[0,1]$ relative to $\{(x,x^*) : x \in (x_*,x^*)\}$ with intensity $x \mapsto -\ln G_{\xi,1,0}(x)Leb_{[0,1]}$ (and multiplicity $(\lambda_\ell)_{\ell\in\mathbb{N}_{\geqslant 1}}$ where $\lambda_1 = 1$).

Now that we know a suitable way to look at the limit process, we could take a step back and use it to look at the $\tilde{N}_n^{(2)}$'s. But everything would be similar, with, however, the terms approaching $-\ln G_{\xi,1,0}$ appearing — in which case the calculation carried out before suffices to aid one intuition.

The review of the classical independent theory is complete. Until the end of this section, we address some related concepts which turn out to be useful to equip the reader for the next section.

The Poisson point process $\tilde{N}^{(1,x)}$ introduced above (on $[0,1]$ with intensity $-\ln G_{\xi,1,0}(x)Leb_{[0,1]}$) can be characterized as the distributional limit of the following sequence of point processes on $[0,1]$:

$$\tilde{N}_n^{(1,x)} := \sum_{i=0}^{n-1} \delta_{\frac{i}{n}} \mathbb{1}_{(x,x^*)}\left((X_i-b_n)/a_n\right),$$

for reasons similar to those of the previous discussion.

Since this one is a homogeneous Poisson point process, we can interpret that its associated inter-arrival times are exponentially distributed with intensity $-\ln G_{\xi,1,0}(x)$.

One can define inter-arrival times associated with the $\tilde{N}_n^{(1,x)}$'s as given by

$$\begin{aligned}\tilde{IA}_j^{n,(1,x)} := \;& \tfrac{1}{n}\inf\left\{i \geqslant 0 : \#\left\{i' \in [0,i] : \tfrac{X_{i'}-b_n}{a_n} > x\right\} = j\right\} \\ & -\tfrac{1}{n}\inf\left\{i \geqslant 0 : \#\left\{i' \in [0,i] : \tfrac{X_{i'}-b_n}{a_n} > x\right\} = j-1\right\}\end{aligned} \quad (j \geqslant 1),$$

and each $(\tilde{IA}_j^{n,(1,x)})_{j\in\mathbb{N}_{\geqslant 1}}$ will be iid with underlying distribution converging vaguely to one of the previous paragraph, when $n \to \infty$.

The analysis developed in theorem 2.2.7 has a counterpart with respect to the alternative normalization introduced in proposition 2.2.4. We highlight what interests us the most.

Theorem 2.2.8 (Adapted Pickands, Balkema, de Haan, Resnick)**.** *Let $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ be a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega,\mathscr{F},\mathbb{P})$. Then items 1, 2 and 3 below are equivalent.*

1. $\exists u : (n, \tau) \in \mathbb{N}_{\geqslant 0} \times \mathbb{R}_{>0} \mapsto u_n(\tau) \in \mathbb{R}$ *continuous and strictly decreasing (increasing) in* τ *(in* n*) satisfying:*

$$\mathbb{P}(M_n \leqslant u_n(\tau)) \underset{n\to\infty}{\longrightarrow} e^{-\tau}, \forall \tau \in \mathbb{R}_{>0},$$

2. $\exists u : (n, \tau) \in \mathbb{N}_{\geqslant 0} \times \mathbb{R}_{>0} \mapsto u_n(\tau) \in \mathbb{R}$ *continuous and strictly decreasing (increasing) in* τ *(in* n*) satisfying: (see proposition 2.2.4),*

$$n\mathbb{P}(X_0 > u_n(\tau)) \underset{n\to\infty}{\longrightarrow} \tau, \forall \tau \in \mathbb{R}_{>0}.$$

3. $\exists u : (n, \tau) \in \mathbb{N}_{\geqslant 0} \times \mathbb{R}_{>0} \mapsto u_n(\tau) \in \mathbb{R}$ *continuous and strictly decreasing (increasing) in* τ *(in* n*) satisfying:*

$$N_n^{(2)} := \sum_{i=0}^{n-1} \delta_{\left(\frac{i}{n}, u_n^{-1}(X_i)\right)} \xrightarrow[n\to\infty]{d} N^{(2)}, \tag{2.6}$$

where $N^{(2)}$ *is a Poisson point process on* $[0,1] \times \mathbb{R}_{>0}$ *with intensity measure* $\gamma = Leb_{[0,1]} \times Leb_{\mathbb{R}_{>0}}$.

Similarly, if the τ-initial-tail of $N^{(2)}$ is projected into the time coordinate, we get $N^{(1,\tau)}$ a Poisson point process on $[0,1]$ with intensity $r(\tau)Leb_{[0,1]}$, where $r(\tau) = \tau$. Alternatively, $N^{(2)}$ is an $(0,\infty)$-lifted (vacuously compound) Poisson point process on the state space $[0,1]$ relative to $\{[0,\tau) : \tau \in \mathbb{R}_{>0}\}$ with intensity $\tau \mapsto \tau Leb_{[0,1]}$ (and multiplicity $(\lambda_\ell)_{\ell \in \mathbb{N}_{\geqslant 1}}$ where $\lambda_1 = 1$).

Once again, $N^{(1,\tau)}$ can be characterized as the distributional limit of

$$N_n^{(1,\tau)} := \sum_{i=0}^{n-1} \delta_{\frac{i}{n}} \mathbb{1}_{u_n^{-1}(X_i) < \tau}, \tag{2.7}$$

a point process on $[0,1]$.

Analogous comments can be made regarding inter-arrival times. Those of $N^{(1,\tau)}$ are exponential with intensity τ, and those of $N_n^{(1,\tau)}$ converge to the former when $n \to \infty$.

2.2.2 Hitting statistics

Our last topic in this review is hitting statistics, which are very much related to extreme value theory.

We still consider $(X_i)_{i\in\mathbb{N}_{\geqslant 0}}$ a sequence of iid $\mathbb{R}$-valued random variables on $(\Omega, \mathscr{F}, \mathbb{P})$. Consider that $X_{0*}\mathbb{P}$ is fully supported and has no atoms. Let $\Gamma \in \mathscr{B}_{\mathbb{R}}$ be the so-called target set, small in the sense that $(X_{0*}\mathbb{P})(\Gamma) = 0$ and denote $\Gamma_\rho = B_\rho(\Gamma)$. Let $U \in \mathscr{B}_{\mathbb{R}}$ (in practice, we will make it some neighborhood of Γ).

Let us define some working objects.

Definition 2.2.9. The **first hitting time**[1] of $(X_i)_{i\geqslant 0}$ into U is the function

$$\begin{array}{rccl} r_U^1 = r_U : & \Omega & \to & \mathbb{N}_{\geqslant 1} \cup \{\infty\} \\ & \omega & \mapsto & \inf\{i \in \mathbb{N}_{\geqslant 1} : X_i(\omega) \in U\} \end{array}.$$

[1] One could consider return time statistics, studying $r_U|_{X_0\in U}$. But in the iid case this adds nothing deep, since knowing that the starting condition of the process is in some set is irrelevant to what happens later.

The associated **higher-order hitting times** are given, for $\ell \geqslant 2$, by the function

$$\begin{array}{rcccl} r_U^\ell & : & \Omega & \to & \mathbb{N}_{\geqslant \ell} \cup \{\infty\} \\ & & \omega & \mapsto & \inf\{i > r_U^{\ell-1} : X_i(\omega) \in U\} \end{array}.$$

Definition 2.2.10. The **hit counting function** of $(X_i)_{i\geqslant 0}$, for $L \geqslant 1$ are given by

$$\begin{array}{rcccl} Z_{*U}^L & : & \Omega & \to & \mathbb{N}_{\geqslant 0} \\ & & \omega & \mapsto & \sum_{i=1}^{L} \mathbb{1}_U \circ X_i(\omega) \end{array}, \qquad \begin{array}{rcccl} Z_U^L & : & \Omega & \to & \mathbb{N}_{\geqslant 0} \\ & & \omega & \mapsto & \sum_{i=0}^{L-1} \mathbb{1}_U \circ X_i(\omega) \end{array}.$$

These objects are related, for example, in the sense that $\{Z_{*U}^L \geqslant \ell\} = \{r_U^\ell \leqslant L\}$, $\{Z_{*U}^L = \ell\} = \{r_U^\ell \leqslant L < r_U^{\ell+1}\}$.

In the previous section, we were interested in controlling the maximum or large values of the process, as given, respectively, by, say, items one and two of theorem 2.2.7 or 2.2.8. Now the extreme event of interest is the process hitting the target or shrinking vicinities of it. How small we consider these vicinities and/or how long we watch the process evolve is something to adjust. This adjustment is done with a balanced normalization in the spirit proposition 2.2.4 or theorem 2.2.8 (rather than the linear alternative of theorem 2.2.7), but this can be accomplished fine-tuning I) the space component, II) the time component, or III) both.

Using functions

$$\begin{array}{rccc} \rho^* : & \mathbb{N}_{\geqslant 1} \times \mathbb{R}_{>0} & \to & \mathbb{R}_{>0} \\ & (m,\tau) & \mapsto & \rho_m^*(\tau) \end{array}, \quad \begin{array}{rccc} L^* : & \mathbb{N}_{\geqslant 1} \times \mathbb{R}_{>0} & \to & \mathbb{N}_{\geqslant 1} \\ & (m,\tau) & \mapsto & L_m^*(\tau) \end{array}$$

one would like to harmonize how fast radii around the target shrink and how long the time we watch the process in order to obtain that asymptotically

$$L_m^*(\tau)\mathbb{P}(X_0 \in \Gamma_{\rho_m^*(\tau)}) \xrightarrow{m} \tau, \forall \tau \in \mathbb{R}_{>0}, \tag{2.8}$$

or even exactly

$$\begin{array}{rll} & L_m^*(\tau)\mathbb{P}(X_0 \in \Gamma_{\rho_m^*(\tau)}) \overset{\forall m}{\equiv} \tau & , \forall \tau \in \mathbb{R}_{>0} \\ \Leftrightarrow & L_m^*(\tau) \overset{\forall m}{\equiv} \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m^*(\tau)})} & , \forall \tau \in \mathbb{R}_{>0} \\ \Leftrightarrow & L_m^*(\tau) \overset{\forall m}{\equiv} \left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m^*(\tau)})} \right\rfloor & , \forall \tau \in \mathbb{R}_{>0}. \end{array} \tag{2.9}$$

The second condition is apparently more demanding, but actually, i) conditions that guarantee solutions to equation (2.8) exist usually guarantee that solutions to equation (2.9) also exist, and ii) solutions to equation (2.8) which are not exactly solutions to equation (2.9) will not modify the asymptotic statistical statements we will be after. We chose to build this section based on the exact version in equation (2.9). In future sections, if the asymptotic version is used instead, one can easily adapt the concepts we will present in the following.

Now let's look at the different normalization approaches.

I) Space-normalization

One may meet the balancing condition given in equation (2.9) by solving for ρ^* and so adjusting the speed at which space-radii shrink while observation time grows in the plain manner, like $L_m^*(\tau) = m$.

Presenting an explicit solution to $\rho_m^*(\tau)$ given $L_m^*(\tau) = m$ is not immediate, but also not important at this point. So we keep looking at equation (2.9) implicitly. Actually, to make a connection with extreme value theory, $\rho_m^*(\tau)$ is not solved directly, but in terms of $\rho_m^*(\tau) = g^{-1}(u_m(\tau))$, where

a)

$$g : [0,\infty) \to [0,\infty] \tag{2.10}$$

is a continuous function that is strictly decreasing near 0, where it attains a global maximum.

b)

$$\begin{array}{rccc} u: & \mathbb{N}_{\geqslant 1} \times \mathbb{R}_{>0} & \to & \mathbb{R} \\ & (m,\tau) & \mapsto & u_m(\tau) \end{array} \tag{2.11}$$

is a continuous function which is strictly decreasing in τ, strictly increasing in m to $g(0)$ (for any τ).

c)

$$m\mathbb{P}(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}) \overset{\forall m}{\equiv} \tau, \forall \tau \in \mathbb{R}_{>0}, [2] \tag{2.12}$$

where $\Gamma_{g^{-1}(u_m(\tau))} = B(\Gamma, g^{-1}(u_m(\tau))) = \{\varphi > u_m(\tau)\}$. Put differently, one can introduce the process $Y_i = \varphi \circ X_i = g \circ d(X_i, \Gamma)$, so a visit very close of Γ translates to a very high φ observation, and the above condition rewrites in the usual extreme value theory form as

$$m\mathbb{P}(Y_0 > u_m(\tau)) \overset{\forall m}{\equiv} \tau,\ \forall \tau \in \mathbb{R}_{>0}. [3] \tag{2.13}$$

The problem of finding a solution is then moved from fine-tuning $\rho_m^*(\tau)$ to fine-tuning $u_m(\tau)$. Here we could write an "explicit" solution as

$$u_m(\tau) = F_{Y_0}^{-1*}(1 - \tau/m)\ (\tau \in (0,1])$$

where $F_{Y_0}^{-1*}$ is the generalized inverse of the cumulative distribution function of Y_0, provided that, say, $Y_{0*}\mathbb{P}$ is fully supported and has no atoms.

As a consequence of such choice, not only equation (2.9) is verified, but also the following asymptotics:

$$\begin{aligned} \mathbb{P}(r^1_{\Gamma_{g^{-1}(u_m(\tau))}} > m) &= \mathbb{P}(Z^m_{*\Gamma_{g^{-1}(u_m(\tau))}} < 1) = \mathbb{P}(Z^m_{*\Gamma_{g^{-1}(u_m(\tau))}} = 0) \\ &= \mathbb{P}(X_1 \notin \Gamma_{g^{-1}(u_m(\tau))}) \cdot \ldots \cdot \mathbb{P}(X_m \notin \Gamma_{g^{-1}(u_m(\tau))}) \\ &= [1 - \mathbb{P}(X_0 \in \Gamma_{g^{-1}(u_m(\tau))})]^m \xrightarrow{m} e^{-\tau}, \end{aligned}$$

[2]The asymptotic counterpart of equation (2.12) reads as

$$m\mathbb{P}(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}) \xrightarrow{m} \tau, \forall \tau \in \mathbb{R}_{>0} \qquad (2.12').$$

[3]The asymptotic counterpart of equation (2.13) reads as

$$m\mathbb{P}(Y_0 > u_m(\tau)) \xrightarrow{m} \tau, \forall \tau \in \mathbb{R}_{>0} \qquad (2.13').$$

and, similarly,

$$\mathbb{P}(M_m \leqslant u_m(\tau)) = \mathbb{P}(Z^m_{\Gamma_{g^{-1}(u_m(\tau))}} = 0)$$
$$= \mathbb{P}(X_0 \notin \Gamma_{g^{-1}(u_m(\tau))}) \cdot \mathbb{P}(X_1 \notin \Gamma_{g^{-1}(u_m(\tau))}) \cdot \ldots \cdot \mathbb{P}(X_{m-1} \notin \Gamma_{g^{-1}(u_m(\tau))})$$
$$= [1 - \mathbb{P}(X_0 \in \Gamma_{g^{-1}(u_m(\tau))})]^m \xrightarrow{m} e^{-\tau}.$$

The space-normalization defined above induces the following point processes, called rare event point processes (REPP).

Definition 2.2.11. The **two-dimensional REPP with space-normalization** associated to the system $((X_i)_{i\geqslant 0}, \mathbb{P}, \Gamma)$ wrt (g,u) satisfying (a-c) (see equations (2.10-2.12)) is the sequence of point processes

$$N_m^{(2,I)} = {}^4 \sum_{i=0}^{n-1} \delta_{(\frac{i}{m}, u_m^{-1}(Y_i))} \text{ on } [0,1] \times (0,\infty).$$

We consider also the following variation with the same nomenclature

$$N_m^{2,I} = \sum_{i=0}^{\infty} \delta_{(\frac{i}{m}, u_m^{-1}(X_i))} \text{ on } [0,\infty) \times (0,\infty).$$

Definition 2.2.12. The **one-dimensional REPP with space-normalization** associated to the system $((X_i)_{i\geqslant 0}, \mathbb{P}, \Gamma)$ wrt (g,u) satisfying (a-c) (see equations (2.10-2.12)) at scale $\tau \in \mathbb{R}_{>0}$ is the sequence of point processes

$$N_m^{(1,I,\tau)} := {}^5 \sum_{i=0}^{n-1} \delta_{\frac{i}{m}} \mathbb{1}_{u_m^{-1}(Y_i) < \tau} = \sum_{i=0}^{n-1} \delta_{\frac{i}{m}} \mathbb{1}_{X_i \in \Gamma_{g^{-1}(u_m(\tau))}} \text{ on } [0,1].$$

We consider also the following variation with the same nomenclature

$$N_m^{1,I,\tau} := \sum_{i=0}^{\infty} \delta_{\frac{i}{m}} \mathbb{1}_{X_i \in \Gamma_{g^{-1}(u_m(\tau))}} \text{ on } [0,\infty).$$

It is immediate that $N_m^{1,I,\tau}$ coincides with the projection of $N_m^{2,I}\Big|_{[0,\infty)\times(0,\tau)}$ into the first coordinate. In particular, by the continuous mapping theorem, if a distributional limit is found for $N_m^{2,I}$, say $N_\infty^{2,I}$, then $\pi_{1*}\left[N_\infty^{2,I}\Big|_{[0,\infty)\times(0,\tau)}\right]$ is the distributional limit of $N_m^{1,I,\tau}$. Similarly for $N_m^{(1,I,\tau)}$ and $N_m^{(2,I)}\Big|_{[0,1]\times(0,\tau)}$.

II) Time-normalization

On the other hand, one may meet the balancing condition given in equation (2.9) by solving for L^* and so adjusting the time during which the process is observed while space-radii shrinks in a plain manner, like $\rho \to 0$.

[4]Coincides with $N_m^{(2)}$, introduced in equation (2.6), but with the old X_i substituted by the new Y_i.

[5]Coincides with $N_m^{(1,\tau)}$, introduced in equation (2.7), but with the old X_i substituted by the new Y_i.

Abusing notation, it turns out that equation (2.9) is fulfilled with

$$L_\rho^*(\tau) := \left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_\rho)} \right\rfloor, \; \rho^* := \rho$$

or, adhering to the notation presented in the subsequential terms used before, given a sequence $\rho_m \searrow 0$ as $m \to \infty$, it turns out that equation (2.9) is fulfilled with

$$L_m^*(\tau) := \left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor, \rho_m^*(\tau) := \rho_m.$$

As a consequence, not only equation (2.9) is verified, but also the following asymptotics:

$$\mathbb{P}(r^1_{\Gamma_{\rho_m}} > \left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor) = \mathbb{P}(Z_{*\Gamma_{\rho_m}}^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} < 1) = \mathbb{P}(Z_{*\Gamma_{\rho_m}}^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} = 0)$$

$$= \mathbb{P}(X_1 \notin \Gamma_{\rho_m}) \cdot \ldots \cdot \mathbb{P}(X_{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} \notin \Gamma_{\rho_m})$$

$$= [1 - \mathbb{P}(X_0 \in \Gamma_{\rho_m})]^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} \xrightarrow{m} e^{-\tau}, \forall \tau \in \mathbb{R}_{>0},$$

and, similarly,

$$\mathbb{P}(Z_{\Gamma_{\rho_m}}^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} = 0) = \mathbb{P}(X_0 \notin \Gamma_{\rho_m}) \cdot \mathbb{P}(X_1 \notin \Gamma_{\rho_m}) \cdot \ldots \cdot \mathbb{P}(X_{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor - 1} \notin \Gamma_{\rho_m})$$

$$= [1 - \mathbb{P}(X_0 \in \Gamma_{\rho_m})]^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{\rho_m})} \right\rfloor} \xrightarrow{m} e^{-\tau}, \forall \tau \in \mathbb{R}_{>0}.$$

These limits would occur similarly if considering $\rho \to 0$ instead of the subsequence $\rho_m \searrow 0$.

The time-normalization defined above induces the following REPPs.

Definition 2.2.13. The **one-dimensional REPP with time-normalization** associated to the system $((X_i)_{i \geqslant 0}, \mathbb{P}, \Gamma)$ wrt $(\rho_m)_{m \geqslant 1} \searrow 0$ at scale $\tau \in \mathbb{R}_{>0}$ is the sequence of point processes

$$N_m^{(1,II,\tau)} := \sum_{i=0}^{\lfloor \tau/\mathbb{P}(X_0 \in \Gamma_{\rho_m}) \rfloor - 1} \delta_{\frac{i}{\lfloor \tau/\mathbb{P}(X_0 \in \Gamma_{\rho_m}) \rfloor}} \mathbb{1}_{X_i \in \Gamma_{\rho_m}} \text{ on } [0,1].$$

We consider also the following variation with the same nomenclature

$$N_m^{1,II,\tau} := \sum_{i=0}^{\infty} \delta_{\frac{i}{\lfloor \tau/\mathbb{P}(X_0 \in \Gamma_{\rho_m}) \rfloor}} \mathbb{1}_{X_i \in \Gamma_{\rho_m}} \text{ on } [0,\infty).$$

III) Space-time normalization

Using the notation from the first two cases, one can also set

$$\rho_m^*(\tau) = g^{-1}(u_m(\tau))$$

where (a-c) are satisfied, together with

$$L_m^*(\tau) = \left\lfloor \frac{\tau}{\mathbb{P}(X_0 \in \Gamma_{g^{-1}(u_m(\tau))})} \right\rfloor.$$

This will recover case (I) where the general shrinking radii ρ_m happen to be the very special one $\rho_m(\tau) := \rho_m^*(\tau) = g^{-1}(u_m(\tau))$, therefore equation 2.9 is immediately verified. Also, the following asymptotics occur

$$\mathbb{P}\left(Z_{\Gamma_{g^{-1}(u_m(\tau))}}^{\left\lfloor \tau/\mathbb{P}\left(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}\right)\right\rfloor} = 0 \right) \sim \left(1 - \frac{\tau}{n}\right)^{\left\lfloor \frac{\tau}{\tau/n} \right\rfloor} \sim \left(1 - \frac{\tau}{n}\right)^n \xrightarrow{m} e^{-\tau}.$$

The space-time-normalization above induces the following REPPs.

Definition 2.2.14. The **one-dimensional REPP with space-time-normalization** associated to the system $((X_i)_{i \geqslant 0}, \mathbb{P}, \Gamma)$ wrt (g, u) satisfying (a-c) (see equations (2.10-2.12)) at scale $\tau \in \mathbb{R}_{>0}$ is the sequence of point processes

$$\tilde{N}_m^{(1,III,\tau)} := \sum_{i=0}^{\left\lfloor \tau/\mathbb{P}\left(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}\right)\right\rfloor - 1} \delta_{\frac{i}{\left\lfloor \tau/\mathbb{P}\left(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}\right)\right\rfloor}} \mathbb{1}_{X_i \in \Gamma_{g^{-1}(u_m(\tau))}} \text{ on } [0,1].$$

We consider also the following variation with the same nomenclature

$$\tilde{N}_m^{1,III,\tau} := \sum_{i=0}^{\infty} \delta_{\frac{i}{\left\lfloor \tau/\mathbb{P}\left(X_0 \in \Gamma_{g^{-1}(u_m(\tau))}\right)\right\rfloor}} \mathbb{1}_{X_i \in \Gamma_{g^{-1}(u_m(\tau))}} \text{ on } [0,\infty).$$

What we discussed in (I-III) is the standard paradigm in the study of hitting statistics. In principle, one could consider how the normalization used in theorem 2.2.7 would manifest in the present matter. As an exercise to connect as many dots as possible and clear out the picture, this is left to the footnote[6].

Now we inquire about higher-order statistics. For one of those suitably chosen normalization schemes $\rho_m^*(\tau)$ and $L_m^*(\tau)$, we want to evaluate

$$\lim_m \mathbb{P}(Z_{*\Gamma_{\rho_m^*(\tau)}}^{L_m^*(\tau)} < n) = \lim_m \mathbb{P}(r_{\Gamma_{\rho_m^*(\tau)}}^n > L_m^*(\tau)),$$

[6] For the space-normalization. Let $(a_m)_{m \in \mathbb{N}_{\geqslant 0}} \subset \mathbb{R}_{>0}, (b_m)_{m \in \mathbb{N}_{\geqslant 0}} \subset \mathbb{R}$ and $\xi \in \mathbb{R}$ be such that

$$m\mathbb{P}(X_0 \in \Gamma_{g^{-1}(a_m x + b_m)}) \xrightarrow{m} -\ln G_{\xi,1,0}(x),\ \forall x \in cont(G_{\xi,1,0}).$$

Then, with $\rho_m^*(x) := g^{-1}(a_m x + b_m)$ and $L_m^*(x) := m$, similar calculations lead us

$$\mathbb{P}(r_{\Gamma_{g^{-1}(a_m x + b_m)}} > m) \xrightarrow{m} G_{\xi,1,0}(x),\ \forall x \in cont(G_{\xi,1,0}).$$

For the time-normalization counterpart. Notice that we already had a linear normalization of time in (I). What we can then expect is some different normalization which reveals a more refined limit. This would be possible provided that one finds some $s : (m,x) \in \mathbb{N}_{\geqslant 1} \times \mathbb{R} \mapsto s_m(x) \in \mathbb{R}_{>0}$ and $\xi \in \mathbb{R}$ so that, taking $L_m^*(x) := s_m(x)$ and $\rho_m^*(x) := \rho_m$, one has

$$\mathbb{P}(r_{\Gamma_{\rho_m}} > s_m(x)) = (1 - \mathbb{P}(X_0 \in \Gamma_{\rho_m}))^{s_m(x)} \xrightarrow{m} G_{\xi,1,0}(x),\ \forall x \in cont(G_{\xi,1,0}).$$

which boils down to the evaluation of

$$\lim_m \mathbb{P}(Z_{\Gamma_{\rho_m^*(\tau)}}^{L_m^*(\tau)} = n).$$

This is not as easy to calculate directly as the first-order case we evaluated before. But adopting the space-normalization approach wrt (g,u), with $\rho_m^*(\tau) = g^{-1}(u_m(\tau))$ and $L_m^*(\tau) = m$, the condition given in equation (2.13) allows us to apply theorem 2.2.8, whose third item will assist us:

$$\mathbb{P}\left(\tilde{N}_m^{(2,I)}([0,1]\times(0,\tau)) = n\right) \xrightarrow{m} \mathbb{P}\left(\tilde{N}^{(2,I)}([0,1]\times(0,\tau)) = n\right)$$

$$\Rightarrow \mathbb{P}\left(N_m^{(1,I,\tau)}([0,1]) = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}) \Rightarrow \mathbb{P}\left(Z_{\Gamma_{g^{-1}(u_m(\tau))}}^m = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}).$$

Although we do not present further justification, it is then natural to expect that, for the time-normalization approach wrt $(\rho_m) \searrow 0$,

$$\mathbb{P}\left(N_m^{(1,II,\tau)}([0,1]) = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}) \Rightarrow \mathbb{P}\left(Z_{\Gamma_{\rho_m}}^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0\in\Gamma_{\rho_m})}\right\rfloor} = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}),$$

whereas, for the space-time-normalization,

$$\mathbb{P}\left(N_m^{(1,III,\tau)}([0,1]) = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}) \Rightarrow \mathbb{P}\left(Z_{\Gamma_{g^{-1}(u_m(\tau))}}^{\left\lfloor \frac{\tau}{\mathbb{P}(X_0\in\Gamma_{g^{-1}(u_m(\tau))})}\right\rfloor} = n\right) \xrightarrow{m} \mathrm{Poi}_\tau(\{n\}).$$

Chapter 3

State of the art: Compound Poisson distributions for dynamical systems

3.1 Deterministic systems

In this section, we review briefly some takeaways of the extreme value theory and hitting statistics for stochastic processes arising from deterministic dynamical systems, with a special bias toward results in the compound Poisson class. Since this literature can become very intricate, we do not intend to be exhaustive or even to cover many details right now, but to present some representative results and concepts in this literature.

In the context of dynamically defined processes, stationarity is kept while independence is lost. However, weaker versions of independence are still available, usually depending on the type of decay of correlations presented by the dynamical system.

An important takeaway is that the Poisson statistics that appeared in sections 2.2.1 and 2.2.2[1] are consequences of independence. When independence is not present, different limiting behaviors might occur, provided that this dependence produces clusters of visits to regions of interest[2]. If this is not the case, we again observe Poisson statistics. On the other hand, if that is the case (e.g. when such a region consists of a periodic point), we have a mixture of two behaviors: i) eventually the orbit visits the region of interest and then repeated visits occur due to the clustering effect produced by the local dynamics around that locus; ii) eventually the orbit escapes the local dynamics and undergoes an excursion around the rest of the phase space, mostly driven by weak independence of the system, so that after an exponential time (i) re-occurs. In the asymptotic limit, this description results in a compound Poisson distribution, where the number of Poissonian events is related to (ii), and their associated multiplicity distribution is related to (i).

For the rest of this section, we consider the following general setup.

Let M be a compact metric space equipped with its Borel σ-algebra $\mathscr{B}_M$ and $T : M \to M$ be a measurable transformation which leaves invariant $\mu \in \mathcal{P}(M)$, an atomless and fully supported probability. Let $\Gamma \in \mathscr{B}_M$ be a target set, small in the sense that $\mu(\Gamma) = 0$ and denote $\Gamma_\rho = B_\rho(\Gamma)$. Let $U \in \mathscr{B}_M$ (in practice, we will make it some neighborhood of Γ).

[1] Recall that $e^{-\tau} = \mathrm{Poi}_\tau(\{0\})$ and $e^{-t} = \mathrm{Poi}_t(\{0\})$.

[2] Namely, a target set or the set of points maximizing an observable.

We induce the dynamically defined stochastic process

$$X_i := T^i : (M, \mathscr{B}_M, \mu) \to (M, \mathscr{B}_M) \ (i \geqslant 0).$$

Consider that there exists (g,u) satisfying the conditions given in equations (2.10-2.12), to be used when we pursue space or space-time normalization. Then we define $\varphi(\cdot) = g \circ d(\cdot, \Gamma)$ and the associated observed stochastic process

$$Y_i := \varphi \circ X_i = \varphi \circ T^i : (M, \mathscr{B}_M, \mu) \to (\mathbb{R}, \mathscr{B}_\mathbb{R}) \ (i \geqslant 0).$$

Here we will use mutatis mutandis the definitions presented in section 2.2.2. Note that $\mu(Z_U^L = n) = \mu(Z_{*U}^L = n)$, by invariance.

Let us introduce the most basic object one comes across in this context.

Definition 3.1.1. The **extremal index** of (T, μ, Γ) is a number $\alpha \in [0,1]$ satisfying one of the following conditions:

I)

$$\lim_{m \to \infty} \mu(M_m \leqslant u_m(\tau)) = e^{-\alpha\tau}, \forall \tau > 0;$$

II)

$$\lim_{\rho \to 0} \mu(r^1_{\Gamma_\rho} > \lfloor t/\mu(\Gamma_\rho) \rfloor) = e^{-\alpha t}, \forall \tau > 0;$$

III)

$$\lim_{m \to \infty} \mu(r^1_{\Gamma_{g^{-1}(u_m(\tau))}} > \lfloor t/\mu(\Gamma_{g^{-1}(u_m(\tau))}) \rfloor) = e^{-\alpha t}, \forall \tau > 0.$$

The pair (g,u) being used in items (I) and (III) above is anyone satisfying the conditions given in equations (2.10), (2.11), (2.12), either in exact version or its m-asymptotic counterpart. Implicitly, it is considered unimportant the specific choice of such (g,u) among the eligible ones. Moreover, it is also implicitly considered that (I), (II), and (III) are equivalent. We do not pursue such characterizations at this level of generality. See [39] and [69] for results in this direction.

On an intuitive level, the extremal index identifies how much clustering is produced by the system near the target, or how much clustering of extreme values is produced by the system: if $\alpha = 1$ there is no clustering effect (as in the independent case), if $\alpha = 0$ the clustering effect is huge. Put in another way, α measures the proportion of points in the germ around the target that leave a cluster of recurrence to the same germ right after the first dynamical iteration (i.e., they are the last extreme observation in their cluster), thus breaking free from the local behavior and initiating an excursion.

3.1.1 Point processes

The approach we are going to review in this section follows from that introduced by [67], [57] and [73] for stationary stochastic processes, using variants of the conditions known as $D(u)$ and $D'(u)$, which were improved and adapted to the realm of dynamically defined processes, as to pursue, both in the absence and presence of clustering, the description of rare event point processes (REPPs).

The REPPs considered in this section always involve space-normalization (i.e., of type I or III presented in section 2.2.2) but, instead of the balancing condition given by equations (2.12) and (2.13), one adopts the slightly more general m-asymptotic counterpart.

We use [34] as a reference point to discuss this section while bringing results from other papers. In [34] they prove the convergence of one-dimensional, two-dimensional and what they call multi-dimensional REPPs.

Since two-dimensional REPPs already carry (more than) enough information to make a transparent connection with the compound Poisson processes we are after, we will not delve into more informed REPPs, such as their multi-dimensional REPPs. We mention briefly that, as compared to two-dimensional REPPs, their second component does not simply register the rarity of an instance (as measured in terms of the $u_m^{-1}(Y_i)$'s) but it actually registers where exactly (near a target point) such instance landed, in terms of associated tangent space. A similar attitude can be found in [74].

We start with one-dimensional REPPs.

In [41] and [12], a periodic-aperiodic dichotomy for $N_m^{1,I,\tau}$ was established, under the conditions of singleton targets, one-dimensional expanding maps, and regular invariant measures which are absolutely continuous with respect to Lebesgue. In the periodic case, with a p-periodic point ζ comprising the target, it was shown that $N_m^{1,I,\tau}$ converges to a compound Poisson point process on $[0,\infty)$ with intensity $\alpha\tau\mathrm{Leb}$ and multiplicity $\lambda_\ell = \alpha(1-\alpha)^{\ell-1}$ ($\ell \geqslant 1$), where $\alpha = 1 - 1/JT^p(\zeta)$. In the aperiodic case, it is shown that $N_m^{1,I,\tau}$ converges to a Poisson point process on $[0,\infty)$ with intensity $\tau\mathrm{Leb}$, which corresponds/extends the previous case as if α was 1.

Notice that whenever $N_m^{1,I,\tau}$ converges to a compound Poisson process on $[0,\infty)$ with intensity $r(\tau) = \alpha\tau\mathrm{Leb}$ and multiplicity $(\lambda_\ell(\tau))_{\ell\in\mathbb{N}_{\geqslant 1}}$, it follows that

$$\mu(M_m \leqslant u_m(\tau)) = \mu(N_m^{1,I,\tau}([0,1]) = 0) \xrightarrow{m} e^{-r(\tau)} = e^{-a\tau},$$

meaning that α is the extremal index of the associated system. Similar observations hold for $N_m^{1,II,\tau}$ and $N_m^{1,III,\tau}$. Under the same assumption, there is still no general relationship between the extremal index α and the multiplicity distribution $(\lambda_\ell(\tau))_{\ell\in\mathbb{N}_{\geqslant 1}}$, even if $(\lambda_\ell(\tau))_{\ell\in\mathbb{N}_{\geqslant 1}} \overset{\forall\tau>0}{=} (\lambda_\ell)_{\ell\in\mathbb{N}_{\geqslant 1}}$. However, in most cases, especially when targets do not overlap with the parabolic locus of non-uniformly hyperbolic maps, a relationship holds: $\alpha = \left(\sum_{\ell\geqslant 1} \ell\lambda_\ell\right)^{-1}$, i.e., the extremal index is the inverse of the mean multiplicity (also known as mean cluster size). This relationship can fail to hold even when the $\alpha > 0$ and $\sum_{\ell\geqslant 1} \ell\lambda_\ell < \infty$, as we discuss at the end of this section. Before this discussion, the reader need not to be preoccupied and might consider that the expected relationship holds.

By the previous results from [41] and [12], one has that limits of $N_m^{2,II}$ are expected to be $(0,\infty)$-lifted compound Poisson processes on the state space $[0,\infty)$ relative to $\{(0,\tau) : \tau \in \mathbb{R}_{>0}\}$ with intensity $\tau \in \mathbb{R}_{>0} \mapsto \alpha\tau\mathrm{Leb}$ and multiplicity $\tau \in \mathbb{R}_{>0} \mapsto \lambda_\ell = \alpha(1-\alpha)^{\ell-1}$ ($\ell \geqslant 1$), where $\alpha = 1 - 1/JT^p(\zeta)$ when ζ is p-periodic and $\alpha = 1$ otherwise. This is indeed the case, as we will see in the next result, theorem 3.1.2, which provides a finer description of these limits.

Theorem 3.1.2. *[[34], thm 4.3] Let M be a compact Riemannian manifold with Lebesgue measure,* Leb*, and let (T,μ,Γ) be a system as introduced in section 3.1 which is given a normalization (g,u) satisfying the conditions given in equations (2.10), (2.11) and (2.12'). Consider that μ is absolutely continuous with respect to* Leb *with*

$$\lim_{\varepsilon\to 0} \frac{\mu(B_\varepsilon(x))}{\mathrm{Leb}(B_\varepsilon(x))} = \frac{d\mu}{d\mathrm{Leb}}(x) \ (\forall x \in M). \tag{3.1}$$

Consider $\Gamma = \{\zeta\}$, where ζ is a hyperbolic repelling point in whose orbit T is continuous.

Assume that decay of correlation is rich enough in the sense that:

A) There exists a Banach space $\mathcal{C}_1$ of real-valued measurable functions defined on M, containing $\mathbb{1}_B$, for all $B \in \mathscr{B}_M$ and if $(B_n)_{n\geqslant 0} \subset \mathscr{B}_M$ is such that there exists a uniform bound for the number of connected components of all B_n's, then there exists $C > 0$ such that $\|\mathbb{1}_{B_n}\|_{\mathcal{C}_1} \leqslant C$ $(n \geqslant 0)$;

B) For all $\phi \in \mathcal{C}_1, \psi \in L^1(\mu)$ and $n \geqslant 1$, one has:

$$\mathrm{Cor}_\mu(\phi, \psi, n) = \frac{1}{\|\phi\|_{\mathcal{C}_1} \|\psi\|_{L^1(\mu)}} \left| \int_M \phi(\psi \circ T^n) d\mu - \int_M \phi d\mu \int_M \psi d\mu \right| \leqslant c_n,$$

with $\sum_{n\geqslant 1} c_n < \infty$.

Then:

I) If ζ is aperiodic, then $N_m^{2,I} \xrightarrow{d} N_{\infty,\mathrm{aper}}^{2,I}$, where the latter is a simple point process on $[0,\infty) \times (0,\infty)$ given by

$$N_{\infty,\mathrm{aper}}^{2,I} = \sum_{i,j=1}^{\infty} \delta_{(T_{i,j},U_{i,j})},$$

where $T_{i,j} = \sum_{l=1}^{j} \bar{T}_{i,l}$, $(\bar{T}_{i,j})_{i,j}$ a matrix of iid random variables distributed like Exp_1 and $(U_{i,j})_{i,j}$ is another matrix of iid random variables distributed like $U_{i,j} \sim \mathrm{Unif}(i-1,1]$ and so that $(\bar{T}_{i,j})_{i,j} \perp (U_{i,j})_{i,j}$.

II) If ζ is p-periodic and $DT^p{}_\zeta(v) = \vartheta v \in T_\zeta M$ ($|\vartheta| > 1$, $\forall v \in T_\zeta M$), then $N_m^{2,I} \xrightarrow{d} N_{\infty,\mathrm{per}}^{2,II}$, where the latter is a simple point process on $[0,\infty) \times (0,\infty)$ given by

$$N_{\infty,\mathrm{per}}^{2,II} = \sum_{i,j=1}^{\infty} \sum_{l=0}^{\infty} \delta_{(T_{i,j},\vartheta^{l\dim(M)} U_{i,j})},$$

where $T_{i,j} = \sum_{l=1}^{j} \bar{T}_{i,l}$, $(\bar{T}_{i,j})_{i,j}$ a matrix of iid random variables distributed like Exp_α, $\alpha = 1 - \vartheta^{-\dim(M)}$, and $(U_{i,j})_{i,j}$ is another matrix of iid random variables distributed like $U_{i,j} \sim \mathrm{Unif}(i-1,1]$ and so that $(\bar{T}_{i,j})_{i,j} \perp (U_{i,j})_{i,j}$.

As a consequence:

i) If ζ is aperiodic, then $N_m^{1,I} \xrightarrow{d} N_{\infty,\mathrm{aper}}^{1,I}$, where the latter is a Poisson point process on $[0,\infty)$ with intensity $1\,\mathrm{Leb}$.

ii) If ζ is p-periodic and $DT^p{}_\zeta(v) = \vartheta v \in T_\zeta M$ ($|\vartheta| > 1$, $\forall v \in T_\zeta M$), then $N_m^{1,II} \xrightarrow{d} N_{\infty,\mathrm{per}}^{1,II}$, where the latter is a compound Poisson process on $[0,\infty)$ with intensity $\tau\alpha\,\mathrm{Leb}$ and multiplicity $\lambda_\ell = \alpha(1-\alpha)^{\ell-1}$ ($\ell \geqslant 1$) (recall definition 2.1.21), where $\alpha = 1 - 1/\vartheta$.

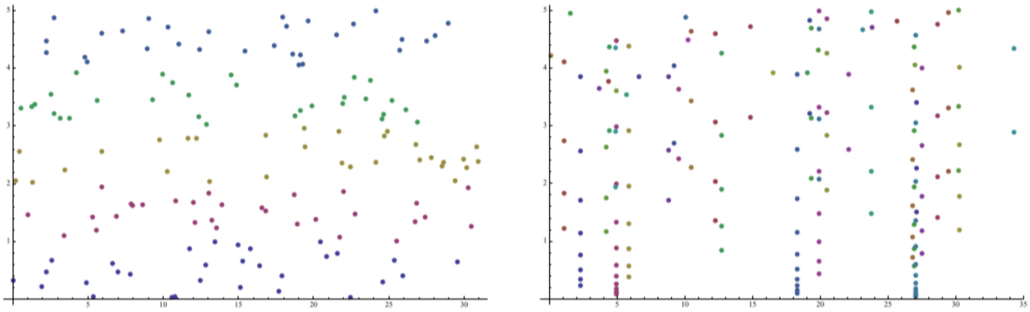


Figure 3.1 Left: Simulation of the limiting process in I). Right: Simulation of the limiting process in II) with $\dim(M) = 1$ and $\vartheta = 3/2$. Source [34]: Freitas, Freitas, Magalhães (2020) (Figure 1).

Intuitively, $N^{2,I}_{\infty,\text{aper}}$ splits the plane $[0,\infty) \times (0,\infty)$ into horizontal stripes of unit size in each of which countably many points realize, with their first coordinates having exponential increments, thereby like a Poisson point process with intensity 1, and their second coordinate, independently, drawn uniformly in the interval prescribing such stripe. On the other hand, $N^{2,I}_{\infty,\text{per}}$ has a similar interpretation, but each point realized as before has countably many "children" pilling above it, i.e., points whose first coordinate coincides with that of its "father" and whose second coordinate scale that of its "father" by successive powers of $\vartheta^{\dim(M)}$.

Right now we are more interested in the main statements and how their objects fit together. But it is worthwhile to consider a rough outline of the proof strategy adopted in [34]:

i) Show that rich enough decay of correlations implies that $(X_i)_{i\geqslant 0}$ satisfies conditions, $\text{Д}^*_q(u)$ and $\text{Д}'_q(u)$, where q controls the size of clusters and in the aperiodic case takes value 0, while in the p-periodic case is taken to be a certain multiple of p.

These conditions, which change slightly throughout the literature, are dynamical versions of the aforementioned $D(u)$ and $D'(u)$. The first one has to do with mixing and is used to show that the occurrence of escapes in a certain cluster is nearly independent from that of next cluster. The second condition is more delicate to check and relates to the short recurrence properties of small vicinities of the target, allowing one to access what time scale makes clusters coherent and what is the importance of those points that escape the cluster.

ii) The conditions obtained in (i) are used in an abstract theorem (their theorem 3.3) which aids an application of Kallenberg's criterion, leading to the desired convergence[3].

At the end of this section, after reviewing some advances in the literature, we look not at actual proofs but at new objects, concepts and formulas that are relevant to this approach. In particular, we address what are clusters and their sizes, which were referred to above but not yet clearly stated.

In [34], they still adapt the technique behind the proof of theorem 3.1.2 to handle situations not covered by the theorem but still in the realm of their techniques. In particular, incorporating ideas from [12] (props. 3.4 and 3.5), they obtain the convergence of two-dimensional REPPs when ζ is a (a)periodic discontinuity point for a piecewise expanding map of the interval (Rychlik map) with Lebesgue measure. In this exact situation, [12] proved the convergence of one-dimensional REPPs (of the type $N^{1,III,\tau}_m$) to compound Poisson distributions whose multiplicity distribution could be different

[3] Notice that simplicity, needed for the application of Kallenberg's criterion, holds for two-dimensional REPPs, but not one-dimensional ones.

from geometric, depending on the combinatorics the ζ. These different distributions are explicitly presented in [12].

In [42], under similar hypotheses to those of theorem 3.1.2, but not restricted to invariant measures absolutely continuous with respect to Lebesgue, and admitting equilibrium states μ_ϕ associated to a Holder potential ϕ of finite topological pressure whose density with respect to the associated conformal measure η_ϕ still satisfies a regularity condition similar to that of equation 3.1 (with Leb substituted by η_ϕ), authors have shown, again for singleton targets, that

$$\mu(M_m \leqslant u_m(\tau)) \xrightarrow{m} e^{-\alpha\tau},$$

where:

$$\alpha = 1 - e^{S_p\phi(\zeta)-pP(\phi)} \text{ if } \zeta \text{ is } p\text{-periodic; 1, otherwise.}$$

We notice that REPPs and compound Poisson statistics were not evaluated in [42]. However, they are available in [12], with similar explicit formulas, for one-dimensional REPPs (of the type $N_m^{1,III,\tau}$) of (dis)continuous (a)periodic targets for the aforementioned Rychlik maps (see [12] prop 3.2 and 3.5).

In the uniformly hyperbolic case with contracting directions, the items (II) and (ii) of theorem 3.1.2 will not apply. However, for linear Anosov automorphisms of the 2-torus equipped with Lebesgue measure, the periodic-aperiodic dichotomy for one-dimensional REPPs (of the type $N_m^{1,III,\tau}$) was shown in [24]. Their strategy was also based on conditions of the type $Д_q^*(u)$ and $Д_q'(u)$, but since Kallenberg's criterion is not available at the one-dimensional level, the approach introduced in [41] was applied. They also point out that limits are dependent on the chosen metric, and, for Euclidean and maximum metrics, explicit formulas are given for the limiting intensity and multiplicity distribution. Again everything depends on the expansion in the unstable direction and the period p, but it is for the maximum metric that we 'usual' formulas are recovered: $\alpha = 1 - 1/|\gamma|^p$ and $\lambda_\ell = \alpha(1-\alpha)^{l-1}$. Interestingly, the λ_ℓ's that appear with the Euclidean metric do not form a geometric distribution.

Theorem 3.1.2 also finds applications with non-uniformly expanding maps with indifferent fixed points such as the Manneville-Pommeau or Liverani-Saussol-Vaienti (LSV) maps, provided that ζ is not itself an indifferent fixed point and avoids the countable set of dynamic discontinuities. These applications rely on the fact that their first return maps are uniformly expanding.

To account for ζ the indifferent fixed point of an LSV map, T_β, with $\beta \in (0, \sqrt{5}-2)$, preserving the absolutely continuous probability μ_β, [44] first notes that taking $u_m(\tau)$'s in usual way,

$$m\mu_\beta(X_0 > u_m(\tau)) \xrightarrow{m} \tau, \text{ or, simply, } m\mu_\beta([0, g^{-1}(u_m(\tau)))) \xrightarrow{m} \tau,$$

implies a degenerate the extreme value law $\mu_\beta(M_m \leqslant u_m(\tau)) \xrightarrow{m} 1$ $(\forall \tau > 0)$, which means that the extremal index α is 0, or, on the other hand, that the $u_m(\tau)$'s grow too fast. To find non-degenerate asymptotics they fix the thresholds requiring that

$$m\mu_\beta\left(\left[T_{\beta,\leftarrow}{}^{-1}[g^{-1}(u_m(\tau))], g^{-1}(u_m(\tau))\right)\right) \xrightarrow{m} \tau,$$

where $T_{\beta,\leftarrow}$ is the left-branch of T_β. Under the new scaling, they find that $\mu_\beta(M_m \leqslant u_m(\tau)) \xrightarrow{m} e^{-\tau}$. The time-normalization counterpart suggested from their solution seems to be

$$L_m^*(\tau) = \left\lfloor \tau/\mu_\beta\left(\left[T_{\beta,\leftarrow}{}^{-1}[g^{-1}(u_m(\tau))], g^{-1}(u_m(\tau))\right)\right)\right\rfloor$$

instead of $L_m^*(\tau) = \lfloor \tau/\mu_\beta([0,g^{-1}(u_m(\tau)))\rfloor$. Curiously, REPPs under the new scaling still present degenerate limits.

When it comes to more general target sets, not limited to singletons, one naturally expects general compound Poisson statistics, precisely because more complicated recurring behavior around the target set can be cooked up. This can be seen most simply in the case of finite targets with pieces of orbits, as evaluated in [13] for one-dimensional REPPs of the type $N_m^{(1,III,\tau)}$. In this direction, see also [56]. More complicated targets, such as countable sets (see [14]), manifolds (see [33] and [23]), and fractal sets (see [37], [38] and [71]), were studied from the perspective of extreme value laws, but not establishing limits for the associated REPPs.

Still using point processes, a more abstract approach to such general target sets was developed in [43], where authors, inspired by ideas in [19], enriched their point processes as much as possible (much more than the aforementioned multidimensional REPPs). This allows them to obtain functional limit theorems for heavy-tailed functions g around general targets that do not overlap the parabolic locus of non-uniformly hyperbolic maps. Another outcome ([43] thm 4.1) is the convergence of their enriched point processes, valid for g's as general as those presented in section 3.1. Their τ-projection into the first coordinate recovers compound Poissonian limits for one-dimension REPPs, whose parameters, however, are not easy to present explicitly for general targets — but are expected to comply with those of [53], whose multiplicity distribution is given by an asymptotic expression (in terms of the T, μ and Γ) denoted by λ_ℓ ($\ell \geqslant 1$) and assumed to exist. Check equation (4.2) to see how the quantities $(\lambda_\ell)_\ell$ appear in the random case according to the approach of chapter 4.

On the other hand, we note that the main hypothesis of theorem 4.1 in [43] is the existence of the so-called pilling process, which is a more abstract (or enriched) counterpart to the assumption of the existence of λ_ℓ's in [53].

We close this section following [3] and [2] to discuss the relationship between the extremal index and the mean cluster size, once a compound Poisson point process is found to be the limit of a one-dimensional REPP. We take the opportunity to clear out some concepts previously invoked in a vague way and introduce some formulas that are relevant to the aforementioned literature but also relevant for comparisons with the alternative approaches discussed in the next sections.

Consider a parameter $q \in \mathbb{N}_{\geqslant 0}$. It will represent the maximum waiting time allowed between consecutive hits to be considered part of the same cluster.

Let us introduce a couple of working sets.

For $n \geqslant 2$, let

$$D_q^{n-1}(U) = \{x \in M : x \in U, \tau_U^1(x) \leqslant q, \tau_U^2(x) - \tau_U^1(x) \leqslant q, \ldots, \tau_U^{n-1}(x) - \tau_U^{n-2}(x) \leqslant q\}$$

be the set of points in U whose q-forward-cluster has at least n entries, and

$$Q_q^{n-1}(U) = \{x \in M : x \in D_q^{n-1}, \tau_U^n(x) - \tau_U^{n-1}(x) > q\}$$

be the set of points in U whose q-forward-cluster has exactly n entries. These definitions are naturally complemented for $n = 1$ with

$$D_q^0(U) = U, \quad Q_q^0(U) = \{x \in M : x \in U, \tau_U^1(x) > q\},$$

which extend the previous interpretations to $n = 1$.

Let also

$$D_q^\infty(U) = \bigcap_{n \geqslant 0} D_q^n(U)$$

be the set of points in U followed by infinitely many visits to itself, all of which at most q units of time apart from its adjacent visits.

Moreover, for $n \geqslant 1$, let

$$H_q^n(U) = \left\{x \in M : x \notin U, \tau_U^1(x) = q, T^q(x) \in Q_q^{n-1}(U)\right\}$$

be the set of points in M which, after exactly q units of time (from iterates 0 to $q-1$) failing to hit U, start a q-cluster that has exactly n entries. To complement this definition with $n = 0$, we put

$$H_q^0(U) = \{x \in M : x \notin U, \tau_U^1(x) = q\},$$

which is the set of points in M which, after exactly q units of time (from iterates 0 to $q-1$) failing to hit U, start a cluster (of any maximum waiting time and size). Note that $H_q^n(U) \subset H_q^0(U)$, for all $n \geqslant 1$.

The following theorems present the basic relationships among the objects above.

Theorem 3.1.3. *[[3] thm 2.1] Consider $n \geqslant 1$ and μ fully supported. Then*

$$\frac{\mu(Q_q^{n-1}(U)) - \mu(Q_q^n(U))}{\mu(Q_q^0(U))} = \frac{\mu(H_q^n(U))}{\mu(H_q^0(U))}.$$

Denote the coinciding quantity in theorem 3.1.3 by $\lambda_q^U(n)$. It represents the finite-time q-cluster size distribution. Finite time is in the sense that U is a frozen neighborhood of Γ.

Theorem 3.1.4. *[[3] thm 2.3] If $\mu(D_q^\infty(U)) = 0$ (which holds when μ is ergodic and fully supported), then*

$$\sum_{n \geqslant 1} n\lambda_q^U(n) = \frac{\mu(D_q^0(U))}{\mu(Q_q^0(U))}.$$

The quantity $\frac{\mu(Q_q^0(U))}{\mu(D_q^0(U))}$ represents the finite-time q-extremal index, since it registers the portion of points in U that terminate their q-clusters, thereby scaping from it. Denote $\frac{\mu(Q_q^0(U))}{\mu(D_q^0(U))} =: \alpha_q^U$. Therefore $\sum_{n \geqslant 1} n\lambda_q^U(n) = 1/\alpha_q^U$.

We pass $U := U_{m,\tau} = \Gamma_{g^{-1}(u_m(\tau))}$ to the objects previously defined and, to make their notation lighter, we substitute "$U_{m,\tau}$" by simply "m, τ".

The following theorem provides general conditions for one-dimensional REPPs to converge and explains how the finite-time objects relate to the quantities of the limit.

Theorem 3.1.5. *[[3] thm 2.5, [41], [35]] Consider a system (T,μ,Γ,g,u) and $q \in \mathbb{N}_{\geqslant 0}$ so that $(X_i)_{i\geqslant 0}$ satisfies conditions $Д_q(u)^*$ and $Д'_q(u)^*$. Assume that $\alpha_q^{m,\tau} \xrightarrow{m} \alpha_q$ ($\forall \tau > 0$) and that, for each $n \geqslant 1$, $\lambda_q^{m,\tau}(n) \xrightarrow{m} \lambda_q(n)$ ($\forall \tau > 0$).*

Then, for all $\tau > 0$, $N_m^{1,I,\tau}$ converges in distribution as $m \to \infty$ to a compound Poisson process on $[0,\infty)$ with intensity $\alpha_q \tau \mathrm{Leb}$ and multiplicity $(\lambda_q(n))_{n\geqslant 1}$.

Adopting the hypotheses of the theorem except for the latter convergence, one still concludes that α_q is the extremal index of the system (see [40] and [42]).

Here we will not discuss further what $Д_q(u)^*$ and $Д'_q(u)^*$ are. It suffices to recall remark (i) after theorem 3.1.2 and be aware that there is appropriate q verifying them in all the systems discussed in this subsection. Similar thing can be said about the remaining hypotheses. We say that q is eligible for the system if $Д_q(u)^*$ and $Д'_q(u)^*$ are satisfied. If q is eligible, any $q' \geqslant q$ will also be, so we naturally search for the minimal eligible one, to be denoted by $q_\star$. It is a subtle matter to find the $q_\star$ of a system (T,μ,Γ,g,u), but some candidates standout, in special because, when $\Gamma = \{\zeta\}$, they recover the period of ζ (zero) if ζ is periodic (aperiodic):

a) (see [2])

$$q'_\star = \inf_{\tau>0} \liminf_{m\to\infty} p'(U_{m,\tau}) \text{ if finite, otherwise } q'_\star = 0,$$

where

$$p'(U) = \inf\left\{k \geqslant 1 \;\middle|\; \mu\left(\{x \in M : x \in U, T^k x \in U\}\right) > 0\right\};$$

b) (see [13])

$$q''_\star = \inf_{\tau>0} \inf\left\{k \geqslant 0 : \lim_{m\to\infty} \inf_{x \in Q_k^0(m,\tau)} r_{Q_k^0(m,\tau)}(x) = \infty\right\}.$$

Also notice that, since the limit in the conclusion is actually independent of q, so any eligible q produces the same $\alpha_q \equiv \alpha$ and $(\lambda_q(n))_{n\geqslant 1} \equiv (\lambda(n))_{n\geqslant 1}$. The said coincidence is not necessarily true for finite-time objects with an m,τ superscript, but for the previous reasons the choice of eligible q has no profound consequences, so in the following discussion we fix $q = q_\star$ and omit q's from the notation, leaving a star instead.

The point we want to make is that the conclusions of theorems 3.1.4 and 3.1.5 tell us that

$$1/\alpha \overset{\forall\tau}{=} \lim_m 1/\alpha_\star^{m,\tau} = \lim_m \sum_{n\geqslant 1} n\lambda_\star^{m,\tau}(n) \tag{3.2}$$

always hold. So the missing piece to guarantee that the extremal index is the inverse of the mean cluster size is the first equality in the following

$$\lim_m \sum_{n\geqslant 1} n\lambda_\star^{m,\tau}(n) \overset{\forall\tau}{=}{}^{?} \sum_{n\geqslant 1} n\lambda_\star(n) = \sum_{n\geqslant 1} n\lambda(n).$$

In [3] authors present an example within the hypotheses of theorems 3.1.3, 3.1.4 and 3.1.5 where the previous equality $\overset{\forall\tau}{=}{}^{?}$ does not hold, actually fails regardless of τ, despite being $\alpha > 0$ and $\sum_{n\geqslant 1} n\lambda(n) < \infty$. It is obtained with a Liverani-Saussol-Vaienti map of the kind discussed in [44] and a target made of the indifferent fixed point and an aperiodic point (it could be another periodic point as well). The authors discuss how such failure can be interpreted as an escape of mass.

Therefore their example produces a situation where the extremal index is not the inverse of the mean cluster size, whereas the finite-time counterpart of this relation holds, as described by theorem 3.1.4 and equation (3.2).

In [2] section 4.2.1, the authors evaluate related questions. They recall that in the previous example the finite-time relation holds for the finite-time mean cluster size given by as in theorem 3.1.3, i.e.,

$$1/\alpha_q^{m,\tau} = \sum_{n\geqslant 1} n\lambda_q^{m,\tau}(n) = \sum_{n\geqslant 1} n\frac{\mu(H_q^n(m,\tau))}{\mu(H_q^0(m,\tau))},$$

where in the events $H_q^n(m,\tau)$ and $H_q^0(m,\tau)$ the first occasion that $U_{m,\tau}$ is visited is the beginning of a q-cluster. As we have seen, the last equation holds in great generality.

However, they show that the latter equation would not hold if, instead, one used, respectively, the events $Q_q^0(U_{m,\tau})$ and $U_{m,\tau}$, in which the first occasion that $U_{m,\tau}$ is visited might not necessarily be the beginning of a q-cluster. Modulo this difference, the following fractions have similar content

$$\lambda'^{m,\tau}_q(n) := \frac{\mu(Q_q^0(U_{m,\tau}))}{\mu(U_{m,\tau})},\ \lambda_q^{m,\tau}(n) = \frac{\mu(H_q^n(m,\tau))}{\mu(H_q^0(m,\tau))}.$$

The expected value of the former quantity,

$$\sum_{n\geqslant 1} n\lambda'^{m,\tau}_q(n) = \sum_{n\geqslant 1} n\frac{\mu(Q_q^0(U_{m,\tau}))}{\mu(U_{m,\tau})} \tag{3.3}$$

is called the finite-time mean sojourn size and, thus, in general, its inverse does not coincide with the finite-time extremal index $\alpha_q^{m,\tau}$. However, in the geometric case, things coincide: finite-time and asymptotic-time statistics, mean cluster size and mean sojourn size.

3.1.2 Spectral methods

The spectral approach to the study of first-hitting time and extreme value laws in the deterministic case was introduced in [60]. It was based on Lasota-Yorke inequalities and the classic perturbative theorems developed in [61] and [62]. These inequalities occur only in the realm of uniformly expanding/hyperbolic systems, with exponential decay of correlations.

In [36], authors considered expanding Lasota-Yorke maps to show conditions of the type Д and establish convergence of two-dimensional REPPs with compound Poisson statistics (see section 3.1.1). Despite the similarity in the starting assumptions, the approach of [36] is fundamentally different from that of [60] and much closer to that described in section 3.1.1. Since [36] handles random dynamical systems, we discuss it in section 3.2.1.

To discuss the spectral approach to the extremal index in the deterministic case, we will basically follow [60].

Let M be improved into a compact Riemannian manifold with Lebesgue measure. Let $(\mathcal{B}, \|\cdot\|)$ be a Banach space of $\mathbb{R}$-valued functions φ on M embedded into a space of distributions acting on a suitable class of functions under integration against φ Leb. In particular, we want to have $1 \in \mathcal{B}$. The dual of $\mathcal{B}$ is also embedded into a space distributions acting on a suitable (possibly different) class of functions. In particular, Leb $\in \mathcal{B}^*$ and, actually $\mathcal{P}(M) \subset \mathcal{B}^*$.

Moreover, assume that if h in the class of functions on which elements of $\mathcal{B}$ can act, then so is $h \circ T$ and let $\mathcal{L}_0$ be the Perron-Frobenius transfer operator given by

$$\begin{array}{llllll} \mathcal{L}_0 : & \mathcal{B} & \to & \mathcal{B} & & \\ & \varphi & \mapsto & \mathcal{L}_0\varphi : & h & \mapsto & \mathcal{L}_0\varphi(h) := \varphi(h \circ T). \end{array}$$

It follows that $\nu_0 :=$ Leb satisfies $\mathcal{L}_0{}^*\nu_0 = \nu_0$. The invariant measures μ that might appear from this approach are absolutely continuous with respect to ν_0. So the target $\Gamma \in \mathscr{B}_M$ is actually taken small so that $\nu_0(\Gamma) = 0$ and, in particular, it would follow that $\mu(\Gamma_0)$. Consider the family of $(\Gamma_\rho)_{\rho \in [0,\rho_0)}$. The system $(T, (\mathcal{B}, \|\cdot\|), \Gamma)$ is assumed to satisfy the following conditions.

1) For all $\varphi \in \mathcal{B}$, it holds that $\mathbb{1}_{\Gamma_\rho}\varphi, \mathbb{1}_{M\setminus\Gamma_\rho}\varphi \in \mathcal{B}$

2) The operators $\mathcal{L}_\rho : (\mathcal{B}, \|\cdot\|) \to (\mathcal{B}, \|\cdot\|)$ given by $\mathcal{L}_\rho(\varphi) = \mathcal{L}_0(\mathbb{1}_{M\setminus\Gamma_\rho}\varphi)$ satisfy:

2.1) $\mathcal{L}_0$ has a spectral gap with leading eigenvalue 1, i.e.,

$$\mathcal{L}_0 = \varphi_0 \otimes \nu_0 + Q_0,$$

$$\mathcal{L}_0\varphi_0 = \varphi_0, \mathcal{L}_0^*\nu_0 = \nu_0, Q_0(\varphi_0) = 0, Q_0^*(\nu_0) = 0, \nu_0(1) = 1, \nu_0(\varphi_0) = 1$$

where $\varphi_0 \otimes \nu_0 : \varphi \mapsto \varphi_0\nu_0(\varphi)$ characterizes the spectral projection in the subspace spanned by φ_0, and $\sigma(Q_0) < 1$.

2.2) $(\mathcal{L}_\rho)_{\rho\in[0,\rho_0)}$ satisfies uniform Lasota-Yorke inequalities, i.e., $\exists \gamma \in (0,1)$, $D > 0$ and a norm (or semi-norm) $|\cdot|_w \leqslant \|\cdot\|$ on $\mathcal{B}$ so that

$$\forall \rho \in [0,\rho_0), \forall \varphi \in \mathcal{B}, \forall n \in \mathbb{N}_{\geqslant 0} : |\mathcal{L}_\rho^n \varphi|_w \leqslant D|\varphi|_w,$$

$$\forall \rho \in [0,\rho_0), \forall \varphi \in \mathcal{B}, \forall n \in \mathbb{N}_{\geqslant 0} : \|\mathcal{L}_\rho^n \varphi\| \leqslant D\gamma^n\|\varphi\| + D|\varphi|_w.$$

2.3) $(\mathcal{L}_\rho)_{\rho\in[0,\rho_0)}$ comprises a triple norm perturbation of $\mathcal{L}_0$, i.e.

$$\forall \rho \in [0,\rho_0) : |||\mathcal{L}_\rho - \mathcal{L}_0||| \leqslant \pi_\rho,$$

where $|||R||| := \sup_{\|\varphi\|\leqslant 1} |R\varphi|_w$, for any linear operator $R : \mathcal{B} \to \mathcal{B}$, and π_ρ is a $[0,\infty)$-valued semi-continuous function so that $\lim_{\rho\to 0} \pi_\rho = 0$.

3) Denoting, for $\rho \in [0,\rho_0)$,

$$A_\rho := \|\nu_0(\mathbb{1}_{\Gamma_\rho}\cdot)\|_{\mathrm{op}((\mathcal{B},\|\cdot\|),(\mathbb{C},|\cdot|))} = \|\nu_0 \circ \mathcal{L}_\rho\|_{\mathrm{op}((\mathcal{B},\|\cdot\|),(\mathbb{C},|\cdot|))},$$

$$B_\rho := \|\mathbb{1}_{\Gamma_\rho}\varphi_0\| \geqslant \|\mathcal{L}_0(\mathbb{1}_{\Gamma_\rho}\varphi_0)\| \text{ and}$$

$$\Delta_\rho := \nu_0 \circ (\mathcal{L}_0 - \mathcal{L}_\rho)(\varphi_0)$$

it holds that

$$\lim_{\rho\to 0} A_\rho = 0 \text{ and } A_\rho B_\rho \leqslant \text{const}\,|\Delta_\rho| \ (\forall \rho \in [0,\rho_0)).$$

Following, [61] and [59], items (2.1-2.3) imply that

i)

$$\forall\rho\in[0,\infty):\sigma_{\text{res}}(\mathcal{L}_\rho)\subset\{z\in\mathbb{C}:|z|\leqslant\gamma\}$$

ii)

$$\lambda_\rho^{-1}\mathcal{L}_\rho=\varphi_\rho\otimes\nu_\rho+Q_\rho\ (\lambda_0=1),$$

$$\mathcal{L}_\rho\varphi_\rho=\lambda_\rho\varphi_\rho,\mathcal{L}_\rho^*\nu_\rho=\lambda_\rho\nu_\rho,Q_\rho(\varphi_\rho)=0,Q_\rho^*(\nu_\rho)=0,\nu_\rho(\varphi_0)=1,\nu_\rho(\varphi_\rho)=1$$

where $\varphi_\rho\otimes\nu_\rho:\varphi\mapsto\varphi_\rho\nu_\rho(\varphi)$ characterizes the spectral projection in the subspace spanned by φ_ρ, and $\sigma(Q_\rho)<\lambda_\rho$ (and $\sigma(Q_\rho)\leqslant\gamma$)

In particular, $\sum_{n\geqslant0}\sup_{\rho\in[0,\rho_0)}\|Q_\rho^n\|_\rho<\infty$.

iii)

$$\exists C>0,\forall\rho\in[0,\rho_0):\|\varphi_\rho\|\leqslant C.$$

Following [62] corollary 2, items (i-iii) and (3) guarantee that: $\forall\rho\in[0,\rho_0),\forall N\geqslant1$ one has

$$\frac{1-\lambda_\rho}{\Delta_\rho}=\frac{\lambda_0-\lambda_\rho}{\Delta_\rho}=\left(\alpha_{N,\rho}+\mathcal{O}_{N\to\infty}((1-\gamma')^N)\right)(1+\mathcal{O}_{\rho\to0}(NA_\rho)) \tag{3.4}$$

where constants associated to both $\mathcal{O}$'s independent of ρ and N, $\gamma'>0$ is a lower bound on the spectral gap of the $\mathcal{L}_\rho$'s, and

$$\alpha_{N,\rho}=1-\sum_{q=0}^{N-1}\lambda_\rho^{-q}\beta_{q,\rho}, \tag{3.5}$$

with

$$\beta_{q,\rho}=\frac{\nu_0\circ(\mathcal{L}_0-\mathcal{L}_\rho)\circ\mathcal{L}_\rho^q\circ(\mathcal{L}_0-\mathcal{L}_\rho)(\varphi_0)}{\Delta_\rho}=\frac{\mu_0(\{x\in M:x\in\Gamma_\rho,\tau_{\Gamma_\rho}^1(x)=q+1\})}{\mu_0(\Gamma_\rho)} \tag{3.6}$$

where $\mu_0(\cdot):=\nu_0(\varphi_0(\cdot))\in\mathcal{P}_T(M)$, and the latter equality in equation (3.6) holds when $\mu_0(\Gamma_\rho)>0$ ($\forall\rho\in(0,\rho_0)$), which is an additional hypothesis assumed until the end of this section.

Notice that equation (3.6) points in the direction of a derivative-type result, that would refine a continuity result, if we had one. But λ_ρ appears in the RHS of equation (3.6) as well, so continuity has to be known upfront we we want to take the $\rho\to0$ limit:

$$\nu_0\circ(\mathcal{L}_0-\mathcal{L}_\rho)(\varphi_\rho)=(\mathcal{L}_0^*\nu_0)(\varphi_\rho)-\nu_0(\mathcal{L}_\rho\varphi_\rho)=(\lambda_0\nu_0)(\varphi_\rho)-\nu_0(\lambda_\rho\varphi_\rho)=\lambda_0-\lambda_\rho,$$

where we have used the normalization $\nu_0(\varphi_\rho)=1$, so, applying assumptions $\lim_{\rho\to0}A_\rho$ (see (3)) and $\|\varphi_\rho\|\leqslant C$ (see (iii)), it follows that $\lim_{\rho\to0}\lambda_\rho=\lambda_0=1$.

The equation (3.6) also encodes why the spectral approach appears in the study of hitting statistics, but in the following, we will present a clearer version of this motivation. Notice also the similitude between fraction appearing in the sojourn size given in equation (3.3) and the RHS in equation (3.6): both consider the portion of points which are initially in certain neighborhood of Γ but break a q-cluster of visits, with the former considering points returning at any later moments $>q$, while the latter considering points returning exactly at time $q+1$.

Now that the perturbation results and their conclusions, equations (3.4-3.6), are clearly stated, it is time to see why they naturally appear in the study of hitting times: for any $n \geqslant 0$,

$$\begin{aligned}\int_{\{r^1_{\Gamma_\rho} \geqslant n\}} \varphi d\nu_0 &= \int_M \prod_{i=0}^{n-1} \mathbb{1}_{M\setminus\Gamma_\rho} \circ T^i \varphi d\nu_0 = \int_M \mathcal{L}_0^n \left(\prod_{i=0}^{n-1} \mathbb{1}_{M\setminus\Gamma_\rho} \circ T^i \varphi \right) d\nu_0 \\ &= \int_M \mathcal{L}_\rho^n(\varphi) d\nu_0 = \int_M \lambda_\rho^n \nu_\rho(\varphi)\varphi_\rho d\nu_0 + \int_M \lambda_\rho^n Q_\rho^n \varphi d\nu_0.\end{aligned}$$

So, passing $\varphi = \varphi_0$ (notice that passing $\varphi = 1$ is also interesting) while using $\nu_\rho(\varphi_0) = 1$ (by design) and $\nu_0(\varphi_\rho) \to 1$ as $\rho \to 0$ (by [62], lemma 6.1), gives

$$\left| \mu_0(r^1_{\Gamma_\rho} \geqslant n) - \lambda_\rho^n \right| \leqslant \lambda_\rho^n \gamma^n \|\varphi_0\|. \tag{3.7}$$

Taking $\rho \to 0$ in equation (3.4) gives

$$\lim_{\rho\to 0} \frac{1-\lambda_\rho}{\mu_0(\Gamma_\rho)} \overset{\forall N}{=} 1 - \sum_{q=0}^{N-1} \beta_q + \mathcal{O}_{N\to\infty}((1-\gamma')^N),$$

provided that $\exists \lim_{\rho\to 0} \beta_{q,\rho} =: \beta_q$, which we assume until the end of this section, and using the previously justified continuity of λ_ρ's. The latter equation, being true for every N, allows us to take the limit $N \to \infty$ to arrive at

$$\begin{aligned}&\lim_{\rho\to 0} \frac{1-\lambda_\rho}{\mu_0(\Gamma_\rho)} = \alpha := 1 - \sum_{q=0}^{\infty} \beta_q \\ &\Rightarrow \lambda_\rho = 1 - \alpha\mu_0(\Gamma_\rho) + o_{\rho\to 0}(\mu_0(\Gamma_\rho)) = e^{-\alpha\mu_0(\Gamma_\rho) + o_{\rho\to 0}(\mu_0(\Gamma_\rho))}.\end{aligned} \tag{3.8}$$

To conclude we look at the time-normalized hitting time:

$$(3.8) \Rightarrow \lambda_\rho^{\left\lfloor \frac{t}{\mu_0(\Gamma_\rho)} \right\rfloor} = e^{-\alpha\mu_0(\Gamma_\rho)\left\lfloor \frac{t}{\mu_0(\Gamma_\rho)} \right\rfloor + o_{\rho\to 0}(t)} \xrightarrow{\rho\to 0} e^{-\alpha t}$$

$$(3.7) \Rightarrow \left| \mu_0(r^1_{\Gamma_\rho} > \lfloor t/\mu_0(\Gamma_\rho) \rfloor) - \lambda_\rho^{\left\lfloor \frac{t}{\mu_0(\Gamma_\rho)} \right\rfloor + 1} \right| \leqslant \lambda_\rho^{\left\lfloor \frac{t}{\mu_0(\Gamma_\rho)} \right\rfloor + 1} \gamma^{\left\lfloor \frac{t}{\mu_0(\Gamma_\rho)} \right\rfloor + 1} \|\varphi_0\|.$$

implying that

$$\lim_{\rho\to 0} \mu_0(r^1_{\Gamma_\rho} > \lfloor t/\mu_0(\Gamma_\rho) \rfloor) = e^{-\alpha t}.$$

Therefore the system (T, μ_0, Γ) has extremal index α, with

$$\alpha = 1 - \sum_{q=0}^{\infty} \beta_q, \text{ with } \beta_q = \lim_{\rho\to 0} \frac{\mu_0(\{x \in M : x \in \Gamma_\rho, \tau^1_{\Gamma_\rho}(x) = q+1\})}{\mu_0(\Gamma_\rho)}.$$

On the extremal value side, given a space-normalization (g, u) as in equations (2.10-2.12), one has

$$\mu_0(M_m \leqslant u_m(\tau)) = \mu_0(r^1_{\Gamma_{g^{-1}(u_m(\tau))}} \geqslant m) \Rightarrow |\mu_0(M_m \leqslant u_m(\tau)) - \lambda^m_{g^{-1}(u_m(\tau))}| \xrightarrow{m} 0,$$

while

$$\lambda^m_{g^{-1}(u_m(\tau))} = e^{-\alpha m\mu_0(\Gamma_{g^{-1}(u_m(\tau))})+mo_{m\to 0}(\mu_0(\Gamma_{g^{-1}(u_m(\tau))}))} \xrightarrow{m} e^{-\alpha\tau},$$

so that

$$\mu_0(M_m \leqslant u_m(\tau)) \xrightarrow{m} e^{-\alpha\tau}.$$

Error terms are explicitly presented in the previous approach (in concrete situations, the terms γ and γ' can be derived from the expansion factor), but they also discuss sharp bounds.

Applications include Rychlik piecewise expanding maps of the interval and higher-dimensional expanding maps, with quite general targets. When the targets are singletons, the results presented in the previous section are recovered. The baker's map was studied with this technique in [11]. It is naturally expected that the approach extends to general equilibrium states. Let us recall that this approach is by design restricted to uniformly hyperbolic situations.

The reader will have noted that the higher-order hitting statistics and compound Poisson distributions were not addressed in the previous paper. Only very recently, in [8], this problem was addressed for random dynamical systems, we will discuss it more carefully in section 3.2.2. To conclude this section we follow [8] (section 2) to review briefly what their approach looks like in the deterministic situation.

Under the normalization

$$m = \left\lfloor \frac{\tau}{\mu_0(\Gamma_{g^{-1}(u_m(\tau))})} \right\rfloor,$$

their objective is to evaluate

$$\lim_{m\to\infty} \mu_0(Z^m_{\Gamma_{g^{-1}(u_m(\tau))}} = n),$$

but this will be approached indirectly: instead of evaluating the distribution of $Z^m_{\Gamma_{g^{-1}(u_m(\tau))}}$, one considers the characteristic function (Fourier transform) of such random variable and studies its pointwise convergence when $\rho \to 0$. Usually, one identifies such limit as the characteristic function of a random variable Z with a known distribution and then concludes that $Z^m_{\Gamma_{g^{-1}(u_m(\tau))}} \underset{\text{wrt } \mu_0}{\overset{d}{\to}} Z$ as $m \to \infty$, using Levy continuity theorem. However, in [8], the approach to the limit will be indirect, as we will see later.

The calculation starts as follows. For $\rho \in [0, \rho_0)$, $s \in \mathbb{R}$ and $L \geqslant 1$, one has

$$\mu_0\left(e^{isZ^L_{\Gamma_\rho}}\right) = \int_M e^{isZ^L_{\Gamma_\rho}} \varphi_0 d\nu_0 = \int_M \mathcal{L}^L_0 \left(e^{isZ^L_{\Gamma_\rho}} \varphi_0\right) d\nu_0 = \int_M \mathcal{L}^L_{\rho,s}(\varphi_0) d\nu_0, \tag{3.9}$$

where $\mathcal{L}_{\rho,s}(\varphi) := \mathcal{L}_0(e^{is\mathbb{1}_{\Gamma_\rho}} \varphi)$ and so $\mathcal{L}^L_{\rho,s}(\varphi) = \mathcal{L}^L_0\left(e^{isZ^L_{\Gamma_\rho}} \varphi\right)$.

Then the spectral perturbation theory developed before is, for each $s \in \mathbb{R}$, applied to $(\mathcal{L}_{\rho,s})_{\rho\in[0,\rho_0)}$, noticing that $\mathcal{L}_{0,s} = \mathcal{L}_0$, as to get

$$\lambda^{-1}_{\rho,s}\mathcal{L}_{\rho,s} = \varphi_{\rho,s} \otimes \nu_{\rho,s} + Q_{\rho,s}$$

$$\mathcal{L}_{\rho,s}\varphi_{\rho,s} = \lambda_{\rho,s}\varphi_{\rho,s}, \mathcal{L}^*_{\rho,s}\nu_{\rho,s} = \lambda_{\rho,s}\nu_{\rho,s}, Q_{\rho,s}(\varphi_{\rho,s}) = 0, Q^*_{\rho,s}(\nu_{\rho,s}) = 0, \nu_{\rho,s}(\varphi_{0,s}) = 1, \nu_{\rho,s}(\varphi_{\rho,s}) = 1,$$

in particular, $\varphi_{0,s} = \varphi_0$, $\nu_{0,s} = \nu_0$ and $\lambda_{0,s} = \lambda_0$, and, ultimately,

$$\lim_{\rho\to 0} \frac{\lambda_0 - \lambda_{\rho,s}}{\Delta_{\rho,s}} = \alpha(s) := 1 - \sum_{q=0}^{\infty} \beta_q(s),$$

where, considering that $\mu(\Gamma_\rho) > 0$,

$$\Delta_{\rho,s} = (1-e^{is})\mu(\Gamma_\rho),$$

$$\begin{aligned}
\beta_q(s) &= \lim_{\rho\to 0} \frac{\nu_0 \circ (\mathcal{L}_0 - \mathcal{L}_{\rho,s}) \circ \mathcal{L}^q_{\rho,s} \circ (\mathcal{L}_0 - \mathcal{L}_{\rho,s})(\varphi_0)}{(1-e^{is})\mu(\Gamma_\rho)} \\
&= \lim_{\rho\to 0} (1-e^{is}) \frac{\displaystyle\int_{\Gamma_\rho \cap T^{-(q+1)}\Gamma_\rho} e^{isZ^q_{*\Gamma_\rho}} \varphi_0 d\nu_0}{\mu_0(\Gamma_\rho)} \\
&= \lim_{\rho\to 0} (1-e^{is}) \sum_{n=0}^{q} e^{ins} \frac{\mu_0(\{x \in M, x \in \Gamma_\rho, T^{q+1}x \in \Gamma_\rho, Z^q_{*\Gamma_\rho}(x) = n\})}{\mu_0(\Gamma_\rho)} \\
&=: (1-e^{is}) \sum_{n=0}^{q} e^{ins} \lim_{\rho\to 0} \kappa^n_q(\rho,s),
\end{aligned}$$

provided that the limits above exist. In particular, it can be shown that

$$\lim_{\rho\to 0} \frac{1-\lambda_{\rho,s}}{\Delta_{\rho,s}} = \alpha(s) := 1 - (1-e^{is}) \sum_{q=0}^{\infty} \sum_{n=0}^{q} e^{ins} \lim_{\rho\to 0} \kappa^n_q(\rho,s), \tag{3.10}$$

where the latter double sum is shown to be finite.

Now we pass to the subsequence of radii $(\rho_m)_{m\geqslant 1} \searrow 0$ given by the space-normalization (g,u), namely, $\rho_m(\tau) = g^{-1}(u_m(\tau))$ $(m \geqslant 1)$.

The last centered equation then implies that

$$\forall s \in \mathbb{R} : \lambda_{g^{-1}(u_m(\tau)),s} \sim_{m\to\infty} 1 - \alpha(s)\Delta_{g^{-1}(u_m(\tau)),s} \tag{3.11}$$

$$= 1 - \alpha(s)(1-e^{is})\mu_0(\Gamma_{g^{-1}(u_m(\tau))}) \sim_{m\to\infty} 1 - \alpha(s)(1-e^{is})\frac{\tau}{m}$$

$$\Rightarrow \lambda^m_{g^{-1}(u_m(\tau)),s} \sim_{m\to\infty} \left[1 - \alpha(s)(1-e^{is})\mu_0(\Gamma_{g^{-1}(u_m(\tau))})\right]^m \sim_{m\to\infty} e^{-\alpha(s)(1-e^{is})}$$

And equation (3.9) becomes

$$\mu_0\left(e^{isZ^m_{\Gamma_{g^{-1}(u_m(\tau))}}}\right) = \int_M \mathcal{L}^m_{g^{-1}(u_m(\tau)),s}(\varphi_0) d\nu_0 \tag{3.12}$$

$$= \int_M \lambda^m_{g^{-1}(u_m(\tau)),s} \nu_{g^{-1}(u_m(\tau)),s}(\varphi_0) \varphi_{g^{-1}(u_m(\tau)),s} d\nu_0$$

$$+ \int_M \lambda^m_{g^{-1}(u_m(\tau)),s} Q^m_{g^{-1}(u_m(\tau)),s} \varphi_0 d\nu_0$$

$$\Rightarrow \left|\mu_0\left(e^{isZ^m_{\Gamma_{g^{-1}(u_m(\tau))}}}\right) - \lambda^m_{g^{-1}(u_m(\tau)),s}\right| \leqslant \lambda^m_{g^{-1}(u_m(\tau)),s} \gamma^m_s \|\varphi_0\|.$$

$$\Rightarrow \mu_0\left(e^{isZ^m_{\Gamma_{g^{-1}(u_m(\tau))}}}\right) \xrightarrow{m\to\infty} e^{-(1-e^{is})\alpha(s)\tau} \ (\forall s\in\mathbb{R}, \forall\tau\in\mathbb{R}_{>0}) \tag{3.13}$$

Since the RHS is continuous at 0, it is the characteristic function of a certain random variable Z. Also, since the sequence of random variables in the LHS is $\mathbb{N}_{\geqslant 0}$-valued, so is the RHS (by the Portmanteau theorem). Finally, Z is infinitely divisible, because $e^{-\alpha(s)(1-e^{is})\tau} = (e^{-\alpha(s)(1-e^{is})\tau/N})^N$. These conditions, due to a result by Feller, imply that Z has a compound Poisson distribution. Moreover, Levy's inversion formula can be applied to get that the mass probability function of compound Poissonian variable Z at $k\in\mathbb{N}_{\geqslant 0}$, thereby leading to the characterization we were initially after:

$$\lim_{m\to\infty}\mu_0(Z^m_{\Gamma_{g^{-1}(u_m(\tau))}} = n) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e^{-isn}e^{-\alpha(s)(1-e^{is})\tau}ds.$$

On the other hand, the underlying limiting Z, defined on an abstract probability space $(\Omega,\mathscr{F},\mathbb{P})$, being compound Poissonian, can be written as $Z(\omega) = \sum_{i=0}^{N(\omega)} X_i(\omega)$, where $N\sim \mathrm{Poi}_\gamma$ ($\gamma\in\mathbb{R}_{>0}$), $\mathbb{P}(X_i=\ell) = \lambda_\ell$ ($i\geqslant 1, \ell\geqslant 1$) and $(X_i)_{i\geqslant 1}$ an independent family, also independent from N. In [8], they point out that $\gamma = t/\sum_{\ell\geqslant 1}\ell\lambda_\ell$ (see their equation (3.38)), while presenting the characteristic function of X_1 (see their equation 2.18):

$$\mathbb{E}(e^{isX_1}) = \frac{\alpha(s)(e^{is}-1)}{\alpha(0)} + 1$$

Therefore, one could apply Levy's inversion formula once again to recover the multiplicity distribution:

$$\lambda_\ell = \lim_{T\to\infty}\int_{-T}^{T} e^{-is\ell}\left(\frac{\alpha(s)(e^{is}-1)}{\alpha(0)}+1\right)ds.$$

3.1.3 Probabilistic approximation with Chen-Stein method

The Chen-Stein method, see [18] and [84], provides a way to compare a given distribution and a given compound Poisson distribution, according to their total variation distance.

In [48], authors use this technique to study hitting statistics of a broad class of measurable maps T on a measurable space M with an intrinsic countable measurable partition C_1 and ϕ-mixing invariant measure μ, i.e. either

i) Left ϕ-mixing:

$$\exists\phi:\mathbb{N}_{\geqslant 1}\to\mathbb{R}, \lim_{k\to\infty}\phi(k)=0, \forall m,m'\geqslant 1, \forall U\in\sigma\left(\bigvee_{i=0}^{m-1}T^{-i}C_1\right), \forall V\in\sigma\left(\bigcup_{n\geqslant 1}\bigvee_{i=0}^{n-1}T^{-i}C_1\right)$$

$$|\mu(U\cap T^{-m-m'}V) - \mu(U)\mu(V)| \leqslant \mu(U)\phi(m'),$$

or

ii) Right ϕ-mixing:

$$\exists\phi:\mathbb{N}_{\geqslant 1}\to\mathbb{R}, \lim_{k\to\infty}\phi(k)=0, \forall m,m'\geqslant 1, \forall U\in\sigma\left(\bigvee_{i=0}^{m-1}T^{-i}C_1\right), \forall V\in\sigma\left(\bigcup_{n\geqslant 1}\bigvee_{i=0}^{n-1}T^{-i}C_1\right)$$

$$|\mu(U\cap T^{-m-m'}V) - \mu(U)\mu(V)| \leqslant \mu(V)\phi(m'),$$

or both.

If instead of the previous two inequalities one considered

$$|\mu(U \cap T^{-m-m'}V) - \mu(U)\mu(V)| \leqslant \mu(U)\mu(V)\psi(m'),$$

it would be defining the so-called ψ-mixing condition, which is a particular case of ϕ-mixing. The decay function ϕ is assumed summable.

For the sake of completeness, let us say that when the inequality

$$|\mu(U \cap T^{-m-m'}V) - \mu(U)\mu(V)| \leqslant \alpha(m'),$$

occurs instead of the previous one, the α-mixing condition is defined.

The theory works for both left and right ϕ-mixing systems, invertible or not. We concentrate the exposition on the left ϕ-mixing non-invertible case.

They consider very general target sets, presented in terms of cylinders. Namely, they consider a nested sequence of sets $U_m \in \sigma\left(\bigvee_{i=0}^{m-1} T^{-i}C_1\right)$ $(m \geqslant 1)$ so that, with $\Gamma := \bigcap_{m\geqslant 1} U_m$, it holds that $\mu(\Gamma) = 0$. Notice that the previous framework includes the case where one starts with a measurable μ-negligible target Γ and considers $U_m := C_m(\Gamma)$, where, for $X \subset M$, $C_m(X) := \bigcup_{\xi \in C_m, \xi \cap X} \xi$. Moreover, it is assumed that they shrink in size fast enough so that: $\exists (a_k)_{k\geqslant 1} \searrow 0, \forall k \geqslant 1, \exists m(k) \geqslant k, \forall m \geqslant m(k):$

$$\sum_{i=k}^{m} \mu(C_i(U_m)) \leqslant a_k. \tag{3.14}$$

If $U_m = C_m(\Gamma)$, then $C_i(U_m) = U_i$ and the previous condition basically reduces to $\sum_{i=k}^{\infty} \mu(U_i) \xrightarrow{k\to\infty} 0$.

Then Chen-Stein method is used to estimate the total variation distance between the distribution of $Z_m^{\lfloor \tau/\mu(U_m)\rfloor} := \sum_{i=0}^{\lfloor \tau/\mu(U_m)\rfloor - 1} \mathbb{1}_{U_m} \circ T^i$ under μ and a compound Poisson distribution $\eta_{L,m}$ with intensity $\tau(\sum_{\ell\geqslant 1} \ell\lambda_\ell(L,m))^{-1}$ and multiplicity distribution

$$\lambda_\ell^{m,L} := \frac{\ell^{-1}\mathbb{E}_\mu(\mathbb{1}_{\sum_{j=i-K}^{i+K} \mathbb{1}_{U_m}\circ T^j = \ell} | \mathbb{1}_{U_m} \circ T^i = 1)}{\sum_{\ell\geqslant 1} \ell^{-1}\mathbb{E}_\mu(\mathbb{1}_{\sum_{j=i-K}^{i+K} \mathbb{1}_{U_m}\circ T^j = \ell} | \mathbb{1}_{U_m} \circ T^i = 1)} \; (\ell \geqslant 1),$$

where the RHS is independent of i, with i taken larger than L.

The key to guarantee that the distributions of $Z_m^{\lfloor \tau/\mu(U_m)\rfloor}$ under μ converge under the limit $\lim_{m\to\infty}$ is to show that the $\lambda_\ell(L,m)$ also converge under the double-limit $\lim_{L\to\infty} \lim_{m\to\infty}$.

In order to do so, they assume that

$$\exists \alpha_\ell(L,m) := \lim_{L\to\infty} \lim_{m\to\infty} \frac{\mu(Z_m^L = \ell, \mathbb{1}_{U_m} = 1)}{\mu(U_m)} \text{ and } \sum_{\ell=1}^{\infty} \ell^2 \alpha_\ell < \infty.$$

Then one shows that these conditions guarantee the desired limits in the parameters of $\eta_{L,m}$, defining a compound Poisson distribution η with intensity $\tau(\sum_{\ell\geqslant 1} \ell\lambda_\ell)^{-1}$ and multiplicity distribution $(\lambda_\ell)_{\ell\geqslant 1}$. This approach was inspired in [53] and therefore also similar to the one developed in chapter 4 (see equation 4.4 and hypothesis (H9)).

It turns out that

$$\|(Z_m^{\lfloor \tau/\mu(U_m)\rfloor})_*\mu - \eta\|_{TV} \overset{\forall L}{\leqslant} \|(Z_m^{\lfloor \tau/\mu(U_m)\rfloor})_*\mu - \eta_{L,m}\|_{TV} + \|\eta_{L,m} - \eta\|_{TV},$$

can be controlled after the m-limit (followed by the L-limit): the second term is controlled with the α_ℓ's and the first is controlled with Chen-Stein, as to produce a bounding term with essentially two parts: a component accounting for long-range interactions (which is further bounded using ϕ-mixing and ϕ-summability), and a component accounting for short-range interactions (further bounded using ϕ-mixing, ϕ-summability, and the approximation condition in equation (3.14)).

3.2 Random systems

Consider M a complete separable metric space, and Ω, the so-called driving space, a complete separable metric space equipped with a measurably-invertible ergodic system $(\theta, \mathbb{P})$.

Consider maps $T_\omega : M \to M$ ($\omega \in \Omega$) which combine to make the a measurable skew product $S : \Omega \times M \to \Omega \times M$, $(\omega, x) \mapsto (\theta\omega, T_\omega x)$. As usual, for higher-order iterates we denote $S^n(\omega, x) = (\theta^n\omega, T_\omega^n(x))$ where $T_\omega^n = T_{\theta^{n-1}\omega} \circ \cdots \circ T_{\theta\omega} \circ T_\omega$ $(n \geqslant 1)$.

Denote

$$\mathcal{P}^{\mathbb{P}}(\Omega \times M) = \{\hat{\mu} \in \mathcal{P}(\Omega \times M) : \hat{\mu}, \pi_{\Omega *}\hat{\mu} = \mathbb{P}\},$$

$$\mathcal{P}_S^{\mathbb{P}}(\Omega \times M) = \{\hat{\mu} \in \mathcal{P}(\Omega \times M) : S_*\hat{\mu} = \hat{\mu}, \pi_{\Omega *}\hat{\mu} = \mathbb{P}\},$$

and

$$\mathcal{P}^{(\mathbb{P})}(M) = \left\{ \begin{array}{c} \mu : \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega \in \mathcal{P}(M) \text{ so that:} \\ \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega(E_\omega) \in [0,1] \text{ is } (\mathscr{B}_\Omega, \mathscr{B}_{[0,1]})\text{-measurable, } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M \end{array} \right\},$$

$$\mathcal{P}_{T_\Omega}^{(\mathbb{P})}(M) = \left\{ \begin{array}{c} \mu : \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega \in \mathcal{P}(M) \text{ so that:} \\ \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega(E_\omega) \in [0,1] \text{ is } (\mathscr{B}_\Omega, \mathscr{B}_{[0,1]})\text{-measurable, } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M \\ T_{\omega *}^n \mu_\omega = \mu_{\theta^n\omega},\ \forall n \geqslant 0,\ \mathbb{P}\text{-a.s.} \end{array} \right\}.$$

Notation. Elements in the latter two sets are will written as $\mu = (\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega}$. We will use the notational device $\in_{\mathbb{P}}$ to identify a family of objects which is defined $\mathbb{P}$-a.s.

A family of measures $(\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega}$ satisfying "$T_{\omega *}^n \mu_\omega = \mu_{\theta^n\omega}$, $\forall n \geqslant 0$, $\mathbb{P}$-a.s." is called a covariant family.

For the next paragraph, we refer to [28] (prop. 3.3) and [5] (sec. 1.4). For $E \in \mathscr{B}_\Omega \times \mathscr{B}_M$, for any $\omega \in \Omega$, its ω-section, $\{x \in M : (\omega, x) \in E\}$, is denoted by E_ω or $E(\omega)$.

Using Rohklin disintegration theorem for $\hat{\mu} \in \mathcal{P}^{\mathbb{P}}(\Omega \times M)$ with respect to the partition $\mathscr{P} = \{[\omega] := \{\omega\} \times M \mid \omega \in \Omega\}$ of $\Omega \times M$, we have that there exists $(\mu_\omega)_\omega$ so that:

$$\text{i) } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M\text{: } \hat{\mu}(E) = \int_\Omega \mu_\omega(E_\omega) d\mathbb{P}(\omega), \text{ ii) } \begin{array}{ccc} \mathcal{P}^{\mathbb{P}}(\Omega \times M) & \hookrightarrow & \mathcal{P}^{(\mathbb{P})}(M) \\ \mathcal{P}_S^{\mathbb{P}}(\Omega \times M) & \hookrightarrow & \mathcal{P}_{T_\Omega}^{(\mathbb{P})}(M) \\ \hat{\mu} & \mapsto & (\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega} \end{array}.$$

Conversely,

$$\begin{array}{rcl} \mathcal{P}^{(\mathbb{P})}(M) & \hookrightarrow & \mathcal{P}^{\mathbb{P}}(\Omega\times M) \\ \mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M) & \hookrightarrow & \mathcal{P}^{\mathbb{P}}_S(\Omega\times M) \\ (\mu_\omega)_{\omega\in_{\mathbb{P}}\Omega} & \mapsto & \hat{\mu}=d\mu_\omega d\mathbb{P}(\omega), \text{ i.e., } \hat{\mu}(E)=\int_\Omega \mu_\omega(E_\omega)d\mathbb{P}(\omega)\ (\forall E\in\mathscr{B}_\Omega\times\mathscr{B}_M) \end{array}.$$

Now, consider a given $\hat{\mu}=d\mu_\omega d\mathbb{P}(\omega)\in\mathcal{P}^{\mathbb{P}}_S(\Omega\times M)$, with the associated $(\mu_\omega)_{\omega\in_{\mathbb{P}}\Omega}\in\mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M)$. Define the marginal measure $\check{\mu}=\pi_{M*}\hat{\mu}=\int_\Omega\mu_\omega\, d\mathbb{P}(\omega)\in\mathcal{P}(M)$. As a implying assumption, in this section we will consider that every measure appearing is atomless and fully supported.

Finally, consider $\Gamma\in\mathscr{B}_\Omega\times\mathscr{B}_M$ so that, $\mathbb{P}$-a.s, $\Gamma(\omega)$ is small in the sense that $\mu_\omega(\Gamma(\omega))=0$. The set Γ is the so-called random target. Denote $\Gamma_\rho(\omega)=B_\rho(\Gamma(\omega))$ ($\rho>0$) and the corresponding ω-collection by Γ_ρ.

The objects considered above comprise what we call a 'targeted random dynamical system', or simply 'system', to be denoted by the tuple $(\theta,\mathbb{P},T_\omega,\mu_\omega,\Gamma)$.

Generally speaking, random dynamical systems are amenable to the so-called annealed or quenched results. Annealed ones are averaged in terms of the ω-noise and they rely on the probabilities $\hat{\mu}\in\mathcal{P}(\Omega\times M)$ or $\check{\mu}\in\mathcal{P}(M)$. Notice that the said average and the measures used will wipe out any dependence on a certain ω. On the other hand, quenched results are valid for ($\mathbb{P}$-generic) fixed ω-realizations and they rely on the probabilities $\mu_\omega\in\mathcal{P}(M)$.

Let's define some working objects. Let $U\in\mathscr{B}_\Omega\times\mathscr{B}_M$ be a set whose ω-sections $U(\omega)\subset M$ have positive μ_ω-measure, $\mathbb{P}$-a.s.. Denoting the latter full $\mathbb{P}$-measure set by Ω', notice that $\Omega''=\bigcap_{n\in\mathbb{Z}}\theta^{-n}\Omega'$ still has full $\mathbb{P}$-measure. So we can consider that $\mu_{\theta^n\omega}(U(\theta^n\omega))>0$, for all $n\in\mathbb{Z}$, $\mathbb{P}$-a.s..

Definition 3.2.1. The **first hitting time** of $(\theta,\mathbb{P},T_\omega,\mu_\omega,U)$ is the family of functions

$$\begin{array}{rccl} r_U^{\omega,1}: & M & \to & \mathbb{N}_{\geqslant 1}\cup\{\infty\} \\ & x & \mapsto & \inf\{i\in\mathbb{N}_{\geqslant 1}: T^i_\omega(x)\in U(\theta^i\omega)\} \end{array}.$$

The associated **higher-order hitting times** are given, for $\ell\geqslant 2$, by the family of functions

$$\begin{array}{rcccl} r_U^{\omega,\ell} & : & M & \to & \mathbb{N}_{\geqslant \ell}\cup\{\infty\} \\ & & x & \mapsto & r_U^{\omega,\ell}(x)=r_U^{\omega,\ell-1}(x)+r_U^{\omega'}(T_\omega^{r_U^{\omega,\ell-1}}(x)) \end{array},$$

where $\omega'=\theta^{r_U^{\omega,\ell-1}(x)}\omega$.

Definition 3.2.2. The **hit counting function** of $(\theta,\mathbb{P},T_\omega,\mu_\omega,U)$, for $L\geqslant 1$ is given by the family of functions

$$\begin{array}{l} Z_{*U}^{\omega,L}:M\to\mathbb{N}_{\geqslant 0} \\ \quad x\mapsto\sum_{i=1}^{L}\mathbb{1}_{U(\theta^i\omega)}\circ T^i_\omega(x) \end{array}, \qquad \begin{array}{l} Z_{U}^{\omega,L}:M\to\mathbb{N}_{\geqslant 0} \\ \quad x\mapsto\sum_{i=0}^{L-1}\mathbb{1}_{U(\theta^i\omega)}\circ T^i_\omega(x) \end{array}.$$

These objects are related, for example, in the sense that $\{Z_{*U}^{\omega,L}\geqslant\ell\}=\{r_U^{\omega,\ell}\leqslant L\}$, $\{Z_{*U}^{\omega,L}=\ell\}=\{r_U^{\omega,\ell}\leqslant L<r_U^{\omega,\ell+1}\}$.

In the next sections, where we review the literature addressing the random case, the so-called random subshifts of finite type (sometimes considered with countable alphabets, though) will appear often. For the convenience of the reader, we will introduce them next. A special feature of these systems is that the concept of periodicity appears much more naturally than for general random systems. They do not fit perfectly in the setup presented above, which is just right for the theory to be developed later in the thesis. However, the adjustments needed are tiny ones.

Let σ be the left shift on $M = {\mathbb{N}_{\geqslant 1}}^{\mathbb{N}_{\geqslant 0}}$, a complete separable metric space under $d(x,y) = 2^{-\min\{n\geqslant 0: x_n \neq y_n\}}$. Let Ω be a complete separable metric space with a measurably-invertible ergodic system $(\theta,\mathbb{P})$. Let $b:\Omega\to\mathbb{N}_{\geqslant 1}$ be measurable with $\mathbb{E}_{\mathbb{P}}(\log b)<\infty$. Let, for each $\omega\in\Omega$, $A(\omega)$ be a $b(\omega)\times b(\theta\omega)$ $\{0,1\}$-valued matrix with at least one 1 in each column and row and whose entries are ω-measurable. Consider the ω-family of closed sets in M

$$M_\omega = \left\{x=(x_0,x_1,\dots) : x_i \in [1,\dots,b(\theta^i\omega)]\cap\mathbb{N}_{\geqslant 1}, A_{x_i,x_{i+1}}(\theta^i\omega)=1, \forall i\geqslant 0\right\}\subset M.$$

Let

$$\begin{array}{rccc}\sigma_\omega: & M_\omega & \to & M_{\theta\omega}\\ & x & \mapsto & \sigma(x),\end{array}$$

with the associated skew map acting on $\mathcal{E}=\{(\omega,x)\in\Omega\times M : x\in M_\omega\}\subset\Omega\times M$ as

$$\begin{array}{rccc}S: & \mathcal{E} & \to & \mathcal{E}\\ & (\omega,x) & \mapsto & (\theta\omega,\sigma_\omega(x)).\end{array}$$

Let $\check{\mu}\in\mathcal{P}_S^{\mathbb{P}}(\mathcal{E})$, $(\mu_\omega)_{\omega\in_{\mathbb{P}}\Omega}$ be the associated disintegration with $\mu_\omega\in\mathcal{P}(M_\omega)$ a.s., and $\check{\mu} = \int_\Omega \mu_\omega d\mathbb{P}(\omega)$. We will consider that every measure appearing is atomless and fully supported.

For these random subshifts, we say that $\zeta\in M_\omega$ is p-periodic (for ω) precisely when it is p-periodic for σ.

Notice that it can be that not only $\sigma_\omega\equiv\sigma$ but also b and A are constant functions on Ω (in which case one could diminish M to $\{1,\dots,b\}^{\mathbb{N}_{\geqslant 0}}$). In particular, it would be M_ω identically M and $\sigma_\omega: M_\omega\to M_{\theta\omega}$ identically $\sigma: M\to M$. Still, $\hat{\mu}$ could be chosen so that the measures $(\mu_\omega)_{\omega\in_{\mathbb{P}}\Omega}$ vary. In this case, measures would be the only place in the system where randomness intervenes. See [20] for an example of this kind, the so-called random Markov shifts.

Let $C_1 = \{[s] : s\in\mathbb{N}_{\geqslant 1}\}$ be the set of 1-cylinders $[s]=\{x\in M : x_0 = s\}$ and, for $n\in\mathbb{N}_{\geqslant 1}\cup\{\infty\}$, $C_n = \bigvee_{j=0}^{n-1}\sigma^{-j}C_1$ be the set of n-cylinders $[s_0,\dots,s_{n-1}] = \{x\in M : x_0 = s_0,\dots x_{n-1}=s_{n-1}\}$ ($s_0,\dots,s_{n-1}\in\mathbb{N}_{\geqslant 1}$). The C_m's comprise a partition of M and, for $m\geqslant 1$, we denote by $C_m(x)$ the element in C_m containing $x\in M$. Let C_* be the σ-algebra generated by $\bigcup_{j\geqslant 1} C_j$. It is assumed that C_* generates $\mathscr{B}_M$ and that C_∞ is comprised of singletons. Notice that the objects in this paragraphs are all deterministic on the entire symbolic space M (and might be restricted to M_ω's when needed).

The targets that will be considered for these systems are given by a (deterministic) nested sequence of sets $U_m\in\sigma\left(\bigvee_{i=0}^{m-1}\sigma^{-i}C_1\right)$[4] ($m\geqslant 1$). The target will be considered small either in the annealed sense that $\check{\mu}(\Gamma)=0$, where $\Gamma := \bigcap_{m\geqslant 1} U_m$, or in quenched sense that $\operatorname{esssup}_\omega \mu_\omega(U_m)\searrow 0$ as $m\to\infty$. Another possibility is to consider singleton targets $\Gamma=\{\zeta\}$ and take $U_m = C_m(\zeta)$, which already carries a notion of smallness.

[4] The σ in the left stands for the σ-algebra generated by a given family of sets. The σ in the right stands for the shift map.

3.2.1 Point processes

Before we consider the literature studying hitting statistics using point processes, let us review some early contributions to related statistical properties for random dynamical systems. In [77], authors established, under an annealed time scaling, a quenched exponential law for first return time to cylinders about non-periodic (typical) points of random subshifts of finite type with super-polynomial decay of correlations, both quenched (for $(\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega}$) and annealed (for $\check{\mu}$). In a similar context, [78] studied quenched extremal indexes associated with p-periodic points ζ, which were found to be a.s.

$$\alpha = \lim_{m\to\infty} \frac{\check{\mu}(C_m(\zeta) \backslash C_{m+p}(\zeta))}{\check{\mu}(C_m(\zeta))},$$

provided that the limit exists.

A couple of references treat the random case using conditions of the type D or Д. Firstly, in [12], authors studied an additive noise absolutely continuous with respect to Lebesgue perturbing the action of a single map, in such a way that the system presents annealed polynomial decay of correlations. For a deterministic singleton target at an arbitrary point ζ, a deterministic function $\varphi(\cdot) = g \circ d(\cdot, \zeta)$, g as in equation (2.10), with the dynamically observed stochastic process defined on the product space as $Y_i(\omega, x) = \varphi(T^i_\omega(x))$ under $\hat{\mu}$ and adopting the annealed time scaling $m\hat{\mu}(\varphi > u_m(\tau)) \to \tau$, they prove that the extremal value is 1 and that a REPP of the type $N_m^{1,III,\tau}$ converges to a Poisson process with intensity τLeb.

Later, using conditions of the type Д, quenched extremal indexes and one-dimensional REPPs were studied, respectively, in [45] and [36]. Since they adopt a similar setup, we chose the latter for a brief discussion.

In [36], authors consider a system $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$ as in section 3.2, with θ a finite full shift, $\mathbb{P}$ an ergodic measure and $\omega \mapsto T_\omega := T_{\pi_0(\omega)}$, where the latter maps are the so-called Lasota-Yorke maps. These maps are piecewise expanding maps of $[0,1]$ with finitely many branches and bounded second derivatives. Expansion, number of branches, and second derivatives are all bounded uniformly in $(\Omega, [0,1])$. Moreover, they satisfy a uniform covering condition and a uniform regularization, for test functions inside a uniform cone. These properties imply they can be equipped with a quasi-invariant family of $(\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega}$ given by $\mu_\omega = h_\omega$Leb, with $\mathcal{L}_\omega(h_\omega) = h_{\theta\omega}$ and $\mathcal{L}^*_\omega$Leb = Leb, $\mathbb{P}$-a.s., where $\mathcal{L}_\omega$ is the transfer operator associated to T_ω acting on BV. This family satisfies a quenched uniform exponential decay of correlations of the form

$$\exists K > 0, \exists \iota \in (0,1), \text{ for } \mathbb{P}\text{-a.s.}, \forall n \geqslant 0, \forall g \in BV, \forall h \in L^1(\text{Leb}) :$$

$$\left| \int_{[0,1]} g \cdot (h \circ T_\omega) d\mu_\omega - \int_{[0,1]} g d\mu_\omega \cdot \int_{[0,1]} h d\mu_{\theta^n\omega} \right| \leqslant K\iota^n \|g\|_{BV} \|h\|_{L^1(\text{Leb})}.$$

In [30], however, authors point out that the above condition follows contingent on $(\theta, \mathbb{P})$ exhibiting at least polynomial decay of correlations.

Moreover, the target Γ is taken to be a deterministic singleton $\{\zeta\}$. The normalization (g, u) is taken with i) g satisfying the condition given equation (2.10) and $\varphi := g \circ d(\cdot, \Gamma)$, ii) u is given by

$$u_m(\tau) = \inf\{u \in \mathbb{R} : \check{\mu}(\varphi \leqslant u) \geqslant 1 - \tau/m\} \ (\tau > 0),$$

which implies

$$\sum_{j=0}^{m-1} \mu_\omega(\varphi \circ T_\omega^j > u_m(\tau)) = \sum_{j=0}^{m-1} \mu_{\theta^j\omega}(\varphi > u_m(\tau)) \xrightarrow{m} \tau,\ \mathbb{P}\text{-a.s. } (\tau > 0).$$

They consider target points that are generic for the Lebesgue measure and conclude that one-dimensional REPPs of the type $N_m^{1,III,\tau}$ converge to a Poisson point process with intensity $\tau\,\mathrm{Leb}$, whose associated extremal index is therefore $\alpha = 1$.

All the previous references in this section adopted an annealed space/time-normalization and rely on, at least, polynomial mixing/decay of correlations of annealed type or in the driving $(\theta, \mathbb{P})$. In [30], authors modify this situation by simultaneously adopting a quenched normalization and relying on a merely ergodic driving $(\theta, \mathbb{P})$, with mixing/decay conditions put only on the fiber maps T_ω's in a quenched form.

They consider random subshifts of finite type with target sets $(U_m)_{m\geqslant 1}$ small in the quenched sense. A certain condition on the family $(U_m)_{m\geqslant 1}$ is assumed and implies that it can not be $U_m = C_m(\zeta)$ for ζ a periodic point. They also assume that

i) $\exists \beta_0, \beta_1 : \mathbb{N}_{\geqslant 0} \to \mathbb{R}$, for $\mathbb{P}$-a.e.$\omega \in \Omega, \forall 0 \leqslant j \leqslant m \leqslant k$:

$$\mu_\omega(U_m \cap \sigma^{-j} U_m) \leqslant \mu_\omega(U_m)\beta_0(j),\ \mu_\omega(U_m \cap \sigma^{-k} U_m) \leqslant \mu_\omega(U_m)\beta_1(m),$$

and

ii) $\exists D : \Omega \to \mathbb{R}_{>0} \in \bigcup_{p\in(0,1]} L^p(\mathbb{P})$ and $\alpha : \mathbb{N}_{\geqslant 1} \to \mathbb{R}_{>0}$ decreasing to $0, \forall A, B \subset M$ measurable, $\forall m \geqslant 1, \forall m' \geqslant 1$:

$$|\mu_\omega(A \cap \sigma^{-m-m'} B) - \mu_\omega(A)\mu_{\theta^{m+m'}\omega}(B)| \leqslant D(\omega)\alpha(m'),$$

or

ii') $\exists \tilde{\alpha} : \mathbb{N}_{\geqslant 1} \to \mathbb{R}_{>0}$ decreasing to $0, \forall A, B \subset M$ measurable, $\forall m \geqslant 1, \forall m' \geqslant 1$:

$$|\mu_\omega(A \cap \sigma^{-m-m'} B) - \mu_\omega(A)\mu_{\theta^{m+m'}\omega}(B)| \leqslant \mu_\omega(A)\tilde{\alpha}(m').$$

They consider the one-dimensional REPP

$$N_m^\omega(x) := \sum_{i=1}^{\infty} \delta_{\sum_{j=1}^{i} \mu_{\theta^j\omega}(A_m)} \mathbb{1}_{U_m} \circ \sigma^i(x),$$

which can be considered a quenched generalization of $N_m^{1,II,1}$ (recall definition 2.2.13) — notice that $\frac{i}{[1/\mu(U)]}$, or $i\mu(U)$, is substituted by $\sum_{j=1}^{i} \mu_{\theta^j\omega}(U)$. With this definition in mind, the following quenched result is found

$$N^\omega_{m\,*}\mu_\omega \xrightarrow{v} N,\ \mathbb{P}\text{-a.s.},$$

where N is a Poisson process on $[0, \infty)$ with intensity Leb. Since the limiting point process is simple, their proof strategy resorts to Kallenberg criteria.

3.2.2 Spectral methods

The spectral approach presented in section 3.1.2 was generalized to random dynamical systems in the quenched sense. Theory in this context was developed for general potentials. Quenched extremal indexes were studied in [10] and quenched hitting statistics were studied in [8]. We follow mostly [8] in this exposition.

Let M be a compact Riemannian manifold with Lebesgue measure and let $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$ be a random system as in section 3.2. In this section, $\mathbb{P}$ is assumed ergodic.

Consider a random potential $\phi_0 : \Omega \times M \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R}$ which is measurable. For $n \geqslant 1$, let $\phi_0^{(n)}$: $\Omega \times M \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R}$ be given by $(\omega, x) \mapsto \sum_{j=0}^{n-1} \phi_0 \circ S^j(\omega, x)$, where the full $\mathbb{P}$-measure set can be taken the same for all n.

Let $(\mathcal{B}_\omega, \|\cdot\|_{\mathcal{B}_\omega})_{\omega \in_{\mathbb{P}} \Omega}$ be a family of Banach spaces and $(\mathcal{B}_\omega^*, \|\cdot\|_{\mathcal{B}_\omega^*})_{\omega \in_{\mathbb{P}} \Omega}$ be the associated family of dual spaces. Let $\mathcal{B} = \prod_{\omega \in_{\mathbb{P}} \Omega} \mathcal{B}_\omega$ be the space of $\mathbb{P}$-a.s. defined measurable maps from Ω to $\bigsqcup_{\omega \in_{\mathbb{P}} \Omega} \mathcal{B}_\omega$ whose image at ω belongs to $\mathcal{B}_\omega$ and, similarly, $\mathcal{B}^* = \prod_{\omega \in_{\mathbb{P}} \Omega} \mathcal{B}_\omega^*$. Assume that the $\mathcal{B}_\omega$'s are formed by $\mathbb{C}$-valued functions φ on M and the $\mathcal{B}_\omega^*$'s contain $\mathcal{P}(M)$. Moreover, assume that, for all $n \geqslant 0$, $e^{\phi_n} \in \mathcal{B}$, i.e., $e^{\phi_n(\omega,\cdot)} \in \mathcal{B}_\omega$, $\mathbb{P}$-a.s..

Assume that $(\mathcal{B}_\omega, \|\cdot\|_{\mathcal{B}_\omega})_{\omega \in_{\mathbb{P}} \Omega}$ accommodates the action of the Ruelle-Perron-Frobenius (RPF) operator

$$\begin{array}{rccccccc} \mathcal{L}_{\omega,0}: & \mathcal{B}_\omega & \to & \mathcal{B}_{\theta\omega} & & & & \\ & \varphi & \mapsto & \mathcal{L}_{\omega,0}\varphi & : & M & \to & \mathbb{R} \\ & & & & & x & \mapsto & \displaystyle\sum_{y \in (T_\omega)^{-1}(\{x\})} \varphi(y) e^{\phi_0(y)} \end{array}$$

$\mathbb{P}$-a.s.. In particular,

$$\begin{array}{rccccccc} \mathcal{L}^n_{\omega,0}: & \mathcal{B}_\omega & \to & \mathcal{B}_{\theta^n\omega} & & & & \\ & \varphi & \mapsto & \mathcal{L}^n_{\omega,0}\varphi & : & M & \to & \mathbb{R} \\ & & & & & x & \mapsto & \displaystyle\sum_{y \in (T^n_\omega)^{-1}(\{x\})} \varphi(y) e^{\phi_0^{(n)}(y)}, \end{array}$$

for all $n \geqslant 0$, and $\mathbb{P}$-a.e.$\omega \in \Omega$. These maps are lifted to the bundles as

$$\begin{array}{rccccccc} \mathcal{L}_0: & \mathcal{B} & \to & \mathcal{B} & & & & \\ & \varphi & \mapsto & \mathcal{L}_0\varphi & : & \Omega & \to & \mathcal{B}_\omega \\ & & & & & \omega & \mapsto & (\mathcal{L}_0\varphi)_\omega = \mathcal{L}_{\theta^{-1}\omega,0}\varphi_{\theta^{-1}\omega} \end{array}$$

$$\begin{array}{rccccccc} \mathcal{L}^n_0: & \mathcal{B} & \to & \mathcal{B} & & & & \\ & \varphi & \mapsto & \mathcal{L}_0\varphi & : & \Omega & \to & \mathcal{B}_\omega \\ & & & & & \omega & \mapsto & (\mathcal{L}_0\varphi)_\omega = \mathcal{L}_{\theta^{-n}\omega,0}\varphi_{\theta^{-n}\omega} \end{array}.$$

Moreover, it is assumed that, for every $\varphi \in \mathcal{B}$, the map $(\omega, x) \in \Omega \times M \overset{\mathbb{P}\text{-a.s.}}{\mapsto} (\mathcal{L}_0\varphi)_\omega(x) \in \mathbb{R}$ is measurable.

Notice that in the spectral approach one does not start with a given $\hat{\mu} = d\mu_\omega d\mathbb{P} \in \mathcal{P}_S^{\mathbb{P}}(\Omega \times M)$ and the associated $(\mu_\omega)_{\omega \in_{\mathbb{P}} \Omega} \in \mathcal{P}_{T_\Omega}^{(\mathbb{P})}(M)$. One considers that there exists

$$\varphi_0 : \Omega \times M \overset{\mathbb{P}\text{-a.s.}}{\to} (0,\infty),\ \nu_0 = \{\nu_{\omega,0}\}_{\omega \in_{\mathbb{P}} \Omega} \in \mathcal{P}^{(\mathbb{P})}(M)$$

and

$$\lambda_0 : \Omega \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R}\backslash\{0\} \text{ in } L^1(\mathbb{P})$$

so that

$$\mathcal{L}_{\omega,0}(\varphi_{\omega,0}) = \lambda_{\omega,0}\varphi_{\theta\omega,0},\ \mathcal{L}^*_{\omega,0}(\nu_{\theta\omega,0}) = \lambda_{\omega,0}\nu_{\theta\omega,0}.$$

They induce $\mu_0 = (\mu_{\omega,0})_{\omega \in_{\mathbb{P}} \Omega} \in \mathcal{P}_{T_\Omega}^{(\mathbb{P})}(M)$, where $\mu_{\omega,0}(\cdot) := \nu_\omega(\varphi_{\omega,0}(\cdot))$.

The target being considered is $\Gamma \in \mathscr{B}_\Omega \times \mathscr{B}_M$ so that, $\mathbb{P}$-a.s, $\Gamma(\omega)$ is small in the sense that $\nu_{\omega,0}(\Gamma(\omega)) = 0$. Moreover, consider $g : \Omega \times [0,\infty) \overset{\mathbb{P}\text{-a.s.}}{\to} [0,\infty)$ so that, for every ω a.s., $g(\omega,\cdot)$ is continuous, strictly decreasing near 0, where they attain a maximum. The normalization considers that given any a random scale

$$\tau : \Omega \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R}_{>0} \text{ in } L^\infty(\mathbb{P})$$

(we will not be varying this one, to make notation lighter), there exists

1. a random sequence of thresholds

$$u_m : \Omega \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R} \text{ in } L^\infty(\mathbb{P})\ (m \geqslant 1)$$

   with $u_m(\omega) \nearrow g_\omega(0)$, as $m \to \infty$, $\mathbb{P}$-a.s.,

2. a constant $W < \infty$ and a random sequence of margins

$$\xi_m : \Omega \overset{\mathbb{P}\text{-a.s.}}{\to} \mathbb{R} \text{ in } L^\infty(\mathbb{P})\ (m \geqslant 1)$$

   with $\lim_{m\to\infty} \xi_{\omega,m} = 0$, $\mathbb{P}$-a.s. and $|\xi_{\omega,m}| \leqslant W$ for all $m \geqslant 1$ and $\mathbb{P}$-a.s.

so that

$$m\mu_{\omega,0}(B_{g_\omega^{-1}(u_m(\omega))}(\Gamma(\omega))) = t_\omega + \xi_{\omega,m}.$$

Notice that radii and scales are random and quenched. Notice also that $\xi_{\omega,m}$ mediates between the exact scaling and the asymptotic one. These conditions make the normalization applied in this theory very general, at least from the quenched randomization point of view.

Instead of the perturbations designed to study quenched extremal index in [10], namely

$$\begin{array}{rccc} \mathcal{L}_{\omega,m} : & \mathcal{B}_\omega & \to & \mathcal{B}_{\theta\omega} \\ & \varphi & \mapsto & \mathcal{L}_{\omega,0}(\mathbb{1}_{M\backslash B_{g_\omega^{-1}(u_m(\omega))}(\Gamma(\omega))}\varphi) \end{array} \quad (m \geqslant 1),$$

to study hitting statistics with the aid of characteristic functions, following [8], one actually introduces the following family of perturbations

$$\begin{array}{rccc} \mathcal{L}_{\omega,m,s} : & \mathcal{B}_\omega & \to & \mathcal{B}_{\theta\omega} \\ & \varphi & \mapsto & \mathcal{L}_{\omega,0}(e^{is\mathbb{1}_{B_{g_\omega^{-1}(u_m(\omega))}}(\Gamma(\omega))}\varphi) \end{array} \quad (s \in \mathbb{R}, m \geqslant 1).$$

$$\begin{array}{rccc} \mathcal{L}^L_{\omega,m,s}: & \mathcal{B}_\omega & \to & \mathcal{B}_{\theta\omega} \\ & \varphi & \mapsto & \mathcal{L}_{\omega,0}(e^{is\tilde{Z}^L_{\omega,m}}\varphi) \end{array} \quad (s \in \mathbb{R}, m \geqslant 1, L \geqslant 1),$$

where $\tilde{Z}^L_{\omega,m}(x) = \sum_{j=0}^{L-1} \mathbb{1}_{B_{g^{-1}_{\theta^j\omega}(u_m(\theta^j\omega))}}(\Gamma(\theta^j\omega)) \circ T^j_\omega(x)$.

We omit details from this point on. The argument follows the structure of that in section 3.1.2, but adapting for the fact that the unperturbed eigenvalue is not anymore 1 and that randomization should be accounted for: in terms of section 3.1.2, quenched versions of conditions (1), (i-iii) and (3) are adopted. Notice that the quenched version (iii) implies quenched exponential decay of correlations for suitable classes of observables.

Based on these conditions, quenched first-order approximations of the perturbed eigenvalues $(\lambda_{\omega,m,s})_{m\geqslant 1}$ are employed. These approximations are based on a quenched perturbative result developed in [10] (theorem 2.1.2). This result is a randomization of [60], however, reading their theorem 2.1.2, one will notice that [10] also adopt hypotheses (1), (i-iii) and (3), whereas, as we have seen in section 3.1.2, [60] adopts (1), (2.1) (included in ii), (2.2), (2.3) and (3), and then argues that (2.1-2.3) implies (i-iii), due to the stability theory developed in [61] and [59].

However, when dealing with applications, both [8] (lemma 4.6) and [10] (2.5.10) use a version of "(2.1-2.3) implies (i-iii)". The quenched version of this implication is due to [29], which randomizes [61]. In this context, one deals with cocycles of operators rather than compositions of a single operator, and many concepts have to be suitably adapted. This translation can be found in a systematic way in the semi-invertible multiplicative ergodic theory literature (see [46] and [50]): e.g., eigenvalues become Lyapunov exponents, eigenspaces become Oseledets spaces, quasi-compactness has to be suitably adapted (see [85]), and so on.

Based on systems satisfying the assumptions reviewed so far and adapting the logic used to get equation (3.13) out of equations (3.11) and (3.12), [8] concludes that for each $s \in \mathbb{R}\backslash\{0\}$ and $\mathbb{P}$-a.e.$\omega \in \Omega$.

$$\lim_{m\to\infty} \mu_{\omega,0}(e^{is\tilde{Z}^m_{\omega,m}}) = \lim_{m\to\infty} \frac{\lambda^m_{\omega,m,s}}{\lambda^m_{\omega,0}} = \exp\left(-(i - e^{is}) \int_\Omega \alpha_\omega(s)\tau_\omega d\mathbb{P}(\omega)\right),$$

where

$$\alpha_\omega(s) = \lim_{m\to\infty} \frac{\lambda_{\omega,0} - \lambda_{\omega,m,s}}{\Delta_{\omega,m}(s)} = \lim_{m\to\infty} \frac{\lambda_{\omega,0} - \lambda_{\omega,m,s}}{\lambda_{\omega,0}(1 - e^{is})\mu_{\omega,0}(\Gamma_{g^{-1}_\omega(u_m(\omega))})}, \tag{3.15}$$

for which there exist more explicit formulas similar to those in equation 3.10 (see [8] equation 3.28).

It is interesting to notice that the limit is of an annealed type despite the very general quenched approach and the driving merely ergodic. The second equality in equation (3.15) is stated in [8] theorem 3.14, whose proof refers to [11] theorem 2.4.5, where, very succinctly, one can notice that $\frac{\lambda_{\omega,m,s}}{\lambda_{\omega,0}}$ can be expressed in a way as to involve iterates of ω under θ and $\frac{\lambda^m_{\omega,m,s}}{\lambda^m_{\omega,0}}$ will sum over these terms, in which case a non-standard ergodic theorem can be applied.

Again they argue that the pointwise limit of the characteristic function found in equation (3.15) is discrete and infinitely divisible, therefore the limiting random variable is compound Poisson distributed. Its distribution can be expressed using Levy's inversion formula, giving

$$\lim_{m\to\infty} \mu_{\omega,0}(\tilde{Z}^m_{\omega,m} = n) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} e^{-isk} \exp\left(-(1 - e^{is}) \int_\Omega \alpha_\omega(s)\tau_\omega d\mathbb{P}(\omega)\right).$$

Once again, as in section 3.1.2, one could derive expressions for the intensity and cluster size distribution of the underlying compound Poissonian limit, using associated characteristic functions (see their equation (3.36) and proposition 3.18).

As it comes to applications, [8] considers certain systems that draw among at most countably many piecewise expanding maps of the interval with finite-many branches. The drawing procedure is driven by $\mathbb{P}$ ergodic, but explicit calculations are carried out for Bernoulli measures. Limit behaviors can either be standard Poisson processes, compound Poisson processes with geometric multiplicity, or compound Poisson processes with non-geometric multiplicity.

3.2.3 Probabilistic approximation

In this section, we briefly comment [52]. They study random subshifts of finite type with ψ-quenched and α-annealed mixing to get, under annealed time-scaling, compound Poissonian hitting statistics to targets at p-periodic points ζ.

The mixing assumptions spell as follows

i) $\exists \psi : \mathbb{N}_{\geqslant 1} \to \mathbb{R}, \lim_{k\to\infty} \psi(k) = 0, \forall m, m' \geqslant 1, \forall U \in C_m, \forall V \in C_*$, for $\mathbb{P}$-a.e.$\omega \in \Omega$:

$$|\mu_\omega(U \cap \sigma^{-m-m'}V) - \mu_\omega(A)\mu_{\theta^{m+m'}\omega}(V) \leqslant \psi(m')\mu_\omega(A)\mu_{\theta^{m+m'}\omega}(V).$$

ii) $\exists \alpha : \mathbb{N}_{\geqslant 1} \to \mathbb{R}, \lim_{k\to\infty} \psi(k) = 0, \forall m, m' \geqslant 1, \forall U \in C_m, \forall V \in C_*$:

$$|\check{\mu}(U \cap \sigma^{-m-m'}V) - \check{\mu}(A)\check{\mu}(V) \leqslant \alpha(m').$$

Moreover, they assume additional conditions to guarantee that cylinders are small enough. Then, using an ad hoc approximation method (which again, as in section 3.1.3, is bounded with short and long-range components), and denoting

$$Z_m^{\lfloor \frac{\tau}{\check{\mu}(C_m(\zeta))} \rfloor} := \sum_{i=1}^{\lfloor \frac{\tau}{\check{\mu}(C_m(\zeta))} \rfloor} \mathbb{1}_{C_m(\zeta)} \circ \sigma^i$$

they conclude that

$$(Z_m^{\lfloor \frac{\tau}{\check{\mu}(C_m(\zeta))} \rfloor})_* \mu_\omega \xrightarrow{v} CPD_{\tau(1-\vartheta),(Geo_{1-\vartheta}(\ell))_{\ell\geqslant 1}},$$

where

$$\vartheta = \lim_{m\to\infty} \frac{\check{\mu}(C_{m+p}(\zeta))}{\check{\mu}(C_m(\zeta))},$$

provided that the limit exists. In this case, the quenched extremal index would be written as

$$\alpha = 1 - \lim_{m\to\infty} \frac{\check{\mu}(C_{m+p}(\zeta))}{\check{\mu}(C_m(\zeta))} = \lim_{m\to\infty} \frac{\check{\mu}(C_m(\zeta) \backslash C_{m+p}(\zeta))}{\check{\mu}(C_m(\zeta))}.$$

Chapter 4

Compound Poisson distributions for random dynamical systems: a probabilistic block-approximation approach

4.1 Assumptions and main results

4.1.1 General setup

Consider M a complete separable metric space, and Ω, the so-called driving space, a complete separable metric space equipped with a measurably-invertible ergodic system $(\theta,\mathbb{P})$.

Consider maps $T_\omega : M \to M$ ($\omega \in \Omega$) which combine to make the a measurable skew product $S : \Omega \times M \to \Omega \times M$, $(\omega,x) \mapsto (\theta\omega, T_\omega x)$. As usual, for higher-order iterates we denote $S^n(\omega,x) = (\theta^n\omega, T^n_\omega(x))$ where $T^n_\omega = T_{\theta^{n-1}\omega} \circ \cdots \circ T_{\theta\omega} \circ T_\omega$ ($n \geqslant 1$).

Denote

$$\mathcal{P}^{\mathbb{P}}(\Omega \times M) = \{\hat{\mu} \in \mathcal{P}(\Omega \times M) : \hat{\mu}, \pi_{\Omega*}\hat{\mu} = \mathbb{P}\},$$

$$\mathcal{P}^{\mathbb{P}}_S(\Omega \times M) = \{\hat{\mu} \in \mathcal{P}(\Omega \times M) : S_*\hat{\mu} = \hat{\mu}, \pi_{\Omega*}\hat{\mu} = \mathbb{P}\},$$

and

$$\mathcal{P}^{(\mathbb{P})}(M) = \left\{ \begin{array}{c} \mu : \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega \in \mathcal{P}(M) \text{ so that:} \\ \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega(E_\omega) \in [0,1] \text{ is } (\mathscr{B}_\Omega, \mathscr{B}_{[0,1]})\text{-measurable, } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M \end{array} \right\},$$

$$\mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M) = \left\{ \begin{array}{c} \mu : \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega \in \mathcal{P}(M) \text{ so that:} \\ \omega \in \Omega \overset{\mathbb{P}\text{-a.s.}}{\mapsto} \mu_\omega(E_\omega) \in [0,1] \text{ is } (\mathscr{B}_\Omega, \mathscr{B}_{[0,1]})\text{-measurable, } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M \\ T^n_{\omega*}\mu_\omega = \mu_{\theta^n\omega}, \forall n \geqslant 0, \mathbb{P}\text{-a.s.} \end{array} \right\}.$$

Notation. Elements in the latter two sets will be written as $\mu = (\mu_\omega)_\omega$, where the outer '$\omega$' subscript (instead of '$\omega \in \Omega$') is to identify that the given family if defined $\mathbb{P}$-a.s.. The underlying full measure subset Ω_0 can be assumed to be forward and backward θ-invariant (otherwise we substitute it by $\bigcap_{n\in\mathbb{Z}} \theta^n\Omega_0$).

A family of measures $(\mu_\omega)_\omega$ satisfying "$T^n_{\omega *}\mu_\omega = \mu_{\theta^n\omega}$, $\forall n \geqslant 0$, $\mathbb{P}$-a.s." is called a covariant family.

For the next paragraph, we refer to [28] (prop. 3.3) and [5] (sec. 1.4). For $E \in \mathscr{B}_\Omega \times \mathscr{B}_M$, for any $\omega \in \Omega$, its ω-section, $\{x \in M : (\omega, x) \in E\}$, is denoted by E_ω or $E(\omega)$.

Using Rohklin disintegration theorem for $\hat{\mu} \in \mathcal{P}^{\mathbb{P}}(\Omega \times M)$ with respect to the partition $\mathscr{P} = \{[\omega] := \{\omega\} \times M \mid \omega \in \Omega\}$ of $\Omega \times M$, we have that there exists $(\mu_\omega)_\omega$ so that:

$$\text{i) } \forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M\text{: } \hat{\mu}(E) = \int_\Omega \mu_\omega(E_\omega) d\mathbb{P}(\omega), \text{ ii) } \begin{array}{ccc} \mathcal{P}^{\mathbb{P}}(\Omega \times M) & \hookrightarrow & \mathcal{P}^{(\mathbb{P})}(M) \\ \mathcal{P}^{\mathbb{P}}_S(\Omega \times M) & \hookrightarrow & \mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M) \\ \hat{\mu} & \mapsto & (\mu_\omega)_\omega \end{array}.$$

Conversely,

$$\begin{array}{ccl} \mathcal{P}^{(\mathbb{P})}(M) & \hookrightarrow & \mathcal{P}^{\mathbb{P}}(\Omega \times M) \\ \mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M) & \hookrightarrow & \mathcal{P}^{\mathbb{P}}_S(\Omega \times M) \\ (\mu_\omega)_\omega & \mapsto & \hat{\mu} = d\mu_\omega d\mathbb{P}(\omega), \text{ i.e., } \hat{\mu}(E) = \int_\Omega \mu_\omega(E_\omega) d\mathbb{P}(\omega) \ (\forall E \in \mathscr{B}_\Omega \times \mathscr{B}_M) \end{array}.$$

Now, consider a given $\hat{\mu} = d\mu_\omega d\mathbb{P}(\omega) \in \mathcal{P}^{\mathbb{P}}_S(\Omega \times M)$, with the associated $(\mu_\omega)_\omega \in \mathcal{P}^{(\mathbb{P})}_{T_\Omega}(M)$. Define the marginal measure $\check{\mu} = \pi_{M*}\hat{\mu} = \int_\Omega \mu_\omega \, d\mathbb{P}(\omega) \in \mathcal{P}(M)$.

Finally, consider $\Gamma \in \mathscr{B}_\Omega \times \mathscr{B}_M$ so that, $\mathbb{P}$-a.s, $\Gamma(\omega)$ is compact and small in the sense that $\mu_\omega(\Gamma(\omega)) = 0$. The set Γ is the so-called random target. Denote $\Gamma_\rho(\omega) = B_\rho(\Gamma(\omega))$ ($\rho > 0$) and the corresponding ω-collection by Γ_ρ.

The objects considered above comprise what we call a 'targeted random dynamical system', or simply 'system', to be denoted by the tuple $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$.

Let's define some working objects. Let $U \in \mathscr{B}_\Omega \times \mathscr{B}_M$ be a set whose ω-sections $U(\omega) \subset M$ have positive μ_ω-measure, $\mathbb{P}$-a.s.. We again can consider that the said underlying set full measure subset is θ-invariant, so we can consider that $\mu_{\theta^n\omega}(U(\theta^n\omega)) > 0$, for all $n \in \mathbb{Z}$, $\mathbb{P}$-a.s..

Definition 4.1.1. The **first hitting time** of $(\theta, \mathbb{P}, T_\omega, \mu_\omega, U)$ is the family of functions

$$\begin{array}{cccl} r_U^{\omega,1} : & M & \to & \mathbb{N}_{\geqslant 1} \cup \{\infty\} \\ & x & \mapsto & \inf\{i \in \mathbb{N}_{\geqslant 1} : T^i_\omega(x) \in U(\theta^i\omega)\} \end{array}.$$

The associated **higher-order hitting times** are given, for $\ell \geqslant 2$, by the family of functions

$$\begin{array}{cccl} r_U^{\omega,\ell} & : M & \to & \mathbb{N}_{\geqslant \ell} \cup \{\infty\} \\ & x & \mapsto & r_U^{\omega,\ell}(x) = r_U^{\omega,\ell-1}(x) + r_U^{\omega'}(T_\omega^{r_U^{\omega,\ell-1}}(x)) \end{array},$$

where $\omega' = \theta^{r_U^{\omega,\ell-1}(x)}\omega$.

Definition 4.1.2. The **hit counting function** of $(\theta,\mathbb{P},T_\omega,\mu_\omega,U)$, for $L \geqslant 1$ is given by the family of functions

$$\begin{aligned} Z_{*U}^{\omega,L} : & M \to \mathbb{N}_{\geqslant 0} \\ & x \mapsto \sum_{i=1}^{L} \mathbb{1}_{U(\theta^i\omega)} \circ T_\omega^i(x) \end{aligned}, \qquad \begin{aligned} Z_{U}^{\omega,L} : & M \to \mathbb{N}_{\geqslant 0} \\ & x \mapsto \sum_{i=0}^{L-1} \mathbb{1}_{U(\theta^i\omega)} \circ T_\omega^i(x) \end{aligned}.$$

These objects are related, for example, in the sense that $\{Z_{*U}^{\omega,L} \geqslant \ell\} = \{r_U^{\omega,\ell} \leqslant L\}$, $\{Z_{*U}^{\omega,L} = \ell\} = \{r_U^{\omega,\ell} \leqslant L < r_U^{\omega,\ell+1}\}$. When $U = \Gamma_\rho$, we write $I_i^{\omega,\rho} = \mathbb{1}_{\Gamma_\rho(\theta^i\omega)} \circ T_\omega^i$.

4.1.2 Working setup

Notation. A $\mathbb{R}$-valued function defined on the product space, $f(\omega,x)$, is often rewritten with the random seed in the sup/subscript, like $f^\omega(x)$ or $f_\omega(x)$, which can be seen as an Ω-family of functions defined on M. And vice versa. When integrating a function, we may simply omit the dummy variable of integration, even if it is a sup/subscript. We leave it for the reader to infer what variables and parameters are being integrated and were omitted. Some examples:

i) $\hat{\mu}(f) = \int_{\Omega\times M} f(\omega,x)d\hat{\mu}(\omega,x) = \int_{\Omega\times M} f_\omega(x)d\hat{\mu}(\omega,x) = \int_\Omega \mu_\omega(f_\omega)d\mathbb{P}(\omega)$, with $\mu_\omega(f_\omega) = \int_M f_\omega(x)d\mu_\omega(x) = \int_M f(\omega,x)d\mu_\omega(x)$;

ii) $\hat{\mu}(Z_{\Gamma_\rho}^L) = \int_{\Omega\times M} Z_{\Gamma_\rho}^{\omega,L}(x)d\hat{\mu}(\omega,x) = \int_\Omega \mu_\omega(Z_{\Gamma_\rho}^{\omega,L})d\mathbb{P}(\omega)$, with $\mu_\omega(Z_{\Gamma_\rho}^{\omega,L}) = \int_M Z_{\Gamma_\rho}^{\omega,L}(x)d\mu_\omega(x)$.

If the aforementioned f is $\{0,1\}$-valued we identify it with the set $F = f^{-1}(\{1\})$, since $f = \mathbb{1}_F$, whereas its partials f_ω are identified with the ω-sections of F, denoted F_ω, since $f_\omega = \mathbb{1}_{F_\omega}$. And vice versa. So, instead of i), we could write i') $\hat{\mu}(F) = \int_{\Omega\times M} \mathbb{1}_F(\omega,x)d\hat{\mu}(\omega,x) = \int_{\Omega\times M} \mathbb{1}_{F_\omega}(x)d\hat{\mu}(\omega,x) = \int_M \mu_\omega(F_\omega)d\mathbb{P}(\omega)$, with $\mu_\omega(F_\omega) = \int_M \mathbb{1}_{F_\omega}(x)d\mu_\omega(x) = \int_M \mathbb{1}_F(\omega,x)d\mu_\omega(x)$.

Notation. We write $\lim_{L\to\infty} \overline{\underline{\lim}}_{\rho\to 0} a(L,\rho)$ for the coinciding value of $\lim_{L\to\infty} \overline{\lim}_{\rho\to 0} a(L,\rho)$ and $\lim_{L\to\infty} \underline{\lim}_{\rho\to 0} a(L,\rho)$, when they do exist and coincide. We also denote $\overset{+}{a}(L) := \overline{\lim}_{\rho\to 0} a(L,\rho)$ and $\overset{-}{a}(L) := \underline{\lim}_{\rho\to 0} a(L,\rho)$ whenever theses are finite quantities.

Notation. Consider non-negative sequences $a(n)$ and $b(n)$ $(n \geqslant 0)$. we will write $a(n) \lesssim_n b(n)$ to mean that there exists a quantity C, independent of n, so that $a(n) \leqslant Cb(n)$ $(\forall n \geqslant 0)$.

When the functions a and b have more than one argument, it is necessary to indicate which of them are controlled uniformly: for example, with $a(n,m)$ and $b(n,m)$ $(n,m \geqslant 1)$,

- $a(n,m) \lesssim_n b(n,m)$ means that there exists quantities C_m $(m \geqslant 0)$, independent of n, so that $a(n,m) \leqslant C_m b(n,m)$ $(\forall n,m \geqslant 0)$;
- $a(n,m) \lesssim_{n,m} b(n,m)$ means that there exists a quantity C, independent of n and m, so that $a(n,m) \leqslant Cb(n,m)$ $(\forall n,m \geqslant 0)$.

In the context where some of the arguments are taken to the limit, we implicitly consider that these are the ones being controlled uniformly and we omit the associated subscript from the $\lesssim$ symbol.

We also employ the usual big-O and little-o notation.

Let's start introducing a couple of new objects: those with a λ will be associated with hitting statistics, and those with an α will be associated with return statistics. Whenever the following limits exist (and the appropriate ones coincide), denote, for $\ell \geqslant 1$ and $\omega \in \Omega$:

I)

$$\lambda_\ell^\omega = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\lambda_\ell^\omega(L,\rho)$$

where

$$\lambda_\ell^\omega(L,\rho) = \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} = \ell | Z_{\Gamma_\rho}^{\omega,L} > 0) = \frac{\mu_\omega(Z_{\Gamma_\rho}^{\omega,L} = \ell)}{\mu_\omega(Z_{\Gamma_\rho}^{\omega,L} > 0)}. \tag{4.1}$$

II) [1]

$$\lambda_\ell = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\lambda_\ell(L,\rho)$$

where

$$\lambda_\ell(L,\rho) = \hat{\mu}(Z_{\Gamma_\rho}^{L} = \ell | Z_{\Gamma_\rho}^{L} > 0) = \frac{\hat{\mu}(Z_{\Gamma_\rho}^{L} = \ell)}{\hat{\mu}(Z_{\Gamma_\rho}^{L} > 0)} \tag{4.2}$$

$$= \frac{\int_\Omega \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} = \ell)d\mathbb{P}(\omega)}{\int_\Omega \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} > 0)d\mathbb{P}(\omega)} = \int_\Omega \lambda_\ell^\omega(L,\rho)\frac{\mu_\omega(Z_{\Gamma_\rho}^{\omega,L} > 0)}{\int_\Omega \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} > 0)d\mathbb{P}(\omega)}d\mathbb{P}(\omega).$$

III)

$$\hat{\alpha}_\ell^\omega = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\hat{\alpha}_\ell^\omega(L,\rho)^2$$

where

$$\hat{\alpha}_\ell^\omega(L,\rho) = \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell | I_0^{\omega,\rho} = 1) = \frac{\mu_\omega(Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell, I_0^{\omega,\rho} = 1)}{\mu_\omega(\Gamma_\rho(\omega))} \tag{4.3}$$

IV)

$$\alpha_\ell^\omega = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\alpha_\ell^\omega(L,\rho)$$

where

$$\alpha_\ell^\omega(L,\rho) = \mu_\omega(Z_{\Gamma_\rho}^{\omega,L} = \ell | I_0^{\omega,\rho} = 1) = \frac{\mu_\omega(Z_{\Gamma_\rho}^{\omega,L} = \ell, I_0^{\omega,\rho} = 1)}{\mu_\omega(\Gamma_\rho(\omega))}. \tag{4.4}$$

Since $\{Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell\} \supset \{Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell+1\}$ and $\{Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell\}\setminus\{Z_{\Gamma_\rho}^{\omega,L} \geqslant \ell+1\} = \{Z_{\Gamma_\rho}^{\omega,L} = \ell\}$, then

$$\hat{\alpha}_\ell^\omega(L,\rho) - \hat{\alpha}_{\ell+1}^\omega(L,\rho) = \alpha_\ell^\omega(L,\rho). \tag{4.5}$$

which entails that the existence of $\hat{\alpha}_\ell^\omega$'s implies that of the α_ℓ^ω's with $\alpha_\ell^\omega = \hat{\alpha}_\ell^\omega - \hat{\alpha}_{\ell+1}^\omega$, because:

$$\lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\hat{\alpha}_\ell^\omega(L,\rho) - \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\hat{\alpha}_{\ell+1}^\omega(L,\rho) \leqslant \lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\alpha_\ell^\omega(L,\rho)$$

$$\lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\alpha_\ell^\omega(L,\rho) \leqslant \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\hat{\alpha}_\ell^\omega(L,\rho) - \lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\hat{\alpha}_{\ell+1}^\omega(L,\rho).$$

[1] A comparison between $(\lambda_\ell)_{\ell\geqslant 1}$ and $(\lambda_q(n))_{n\geqslant 1}$ defined in theorem 3.1.3 is worthwhile.

[2] Notice that, by L-monotonicity, the outer limits always exist provided that the inner ones do.

V)

$$\hat{\alpha}_\ell = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\hat{\alpha}_\ell(L,\rho)^3,$$

where

$$\hat{\alpha}_\ell(L,\rho) = \hat{\mu}(Z^L_{\Gamma_\rho} \geqslant \ell | I^\rho_0 = 1) = \frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant \ell, I^\rho_0 = 1)}{\hat{\mu}(\Gamma_\rho)} \tag{4.6}$$

$$= \frac{\int_\Omega \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} \geqslant \ell, I^{\omega,\rho}_0 = 1)d\mathbb{P}(\omega)}{\int_\Omega \mu_\omega(\Gamma_\rho(\omega))d\mathbb{P}(\omega)} = \int_\Omega \hat{\alpha}^\omega_\ell(L,\rho)\frac{\mu_\omega(\Gamma_\rho(\omega))}{\int_\Omega \mu_\omega(\Gamma_\rho(\omega))d\mathbb{P}(\omega)}d\mathbb{P}(\omega).$$

VI)

$$\alpha_\ell = \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\alpha_\ell(L,\rho)$$

where

$$\alpha_\ell(L,\rho) = \hat{\mu}(Z^L_{\Gamma_\rho} = \ell | I^\rho_0 = 1) = \frac{\hat{\mu}(Z^L_{\Gamma_\rho} = \ell, I^\rho_0 = 1)}{\hat{\mu}(\Gamma_\rho)} \tag{4.7}$$

$$= \frac{\int_\Omega \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} = \ell, I^{\omega,\rho}_0 = 1)d\mathbb{P}(\omega)}{\int_\Omega \mu_\omega(\Gamma_\rho(\omega))d\mathbb{P}(\omega)} = \int_\Omega \alpha^\omega_\ell(L,\rho)\frac{\mu_\omega(\Gamma_\rho(\omega))}{\int_\Omega \mu_\omega(\Gamma_\rho(\omega))d\mathbb{P}(\omega)}d\mathbb{P}(\omega).$$

Since $\{Z^L_{\Gamma_\rho} \geqslant \ell\} \supset \{Z^L_{\Gamma_\rho} \geqslant \ell+1\}$ and $\{Z^L_{\Gamma_\rho} \geqslant \ell\}\backslash\{Z^L_{\Gamma_\rho} \geqslant \ell+1\} = \{Z^L_{\Gamma_\rho} = \ell\}$, then

$$\hat{\alpha}_\ell(L,\rho) - \hat{\alpha}_{\ell+1}(L,\rho) = \alpha_\ell(L,\rho). \tag{4.8}$$

which entails that the existence of $\hat{\alpha}_\ell$'s implies that of the α_ℓ's with $\alpha_\ell = \hat{\alpha}_\ell - \hat{\alpha}_{\ell+1}$, because:

$$\lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\hat{\alpha}_\ell(L,\rho) - \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\hat{\alpha}_{\ell+1}(L,\rho) \leqslant \lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\alpha_\ell(L,\rho)$$

$$\lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\alpha_\ell(L,\rho) \leqslant \lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\,\hat{\alpha}_\ell(L,\rho) - \lim_{L\to\infty}\underline{\lim_{\rho\to 0}}\,\hat{\alpha}_{\ell+1}(L,\rho).$$

Now we upgrade the general setup of section 4.1.1. To optimize for generality, we present in abstract terms the conditions which are required from the systems we will work with. In concrete examples, these conditions need to be verified, but one should keep in mind that they are conceived to accommodate non-uniformly expanding behavior and random targets that do not overlap very badly with the regions where uniformity breaks.

On top of the features prescribed to the objects in our system throughout section 4.1.1, we will consider the following hypotheses.

H1 (Ambient)**.** Let M be a compact Riemannian manifold and Ω a compact metric space.

H2 (Invertibility features)**.**

[3] See footnote 2.

2.1 (Degree)**.** $\forall\omega\in\Omega, \forall n\geqslant 1, \forall x\in M : \#(T_\omega^n)^{-1}(\{x\})<\infty$ with

$$\sup_{n\geqslant 0}\#(T_\omega^n)^{-1}(\{x\})\leqslant\infty\ (\forall\omega,x),\ \sup_{\omega\in\Omega}\#(T_\omega^n)^{-1}(\{x\})\leqslant\infty\ (\forall n,x),\ \sup_{x\in M}\#(T_\omega^n)^{-1}(\{x\})<\infty\ (\forall\omega,n).$$

2.2 (Covering)**.** $\exists R>0, \mathcal{N}\geqslant 1, \forall\omega\in\Omega, \forall n\geqslant 1, \exists(y_k^{\omega,n})_{k\in K_{\omega,n}}\subset M$ with $\#K_{\omega,n}<\infty$ so that $(B_R(y_k^{\omega,n}))_{k\in K_{\omega,n}}$ has at most $\mathcal{N}$ overlaps.

Terminology suggests that $(B_R(y_k^{\omega,n}))_{k\in K_{\omega,n}}$ covers M entirely, but a small defect is allowed, in the sense of (H2.5) below.

2.3 (Inverse branches)**.** $\forall\omega\in\Omega, \forall n\geqslant 1, \forall k\in K_{\omega,n}$,

$$\mathrm{IB}_k^{\omega,n}=\{\varphi : B_R(y_k^{\omega,n})\to M \text{ diffeomorphic onto its image with } T_\omega^n\circ\varphi=\mathrm{id}\}$$

is non-empty, finite[4] and so that $\varphi,\psi\in\mathrm{IB}_k^{\omega,n}, \varphi\neq\psi\Rightarrow\varphi(\mathrm{dom}(\varphi))\cap\psi(\mathrm{dom}(\psi))=\varnothing$. In particular, the set $\mathrm{IB}(T_\omega^n)=\bigcup_{k\in K_{\omega,n}}\mathrm{IB}_k^{\omega,n}$ is finite and so that $\varphi,\psi\in\mathrm{IB}(T_\omega^n), \mathrm{dom}(\varphi)\cap\mathrm{dom}(\psi)=\varnothing\Rightarrow\varphi(\mathrm{dom}(\varphi))\cap\psi(\mathrm{dom}(\psi))=\varnothing$.

The following item is a consequence of the previous ones, but we list it here for convenience.

2.4 (Cylinders)**.** $\forall\omega\in\Omega, \forall n\geqslant 1$, $C_n^\omega=\{\xi=\varphi(\mathrm{dom}(\varphi)) : \varphi\in\mathrm{IB}(T_\omega^n)\}$ is finite and has at most $\mathcal{N}$ overlaps.

2.5 (Large covering)**.** For $\mathbb{P}$-a.e. $\omega\in\Omega$, $\forall n\geqslant 1$, $\mu_\omega\left(M\backslash\bigcup_{\xi\in C_n^\omega}\xi\right)=0$.

2.6 (Non-degenerancy)**.** $\exists\iota>0$ so that

$$\operatorname*{ess\,inf}_{\omega\in\Omega}\inf_{n\geqslant 1}\inf_{k\in K_{\omega,n}}\mu_{\theta^n\omega}(B_R(y_k^{\omega,n}))>\iota.$$

Next, we consider that the aforementioned (plain) cylinders are refined enough as to split and distinguish regions with different hyperbolic behavior.

H3 (Hyperbolicity and cylinders)**.** Plain cylinders split into acceptable (and unacceptable) cylinders, whereas acceptable cylinders subsplit into good (and bad) cylinders.
Namely: $\forall\omega\in\Omega, \forall n\geqslant 1 : C_\omega^n=\overset{+}{C_n^\omega}\sqcup\overset{-}{C_n^\omega}, \overset{+}{C_n^\omega}=\overset{++}{C_n^\omega}\sqcup\overset{+-}{C_n^\omega}$, making measurable

$$\overset{*}{\mathcal{C}_n}(\omega,x)=\begin{cases}1, x\in\bigcup_{\xi\in\overset{*}{C_n^\omega}}\xi,\\ 0,\ \text{otherwise}\end{cases}\qquad(*\in\{+,-,++,+-\}).$$

Notation. For $*\in\{+,-,++,+-\}$, write $\overset{*}{\mathrm{IB}}(T_\omega^n)=\{\varphi\in\mathrm{IB}(T_\omega^n) : \xi=\varphi(\mathrm{dom}(\varphi))\in\overset{*}{C_n^\omega}\}$.

This splitting distinguishes hyperbolic behavior in the sense of satisfying:

3.1 (Weak hyperbolicity on plain cylinders)**.** $\forall n\geqslant 1$:

$$1\leqslant\inf_{\omega\in\Omega}\inf_{\xi\in C_n^\omega}\inf_{\substack{v\in T_xM\\ \|v\|=1}}|DT_\omega^n(x)v|\leqslant\sup_{\omega\in\Omega}\sup_{\xi\in C_n^\omega}\sup_{x\in\xi}\sup_{\substack{v\in T_xM\\ \|v\|=1}}|DT_\omega^n(x)v|\leqslant\infty.$$

[4]Cardinalities behave as in (H2.1).

3.2 (Bounded derivatives on acceptable cylinders)**.** $\forall n \geqslant 1$:

$$\sup_{\omega\in\Omega}\sup_{\xi\in C_n^\omega}\sup_{x\in\xi}\sup_{\substack{v\in T_xM\\ \|v\|=1}} |DT_\omega^n(x)v| =: a_n < \infty.$$

3.3 (Distortion on good cylinders)**.** $\exists \mathfrak{d} \geqslant 0, \exists C > 1, \forall n \geqslant 1$: (denoting $\xi = \varphi(\mathrm{dom}(\varphi))$)

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\ \sup_{\varphi\in \overset{++}{\mathrm{IB}}(T_\omega^n)}\ \sup_{x,y\in\xi}\frac{J_\varphi(x)}{J_\varphi(y)} \leqslant Cn^{\mathfrak{d}},$$

where

$$J_\varphi(x) = \frac{d\varphi_*\left[\mu_{\theta^n\omega}|_{\mathrm{dom}(\varphi)}\right]}{d\mu_\omega|_{\varphi(\mathrm{dom}(\varphi))}}(x) = \frac{d\varphi_*\left[\mu_{\theta^n\omega}|_{T_\omega^n\xi}\right]}{d\mu_\omega|_\xi}(x).$$

3.4 (Backward contraction on good cylinders)**.** $\exists\kappa > 1, \exists D > 1, \forall n \geqslant 1$: (denoting $\xi = \varphi(\mathrm{dom}(\varphi))$)

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\ \sup_{\varphi\in \overset{++}{\mathrm{IB}}(T_\omega^n)}\ \sup_{z\in\mathrm{dom}(\varphi)}\ \sup_{\substack{v\in T_zM\\ \|v\|=1}} |D\varphi(z)v| \leqslant Dn^{-\kappa}, \text{ i.e., } Dn^\kappa \leqslant \operatorname*{ess\,inf}_{\omega\in\Omega}\ \inf_{\varphi\in \overset{++}{\mathrm{IB}}(T_\omega^n)}\ \inf_{x\in\xi}\ \inf_{\substack{v\in T_xM\\ \|v\|=1}} |DT_\omega^n(x)v|,$$

and, in particular,

$$\operatorname*{ess\,sup}_{\omega\in\Omega}\ \sup_{\varphi\in \overset{++}{\mathrm{IB}}(T_\omega^n)}\ \mathrm{diam}(\xi) \leqslant Dn^{-\kappa}.$$

H4 (Target position)**.**

4.1 (Uniform inclusion in adequate set)**.** $\forall L \geqslant 1, \exists \rho_{\mathrm{sep}}(L) > 0, \forall \rho \leqslant \rho_{\mathrm{sep}}(L), \forall \omega \in \Omega$:

$$\forall 1 \leqslant L' \leqslant L, \forall 0 \leqslant j \leqslant L'-1 : (T_\omega^j)^{-1}\Gamma_{3/2\rho}(\theta^j\omega) \subset \overset{+}{\mathcal{C}}{}_{L'-1}^{\omega}.$$

4.2 (Quenched separation from non-good set)**.** It holds that[5]

$$\lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\sum_{n=L}^{\infty}\sum_{i=0}^{\lfloor 1/\hat{\mu}(\Gamma_\rho)\rfloor}\mu_{\theta^i\omega}\left(\Gamma(\theta^i\omega)\cap\left[\overset{+-}{\mathcal{C}}{}_n^{\theta^i\omega}\cup\overset{-}{\mathcal{C}}{}_n^{\theta^i\omega}\right]\right) = 0,\ \mathbb{P}\text{-a.s.}.$$

H5 (Lipschitz regularities)**.**

5.1 (Map)**.** $\sup_{x\in M}\mathrm{Lip}(T_\cdot(x) : \Omega \to M) < \infty$.

5.2 (Driving)**.** $\mathrm{Lip}(\theta) < \infty$.

5.3 (Target)**.** $\mathrm{Lip}(\Gamma : \Omega \to \mathscr{P}(M)) < \infty$,where $\mathscr{P}(M) = \{A \subset M, A \text{ compact}, A \neq \varnothing\}$ is equipped with the Hausdorff distance $d_H(A,B) = \sup_{x\in A}\inf_{y\in B} d(x,y) \vee \sup_{y\in B}\inf_{X\in A} d(x,y)$, which makes it a compact metric space.

[5] It is expected that this condition can be substituted by an averaged counterpart

$$\lim_{L\to\infty}\overline{\lim_{\rho\to 0}}\sum_{n=L}^{\infty}\frac{\hat{\mu}\left(\left[\overset{+-}{\mathcal{C}}_n\cup\overset{-}{\mathcal{C}}_n\right]\cap\Gamma_\rho\right)}{\hat{\mu}(\Gamma_\rho)} = 0.$$

.

H6 (Measure regularity)**.**

6.1 (Ball regular)**.** $\exists 0 < d_0 \leqslant d_1 < \infty, \exists C_0, C_1 > 0, \exists \rho_{\dim} \leqslant 1, \forall \rho \leqslant \rho_{\dim}$, for $\mathbb{P}$-a.e. $\omega \in \Omega$:

$$C_1 \rho^{d_1} \leqslant \mu_\omega(\Gamma_\rho(\omega)) \leqslant C_0 \rho^{d_0}.$$

6.2 (Annulus regular)**.** $\exists \eta \geqslant \beta > 0, \exists E > 0, \exists \rho_{\dim} \leqslant 1, \forall \rho \leqslant \rho_{\dim}, \forall r \in (0, \rho/2)$, for $\mathbb{P}$-a.e. $\omega \in \Omega$:

$$\frac{\mu_\omega(\Gamma_{\rho+r}(\omega) \backslash \Gamma_{\rho-r}(\omega))}{\mu_\omega(\Gamma_\rho(\omega))} \leqslant E \frac{r^\eta}{\rho^\beta}.$$

H7 (Decay of correlations)**.** $\exists \mathfrak{p} > 1$ so that

7.1 (Quenched)**.** For $\mathbb{P}$-a.e. $\omega \in \Omega$, $\forall G \in \mathrm{Lip}_{d_M}(M, \mathbb{R}), \forall H \in L^\infty(M, \mathbb{R}), \forall n \geqslant 1$:

$$\left| \int_M G \cdot (H \circ T_\omega^n) d\mu_\omega - \mu_\omega(G) \mu_{\theta^n \omega}(H) \right| \lesssim n^{-\mathfrak{p}} \|G\|_{\mathrm{Lip}_{d_M}} \|H\|_\infty.$$

7.2 (Annealed)**.** $\forall G \in \mathrm{Lip}_{d_{\Omega \times M}}(\Omega \times M, \mathbb{R}), \forall H \in L^\infty(\Omega \times M, \mathbb{R}), \forall n \geqslant 1$:

$$\left| \int_{\Omega \times M} G \cdot (H \circ S^n) d\hat{\mu} - \hat{\mu}(G) \hat{\mu}(H) \right| \lesssim n^{-\mathfrak{p}} \|G\|_{\mathrm{Lip}_{d_{\Omega \times M}}} \|H\|_\infty.$$

H8 (Hitting regular)**.**

$$\exists (\lambda_\ell)_{\ell \geqslant 1}, \sum_{\ell=1}^\infty \lambda_\ell = 1, \sum_{\ell=1}^\infty \ell^3 \lambda_\ell < \infty.$$

H9 (Return regular)**.**

$$\exists (\alpha_\ell)_{\ell \geqslant 1}, \alpha_1 > 0, \sum_{\ell=1}^\infty \alpha_\ell = 1, \sum_{\ell=1}^\infty \ell^2 \alpha_\ell < \infty.$$

We call α_1 the extremal index.

H9' (Pre return regular)**.** *It holds that*

$$\exists (\hat{\alpha}_\ell)_{\ell \geqslant 1}, \hat{\alpha}_1 - \hat{\alpha}_2 > 0, \sum_{\ell=1}^\infty \ell \hat{\alpha}_\ell < \infty.$$

Using the final implication of item VI), it is immediate that (H9') $\Rightarrow$ (H9), because $\alpha_1 = \hat{\alpha}_1 - \hat{\alpha}_2 > 0$, $\sum_{\ell=1}^\infty \alpha_\ell = \hat{\alpha}_1 = 1$, and $\sum_{\ell=1}^\infty \ell^2 \alpha_\ell \leqslant 2 \sum_{\ell=1}^\infty \ell \hat{\alpha}_\ell < \infty$.

Moreover, for technical conditions, we assume that the quantities appearing in the previous hypotheses harmonize so that the following constraints hold. Mostly, they hold when (polynomial) decay is sufficiently fast.

H10 (Parametric constraints)**.** It holds that

10.1. $d_0(\mathfrak{p} - 1) > \frac{2\left(\frac{\beta + d_1}{\eta} \vee 1\right) + d_1}{d_0/d_1}$,

10.2. $\frac{d_0}{\mathfrak{d}+1} \mathfrak{p} > 2\left(\frac{\beta + d_1}{\eta} \vee 1\right) + d_1$,

10.3. $\mathfrak{d} < \kappa d_0 - 1$.

4.1.3 Main results

The first result to be presented, theorem 4.1.3, valid in the general setup of section 4.1.1, expresses hitting statistics (λ_ℓ's) in terms of return statistics (α_ℓ's).

Theorem 4.1.3. *Let $(\theta,\mathbb{P},T_\omega,\mu_\omega,\Gamma)$ be a system as described in section 4.1.1, with $(\theta,\mathbb{P})$ only assumed invariant.*

Then

$$\text{(H9')} \Rightarrow \lambda_\ell = \frac{\alpha_\ell - \alpha_{\ell+1}}{\alpha_1} \ (\ell \geqslant 1) \text{ and (H8).}$$

The essential part[6] of this theorem is to conclude the equality, which will be proven in section 4.2. It implies that $\alpha_1 = (\sum_{\ell=1}^\infty \ell\lambda_\ell)^{-1}$.

Although important on its own, it actually plays an auxiliary role within the paper, serving the following two purposes (the essential one being the second):

1) Technical: help the proof of our main result, theorem 4.1.4, via its use in the proof of lemma 4.4.1. Notice that this lemma can be restated assuming (H8) and not invoking α_ℓ's at all, but relying only on λ_ℓ's. However, we still use α_ℓ's (and theorem 4.1.3) in the proof of the said lemma. We believe one could bypass this intricacy and write a spin-off version of lemma 4.4.1 with an associated proof reformulation where α_ℓ's and theorem 4.1.3 do not play any role. The consequence would be an associated spin-off version of theorem 4.1.4 written completely free of α_ℓ's and relying solely on λ_ℓ's.

2) Examples: to handle examples, one will always need theorem 4.1.3 to compute the λ_ℓ's appearing in theorem 4.1.4 (or its hypothetical spin-off). This is because return statistics are generally much easier to compute than hitting statistics, so, whenever facing a concrete example, we calculate the α_ℓ's to obtain the λ_ℓ's.

Let us now formulate our main result. It says that the targeted random dynamical systems being considered have quenched limit entry distributions in the compound Poisson class.

Theorem 4.1.4. *Let $(\theta,\mathbb{P},T_\omega,\mu_\omega,\Gamma)$ be a system satisfying hypotheses (H1)-(H7),(H9') (so (H8), by theorem 4.1.3) with the parametric constraints (H10.1)-(H10.3).*

Then: $\forall t>0, \forall n \geqslant 0, \forall (\rho_m)_{m\geqslant 1} \searrow 0$ with $\sum_{m\geqslant 1} {\rho_m}^q < \infty$ (for some $0<q<q(d_0,d_1,\eta,\beta,\mathfrak{p})$[7]) one has

$$\mu_\omega(Z_{\Gamma_{\rho_m}}^{\omega,\lfloor t/\hat{\mu}(\Gamma_{\rho_m})\rfloor} = n) \xrightarrow[m\to\infty]{\mathbb{P}\text{-a.s.}} \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\}), \tag{4.9}$$

where $\mathrm{CPD}_{s,(\lambda_\ell)_\ell}$ is the compound Poisson distribution with intensity s and multiplicity distribution $(\lambda_\ell)_\ell$ (see definition 2.1.11[8]).

Remark 3. If the system has exponential asymptotics in (H7) and (H3.4), the previous conclusion is still true, but, actually, with fewer parametric conditions being required: instead of (H10.1)-(H10.3), only $\kappa d_0 > 1$ is needed.

---

[6](H8) follows from the previous equality and (H9) because $\sum_{\ell=1}^\infty \lambda_\ell = \frac{\sum_{\ell=1}^\infty \alpha_\ell - \alpha_{\ell+1}}{\alpha_1} = 1$ and $\sum_{\ell=1}^\infty \ell^3\lambda_\ell = (\alpha_1)^{-1}\sum_{\ell=1}^\infty \ell^3(\alpha_\ell - \alpha_{\ell+1}) = (\alpha_1)^{-1}\left(\alpha_1 + \sum_{\ell=2}^\infty \ell^3\alpha_\ell - \sum_{\ell=2}^\infty (\ell-1)^3\alpha_\ell\right) = (\alpha_1)^{-1}\left(\alpha_1 + \sum_{\ell=2}^\infty (3\ell^2 - 3\ell + 1)\alpha_\ell\right) \leqslant (\alpha_1)^{-1}\left(1 + \sum_{\ell=1}^\infty 7\ell^2\alpha_\ell\right) < \infty$.

[7]A quantity to be introduced in lemma 4.4.2.

[8]Please note that chapters are independent. So despite the notational coincidence, the probability space in definition 2.1.11 is an abstract one, not the same as the random driving of the random dynamical system.

The rest of the thesis is organized into two parts:

I) Theory: Until section 4.5 we work to prove theorem 4.1.4, accomplishing the required auxiliary results, among which we highlight theorem 4.1.3, theorem 4.3.1, and lemmas 4.4.1, 4.4.2 and 4.4.3. Let us briefly recall the point of these auxiliary results and provide a blueprint of the proof.

Theorem 4.1.3 calculates hitting statistics in terms of return statistics. Its role was already discussed above.

Theorem 4.3.1 provides the skeleton of the proof of theorem 4.1.4, describing the asymptotics we are after with a leading term and an error. The leading term appears from spliting the ω-quenched hit-counting function into equally time-sized blocks and mimicking them with an independency of random variables distributed just like each of them[9]. The errors have a structure, basically being divided into two parts[10]. The first one (comprised of terms $\mathcal{R}^1$ and $\tilde{\mathcal{R}}^1$) accounts for long-range effects and will be controlled using weak hyperbolicity features (H3.1,H3.2), the target uniform inclusion in the adequate set (H4.1), the annulus regularity (H6.2) and quenched decay (H7.1). The second one (comprised of terms $\mathcal{R}^2$ and $\mathcal{R}^3$) accounts for short-range interactions and will be controlled using the structure of the covering system (H2), distortion (H3.3), strong hyperbolicity features (H3.4) and ball regularity (H6.1). Notice no annealed decay was used yet.

After proceeding as above, $\mathcal{R}^1$ and $\tilde{\mathcal{R}}^1$ turn out to be ω-uniformly bounded, due to the ω-uniform constants throughout the hypotheses. But $\mathcal{R}^2$ and $\mathcal{R}^3$ turn out to still be bounded with ω-dependence. The leading term is also ω-dependent.

Lemmas 4.4.1, 4.4.2 and 4.4.3 will used in the proof of theorem 4.1.4 to tame the above-mentioned ω-dependent leading and error terms under the limit, in an almost sure manner, allowing for a quenched limit theorem: the error will go to zero and the leading term to a compound Poisson. Lemma 4.4.2 does a certain variance control and its proof uses the annealed decay of correlation (H7.2). This is the only place where annealed decay is used. Lemma 4.4.3 is ultimately the artifact allowing the almost sure result. It applies the Chebychev inequality, with variance coming from lemma 4.4.2, followed by the Borel-Cantelli lemma, in such a way as to show that a certain sequence of variables converges a.s. to its average, whose asymptotics come from lemma 4.4.1.

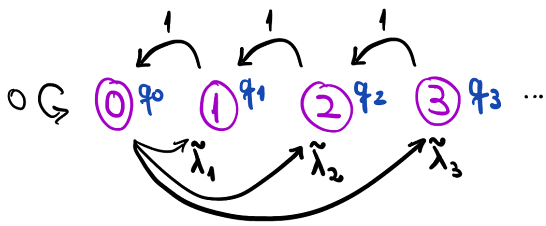
II) Examples: We consider certain random piecewise expanding one-dimensional systems, casting new light on the well-known deterministic dichotomy between periodic and aperiodic points, their typical extremal index formula $\mathrm{EI} = 1 - 1/JT^p(\zeta)$, and recovering the geometric case for general Bernoulli-driven systems, but distinct behavior otherwise.

4.2 Proof of theorem 4.1.3

Remark 4. Before discussing the proof, to aid one's intuition on the result itself (but not directly on the proof strategy to be pursued) it is useful to consider that the algebraic relationship in the conclusion of theorem 4.1.3 also appears from a Markov chain of the form below

[9]A comparison with the roles played by conditions $Д_q'^*(u)$ in section 3.1.1 is worthwhile. See item (i) after theorem 3.1.2.

[10]A comparison with how the errors in the probabilistic approximation sections 3.1.3 and 3.2.3 are controlled is worthwhile.



where the q_i's represent the stationary mass distribution of the chain. By letting $\tilde{\alpha}_\ell := q_\ell / \sum_{j\geqslant 1} q_j$ ($\ell \geqslant 1$), it is possible to verify that $\tilde{\lambda}_\ell = \frac{\tilde{\alpha}_\ell - \tilde{\alpha}_{\ell+1}}{\alpha_1}$.

One can interpret that each non-zero state s refers to the space of points in a germ of the target set which undergo s visits (present included) to the same germ before escaping from it (i.e., escaping from its local recurrence behavior), while the state zero represents points outside the germ (and its local recurrence behavior). The fact that the transition probabilities on the top and the left are 1 and 0 can be thought of as an acceleration of the system. In this case, $\tilde{\lambda}_\ell$'s can be understood as cluster size statistics and the $\tilde{\alpha}_\ell$ are sometimes called sojourn statistics.

Let us note that theorem 4.1.3 generalizes theorem 2 from [53]. The proof is very similar but presented here for the convenience of the reader and to fix some inaccuracies found in the latter.

In the scope of this section, let arbitrarily chosen $\ell \in \mathbb{N}_{\geqslant 1}$ and $\gamma \in \mathbb{N}_{>1}$ be fixed. These will be used in the forthcoming proof of theorem 4.1.3 and lemma 4.2.1.

Lemma 4.2.1. *It holds that*

I) $\forall \eta > 0, \exists L_2(\eta), \forall L' > L \geqslant L_2(\eta), \exists \rho_2(\eta, L, L'), \forall \rho \leqslant \rho_2(\eta, L, L')$:

$$\hat{\mu}\left(Z_{\Gamma_\rho}^{L'-L} \circ S^L > 0, I_0^\rho = 1\right) \leqslant \eta \hat{\mu}(\Gamma_\rho).$$

II) $\forall \eta > 0, \exists L_2(\eta), \forall L', L$ *so that* $L' - L \geqslant L_2(\eta)$*,* $\exists \rho_2(\eta, L, L') > 0, \forall \rho \leqslant \rho_2(\eta, L, L')$:

$$\hat{\mu}\left(Z_{\Gamma_\rho}^{L} > 0, I_{L'}^\rho = 1\right) \leqslant \eta \hat{\mu}(\Gamma_\rho).$$

Proof. Manipulating the definitions of $\hat{\alpha}_\ell$'s, the associated limits and the finiteness of the series in the hypothesis, it can be shown that: $\forall \varepsilon > 0$

i) $\exists k_0(\varepsilon) \geqslant 1$ so that $\sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k \leqslant \varepsilon$.

ii) $\exists L_0(\varepsilon) \geqslant [2k_0(\varepsilon) \vee 2\ell]\,(\varepsilon^{-1})^\gamma, \forall L \geqslant L_0(\varepsilon), \exists \rho_0(\varepsilon, L) > 0, \forall \rho \leqslant \rho_0(\varepsilon, L)$ one has

   a) $\forall H \in [L, 2L^\gamma] : \sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k(H,\rho) \leqslant 2\varepsilon, \left|\sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k - \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k(H,\rho)\right| \leqslant 2\varepsilon.$

   b) $\forall q \in [1, 2k_0(\varepsilon) \vee 2\ell], \forall H \in [L, 2L^\gamma] : |\hat{\alpha}_q - \hat{\alpha}_q(H,\rho)| \leqslant \varepsilon, |\alpha_q - \alpha_q(H,\rho)| \leqslant \varepsilon.$

iii) $\exists L_1(\varepsilon) \geqslant L_0(\varepsilon), \forall L' > L \geqslant L_1(\varepsilon), \exists \rho_1(\varepsilon, L, L') \in \left(0, \bigwedge_{H=L}^{L'} \rho_0(\varepsilon, H)\right), \forall \rho \leqslant \rho_1(\varepsilon, L, L')$ one has $\sum_{k=1}^{\infty} |\hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho)| \leqslant 6\varepsilon$.

To justify i) use that $\sum_{k=1}^{\infty} k\hat{\alpha}_k < \infty$.

To justify ii.a) start noticing that $\forall \varepsilon > 0, \exists L_0(\varepsilon), \forall L \geqslant L_0(\varepsilon)$:

$$0 \leqslant \hat{\alpha}_k - \overset{\pm}{\hat{\alpha}}_k(H) \leqslant \frac{\varepsilon}{k_0(\varepsilon)} \ (\forall k \in [1, k_0(\varepsilon)], \forall H \in [L, 2L^\gamma]),$$

$$\Rightarrow 0 \leqslant \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k - \sum_{k=1}^{k_0(\varepsilon)} \overset{\pm}{\hat{\alpha}}_k(H) \leqslant \varepsilon \ (\forall H \in [L, 2L^\gamma]),$$

because $\hat{\alpha}_k = \lim_L \overset{\pm}{\hat{\alpha}}_k(L)$ occurs monotonically increasing in L.

Then consider that $\forall \varepsilon > 0, \forall L \geqslant 0, \exists \rho_0(\varepsilon, L), \forall \rho \leqslant \rho_0(\varepsilon, L)$:

$$\overset{-}{\hat{\alpha}}_k(H) - \frac{\varepsilon}{k_0(\varepsilon)^2(2L^\gamma)^2} \leqslant \hat{\alpha}_k(H, \rho) \leqslant \overset{+}{\hat{\alpha}}_k(H) + \frac{\varepsilon}{k_0(\varepsilon)^2(2L^\gamma)^2}$$
$$(\forall k \in [1, k_0(\varepsilon)^2(2L^\gamma)^2], \forall H \in [L, 2L^\gamma])$$

implying

$$\sum_{k=1}^{k_0(\varepsilon)} \overset{-}{\hat{\alpha}}_k(H) - \varepsilon \leqslant \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k(H, \rho) \leqslant \sum_{k=1}^{k_0(\varepsilon)} \overset{+}{\hat{\alpha}}_k(H) + \varepsilon \ (\forall H \in [L, 2L^\gamma]),$$

and

$$\begin{aligned}\sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k(H, \rho) = \sum_{k=k_0(\varepsilon)}^{H} k\hat{\alpha}_k(H, \rho) &\leqslant \sum_{k=k_0(\varepsilon)}^{H} k \overset{+}{\hat{\alpha}}_k(H) + \sum_{k=k_0(\varepsilon)}^{H} k \frac{\varepsilon}{k_0(\varepsilon)^2(2L^\gamma)^2} \\ &\leqslant \sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k + (2L^\gamma)^2 \frac{\varepsilon}{k_0(\varepsilon)^2(2L^\gamma)^2} \\ &\leqslant \varepsilon + \varepsilon = 2\varepsilon \ (\forall H \in [L, 2L^\gamma]).\end{aligned}$$

Finally, combining the conditions and conclusions of the two previous paragraphs: $\forall \varepsilon > 0, \exists L_0(\varepsilon), \forall L \geqslant L_0(\varepsilon), \exists \rho_0(\varepsilon, L) > 0, \forall \rho \leqslant \rho_0(\varepsilon, L), \forall H \in [L, 2L^\gamma]$:

$$\sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k(H, \rho) \leqslant 2\varepsilon \text{ and } \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k - 2\varepsilon \leqslant \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k(H, \rho) \leqslant \sum_{k=1}^{k_0(\varepsilon)} \hat{\alpha}_k + 2\varepsilon,$$

as desired.

To justify ii.b) one can adapt the argument used above to show the second inequality in ii.a).

Finally, to justify iii) start noticing that $\forall \varepsilon > 0, \exists L_1(\varepsilon) \geqslant L_0(\varepsilon), \forall L' > L \geqslant L_1(\varepsilon)$:

$$\left| \overset{\circledast}{\hat{\alpha}}_k(L') - \overset{*}{\hat{\alpha}}_k(L) \right| \leqslant \frac{\varepsilon}{4k_0(\varepsilon)} \ (\forall \circledast, * \in \{+, -\}),$$

and therefore $\forall \varepsilon > 0, \exists L_1(\varepsilon) \geqslant L_0(\varepsilon), \forall L' > L \geqslant L_1(\varepsilon), \exists \rho_1(\varepsilon, L, L') \in \left(0, \bigwedge_{H=L}^{L'} \rho_0(\varepsilon, H)\right)$, $\forall \rho \leqslant \rho_1(\varepsilon, L, L'), \forall k \in [1, k_0(\varepsilon)]$:

$$\overset{-}{\hat{\alpha}}_k(L') - \frac{\varepsilon}{4k_0(\varepsilon)} \leqslant \hat{\alpha}_k(L', \rho) \leqslant \overset{+}{\hat{\alpha}}_k(L') + \frac{\varepsilon}{4k_0(\varepsilon)} \ (\forall k \in [1, k_0(\varepsilon)])$$

$$\bar{\hat{\alpha}}_k(L) - \frac{\varepsilon}{4k_0(\varepsilon)} \leqslant \hat{\alpha}_k(L,\rho) \leqslant \overset{+}{\hat{\alpha}}_k(L) + \frac{\varepsilon}{4k_0(\varepsilon)} \; (\forall k \in [1, k_0(\varepsilon)])$$

$$\text{so that } -\frac{\varepsilon}{k_0(\varepsilon)} \leqslant \hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho) \leqslant \frac{\varepsilon}{k_0(\varepsilon)} \; (\forall k \in [1, k_0(\varepsilon)])$$

$$\Rightarrow \sum_{k=1}^{k_0(\varepsilon)} |\hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho)| \leqslant \varepsilon,$$

and, since the quantifiers were subordinated to those of (ii), we actually get that

$$\begin{aligned}\textstyle\sum_{k=1}^{\infty} |\hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho)| \quad &\leqslant \textstyle\sum_{k=1}^{k_0(\varepsilon)} |\hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho)| + \sum_{k=k_0(\varepsilon)}^{\infty} |\hat{\alpha}_k(L',\rho) - \hat{\alpha}_k(L,\rho)| \\ &\leqslant \varepsilon + \textstyle\sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k(L',\rho) + \sum_{k=k_0(\varepsilon)}^{\infty} k\hat{\alpha}_k(L',\rho) \\ &\leqslant \varepsilon + 2\varepsilon + 2\varepsilon \leqslant 6\varepsilon,\end{aligned}$$

as desired

Now we prove I).

Let $L_2(\eta) := L_1(\eta/6)$ and $\rho_2(\eta,L,L') := \rho_1(\eta/6,L,L')$. Consider $L' > L \geqslant L_2(\eta)$ and $\rho \leqslant \rho_2(\eta,L,L')$. Then

$$\hat{\mu}\left(Z_{\Gamma_\rho}^{L'-L} \circ S^L > 0, I_0^\rho = 1\right)$$

$$= \int_\Omega \sum_{k=1}^{\infty} \mu_\omega \left(Z_{\Gamma_\rho}^{\theta^L\omega, L'-L} \circ T_\omega^L > 0, I_0^{\omega,\rho} = 1, Z_{\Gamma_\rho}^{\omega,L'} = k\right) d\mathbb{P}(\omega)$$

$$\leqslant \int_\Omega \sum_{k=1}^{\infty} \mu_\omega \left(\Gamma_\rho(\omega) \cap \left[\{Z_{\Gamma_\rho}^{\omega,L'} \geqslant k\} \backslash \{Z_{\Gamma_\rho}^{\omega,L} \geqslant k\}\right]\right) d\mathbb{P}(\omega)$$

$$= \sum_{k=1}^{\infty} \left[\int_\Omega \mu_\omega(\Gamma_\rho(\omega))\hat{\alpha}_k^\omega(L',\rho) d\mathbb{P}(\omega) - \int_\Omega \mu_\omega(\Gamma_\rho(\omega))\hat{\alpha}_k^\omega(L,\rho) d\mathbb{P}(\omega)\right]$$

$$= \sum_{k=1}^{\infty} \hat{\alpha}_k(L',\rho)\hat{\mu}(\Gamma_\rho) - \hat{\alpha}_k(L,\rho)\hat{\mu}(\Gamma_\rho) \leqslant 6\eta/6\,\hat{\mu}(\Gamma_\rho) = \eta\hat{\mu}(\Gamma_\rho), \text{ by (iii).}$$

This concludes the proof of I).

Now we prove II).

Again let $L_2(\eta) := L_1(\eta/6)$ and $\rho_2(\eta,L,L') := \rho_1(\eta/6,L,L')$. Consider L', L so that $L' - L \geqslant L_2(\eta)$ and $\rho \leqslant \rho_2(\eta,L,L')$. Then

$$\hat{\mu}\left(Z_{\Gamma_\rho}^L > 0, I_{L'}^\rho = 1\right) \leqslant \int_\Omega \sum_{i=0}^{L-1} \mu_\omega \left(I_i^{\omega,\rho}{=}1, I_{L'}^{\omega,\rho}{=}1\right) d\mathbb{P}(\omega) = \int_\Omega \sum_{i=0}^{L-1} \mu_\omega \left(I_0^{\omega,\rho}{=}1, I_{L'-i}^{\omega,\rho}{=}1\right) d\mathbb{P}(\omega)$$

$$= \int_\Omega \mathbb{E}_{\mu_\omega}\left(I_0^{\omega,\rho} \sum_{i=L'-L+1}^{L'} I_i^{\omega,\rho}\right) d\mathbb{P}(\omega) = \int_\Omega \mathbb{E}_{\mu_\omega}\left(I_0^{\omega,\rho}\left[Z_{\Gamma_\rho}^{\omega,L'+1} - Z_{\Gamma_\rho}^{\omega,L'-L+1}\right]\right) d\mathbb{P}(\omega)$$

$$= \int_\Omega \sum_{k=0}^{\infty} k\left[\mu_\omega\left(Z_{\Gamma_\rho}^{\omega,L'+1} = k, I_0^{\omega,\rho} = 1\right) - \mu_\omega\left(Z_{\Gamma_\rho}^{\omega,L'-L+1} = k, I_0^{\omega,\rho} = 1\right)\right] d\mathbb{P}(\omega)$$

$$= \int_\Omega \sum_{k=0}^{\infty} k\left[\alpha_k^\omega(L'+1,\rho)\mu_\omega(\Gamma_\rho(\omega)) - \alpha_k^\omega(L'-L+1,\rho)\mu_\omega(\Gamma_\rho(\omega))\right] d\mathbb{P}(\omega),$$

but $\sum_{k=1}^\infty k\alpha_k^\omega(T,\rho) = \sum_{k=1}^\infty k[\hat{\alpha}_k^\omega(T,\rho) - \hat{\alpha}_{k+1}^\omega(T,\rho)] = \sum_{k=1}^\infty \hat{\alpha}_k^\omega(T,\rho)$, so

$$= \int_\Omega \Big[\sum_{k=0}^{\infty} \hat{\alpha}_k^\omega(L'+1,\rho) - \hat{\alpha}_k^\omega(L'-L+1,\rho)\Big]\mu_\omega(\Gamma_\rho(\omega))d\mathbb{P}(\omega)$$

$$= \sum_{k=0}^{\infty}\left[\hat{\alpha}_k(L'+1,\rho) - \hat{\alpha}_k(L'-L+1,\rho)\right]\hat{\mu}(\Gamma_\rho) \leqslant 6\eta/6\hat{\mu}(\Gamma_\rho) = \eta\hat{\mu}(\Gamma_\rho), \text{ by (iii).}$$

This concludes the proof of II). ■

Proof of theorem 4.1.3. Let $\varepsilon \in (0,\alpha_1/26)$ and consider a function $\eta(\varepsilon)$ to be chosen in due time.

Set $L_3(\varepsilon)$ be large so that $L \geqslant L_3(\varepsilon) \Rightarrow L^{\gamma-1} \geqslant (k_0(\varepsilon)+1)\varepsilon^{-1}$ and $L^\gamma - L > L \geqslant 2\left[(L_1(\varepsilon) \vee L_2(\eta(\varepsilon))) + 1\right]$, where $k_0(\varepsilon)$ is that found in the proof of 4.2.1 item i), $L_1(\varepsilon)$ found in item iii) of the same proof and again $L_2(\eta) = L_1(\eta/6)$.

we adopt the notation in the statement and proof of the previous lemma

Set $\rho_3(\varepsilon, L)$ to

$$\min\left\{\rho_1(\varepsilon,L,L^\gamma), \rho_1(\varepsilon,L^\gamma-L,L^\gamma), \min_{i\in[L,2L^\gamma-L-1]}\left\{\begin{matrix}\rho_2(\eta(\varepsilon),L,2L^\gamma-i), \rho_2(\eta(\varepsilon),L/2,2L^\gamma-i)\\ \rho_2(\eta(\varepsilon),L/2,i), \rho_2(\eta(\varepsilon),i-L/2+1,i)\end{matrix}\right\}\right\}.$$

Consider $L \geqslant L_3(\varepsilon)$ and $\rho \leqslant \rho_3(\varepsilon, L)$. We evaluate the numerator that appears when expanding λ_ℓ.

$$\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma} = \ell\right) = \int_\Omega \ell^{-1}\sum_{i=0}^{2L^\gamma-1}\mathbb{E}_{\mu_\omega}\left(\mathbb{1}_{\left\{Z_{\Gamma_\rho}^{\omega,2L^\gamma}=\ell\right\}} I_i^{\omega,\rho}\right)d\mathbb{P}(\omega)$$

$$\Rightarrow \left|\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma} = \ell\right) - \ell^{-1}\sum_{i=L}^{2L^\gamma-L-1}\int_\Omega \mu_\omega\left(Z_{\Gamma_\rho}^{\omega,2L^\gamma} = \ell, I_i^{\omega,\rho} = 1\right)d\mathbb{P}(\omega)\right|$$
$$\leqslant \ell^{-1}2L\hat{\mu}(\Gamma_\rho). \tag{4.10}$$

We establish some notation. When $i \in [L, 2L^\gamma - L - 1]$ and $k \in [0, \ell-1]$, write:

$$D_{\omega,i}^{L,L^\gamma} = \left\{\sum_{u=i+L}^{2L^\gamma-1} I_u^{\omega,\rho} > 0, I_i^{\omega,\rho} = 1\right\};$$

$$F_{\omega,i}^{L/2} = \left\{\sum_{u=0}^{L/2-1} I_u^{\omega,\rho} > 0, I_i^{\omega,\rho} = 1\right\};$$

$$R_{\omega,\ell,k}^{i,L} = \left\{\sum_{u=0}^{i-1} I_u^{\omega,\rho} = k, \sum_{u=i}^{i+L-1} I_u^{\omega,\rho} = \ell-k, I_i^{\omega,\rho} = 1\right\};$$

$$R_{\omega,\ell,k}^{i,L}(j) = R_{\omega,\ell,k}^{i,L} \cap \left\{I_j^{\omega,\rho} = 1, I_a^{\omega,\rho} = 0\ \forall a \in [0,j)\right\}, \text{ for } j \in [0,i];$$

$$S_{\omega,\ell,k}^{i,L}(j) = R_{\omega,\ell,k}^{i,L} \cap \left\{I_j^{\omega,\rho} = 1, I_b^{\omega,\rho} = 0\ \forall b \in (j,i)\right\}, \text{ for } j \in [0,i-1].$$

To update the estimate in equation (4.10), we will apply many approximation steps, to be identified with uppercase roman letters and justified only at the very end.

A) For $i \in [L, 2L^\gamma - L - 1]$:

$$\left| \int_\Omega \mu_\omega \left(Z_{\Gamma_\rho}^{\omega, 2L^\gamma} = \ell, I_i^{\omega,\rho} = 1 \right) d\mathbb{P}(\omega) - \int_\Omega \mu_\omega \left(Z_{\Gamma_\rho}^{\omega, i+L} = \ell, I_i^{\omega,\rho} = 1 \right) d\mathbb{P}(\omega) \right| \leqslant \eta(\varepsilon) \hat{\mu}(\Gamma_\rho)$$

Combining equation (4.10) with approximation (A), while observing that second integrand above equals $\mu_\omega(\bigcup_{k=0}^{\ell-1} R^{i,L}_{\omega,\ell,k})$, gives

$$\begin{aligned} &\left| \hat{\mu} \left(Z_{\Gamma_\rho}^{2L^\gamma} = \ell \right) - \ell^{-1} \sum_{i=L}^{2L^\gamma - L - 1} \sum_{k=0}^{\ell-1} \int_\Omega \mu_\omega (R^{i,L}_{\omega,\ell,k}) d\mathbb{P}(\omega) \right| \\ &\qquad \leqslant \ell^{-1} [2L + (2L^\gamma - 2L)\eta(\varepsilon)] \hat{\mu}(\Gamma_\rho). \end{aligned} \tag{4.11}$$

B) For $i \in [L, 2L^\gamma - L - 1]$, $k \in [0, \ell - 1]$:

$$\left| \int_\Omega \mu_\omega (R^{i,L}_{\omega,\ell,0}) d\mathbb{P}(\omega) - \int_\Omega \mu_\omega (R^{i,L}_{\omega,\ell,k}) d\mathbb{P}(\omega) \right| \leqslant 3\eta(\varepsilon) \hat{\mu}(\Gamma_\rho).$$

C) For $i \in [L, 2L^\gamma - L - 1]$:

$$\left| \int_\Omega \mu_\omega \left(R^{i,L}_{\omega,\ell,0} \right) d\mathbb{P}(\omega) - \int_\Omega \mu_\omega \left(R^{L,L}_{\omega,\ell,0} \right) d\mathbb{P}(\omega) \right| \leqslant \eta(\varepsilon) \hat{\mu}(\Gamma_\rho).$$

Combining equation (4.11) with approximations (B) and (C), gives

$$\begin{aligned} &\left| \hat{\mu} \left(Z_{\Gamma_\rho}^{2L^\gamma} = \ell \right) - (2L^\gamma - 2L) \int_\Omega \mu_\omega \left(R^{L,L}_{\omega,\ell,0} \right) d\mathbb{P}(\omega) \right| \\ &\qquad \leqslant [2L + 5(2L^\gamma - 2L)\eta(\varepsilon)] \hat{\mu}(\Gamma_\rho). \end{aligned} \tag{4.12}$$

Now we look to the other side of the equality we are trying to prove.

Notice that

$$\begin{aligned} \alpha_\ell(L,\rho) - \alpha_{\ell+1}(L,\rho) &= \hat{\mu}(\Gamma_\rho)^{-1} \int_\Omega \left[\begin{array}{c} \mu_\omega \left(\sum_{u=L}^{2L-1} I_u^{\omega,\rho} = \ell, I_L^{\omega,\rho} = 1 \right) \\ -\mu_\omega \left(\sum_{u=L}^{2L-1} I_u^{\omega,\rho} = \ell + 1, I_L^{\omega,\rho} = 1 \right) \end{array} \right] d\mathbb{P}(\omega) \\ &= \hat{\mu}(\Gamma_\rho)^{-1} \left[\begin{array}{c} \sum_{k=0}^{k_0(\varepsilon)} \int_\Omega \mu_\omega (R^{L,L}_{\omega,\ell+k,k}) - \mu_\omega (R^{L,L}_{\omega,\ell+k+1,k}) d\mathbb{P}(\omega) \\ + \sum_{k=k_0(\varepsilon)+1}^{\infty(L)} \int_\Omega \mu_\omega (R^{L,L}_{\omega,\ell+k,k}) - \mu_\omega (R^{L,L}_{\omega,\ell+k+1,k}) d\mathbb{P}(\omega) \end{array} \right] \qquad (4.13) \\ &\overset{(*)}{=} \hat{\mu}(\Gamma_\rho)^{-1} \left[\begin{array}{c} \sum_{k=0}^{k_0(\varepsilon)} \int_\Omega \mu_\omega (R^{L,L}_{\omega,\ell+k,k}) - \left(\mu_\omega (R^{L,L}_{\omega,\ell+k+1,k+1}) + \mathcal{O}\left(\eta(\varepsilon) \hat{\mu}(\Gamma_\rho) \right) \right) d\mathbb{P}(\omega) \\ + \sum_{k=k_0(\varepsilon)+1}^{\infty(L)} \int_\Omega \mu_\omega (R^{L,L}_{\omega,\ell+k,k}) - \mu_\omega (R^{L,L}_{\omega,\ell+k+1,k}) d\mathbb{P}(\omega) \end{array} \right] \\ &\Rightarrow \left| (\alpha_\ell(L,\rho) - \alpha_{\ell+1}(L,\rho)) \hat{\mu}(\Gamma_\rho) - \int_\Omega \mu_\omega \left(R^{L,L}_{\omega,\ell,0} \right) d\mathbb{P}(\omega) \right| \end{aligned}$$

$$\leqslant 2k_0(\varepsilon)\mathcal{O}\left(\eta(\varepsilon)\hat{\mu}(\Gamma_\rho)\right) + 2\sum_{k=k_0(\varepsilon)+1}^{\infty(L)} \int_\Omega \mu_\omega(R^{L,L}_{\omega,\ell+k,k}) + \mu_\omega(R^{L,L}_{\omega,\ell+k+1,k})d\mathbb{P}(\omega),$$

$$\overset{(\circledast)}{\leqslant} 2k_0(\varepsilon)\mathcal{O}\left(\eta(\varepsilon)\hat{\mu}(\Gamma_\rho)\right) + 7(k_0(\varepsilon)+\ell)\hat{\alpha}_{k_0(\varepsilon)+\ell}(2L,\rho)\hat{\mu}(\Gamma_\rho) \tag{4.14}$$

where steps $(\ast)$ and $(\circledast)$ are justified, respectively, with the following two approximations.

D) For $i \in [L, 2L^\gamma - L - 1]$, $k \in [1, \ell-1]$ (for other k's, zeroes pop up):

$$\left|\int_\Omega \mu_\omega\left(R^{i,L}_{\omega,\ell,k-1}\right)d\mathbb{P}(\omega) - \int_\Omega \mu_\omega\left(R^{i,L}_{\omega,\ell,k}\right)d\mathbb{P}(\omega)\right| \leqslant 3\eta(\varepsilon)\hat{\mu}(\Gamma_\rho).$$

E)

$$\sum_{k=k_0(\varepsilon)+1}^{\infty(L)} \int_\Omega \mu_\omega(R^{L,L}_{\omega,\ell+k,k})d\mathbb{P}(\omega) \leqslant (k_0(\varepsilon)+\ell)\hat{\alpha}_{k_0(\varepsilon)+\ell}(2L,\rho)\hat{\mu}(\Gamma_\rho).$$

Now choose $\eta(\varepsilon) = \varepsilon/(k_0(\varepsilon)+1)$. Combining equations (4.12) & (4.14), using (ii.a) from the proof of lemma 4.2.1, and factoring L^γ out (notice $L^{1-\gamma} \leqslant \varepsilon$), gives:

$$\left|\hat{\mu}\left(Z^{2L^\gamma}_{\Gamma_\rho} = \ell\right) - 2L^\gamma\big(\alpha_\ell(L,\rho) - \alpha_{\ell+1}(L,\rho)\big)\hat{\mu}(\Gamma_\rho)\right| \leqslant 84L^\gamma\varepsilon\hat{\mu}(\Gamma_\rho). \tag{4.15}$$

We can finally evaluate the denominator which appears when expanding λ_ℓ.

$$\hat{\mu}\left(Z^{2L^\gamma}_{\Gamma_\rho} > 0\right) = \sum_{h=1}^{k_0(\varepsilon)+1}\left[\hat{\mu}\left(Z^{2L^\gamma}_{\Gamma_\rho} = h\right)\right] + \hat{\mu}\left(Z^{2L^\gamma}_{\Gamma_\rho} \geqslant k_0(\varepsilon)+2\right)$$

$$= \sum_{h=1}^{k_0(\varepsilon)+1}\left[\int_\Omega h^{-1}\mathbb{E}_{\mu_\omega}\left(\mathbb{1}_{\left\{Z^{\omega,2L^\gamma}_{\Gamma_\rho}=h\right\}} Z^{\omega,2L^\gamma}_{\Gamma_\rho}\right)d\mathbb{P}(\omega)\right] + \hat{\alpha}_{k_0(\varepsilon)+2}(2L^\gamma,\rho)\hat{\mu}(\Gamma_\rho)$$

$$= \sum_{i=0}^{2L^\gamma-1}\sum_{h=1}^{k_0(\varepsilon)+1}\left[h^{-1}\sum_{k=0}^{h-1}\int_\Omega \mu_\omega\left(R^{i,2L^\gamma-i}_{\omega,h,k}\right)d\mathbb{P}(\omega)\right] + \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)), \tag{4.16}$$

where the last line applied (ii.a) from the proof of lemma 4.2.1.

Then we consider the following approximation.

F) For $h \in [1, k_0(\varepsilon)+1]$, $i \in [L, 2L^\gamma - L - 1]$, $k \in [0, h-1]$:

$$\left|\int_\Omega \mu_\omega\left(R^{i,2L^\gamma-i}_{\omega,h,k}\right)d\mathbb{P}(\omega) - \int_\Omega \mu_\omega\left(R^{i,L}_{\omega,h,k}\right)d\mathbb{P}(\omega)\right| \leqslant \eta(\varepsilon)\hat{\mu}(\Gamma_\rho).$$

Starting from equation (4.16), splitting the i-sum into middle ($i \in [L, 2L^\gamma - L - 1]$) plus tail terms and applying (F,B,C,B) to the middle ones, gives

$$\hat{\mu}\left(Z^{2L^\gamma}_{\Gamma_\rho} > 0\right) = (2L^\gamma - 2L)\left[\sum_{h=1}^{k_0(\varepsilon)+1}\left(\int_\Omega \mu_\omega\left(R^{L,L}_{\omega,h,h-1}\right)d\mathbb{P}(\omega)\right) + \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho))\right]$$

$$+\mathcal{O}\Big(4L(k_0(\varepsilon)+1)\varepsilon^{-1}\varepsilon\hat{\mu}(\Gamma_\rho)\Big) + \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho))$$

$$\Rightarrow \left| \hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma} > 0\right) - 2L^\gamma \sum_{h=1}^{k_0(\varepsilon)+1} \int_\Omega \mu_\omega\left(R_{\omega,h,h-1}^{L,L}\right) d\mathbb{P}(\omega) \right|$$

$$\leqslant 3L^\gamma \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)) + \mathcal{O}(4L\cdot L^{\gamma-1}\varepsilon\hat{\mu}(\Gamma_\rho)) + \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)) + 4L \sum_{h=1}^{k_0(\varepsilon)+1} \int_\Omega \mu_\omega\left(R_{\omega,h,h-1}^{L,L}\right) d\mathbb{P}(\omega)$$

$$\leqslant 8L^\gamma \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)) + 4L \sum_{h=1}^{k_0(\varepsilon)+1} \int_\Omega \mu_\omega\left(R_{\omega,h,h-1}^{L,L}\right) d\mathbb{P}(\omega),$$

where it was used that $(k_0(\varepsilon)+1)\varepsilon^{-1} \leqslant L^{\gamma-1}$.

To take care of the summations on both sides of the previous inequality we observe that, when $\ell = 1$ is given to the "$\alpha_\ell(L,\rho)$ side of" equation (4.13), one gets

$$\left| \alpha_1(L,\rho)\hat{\mu}(\Gamma_\rho) - \sum_{h=1}^{k_0(\varepsilon)+1} \int_\Omega \mu_\omega(R_{\omega,h,h-1}^{L,L}) d\mathbb{P}(\omega) \right| = \sum_{h=k_0(\varepsilon)+1}^{\infty(L)} \int_\Omega \mu_\omega(R_{\omega,1+h,h}^{L,L}) d\mathbb{P}(\omega)$$

$$\leqslant (k_0(\varepsilon)+1)\hat{\alpha}_{k_0(\varepsilon)+1}(2L,\rho)\hat{\mu}(\Gamma_\rho) \leqslant 2\varepsilon\hat{\mu}(\Gamma_\rho),$$

where (E) and (ii.a) from the proof of lemma 4.2.1 are applied.

Therefore

$$\left| \hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma} > 0\right) - 2L^\gamma \alpha_1(L,\rho)\hat{\mu}(\Gamma_\rho) \right|$$

$$\leqslant 8L^\gamma \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)) + 4L\big(\alpha_1(L,\rho)\varepsilon^{-1}\varepsilon\hat{\mu}(\Gamma_\rho) + 2\varepsilon\hat{\mu}(\Gamma_\rho)\big) + 4L^\gamma \varepsilon\hat{\mu}(\Gamma_\rho)$$

$$\leqslant 20L^\gamma \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)) + 4\alpha_1(L,\rho)LL^{\gamma-1}\varepsilon\hat{\mu}(\Gamma_\rho) \leqslant 12\cdot 2L^\gamma \mathcal{O}(\varepsilon\hat{\mu}(\Gamma_\rho)). \tag{4.17}$$

Since $\alpha_1(L,\rho)$ is ε-close to α_1 and $\varepsilon \in (0, \alpha_1/26)$, the previous equation sets $\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma} > 0\right)$ far from zero, so we can really work with $\hat{\mu}(Z_{\Gamma_\rho}^{2L^\gamma} > 0)$ as a denominator.

Combining the estimates given in equations (4.15) and (4.17) gives:

$$\left|\frac{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}=\ell\right)}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}-\frac{\alpha_\ell-\alpha_{\ell+1}}{\alpha_1}\right|$$

$$\leqslant\left|\frac{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}=\ell\right)-\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}\right|+\left|\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}-\frac{\alpha_\ell-\alpha_{\ell+1}}{\alpha_1}\right|$$

$$\leqslant\frac{\left|\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}=\ell\right)-2L^\gamma\hat{\mu}(\Gamma_\rho)(\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho))\right|+\left|2L^\gamma\hat{\mu}(\Gamma_\rho)(\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho))-\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)\right|}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}$$
$$+\left|\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}-\frac{\alpha_\ell-\alpha_{\ell+1}}{\alpha_1}\right|$$

$$\underset{(4.17)}{\overset{(4.15)}{\leqslant}}\frac{42\cdot 2L^\gamma\varepsilon\hat{\mu}(\Gamma_\rho)+\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}12\cdot 2L^\gamma\varepsilon\hat{\mu}(\Gamma_\rho)}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}+\left|\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}-\frac{\alpha_\ell-\alpha_{\ell+1}}{\alpha_1}\right|$$

$$\leqslant 42\varepsilon\left(\frac{1}{\alpha_1}+\frac{\mathcal{O}(\varepsilon)}{(\alpha_1-\mathcal{O}(\varepsilon))^2}\right)+\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}12\varepsilon\left(\frac{1}{\alpha_1}+\frac{\mathcal{O}(\varepsilon)}{(\alpha_1-\mathcal{O}(\varepsilon))^2}\right)$$
$$+\left|\frac{\alpha_\ell(L,\rho)-\alpha_{\ell+1}(L,\rho)}{\alpha_1(L,\rho)}-\frac{\alpha_\ell-\alpha_{\ell+1}}{\alpha_1}\right| \tag{4.18}$$

where the last inequality applied the control

$$\left|\frac{2L^\gamma\hat{\mu}(\Gamma_\rho)}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}-\frac{1}{\alpha_1}\right|\leqslant\left(\sup_{z\geqslant\alpha_1-\mathcal{O}(\varepsilon)}z^{-2}\right)\left|\frac{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}{2L^\gamma\hat{\mu}(\Gamma_\rho)}-\alpha_1\right|\leqslant\frac{\mathcal{O}(\varepsilon)}{(\alpha_1-\mathcal{O}(\varepsilon))^2},$$

which is due to equation (4.17).

Passing $\lim_{\varepsilon\to 0}\overline{\lim}_{L\to\infty}\overline{\lim}_{\rho\to 0}$ over equation (4.18), we observe the RHS going to zero and we find that

$$\overline{\lim_{L\to\infty}}\,\overline{\lim_{\rho\to 0}}\,\frac{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}=\ell\right)}{\hat{\mu}\left(Z_{\Gamma_\rho}^{2L^\gamma}>0\right)}=\frac{\alpha_l-\alpha_{l+1}}{\alpha_1}.$$

Alternating between lim sup's and lim inf's lets us reach the desired conclusion.

Now we prove each of the approximations used above. Many of them rely on initial inclusions which are indicated and whose justification is left to the reader.

Proof of A) One can check that

$$\left\{Z_{\Gamma_\rho}^{\omega,i+L}=\ell,I_i^{\omega,\rho}=1\right\}\cap(D_{\omega,i}^{L,L^\gamma})^c\subset\left\{Z_{\Gamma_\rho}^{\omega,2L^\gamma}=\ell,I_i^{\omega,\rho}=1\right\}$$

$$\subset\left\{Z_{\Gamma_\rho}^{\omega,i+L}=\ell,I_i^{\omega,\rho}=1\right\}\cup D_{\omega,i}^{L,L^\gamma},$$

which gives an inequality that, after integration, can bound the LHS of A) by $\int_\Omega \mu_\omega(D^{L,L^\gamma}_{\omega,i})d\mathbb{P}(\omega)$.

Then we calculate

$$\int_\Omega \mu_\omega(D^{L,L^\gamma}_{\omega,i})d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(Z^{\theta^L\omega,2L^\gamma-i-L}_{\Gamma_\rho}\circ T^L_\omega > 0, I^{\omega,\rho}_0 = 1\right)d\mathbb{P}(\omega),$$

but $2L^\gamma - i > L \Leftrightarrow 2L^\gamma - L - 1 \geqslant i$, $L \geqslant L_3(\varepsilon) \geqslant L_2(\eta(\varepsilon))$ and $\rho \leqslant \rho_3(\varepsilon, L) \leqslant \rho_2(\eta(\varepsilon), L, 2L^\gamma - i)$, so (I, lemma 4.2.1) applies, implying that the last integral is bounded above by $\eta(\varepsilon)\hat{\mu}(\Gamma_\rho)$. □

Proof of B) One can check that

$$R^{i,L}_{\omega,l,0}\cap(D^{L/2,L^\gamma}_{\omega,i}\cup F^{L/2}_{\omega,i})^c \subset \bigcup_{j=i-L/2+1}^{i}(T^{i-j}_\omega)^{-1}R^{i,L}_{\theta^{i-j}\omega,l,k}(j)\subset R^{i,L}_{\omega,l,0}\cup(D^{L/2,L^\gamma}_{\omega,i}\cup F^{L/2}_{\omega,i}).$$

Then we calculate

$$\int_\Omega \mu_\omega\left(F^{L/2}_{\omega,i}\right)d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(Z^{\omega,L/2}_{\Gamma_\rho} > 0, I^{\omega,\rho}_i = 1\right)d\mathbb{P}(\omega) \leqslant \eta(\varepsilon)\hat{\mu}(\Gamma_\rho),$$

where the inequality follows from an application of (II, lemma 4.2.1), since $i - L/2 \geqslant L - L/2 = L/2 \geqslant L_2(\eta(\varepsilon))$ and $\rho \leqslant \rho_3(\varepsilon, L) \leqslant \rho_2(\eta(\varepsilon), L/2, i)$.

And also

$$\int_\Omega \mu_\omega\left(D^{L/2,L^\gamma}_{\omega,i}\right)d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(Z^{\theta^{L/2}\omega,2L^\gamma-i-L/2}_{\Gamma_\rho}\circ T^{L/2}_\omega > 0, I^{\omega,\rho}_0 = 1\right)d\mathbb{P}(\omega) \leqslant \eta(\varepsilon)\hat{\mu}(\Gamma_\rho),$$

where the inequality follows from an application of (I, lemma 4.2.1): $2L^\gamma - i > L/2 \Leftrightarrow 2L^\gamma - L/2 - 1 \geqslant i$, $L/2 \geqslant L_2(\eta(\varepsilon))$ and $\rho \leqslant \rho_3(\varepsilon, L) \leqslant \rho_2(\eta(\varepsilon), L/2, 2L^\gamma - i)$.

The consequence is the approximation

$$\left|\int_\Omega \mu_\omega\left(R^{i,L}_{\omega,l,0}\right)d\mathbb{P}(\omega) - \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i}(T^{i-j}_\omega)^{-1}R^{i,L}_{\theta^{i-j}\omega,l,k}(j)\right)d\mathbb{P}(\omega)\right| \leqslant 2\eta(\varepsilon)\hat{\mu}(\Gamma_\rho).$$

However,

$$\int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i}(T^{i-j}_\omega)^{-1}R^{i,L}_{\theta^{i-j}\omega,l,k}(j)\right)d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i}R^{i,L}_{\omega,l,k}(j)\right)d\mathbb{P}(\omega)$$

$$\leqslant \int_\Omega \mu_\omega\left(R^{i,L}_{\omega,l,k}\right)d\mathbb{P}(\omega) \leqslant \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i}R^{i,L}_{\omega,l,k}(j)\right)d\mathbb{P}(\omega) + \eta(\varepsilon)\hat{\mu}(\Gamma_\rho),$$

where the last inequality follows from the inclusion

$$R^{i,L}_{\omega,l,k}\subset \bigcup_{j=i-L/2+1}^{i}R^{i,L}_{\omega,l,k}(j)\cup F^{i-\frac{L}{2}}_{\omega,i}$$

and the estimate

$$\int_\Omega \mu_\omega\left(F_{\omega,i}^{i-L/2}\right)d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(Z_{\Gamma_\rho}^{\omega,i-L/2+1} > 0, I_i^{\omega,\rho} = 1\right)d\mathbb{P}(\omega) \leqslant \eta(\varepsilon)\hat{\mu}(\Gamma_\rho),$$

with the inequality following from another application of (II, lemma 4.2.1): $i-(i-L/2+1) = L/2-1 \geqslant L_2(\eta(\varepsilon))$ and $\rho \leqslant \rho_3(\varepsilon, L) \leqslant \rho_2(\eta(\varepsilon), i - L/2 + 1, i)$.

The conclusion follows from aforementioned approximation and the previous control. □

Proof of C) One can check that

$$R_{\omega,l,0}^{i,L} \subset (T_\omega^{i-L})^{-1}R_{\theta^{i-L}\omega,l,0}^{L,L},\ (T_\omega^{i-L})^{-1}R_{\theta^{i-L}\omega,l,0}^{L,L}\backslash R_{\omega,l,0}^{i,L} \subset F_{\omega,i}^{i-L/2},$$

which gives an inequality that, after integration, can bound the LHS of the expression we need to control by $\int_\Omega \mu_\omega(F_{\omega,i}^{i-L/2})d\mathbb{P}(\omega) \leqslant \eta(\varepsilon)\hat{\mu}(\Gamma_\rho)$, where the later estimate is identical to that obtained at the end of the proof of B). □

Proof of D) One can check that

$$R_{\omega,l,k-1}^{i,L} \cap (D_{\omega,i}^{L/2,L^\gamma} \cup F_{\omega,i}^{L/2})^c \subset \bigsqcup_{j=i-L/2+1}^{i-1} (T_\omega^{i-j})^{-1} S_{\theta^{i-j}\omega,l,k}^{i,L}(j) \subset R_{\omega,l,k-1}^{i,L} \cup (D_{\omega,i}^{L/2,L^\gamma} \cup F_{\omega,i}^{L/2}). \quad (4.19)$$

The corrective sets $F_{\omega,i}^{L/2}$ and $D_{\omega,i}^{L/2,L^\gamma}$ are treated again as in the proof of B), implying that

$$\left|\int_\Omega \mu_\omega\left(R_{\omega,l,k-1}^{i,L}\right)d\mathbb{P}(\omega) - \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2}^{i-1} (T_\omega^{i-j})^{-1}S_{\theta^{i-j}\omega,l,k}^{i,L}(j)\right)d\mathbb{P}(\omega)\right| \leqslant 2\eta(\varepsilon)\hat{\mu}(\Gamma_\rho). \quad (4.20)$$

However,

$$\int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i-1} (T_\omega^{i-j})^{-1}S_{\theta^{i-j}\omega,l,k}^{i,L}(j)\right)d\mathbb{P}(\omega) = \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i-1} S_{\omega,l,k}^{i,L}(j)\right)d\mathbb{P}(\omega)$$

$$\leqslant \int_\Omega \mu_\omega\left(R_{\omega,l,k}^{i,L}\right)d\mathbb{P}(\omega) \leqslant \int_\Omega \mu_\omega\left(\bigsqcup_{j=i-L/2+1}^{i} S_{\omega,l,k}^{i,L}(j)\right)d\mathbb{P}(\omega) + \eta(\varepsilon)\hat{\mu}(\Gamma_\rho),$$

where the last inequality follows from the inclusion

$$R_{\omega,l,k}^{i,L} \subset \bigsqcup_{j=i-L/2+1}^{i} S_{\omega,l,k}^{i,L}(j) \cup F_{\omega,i}^{i-\frac{L}{2}},$$

and the respective estimate of the corrective set using (II, lemma 4.2.1), precisely as in the end of the proof of B). □

Proof of E) One can check that

$$\bigcup_{j=1}^{L}\bigcup_{k=k_0(\varepsilon)}^{\infty(L)}(T^j_\omega)^{-1}R^{L,L}_{\theta^j\omega,l+k,k}(L-j)\subset(T^L_\omega)^{-1}\left\{Z^{\theta^L\omega,2L}_{\Gamma_\rho}\geqslant k_0(\varepsilon)+\ell,I_0^{\theta^L\omega,\rho}=1\right\}.$$

Therefore, integrating, manipulating and using invariance over and over:

$$\int_\Omega\mu_\omega\left(\bigcup_{j=1}^{L}\bigcup_{k=k_0(\varepsilon)}^{\infty(L)}(T^j_\omega)^{-1}R^{L,L}_{\theta^j\omega,l+k,k}(L-j)\right)d\mathbb{P}(\omega)=\sum_{k=k_0(\varepsilon)}^{\infty(L)}\int_\Omega\mu_\omega\left(R^{L,L}_{\omega,l+k,k}\right)d\mathbb{P}(\omega)$$

$$\leqslant\int_\Omega\mu_\omega\left(Z^{\omega,2L}_{\Gamma_\rho}\geqslant k_0(\varepsilon)+\ell,I_0^{\omega,\rho}=1\right)d\mathbb{P}(\omega)$$

$$=\hat\alpha_{k_0(\varepsilon)+\ell}(2L,\rho)\hat\mu(\Gamma_\rho)\leqslant(k_0(\varepsilon)+\ell)\hat\alpha_{k_0(\varepsilon)+\ell}(2L,\rho)\hat\mu(\Gamma_\rho).$$

□

Proof of F) One can check that

$$R^{i,L}_{\omega,h,k}\cap(D^{L,L^\gamma}_{\omega,i})^c\subset R^{i,2L^\gamma-i}_{\omega,h,k}\subset R^{i,L}_{\omega,h,k}\cup D^{L,L^\gamma}_{\omega,i}.$$

It remains to reapply the justification used in the proof of (A) to get that $\int_\Omega\mu_\omega(D^{L,L^\gamma}_{\omega,i})d\mathbb{P}(\omega)\leqslant\eta(\varepsilon)\hat\mu(\Gamma_\rho)$. □

■

4.3 An abstract approximation theorem

The following theorem approximates the probability distribution of an arbitrary sum of binary variables in terms of the distribution of a suitable sum of independent random variables. More precisely, to build the 'suitable' independent random variables, one splits the first sum into smaller block-sums, and each of them is distributionally mimicked by a new random variable, with the collection of new ones being taken to be independent.

Theorem 4.3.1. *Consider $n\geqslant 0$, $L\geqslant n$, $N\in\mathbb{N}_{\geqslant 3}$ large enough so that $L\leqslant\lfloor\frac{N}{3}\rfloor$, and $(X_i)_{i=0}^{N-1}$ arbitrary $\{0,1\}$-valued random variables on $(\mathcal{X},\mathscr{X},\mathbb{Q})$. Denote $N'_L:=\frac{N}{L}\in\mathbb{N}_{\geqslant 3}$*[11] *and $(Z^L_j)_{j=0}^{N'_L-1}$ given by $Z^L_j:=\sum_{i=jL}^{(j+1)L-1}X_i$*[12].

[11] Although L need not divide N, we pretend this is the case, for simplification purposes, i.e. to neglect possible remainder terms associated with the fractional part — which should not play a role in the asymptotics (of either the error and leading terms).

[12] This is the first instance in the text where the letter L is used to measure block-size, where L is fixed. Before, L was iteration-time and eventually sent to infinity, as in definitions (I-VI) of section 4.1.2. However, through the text, these use cases merge, in the sense that the block of size L iterates the maps L times: much before the L-limit we are inclined to see L as a fixed block-size, much closer to the L-limit we are inclined to see it as growing iteration-time. It is implicit that every time this merger occurs we will eventually want to take the L-limit. There is a special situation where this is not the case, to be seen in lemma 4.4.2 and lemma 4.4.3 item (3), where we will be equally interested in $L=1$ and $L\to\infty$.

Let $(\tilde{Z}_j^L)_{j=0}^{N_L'-1}$ *be an independency of* $\mathbb{N}_{\geqslant 0}$*-valued random variables on* $(X,\mathcal{X},\mathbb{Q})$[13] *satisfying* $\tilde{Z}_j^L \sim Z_j^L$ *(* $j=0,\dots,N_L'-1$*) and* $(\tilde{Z}_j^L)_{j=0}^{N_L'-1} \perp (Z_j^L)_{j=0}^{N_L'-1}$.

Denote $\tilde{W}_{a,b}^L := \sum_{j=a}^{b} \tilde{Z}_j^L$ $(0 \leqslant a \leqslant b \leqslant N_L'-1)$ *and* $\tilde{W}^L := \tilde{W}_{0,N_L'-1}^L$*. Similarly notation with* $\sim$*'s erased is adopted, in which case* W^L *coincides with* $W := \sum_{i=0}^{N-1} X_i$.

Then:

$$\left|\mathbb{Q}(W=n) - \mathbb{Q}(\tilde{W}^L=n)\right| \lesssim \left(\tilde{\mathcal{R}}^1(N,L) + \mathcal{R}^1(N,L,\Delta) + \mathcal{R}^2(N,L,\Delta) + \mathcal{R}^3(N,L,\Delta)\right),$$

where

$$\tilde{\mathcal{R}}^1(N,L,\Delta) = \sum_{j=0}^{N_L'-1} \max_{q\in[0,n]} \left|\mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+\Delta,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+\Delta,N_L'-1}^L = q)\right|,$$

$$\mathcal{R}^1(N,L,\Delta) = \sum_{j=0}^{N_L'-1} \max_{q\in[1,n]} \sum_{u=1}^{q} \left|\mathbb{Q}\Big(Z_j^L = u, W_{j+\Delta,N_L'-1}^L = q-u\Big) - \mathbb{Q}\Big(Z_j^L = u\Big)\mathbb{Q}\Big(W_{j+\Delta,N_L'-1}^L = q-u\Big)\right|,$$

$$\mathcal{R}^2(N,L,\Delta) = \sum_{j=0}^{N_L'-1} \mathbb{Q}\left(Z_j^L \geqslant 1, W_{j+1,j+\Delta-1}^L \geqslant 1\right) \text{ and}$$

$$\mathcal{R}^3(N,L,\Delta) = \sum_{i=0}^{N} \sum_{q=0\vee(i-\Delta L)}^{i} \mathbb{Q}(X_i = 1)\mathbb{Q}(X_q = 1),$$

with the convention that, for $b > a$, $W_{b,a}^L \equiv 0$ *and* $\mathbb{Q}(W_{b,a}^L \geqslant 1) = 0$.

Proof. Notice that by using a telescopic sum and the given independence, one has

$$\left|\mathbb{Q}(W=n) - \mathbb{Q}(\tilde{W}^L=n)\right| \leqslant \sum_{j=0}^{N_L'-1} \left|\mathbb{Q}(\tilde{W}_{0,j-1}^L + W_{j,N_L'-1}^L = n) - \mathbb{Q}(\tilde{W}_{0,j}^L + W_{j+1,N_L'-1}^L = n)\right|$$

$$\leqslant \sum_{j=0}^{N_L'-1} \sum_{l=0}^{n} \mathbb{Q}(\tilde{W}_{0,j-1}^L = l) \left|\mathbb{Q}(W_{j,N_L'-1}^L = n-l) - \mathbb{Q}(\tilde{Z}_j^L + W_{j+1,N_L'-1}^L = n-l)\right|.$$

We now estimate

$$\left|\mathbb{Q}(W_{j,N_L'-1}^L = q) - \mathbb{Q}(\tilde{Z}_j^L + W_{j+1,N_L'-1}^L = q)\right|$$

$$\leqslant \sum_{u=0}^{q} \left|\mathbb{Q}(Z_j^L = u, W_{j+1,N_L'-1}^L = q-u) - \mathbb{Q}(\tilde{Z}_j^L = u, W_{j+1,N_L'-1}^L = q-u)\right|$$

$$= \sum_{u=0}^{q} \left|\mathbb{Q}(Z_j^L = u, W_{j+1,N_L'-1}^L = q-u) - \mathbb{Q}(Z_j^L = u)\mathbb{Q}(W_{j+1,N_L'-1}^L = q-u)\right| =: \sum_{u=0}^{q} |\mathcal{R}_j(q,u)|.$$

[13] For the statement of the theorem, it seems unimportant that the domain of the mimicking random variables is that of the original ones, but this is used in the proof. Of course, $(\mathcal{X},\mathscr{X},\mathbb{Q})$ then has to be a rich enough space in order to accommodate the existence of such mimicking random variables. This will not be an issue in our application.

We single out $u = 0$ from the previous sum,

$$\begin{aligned}
|\mathcal{R}_j(q,0)| &= \Big|\mathbb{Q}(Z_j^L = 0, W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L = 0)\mathbb{Q}(W_{j+1,N_L'-1}^L = q)\Big| \\
&= \Big|\Big(\mathbb{Q}(W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,N_L'-1}^L = q)\Big) \\
&\quad - \Big(\mathbb{Q}(W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+1,N_L'-1}^L = q)\Big)\Big| \\
&= \Big|\mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,N_L'-1}^L = q)\Big|.
\end{aligned}$$

It follows that

$$\big|\mathbb{Q}(W = n) - \mathbb{Q}(\tilde{W}^L = n)\big| \leqslant \sum_{j=0}^{N_L'-1}\sum_{q=0}^{n}\sum_{u=0}^{q}|\mathcal{R}_j(q,u)|$$

$$\begin{aligned}
&\leqslant n\sum_{j=0}^{N_L'-1}\max_{q\in[0,n]}\Big|\mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,N_L'-1}^L = q)\Big| + \sum_{j=0}^{N_L'-1}\sum_{q=0}^{n}\sum_{u=1}^{q}|\mathcal{R}_j(q,u)| \\
&\lesssim \sum_{j=0}^{N_L'-1}\max_{q\in[0,n]}\Big|\mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+1,N_L'-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,N_L'-1}^L = q)\Big| + \sum_{j=0}^{N_L'-1}\sum_{q=0}^{n}\sum_{u=1}^{q}|\mathcal{R}_j(q,u)|.
\end{aligned}$$

The first summation will be kept on hold. We deal with the second one now, which will be split into three terms, as follows.

For $u = 1,\ldots,q$, we expand $|\mathcal{R}_j(q,u)|$ using the triangular inequality, where we include intermediate terms using the time gap Δ, to get the following three components

$$\begin{aligned}
|\mathcal{R}_j(q,u)| &\leqslant \Big|\mathbb{Q}(Z_j^L = u, W_{j+1,N_L'-1}^L = q-u) - \mathbb{Q}(Z_j^L = u, W_{j+\Delta,N_L'-1}^L = q-u)\Big| \\
&+ \Big|\mathbb{Q}(Z_j^L = u, W_{j+\Delta,N_L'-1}^L = q-u) - \mathbb{Q}(Z_j^L = u)\mathbb{Q}(W_{j+\Delta,N_L'-1}^L = q-u)\Big| \\
&+ \Big|\mathbb{Q}(Z_j^L = u)\mathbb{Q}(W_{j+\Delta,N_L'-1}^L = q-u) - \mathbb{Q}(Z_j^L = u)\mathbb{Q}(W_{j+1,N_L'-1}^L = q-u)\Big|,
\end{aligned}$$

where the entries in the RHS are denoted, respectively, by $|\mathcal{R}_j^2(q,u)|$, $|\mathcal{R}_j^1(q,u)|$ and $|\mathcal{R}_j^3(q,u)|$ (note the unusual order).

Then the following three terms bound the later triple sum.

First:

$$\sum_{j=0}^{N_L'-1}\sum_{q=0}^{n}\sum_{u=1}^{q}|\mathcal{R}_j^1(q,u)| \lesssim \sum_{j=0}^{N_L'-1}\max_{q\in[1,n]}\sum_{u=1}^{q}|\mathcal{R}_j^1(q,u)| = \mathcal{R}^1(N,L,\Delta).$$

Second:

$$\begin{aligned}
\sum_{j=0}^{N_L'-1}\sum_{q=0}^{n}\sum_{u=1}^{q}|\mathcal{R}_j^2(q,u)| &\lesssim \sum_{j=0}^{N_L'-1}\max_{q\in[1,n]}\sum_{u=1}^{q}|\mathcal{R}_j^2(q,u)| \\
&\lesssim \sum_{j=0}^{N_L'-1}\mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,j+\Delta-1}^L \geqslant 1) = \mathcal{R}^2(N,L,\Delta),
\end{aligned}$$

where the step used that

$$A_u := \{Z_j^L = u, W_{j+1,N_L'-1}^L = q-u\}, B_u := \{Z_j^L = u, W_{j+\Delta,N_L'-1}^L = q-u\}$$

$$\Rightarrow A_u \backslash B_u, B_u \backslash A_u \subset \{Z_j^L = u, W_{j+1,j+\Delta-1}^L \geqslant 1\}$$

$$\begin{aligned}
\Rightarrow \sum_{u=1}^{q} |\mathcal{R}_j^2(q,u)| &= \sum_{u=1}^{q} |\mathbb{Q}(A_u) - \mathbb{Q}(B_u)| \leqslant \sum_{u=1}^{q} \mathbb{Q}(Z_j^L = u, W_{j+1,j+\Delta-1}^L \geqslant 1) \\
&\leqslant \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,j+\Delta-1}^L \geqslant 1).
\end{aligned}$$

Third:

$$\sum_{j=0}^{N_L'-1} \sum_{q=0}^{n} \sum_{u=1}^{q} |\mathcal{R}_j^3(q,u)| \lesssim \sum_{j=0}^{N_L'-1} \max_{q \in [1,n]} \sum_{u=1}^{q} |\mathcal{R}_j^3(q,u)|$$

$$\lesssim \sum_{j=0}^{N_L'-1} \sum_{l=jL}^{(j+1)L-1} \sum_{i=(j+1)L}^{(j+\Delta+1)L-1} \mathbb{Q}(X_i{=}1)\mathbb{Q}(X_l{=}1) = \sum_{i=0}^{N+\Delta L+L} \sum_{l=0 \vee (i-L-\Delta L)}^{i-L} \mathbb{Q}(X_l{=}1)\mathbb{Q}(X_i{=}1)$$

$$\leqslant \sum_{i=0}^{N} \sum_{l=0 \vee (i-\Delta L)}^{i} \mathbb{Q}(X_l = 1)\mathbb{Q}(X_i = 1) = \mathcal{R}^3(N, L, \Delta),$$

where the second $\lesssim$ step used the following: (with $q' = q - u$)

$$\begin{aligned}
\mathbb{Q}(W_{j+1,N_L'-1}^L{=}q') &= \mathbb{Q}(Z_{j+1}^L{\geqslant}1, W_{j+1,N_L'-1}^L{=}q') + \mathbb{Q}(Z_{j+1}^L{=}0, W_{j+1,N_L'-1}^L{=}q') \\
\mathbb{Q}(Z_{j+1}^L = 0, W_{j+1,N_L'-1}^L = q') &= \mathbb{Q}(Z_{j+1}^L = 0, W_{j+2,N_L'-1}^L = q') \\
&= \mathbb{Q}(W_{j+2,N_L'-1}^L = q') - \mathbb{Q}(Z_{j+1}^L \geqslant 1, W_{j+2,N_L'-1}^L = q') \\
\Rightarrow \quad \begin{array}{c} |\mathbb{Q}(W_{j+1,N_L'-1}^L = q') \\ -\mathbb{Q}(W_{j+2,N_L'-1}^L = q')| \end{array} &= \begin{array}{c} |\mathbb{Q}(Z_{j+1}^L \geqslant 1, W_{j+1,N_L'-1}^L = q') \\ -\mathbb{Q}(Z_{j+1}^L \geqslant 1, W_{j+2,N_L'-1}^L = q')| \end{array}
\end{aligned}$$

but, with $A := \{Z_{j+1}^L \geqslant 1, W_{j+1,N_L'-1}^L = q'\}$ and $B := \{Z_{j+1}^L \geqslant 1, W_{j+2,N_L'-1}^L = q'\}$, one has $A \backslash B, B \backslash A \subset \{Z_{j+1}^L \geqslant 1\}$, implying

$$|\mathbb{Q}(W_{j+1,N_L'-1}^L = q') - \mathbb{Q}(W_{j+2,N_L'-1}^L = q')| \leqslant \mathbb{Q}(Z_{j+1}^L \geqslant 1)$$

$$\Rightarrow |\mathbb{Q}(W_{j+l,N_L'-1}^L = q') - \mathbb{Q}(W_{j+l+1,N_L'-1}^L = q')| \leqslant \mathbb{Q}(Z_{j+l}^L \geqslant 1) \leqslant \sum_{i=(j+l)L}^{(j+l+1)L-1} \mathbb{Q}(X_i = 1)$$

$$\Rightarrow |\mathbb{Q}(W_{j+1,N_L'-1}^L = q') - \mathbb{Q}(W_{j+\Delta,N_L'-1}^L = q')| \leqslant \sum_{l=1}^{\Delta-1} \mathbb{Q}(Z_{j+l}^L \geqslant 1) \leqslant \sum_{i=(j+1)L}^{(j+\Delta)L-1} \mathbb{Q}(X_i = 1)$$

$$\begin{aligned}
\Rightarrow \sum_{u=1}^{q} |\mathcal{R}_j^3(q,u)| &\leqslant \sum_{u=1}^{q} \mathbb{Q}(Z_j^L = u) \sum_{i=(j+1)L}^{(j+\Delta)L-1} \mathbb{Q}(X_i = 1) \\
&\leqslant \sum_{l=jL}^{(j+1)L-1} \sum_{i=(j+1)L}^{(j+\Delta)L-1} \mathbb{Q}(X_l = 1)\mathbb{Q}(X_i = 1).
\end{aligned}$$

Now we should deal with the summation we left on hold, coming from the singled-out term with $u = 0$, namely,

$$\sum_{j=0}^{N'_L-1} \max_{q\in[0,n]} \left| \mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+1,N'_L-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+1,N'_L-1}^L = q) \right|.$$

Using an analogous triangular inequality trick, by adding two mixed terms that have a gap Δ, we expand the previous absolute value into three parts. These three parts are named, in analogy to the previous terms, by $\tilde{\mathcal{R}}^1(N,L,\Delta)$, $\tilde{\mathcal{R}}^2(N,L,\Delta)$ and $\tilde{\mathcal{R}}^3(N,L,\Delta)$. The exact same procedures applied above to bound $\mathcal{R}^2$ and $\mathcal{R}^3$ reapply and it follows that $\tilde{\mathcal{R}}^2(N,L,\Delta) \leqslant \mathcal{R}^2(N,L,\Delta)$ and $\tilde{\mathcal{R}}^3(N,L,\Delta) \leqslant \mathcal{R}^3(N,L,\Delta)$. So we are just left with an extra term given by

$$\tilde{\mathcal{R}}^1(N,L,\Delta) = \sum_{j=0}^{N'_L-1} \max_{q\in[0,n]} \left| \mathbb{Q}(Z_j^L \geqslant 1)\mathbb{Q}(W_{j+\Delta,N'_L-1}^L = q) - \mathbb{Q}(Z_j^L \geqslant 1, W_{j+\Delta,N'_L-1}^L = q) \right|,$$

as desired. ∎

4.4 Borel-Cantelli type lemmata

The objective of this section is its final lemma 4.4.3, which will be used in the proof of theorem 4.1.4. This lemma and its proof strategy was inspired in [77] (lemma 9). To implement the said proof, we need to rely on the *ad-hoc* lemmas 4.4.1 and 4.4.2. Lemmas 4.4.1 and 4.4.2 are essentially independent, although lemma 4.4.1 uses return statistics in its hypothesis and relies on theorem 4.1.3 in its proof. We believe that the dependencies in the last sentence might not be intrinsic and could be untied.

Lemma 4.4.1. *Let $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$ be a system satisfying hypothesis (H9') (so (H8), by theorem 4.1.3) . Then:*

$$\lim_{L\to\infty} \overline{\lim_{\rho\to 0}} \frac{\hat{\mu}(Z_{\Gamma_\rho}^L \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} = \left(\textstyle\sum_{\ell=1}^{\infty} \ell\lambda_\ell\right)^{-1} = \alpha_1 \tag{4.21}$$

and

$$\lim_{L\to\infty} \overline{\lim_{\rho\to 0}} \frac{\hat{\mu}(Z_{\Gamma_\rho}^L = n)}{L\hat{\mu}(\Gamma_\rho)} = \left(\textstyle\sum_{\ell=1}^{\infty} \ell\lambda_\ell\right)^{-1}\lambda_n = \alpha_1\lambda_n \ (n \geqslant 1) \tag{4.22}$$

Proof. Using (H9') (for the following items (i.b-ii)) and (H8) (items (i.a,iii-iv)), it holds that: $\forall\varepsilon > 0$

i) $\exists \ell_0(\varepsilon) \geqslant 1$ so that

  a) $\sum_{\ell=\ell_0(\varepsilon)}^{\infty} \ell^3\lambda_\ell \leqslant \varepsilon$,

  b) $\forall L \geqslant 1$: $\displaystyle\sum_{\ell=\ell_0(\varepsilon)}^{\infty} \ell\overset{+}{\hat{\alpha}}_\ell(L) \leqslant \sum_{\ell=\ell_0(\varepsilon)}^{\infty} \ell\hat{\alpha}_\ell \leqslant \varepsilon$.

ii) $\forall L \geqslant 1, \exists \rho_1(\varepsilon, L), \forall \rho \leqslant \rho_1(\varepsilon, L)$:

$$\overset{-}{\hat{\alpha}}_\ell(L) - \varepsilon/(L^2) \leqslant \hat{\alpha}_\ell(L,\rho) \leqslant \overset{+}{\hat{\alpha}}_\ell(L) + \varepsilon/(L^2) \ (\forall \ell = 1, \dots, L)$$

$$\Rightarrow \sum_{\ell=\ell_0(\varepsilon)}^{L} \ell \hat{\alpha}_\ell(L,\rho) \leqslant \sum_{\ell=\ell_0(\varepsilon)}^{L} \ell \left(\overset{+}{\hat{\alpha}}_\ell(L) + \varepsilon/(L^2) \right) \leqslant 2\varepsilon \text{ by (i).}$$

iii) $\forall L \geqslant 1, \exists \rho_3(\varepsilon, L), \forall \rho \leqslant \rho_3(\varepsilon, L)$:

$$\overset{-}{\lambda}_\ell(L) - \varepsilon/(\ell_0(\varepsilon))^2 \leqslant \lambda_\ell(L,\rho) \leqslant \overset{+}{\lambda}_\ell(L) + \varepsilon/(\ell_0(\varepsilon))^2 \ (\forall \ell = 1, \dots, \ell_0(\varepsilon)).$$

iv) $\exists L_0(\varepsilon) > \ell_0(\varepsilon), \forall L \geqslant L_0(\varepsilon)$:

$$|\lambda_\ell - \overset{-}{\lambda}_\ell(L)| \leqslant \varepsilon/(\ell_0(\varepsilon))^2, \ |\lambda_\ell - \overset{+}{\lambda}_\ell(L)| \leqslant \varepsilon/(\ell_0(\varepsilon))^2 \ (\forall \ell = 1, \dots, \ell_0(\varepsilon))$$

$$\Rightarrow | \overset{+}{\lambda}_\ell(L) - \overset{-}{\lambda}_\ell(L)| \leqslant 2\varepsilon/(\ell_0(\varepsilon))^2 \ (\forall \ell = 1, \dots, \ell_0(\varepsilon)).$$

v) (due to items (iv-v)) $\exists L_0(\varepsilon), \forall L \geqslant L_0(\varepsilon), \exists \rho_3(\varepsilon, L), \forall \rho \leqslant \rho_3(\varepsilon, L)$:

$$|\lambda_\ell(L,\rho) - \overset{*}{\lambda}_\ell(L)| \leqslant 3\varepsilon/(\ell_0(\varepsilon))^2 \ (\forall \ell = 1, \dots, \ell_0(\varepsilon), \forall * \in \{-,+\})$$

$$\Rightarrow |\lambda_\ell(L,\rho) - \lambda_\ell| \leqslant 4\varepsilon/(\ell_0(\varepsilon))^2 \ (\forall \ell = 1, \dots, \ell_0(\varepsilon), \forall * \in \{-,+\})$$

$$\Rightarrow \left| \sum_{\ell=1}^{\ell_0(\varepsilon)} (\ell-1)\lambda_\ell(L,\rho) - \sum_{\ell=1}^{\ell_0(\varepsilon)} (\ell-1)\lambda_\ell \right| \leqslant \sum_{\ell=1}^{\ell_0(\varepsilon)} \ell_0(\varepsilon) 4\varepsilon/(\ell_0(\varepsilon))^2 \leqslant 4\varepsilon.$$

Now, considering any $\varepsilon < 1/5 \sum_{\ell=1}^{\infty} \ell \lambda_\ell$, $L \geqslant L_0(\varepsilon)$ and $\rho \leqslant \rho_1(\varepsilon, L) \wedge \rho_2(\varepsilon) \wedge \rho_3(\varepsilon, L)$, we evaluate the quantity of interest, $\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)/L\hat{\mu}(\Gamma_\rho)$, starting with its numerator:

$$\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1) = \int_\Omega \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} \geqslant 1) d\mathbb{P}(\omega) = \int_\Omega \mu_\omega \left(\bigcup_{j=0}^{L-1} (T^j_\omega)^{-1} \Gamma_\rho(\theta^j \omega) \right) d\mathbb{P}(\omega)$$

$$\overset{(\star)}{=} \int_\Omega \sum_{j=0}^{L-1} \mu_\omega((T^j_\omega)^{-1} \Gamma_\rho(\theta^j \omega)) d\mathbb{P}(\omega) - \int_\Omega \sum_{\ell=0}^{L-1} \ell \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} = \ell+1) d\mathbb{P}(\omega)$$

$$= L\hat{\mu}(\Gamma_\rho) - \int_\Omega \left(\sum_{\ell=0}^{L-1} \ell \lambda^\omega_{\ell+1}(L,\rho) \right) \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} > 0) d\mathbb{P}(\omega)$$

$$\begin{aligned} = L\hat{\mu}(\Gamma_\rho) \quad & - \int_\Omega \left(\sum_{\ell=0}^{\ell_0(\varepsilon)-1} \ell \lambda^\omega_{\ell+1}(L,\rho) \right) \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} > 0) d\mathbb{P}(\omega) \\ & - \int_\Omega \left(\sum_{\ell=\ell_0(\varepsilon)}^{\infty} \ell \lambda^\omega_{\ell+1}(L,\rho) \right) \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} > 0) d\mathbb{P}(\omega) \end{aligned}$$

$$= L\hat{\mu}(\Gamma_\rho) - \sum_{\ell=0}^{\ell_0(\varepsilon)-1} \ell \hat{\mu}(Z^L_{\Gamma_\rho} = \ell+1) - \int_\Omega \left(\sum_{\ell=\ell_0(\varepsilon)}^{\infty} \ell \lambda^\omega_{\ell+1}(L,\rho) \right) \mu_\omega(Z^{\omega,L}_{\Gamma_\rho} > 0) d\mathbb{P}(\omega)$$

$$
\begin{aligned}
= L\hat{\mu}(\Gamma_\rho) \quad &- \quad \left(\sum_{\ell=1}^{\ell_0(\varepsilon)}(\ell-1)\lambda_\ell(L,\rho)\right)\hat{\mu}(Z^L_{\Gamma_\rho}>0)\\
&- \quad \int_\Omega\left(\sum_{\ell=\ell_0(\varepsilon)+1}^{\infty}(\ell-1)\lambda^\omega_\ell(L,\rho)\right)\mu_\omega(Z^{\omega,L}_{\Gamma_\rho}>0)d\mathbb{P}(\omega)
\end{aligned}
$$

where $(\star)$ applied a typical Venn diagram argument using overcounting and correction.

Then we consider the following two estimates.

First, we have that:

$$
\sum_{\ell=1}^{\ell_0(\varepsilon)}(\ell-1)\lambda_\ell(L,\rho)\overset{(v)}{\leqslant}\sum_{\ell=1}^{\ell_0(\varepsilon)}(\ell-1)\lambda_\ell+4\varepsilon\leqslant\sum_{\ell=1}^{\infty}(\ell-1)\lambda_\ell+5\varepsilon \text{ and}
$$

$$
\begin{aligned}
\sum_{\ell=1}^{\ell_0(\varepsilon)}(\ell-1)\lambda_\ell(L,\rho) \quad &\overset{(v)}{\geqslant} \quad \sum_{\ell=1}^{\ell_0(\varepsilon)}(\ell-1)\lambda_\ell-4\varepsilon=\sum_{\ell=1}^{\infty}(\ell-1)\lambda_\ell-\sum_{\ell=\ell_0(\varepsilon)+1}^{\infty}(\ell-1)\lambda_\ell-4\varepsilon\\
&\overset{(i.a)}{\geqslant} \quad \sum_{\ell=1}^{\infty}(\ell-1)\lambda_\ell-5\varepsilon.
\end{aligned}
$$

Second, with $\upsilon^\omega_{\Gamma_\rho}(x)=\inf\{j\geqslant 0: T^j_\omega\in\Gamma_\rho(\theta^j\omega)\}$, we have that:

$$
\begin{aligned}
0 &\leqslant \int_\Omega\left(\sum_{\ell=\ell_0(\varepsilon)+1}^{\infty}(\ell-1)\lambda^\omega_\ell(L,\rho)\right)\mu_\omega(Z^{\omega,L}_{\Gamma_\rho}>0)d\mathbb{P}(\omega)\leqslant\sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\hat{\mu}(Z^L_{\Gamma_\rho}=\ell)\\
&= \sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\sum_{j=0}^{L-1}\hat{\mu}(Z^L_{\Gamma_\rho}=\ell,\upsilon_{\Gamma_\rho}=j)\leqslant\sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\sum_{j=0}^{L-1}\hat{\mu}(Z^{L-j}_{\Gamma_\rho}\circ S^j=\ell,(S^j)^{-1}\Gamma_\rho)\\
&= \sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\sum_{j=0}^{L-1}\alpha_\ell(L-j,\rho)\hat{\mu}(\Gamma_\rho)=\left(\sum_{j=0}^{L-1}\sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\alpha_\ell(L-j,\rho)\right)\hat{\mu}(\Gamma_\rho)\\
&\leqslant \left[\sum_{j=0}^{L-1}\left(\sum_{\ell=\ell_0(\varepsilon)+1}^{L}\hat{\alpha}_\ell(L-j,\rho)\right)+\ell_0(\varepsilon)\hat{\alpha}_{\ell_0(\varepsilon)}(L-j,\rho)-L\hat{\alpha}_{L+1}(L-j,\rho)\right]\hat{\mu}(\Gamma_\rho)\\
&\leqslant \left[\sum_{j=0}^{L-1}\sum_{\ell=\ell_0(\varepsilon)+1}^{L}\ell\hat{\alpha}_\ell(L-j,\rho)\right]\hat{\mu}(\Gamma_\rho)\overset{(ii)}{\leqslant}2\varepsilon L\hat{\mu}(\Gamma_\rho)
\end{aligned}
$$

Combining what we got so far, it follows that:

$$\begin{aligned}
\frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} &\leqslant \frac{L\hat{\mu}(\Gamma_\rho) - \left(\sum_{\ell=1}^{\infty}(\ell-1)\lambda_\ell - 5\varepsilon\right)\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} \\
&= 1 - \left(\sum_{\ell=1}^{\infty} \ell\lambda_\ell - 1 - 5\varepsilon\right)\frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} \\
\Rightarrow \frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} &\leqslant \frac{1}{\sum_{\ell=1}^{\infty} \ell\lambda_\ell - 5\varepsilon}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} &\geqslant \frac{L\hat{\mu}(\Gamma_\rho) - \left(\sum_{\ell=1}^{\infty}(\ell-1)\lambda_\ell + 5\varepsilon\right)\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1) - 2\varepsilon L\hat{\mu}(\Gamma_\rho)}{L\hat{\mu}(\Gamma_\rho)} \\
&= 1 - \left(\sum_{\ell=1}^{\infty} \ell\lambda_\ell - 1 + 5\varepsilon\right)\frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} - 2\varepsilon \\
\Rightarrow \frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} &\geqslant \frac{1-2\varepsilon}{\sum_{\ell=1}^{\infty} \ell\lambda_\ell + 5\varepsilon}
\end{aligned}$$

Considering the final two inequalities and passing $\lim_{\varepsilon\to 0}\overline{\lim}_{L\to\infty}\overline{\lim}_{\rho\to 0}$ we observe that

$$\overline{\lim_{L\to\infty}}\,\overline{\lim_{\rho\to 0}}\,\frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)} = \left(\sum_{k=1}^{\infty} k\lambda_k\right)^{-1} = \alpha_1$$

Alternating between lim sup's and lim inf's lets us reach the first desired conclusion.

Finally, to take care of the second desired conclusion, it suffices to note that

$$\frac{\hat{\mu}(Z^L_{\Gamma_\rho} = n)}{L\hat{\mu}(\Gamma_\rho)} = \frac{\hat{\mu}(Z^L_{\Gamma_\rho} \geqslant 1)}{L\hat{\mu}(\Gamma_\rho)}\,\frac{\hat{\mu}(Z^L_{\Gamma_\rho} = n)}{\hat{\mu}(Z^L_{\Gamma_\rho} > 0)},$$

then take the appropriate limits and apply the first conclusion we have just proved (to obtain α_1), together with the definition of λ_n. ■

Lemma 4.4.2. *Let $(\theta,\mathbb{P},T_\omega,\mu_\omega,\Gamma)$ be a system satisfying hypotheses (H1), (H3.1), (H4.1), (H5), (H6.1), (H6.2), (H7.1) and (H7.2) with the parametric constraint (H10.1).*

Then: $\forall t>0, \forall n\geqslant 1$, $\forall L\geqslant 1$[14], $\exists\rho_{\mathrm{var}}(L)>0$, $\forall\rho\leqslant\rho_{\mathrm{var}}(L)$ small enough so that $N_\rho := \lfloor\frac{t}{\hat{\mu}(\Gamma_\rho)}\rfloor \geqslant 3$ and $N'_{\rho,L} := \frac{N_\rho}{L} \in \mathbb{N}_{\geqslant 3}$[15], one has:

$$\mathrm{var}_{\mathbb{P}}(\mathfrak{W}^{L,n}_\rho) \leqslant C_{t,L}\cdot\rho^q,\ \forall q\in\big(0, q(d_0,d_1,\eta,\beta,\mathfrak{p})\big),$$

[14] See footnote 12.
[15] See footnote 11.

where[16]

$$\mathfrak{W}_\rho^{L,n}(\omega) := \sum_{j=0}^{N'_{\rho,L}-1} \mu_\omega(Z_j^{\omega,\rho,L} = n), Z_j^{\omega,\rho,L} := \sum_{l=jL}^{(j+1)L-1} I_l^{\omega,\rho}, I_l^{\omega,\rho} := \mathbb{1}_{\Gamma_\rho(\theta^l\omega)} \circ T_\omega^l$$

and $q(d_0, d_1, \eta, \beta, \mathfrak{p})$ *is a positive quantity to be presented in the proof (which can be written explicitly).*

Proof. Let t, n and L be as in the statement. Fix $\alpha \in (0,1)$. Set $\rho_{\text{var}}(L) \leqslant \rho_{\text{sep}}(L) \wedge \rho_{\text{dim}}$ small enough so that $N_\rho{}^\alpha < N'_{\rho,L}$. Consider $\rho \leqslant \rho_{\text{var}}(L)$ as in the statement.

For a given $j \in [0, N'_{\rho,L} - 1]$, write $\omega' = \theta^{jL}\omega$ and notice that

$$\begin{aligned}
\mathbb{E}_\mathbb{P}(\mathfrak{W}_\rho^{L,n}) &= \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}\left(\mu_\omega(Z_j^{\omega,\rho,L} = n)\right) = \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}\left(\mu_\omega(\textstyle\sum_{l=jL}^{(j+1)L-1} \mathbb{1}_{\Gamma_\rho(\theta^l\omega)} \circ T_\omega^l = n)\right) \\
&= \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}\left(\mu_\omega(\textstyle\sum_{i=0}^{L-1} \mathbb{1}_{\Gamma_\rho(\theta^i\omega')} \circ T_{\omega'}^i \circ T_\omega^{jL} = n)\right) = \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}\left(\mu_{\omega'}(\textstyle\sum_{i=0}^{L-1} \mathbb{1}_{\Gamma_\rho(\theta^i\omega')} \circ T_{\omega'}^i = n)\right) \\
&= \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}\left(\mu_\omega(\textstyle\sum_{i=0}^{L-1} \mathbb{1}_{\Gamma_\rho(\theta^i\omega)} \circ T_\omega^i = n)\right) = \sum_{j=0}^{N'_{\rho,L}-1} \mathbb{E}_\mathbb{P}(\mu_\omega(Z_0^{\omega,\rho,L} = n)) \\
&= \sum_{j=0}^{N'_{\rho,L}-1} \hat{\mu}(Z_0^{\rho,L} = n) = N'_{\rho,L}\hat{\mu}(Z_0^{\rho,L} = n).
\end{aligned}$$

Now fix $\Delta := N_\rho{}^\alpha < N'_{\rho,L}$. Then:

$$\begin{aligned}
\mathbb{E}_\mathbb{P}((\mathfrak{W}_\rho^{L,n})^2) &= \sum_{i,j=0}^{N'_{\rho,L}-1} \int_\Omega \mu_\omega(Z_i^{\omega,\rho,L} = n)\mu_\omega(Z_j^{\omega,\rho,L} = n)d\mathbb{P}(\omega) \\
&= 2\sum_{i=0}^{N'_{\rho,L}-1} \sum_{j=i}^{(i+\Delta)\wedge(N'_{\rho,L}-1)} \int_\Omega \mu_\omega(Z_i^{\omega,\rho,L} = n)\mu_\omega(Z_j^{\omega,\rho,L} = n)d\mathbb{P}(\omega) \\
&+ 2\sum_{i=0}^{N'_{\rho,L}-1} \sum_{j=(i+\Delta)\wedge(N'_{\rho,L}-1)+1}^{N'_{\rho,L}-1} \int_\Omega \mu_\omega(Z_i^{\omega,\rho,L} = n)\mu_\omega(Z_j^{\omega,\rho,L} = n)d\mathbb{P}(\omega) \\
&=: (I) + (II).
\end{aligned}$$

Immediately we get that

$$\mu_\omega(Z_j^{\omega,\rho,L} = n) \leqslant \mu_\omega(Z_j^{\omega,\rho,L} \geqslant 1) \leqslant \sum_{l=jL}^{(j+1)L-1} \mu_{\theta^l\omega}(\Gamma_\rho(\theta^l\omega)) \overset{\text{(H6.2)}}{\lesssim} L\rho^{d_0}$$

$$\Rightarrow (I) \lesssim L\rho^{d_0}\Delta\mathbb{E}_\mathbb{P}(\mathfrak{W}_\rho^{L,k}) = \Delta\rho^{d_0}N_\rho\hat{\mu}(Z_0^{\rho,L} = n).$$

Most of the remaining work is to control component (II).

[16] The notation $Z_j^{\omega,\rho,L}$ is in parallel to that of Z_j^L in theorem 4.3.1. They, on purpose, resemble that $Z_{\Gamma_\rho}^{\omega,L}$ introduced in definition 4.1.2.

Fix $\omega \in \Omega$ and, for a given $i \in [0, N'_{\rho,L} - 1]$, write $\omega' = \theta^{iL}\omega$. Moreover, consider $r \in (0, \rho/2)$, $v \in [0, L-1]$ and denote by

$$U_{v,\omega'} = \Gamma_\rho(\theta^v \omega'), \ \overset{-}{U}_{v,r,\omega'} = B_r({U_{v,\omega'}}^c)^c, \ \overset{+}{U}_{v,r,\omega'} = B_r(U_{v,\omega'}), \tag{4.23}$$

respectively, the ρ-sized target with seed ω' v-steps ahead; its diminishment by radius r; and its enlargment by radius r. They relate as $\overset{-}{U}_{v,r,\omega'} \subset U_{v,\omega'} \subset \overset{+}{U}_{v,r,\omega'}$.

Moreover, dynamical counterparts of those in equation (4.23) are denote by

$$\begin{aligned}
\{Z_0^{\omega',\rho,L} = n\} \quad = \quad \mathcal{U}_{\omega'} \quad &= \bigsqcup_{0 \leqslant v_1 < \ldots < v_n \leqslant L-1} \left(\bigcap_{l=1}^{n} (T_{\omega'}^{v_l})^{-1} U_{v_l,\omega'} \ \cap \bigcap_{\substack{v \in [0,L-1] \\ \setminus \{v_l : l = 1,\ldots,n\}}} (T_{\omega'}^{v})^{-1} {U_{v,\omega'}}^c \right), \\
\overset{-}{\mathcal{U}}_{r,\omega'} \quad &= \bigsqcup_{0 \leqslant v_1 < \ldots < v_n \leqslant L-1} \left(\bigcap_{l=1}^{n} (T_{\omega'}^{v_l})^{-1} \overset{-}{U}_{v_l,\omega'} \ \cap \bigcap_{\substack{v \in [0,L-1] \\ \setminus \{v_l : l = 1,\ldots,n\}}} (T_{\omega'}^{v})^{-1} {\overset{+}{U}_{v,\omega'}}^c \right), \\
\overset{+}{\mathcal{U}}_{r,\omega'} \quad &= \bigsqcup_{0 \leqslant v_1 < \ldots < v_n \leqslant L-1} \left(\bigcap_{l=1}^{n} (T_{\omega'}^{v_l})^{-1} \overset{+}{U}_{v_l,\omega'} \ \cap \bigcap_{\substack{v \in [0,L-1] \\ \setminus \{v_l : l = 1,\ldots,n\}}} (T_{\omega'}^{v})^{-1} {\overset{-}{U}_{v,\omega'}}^c \right),
\end{aligned}$$

describing

- the locus of points which hit the ρ-sized target exactly k times during the time interval $[0, L-1]$ when given the random seed ω';

- the diminishment of the first by radius r, in the sense that hits are considered in a r-stringent way (at least r-inside the ρ-sized target) and non-hits are considered in a r-stringent way (at least r-away from the ρ-sized target);

- the enlargment of the first by radius r, in the sense that hits are considered in a r-permissive way (at most r-away from the ρ-sized target) and non-hits are considered in a r-permissive way (at most r-inside the ρ-sized target).

They relate as $\overset{-}{\mathcal{U}}_{r,\omega'} \subset \mathcal{U}_{\omega'} \subset \overset{+}{\mathcal{U}}_{r,\omega'}$.

Finally, define

$$\overset{-}{\phi}{}_r^{\omega'}(x) = \begin{cases} 1, x \in \overset{-}{\mathcal{U}}_{r,\omega'} \\ 0, x \in {\mathcal{U}_{\omega'}}^c \\ \dfrac{d_M(x, {\mathcal{U}_{\omega'}}^c)}{d_M(x, {\mathcal{U}_{\omega'}}^c) + d_M(x, \overset{-}{\mathcal{U}}_{r,\omega'})}, x \in {\mathcal{U}_{\omega'}}^c \setminus \overset{-}{\mathcal{U}}_{r,\omega'} \end{cases} \qquad \overset{+}{\phi}{}_r^{\omega'}(x) = \begin{cases} 1, x \in \mathcal{U}_{\omega'} \\ 0, x \in {\overset{+}{\mathcal{U}}_{r,\omega'}}^c \\ \dfrac{d_M(x, {\overset{+}{\mathcal{U}}_{r,\omega'}}^c)}{d_M(x, {\overset{+}{\mathcal{U}}_{r,\omega'}}^c) + d_M(x, \mathcal{U}_{\omega'})}, x \in \overset{+}{\mathcal{U}}_{r,\omega'} \setminus \mathcal{U}_{\omega'} \end{cases}.$$

They relate as $\overset{-}{\phi}{}_r^{\omega'} \leqslant \mathbb{1}_{\mathcal{U}_{\omega'}} \leqslant \overset{+}{\phi}{}_r^{\omega'}$.

Using that $\mathrm{Lip}_{d_M}\left(d_M(x,\overset{+}{\mathcal{U}}_{r,\omega'})\right), \mathrm{Lip}_{d_M}\left(d_M(x,\mathcal{U}_{\omega'})\right) \leqslant 1$, it can be checked that

$$\mathrm{Lip}_{d_M}(\overset{+}{\phi}{}_r^{\omega'}) \leqslant \frac{6\,\mathrm{diam}(M)}{\left(\min_{x\in M}[d_M(x,\overset{+}{\mathcal{U}}_{r,\omega'})+d_M(x,\mathcal{U}_{\omega'})]\right)^2} \leqslant \frac{6\,\mathrm{diam}(M)}{d_{min}\left(\mathcal{U}_{\omega'},\overset{+}{\mathcal{U}}_{r,\omega'}{}^c\right)^2},$$

where $d_{min}\left(\mathcal{U}_{\omega'},\overset{+}{\mathcal{U}}_{r,\omega'}{}^c\right) := \inf\{d_M(x,y) : x\in\mathcal{U}_{\omega'}, y\in\overset{+}{\mathcal{U}}_{r,\omega'}{}^c\}$.

Notice that for a point $x\in\mathcal{U}_{\omega'}$ to be minimally-displaced in such a way as to reach $\overset{+}{\mathcal{U}}_{r,\omega'}{}^c$, either: a) some of the hits in its finite-orbit is consequently-displaced to an extent which now makes it at least r-away from associated ρ-sized target, or b) some of the non-hits in its finite-orbit is consequently-displaced to an extent which now makes it at least r-inside the associated ρ-sized target. In either case, the associated image point of x has to be consequently-displaced by distance at least r. When the said image point being consequently-displaced happens to be the last one in the orbit of x, i.e., its $L-1$ iterate, by the expanding feature of the system (H3.1) and since $\mathcal{U}_{\omega'} \subset \bigcup_{j=0}^{L-1}(T_{\omega'}^j)^{-1}\Gamma_{3/2\rho}(\theta^j\omega') \overset{\text{(H4.1)}}{\subset} \overset{+}{\mathcal{C}}{}^{\omega'}_{L-1}$, this is when x would have to be displaced the least: no more than r/a_{L-1} (use (H4.1) and (H3.2)). Therefore $r/a_{L-1} \leqslant d_{min}\left(\mathcal{U}_{\omega'},\overset{+}{\mathcal{U}}_{r,\omega'}{}^c\right)$, and so

$$\mathrm{Lip}_{d_M}(\overset{+}{\phi}{}_r^{\omega'}) \leqslant 6\,\mathrm{diam}(M)a_{L-1}{}^2/r^2$$
$$\|\overset{+}{\phi}{}_r^{\omega'}\|_{Lip_{d_M}} = \|\overset{+}{\phi}{}_r^{\omega'}\|_\infty \vee \mathrm{Lip}_{d_M}(\overset{+}{\phi}{}_r^{\omega'}) = 1 \vee \mathrm{Lip}_{d_M}(\overset{+}{\phi}{}_r^{\omega'}) = \mathrm{Lip}_{d_M}(\overset{+}{\phi}{}_r^{\omega'}) \leqslant 6\,\mathrm{diam}(M)a_{L-1}{}^2/r^2$$

where the last equality follows from ρ sufficiently small.

Now we start looking at (II) directly:

$$\left|\int_\Omega \mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_\omega(Z_i^{\omega,\rho,L}=n)d\mathbb{P}(\omega) - \int_\Omega \mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega)\right|$$
$$= \left|\int_\Omega \mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_{\omega'}(\mathbb{1}_{\mathcal{U}_{\omega'}})d\mathbb{P}(\omega) - \int_\Omega \mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega)\right|$$
$$\lesssim \int_\Omega \mu_{\theta^j\omega}(Z_0^{\theta^j\omega,\rho,L}=n)L\frac{r^\eta}{\rho^\beta}d\mathbb{P}(\omega) = L\frac{r^\eta}{\rho^\beta}\hat{\mu}(Z_0^{\rho,L}=n),$$

where the last inequality is because

$$\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'}) \leqslant \mu_{\omega'}(\overset{+}{\mathcal{U}}_{r,\omega'}\setminus\overset{-}{\mathcal{U}}_{r,\omega'}) \leqslant \sum_{v=0}^{L-1}\mu_{\theta^v\omega}(\overset{+}{U}_{v,r,\omega'}\setminus\overset{-}{U}_{v,r,\omega'}) \overset{\text{(H6.2)}}{\lesssim} L\frac{r^\eta}{\rho^\beta}.$$

The approximating term that appeared above is transformed as follows:

$$\left|\int_\Omega \mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega) - \int_\Omega \mu_{\omega'}(\mathbb{1}_{\{Z_{j-i}^{\omega',\rho,L}=n\}}\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega)\right|$$
$$= \left|\int_\Omega \mu_{\omega'}(Z_{j-i}^{\omega',\rho,L}=n)\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega) - \int_\Omega \mu_{\omega'}(\mathbb{1}_{\{Z_0^{\theta^{(j-i)L}\omega',\rho,L}\circ T_{\omega'}^{(j-i)L}=n\}}\overset{+}{\phi}{}_r^{\omega'})d\mathbb{P}(\omega)\right|$$
$$= \int_\Omega \left|\mu_{\theta^{(j-i)L}\omega'}(Z_0^{\theta^{(j-i)L}\omega',\rho,L}=n)\mu_{\omega'}(\overset{+}{\phi}{}_r^{\omega'}) - \mu_{\omega'}(\mathbb{1}_{\{Z_0^{\theta^{(j-i)L}\omega',\rho,L}=n\}}\circ T_{\omega'}^{(j-i)L}\,\overset{+}{\phi}{}_r^{\omega'})\right|d\mathbb{P}(\omega)$$

$$\overset{(H7.1)}{\lesssim} \int_\Omega ((j-i)L)^{-\mathfrak{p}} \|\overset{+}{\phi}{}_r^{\omega'}\|_{\mathrm{Lip}_{d_M}} 1 d\mathbb{P}(\omega) \lesssim^{17} ((j-i)L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2}.$$

Whereas the new approximating term which appeared above is transformed as follows:

$$\int_\Omega \mu_{\omega'}(\mathbb{1}_{\{Z_{j-i}^{\omega',\rho,L}=n\}} \overset{+}{\phi}{}_r^{\omega'}) d\mathbb{P}(\omega) = \int_\Omega \mu_\omega(\mathbb{1}_{\{Z_{j-i}^{\omega,\rho,L}=n\}} \overset{+}{\phi}{}_r^{\omega}) d\mathbb{P}(\omega)$$

$$= \int_{\Omega\times M} \mathbb{1}_{\{Z_{j-i}^{\rho,L}=n\}} \overset{+}{\phi}{}_r d\hat{\mu} = \int_{\Omega\times M} \mathbb{1}_{\{Z_0^{\rho,L}=n\}} \circ S^{(j-i)L} \overset{+}{\phi}{}_r d\hat{\mu} \tag{4.24}$$

and

$$\left| \int_{\Omega\times M} \mathbb{1}_{\{Z_0^{\rho,L}=n\}} \circ S^{(j-i)L} \overset{+}{\phi}{}_r d\hat{\mu} - \int_{\Omega\times M} \mathbb{1}_{\{Z_0^{\rho,L}=n\}} d\hat{\mu} \cdot \int_{\Omega\times M} \overset{+}{\phi}{}_r d\hat{\mu} \right|$$

$$\overset{(H7.2)}{\lesssim} ((j-i)L)^{-\mathfrak{p}} \|\overset{+}{\phi}{}_r\|_{\mathrm{Lip}_{d_{\Omega\times M}}} \lesssim ((j-i)L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2},$$

where, recalling that $\overset{+}{\phi}{}_r^{\omega} = \overset{+}{\phi}{}_r^{\omega}(n,L,\rho)$, we have used that

$$\begin{aligned}
\mathrm{Lip}_{d_{\Omega\times M}}(\overset{+}{\phi}{}_r) &= \sup_{(\omega_1,x_1)\neq(\omega_2,x_2)} \frac{|\overset{+}{\phi}{}_r^{\omega_1}(x_1) - \overset{+}{\phi}{}_r^{\omega_2}(x_2)|}{d_\Omega(\omega_1,\omega_2) \vee d_M(x_1,x_2)} \\
&\leqslant \sup_{x_1} \sup_{\omega_1\neq\omega_2} \frac{|\overset{+}{\phi}{}_r^{\omega_1}(x_1) - \overset{+}{\phi}{}_r^{\omega_2}(x_1)|}{d_\Omega(\omega_1,\omega_2)} + \sup_{\omega_2} \sup_{x_1\neq x_2} \frac{|\overset{+}{\phi}{}_r^{\omega_2}(x_1) - \overset{+}{\phi}{}_r^{\omega_2}(x_2)|}{d_M(x_1,x_2)} \\
&\leqslant^{18} \frac{{a_{L-1}}^2}{r^2} + \sup_{x} \sup_{\omega_1\neq\omega_2} \frac{|\overset{+}{\phi}{}_r^{\omega_1}(x) - \overset{+}{\phi}{}_r^{\omega_2}(x)|}{d_\Omega(\omega_1,\omega_2)} \\
&\overset{(\star)}{\leqslant} \frac{{a_{L-1}}^2}{r^2} + \frac{(\alpha^L\beta+\gamma){a_{L-1}}^2}{r^2} \lesssim \frac{\alpha^L {a_{L-1}}^2}{r^2},
\end{aligned}$$

with $(\star)$ following from $\omega \mapsto \overset{+}{\phi}{}_r^{\omega}(x)$ being a locally Lipschitz function whose associated local Lipschitz constants are bounded by $\frac{(\alpha^L\beta+\gamma){a_{L-1}}^2}{r^2}$, where $\alpha = \mathrm{Lip}(\theta) \vee 1$, $\beta = \mathrm{Lip}(\Gamma) \vee 1$, $\gamma = \sup_{x\in M} \mathrm{Lip}(T_\cdot(x) : \Omega \to M)$. This is verified in the following paragraph.

Fix $x \in M$ and consider $\omega \in \Omega$. In case $x \in \mathrm{int}(\mathcal{U}_\omega)$ (or $\mathrm{int}({\overset{+}{\mathcal{U}}_{r,\omega}}^c)$), there is $u_x(\omega) > 0$ so that $\tilde{\omega} \in B_{u_x(\omega)}(\omega)$ implies $x \in \mathrm{int}\,\mathcal{U}_{\tilde{\omega}}$ (or $\mathrm{int}\,{\overset{+}{\mathcal{U}}_{r,\tilde{\omega}}}^c$), so the function of interest is locally constant. In case $x \in \mathrm{int}({\overset{+}{\mathcal{U}}_{r,\omega}}^c \setminus \mathcal{U}_\omega)$, it boils down to understand how the linear interpolation within $\overset{+}{\phi}{}_r$ varies with $\tilde{\omega} \in B_{u'_x(\omega)}(\omega)$, where $u'_x(\omega)$ is that for which $\tilde{\omega} \in B_{u'_x(\omega)}(\omega)$ implies $x \in \mathrm{int}({\overset{+}{\mathcal{U}}_{r,\tilde{\omega}}}^c \setminus \mathcal{U}_{\tilde{\omega}})$. For this purpose, we first evaluate the Lipschitz constant of $\tilde{\omega} \in B_{u'_x(\omega)}(\omega) \mapsto d(x, \mathcal{U}_{\tilde{\omega}})$ and $\tilde{\omega} \in B_{u'_x(\omega)}(\omega) \mapsto d(x, {\overset{+}{\mathcal{U}}_{r,\tilde{\omega}}}^c)$:

---

[17] Notice that $\|\overset{+}{\phi}{}_r^{\omega'}\|_{\mathrm{Lip}_{d_M}} \lesssim {a_{L-1}}^2/r^2$ a.s. is enough to justify the above inequality. However, our hypotheses imply this is true for every ω. This might seem an excess, but later in the proof we will need the inequality for every ω. See the next footnote.

[18] Here one needs $\|\overset{+}{\phi}{}_r^{\omega'}\|_{\mathrm{Lip}_{d_M}} \leqslant {a_{L_1}}^2/r^2$ for every ω. See the previous footnote.

i)

$$\begin{aligned} &|d(x,\mathcal{U}_\omega)-d(x,\mathcal{U}_{\tilde\omega})| \leqslant d_H(\mathcal{U}_\omega,\mathcal{U}_{\tilde\omega}) \\ \leqslant &\begin{pmatrix} (\sup_\omega \operatorname{Lip}(T_\omega^{-1}:\mathscr{P}(M)\to\mathscr{P}(M))\vee 1)^L\cdot(\operatorname{Lip}(\Gamma)\vee 1)\cdot(\operatorname{Lip}(\theta)\vee 1)^L \\ +\sup_{A\in\mathscr{P}(M)}\operatorname{Lip}(T_\cdot^{-1}A:\Omega\to\mathscr{P}(M)) \end{pmatrix} d_\Omega(\omega,\tilde\omega) \\ &\leqslant (\alpha^L\beta+\gamma)d_\Omega(\omega,\tilde\omega). \end{aligned}$$

since

$$\operatorname{Lip}\left(\bigcap:\mathscr{P}(M)\times\mathscr{P}(M)\to\mathscr{P}(M)\right)\leqslant 1,$$

$$\operatorname{Lip}\left(\bigcup:\mathscr{P}(M)\times\mathscr{P}(M)\to\mathscr{P}(M)\right)\leqslant 1,$$

$$\operatorname{Lip}\left(B_\rho:\mathscr{P}(M)\to\mathscr{P}(M)\right)\leqslant 1,$$

$$\sup_\omega \operatorname{Lip}\left(T_\omega^{-1}:\mathscr{P}(M)\to\mathscr{P}(M)\right)\leqslant 1/\inf_{\omega\in\Omega}\inf_{\xi\in C_1^\omega}\operatorname{CoLip}(T_\omega|_\xi:\xi\to M)\leqslant 1$$

and

$$\sup_{A\in\mathscr{P}(M)}\operatorname{Lip}\left(T_\cdot^{-1}A:\Omega\to\mathscr{P}(M)\right)\leqslant\frac{\sup_{x\in M}\operatorname{Lip}(T_\cdot(x):\Omega\to M)}{\inf_{\omega\in\Omega}\inf_{\xi\in C_1^\omega}\operatorname{CoLip}(T_\omega|_\xi:\xi\to M)}\leqslant\sup_{x\in M}\operatorname{Lip}(T_\cdot(x):\Omega\to M),$$

where $\operatorname{CoLip(T)}=\inf_{x\neq y}\frac{d(Tx,Ty)}{d(x,y)}$.
ii) Similarly,

$$|d(x,\overset{+}{\mathcal{U}}_{r,\omega}{}^c)-d(x,\overset{+}{\mathcal{U}}_{r,\tilde\omega}{}^c)|\leqslant(\alpha^L\beta+\gamma)d_\Omega(\omega,\tilde\omega),$$

since also $\operatorname{Lip}(B_r:\mathscr{P}(M)\to\mathscr{P}(M))\leqslant 1$.

To conclude justifying $(\star)$, one repeats the calculations for the Lipschitz constant of a quotient and applies (i) and (ii) to get that

$$\begin{aligned}\operatorname{Lip}_{d_\Omega}\left(\frac{d_M(x,\overset{+}{\mathcal{U}}_{r,\omega'}{}^c)}{d_M(x,\overset{+}{\mathcal{U}}_{r,\omega'}{}^c)+d_M(x,\mathcal{U}_{\omega'})}\right) &\leqslant \frac{4\operatorname{diam}(M)(\alpha^L\beta+\gamma)}{d_{min}(\overset{+}{\mathcal{U}}_{r,\omega'}{}^c,\mathcal{U}_{\omega'})^2}d_\Omega(\omega,\tilde\omega) \\ &\lesssim \frac{(\alpha^L\beta+\gamma)a_{L-1}{}^2}{r^2}d_\Omega(\omega,\tilde\omega).\end{aligned}$$

Finally, we notice that

$$\left|\hat\mu(Z_0^{\rho,L}=n)\hat\mu(\overset{+}{\phi}_r)-\hat\mu(Z_0^{\rho,L}=n)^2\right|\leqslant\hat\mu(Z_0^{\rho,L}=n)\int_\Omega\mu_\omega(\overset{+}{\phi}{}_r^\omega-\mathbb{1}_{\mathcal{U}_\omega})d\mathbb{P}(\omega)\overset{\text{(H6.2)}}{\lesssim}L\frac{r^\eta}{\rho^\beta}\hat\mu(Z_0^{\rho,L}=n).$$

Combining the previous four steps, we arrive at

$$\left|\int_\Omega\mu_\omega(Z_j^{\omega,\rho,L}=n)\mu_\omega(Z_i^{\omega,\rho,L}=n)d\mathbb{P}(\omega)-\hat\mu(Z_0^{\rho,L}=n)^2\right|$$

$$\lesssim L\frac{r^\eta}{\rho^\beta}\hat\mu(Z_0^{\rho,L}=n)+((j-i)L)^{-\mathfrak{p}}\frac{\alpha^L a_{L-1}{}^2}{r^2},$$

which implies

$$
\begin{aligned}
(II) \lesssim & \sum_{i=0}^{N'_{\rho,L}-1} \sum_{j=(i+\Delta)\wedge(N'_{\rho,L}-1)+1}^{N'_{\rho,L}-1} \left(\hat{\mu}(Z_0^{\rho,L}=n)^2 + L\frac{r^\eta}{\rho^\beta}\hat{\mu}(Z_0^{\rho,L}=n) + ((j-i)L)^{-\mathfrak{p}} \frac{\alpha^L {a_{L-1}}^2}{r^2} \right) \\
\lesssim & N'_{\rho,L}(N'_{\rho,L}-\Delta) \left(\hat{\mu}(Z_0^{\rho,L}=n)^2 + L\frac{r^\eta}{\rho^\beta}\hat{\mu}(Z_0^{\rho,L}=n) \right) + N'_{\rho,L} \frac{\alpha^L {a_{L-1}}^2}{r^2} (\Delta L)^{-\mathfrak{p}+1}.
\end{aligned}
$$

Then we can conclude the following about the variance:

$$
\begin{aligned}
\operatorname{var}_{\mathbb{P}}(\mathfrak{W}_\rho^{L,n}) &= \mathbb{E}_{\mathbb{P}}((\mathfrak{W}_\rho^{L,n})^2) - (\mathbb{E}_{\mathbb{P}}(\mathfrak{W}_\rho^{L,n}))^2 \\
&\lesssim \Delta\rho^{d_0} N_\rho \hat{\mu}(Z_0^{\rho,L}=n) \\
&+ N'_{\rho,L}(N'_{\rho,L}-\Delta) \left(\hat{\mu}(Z_0^{\rho,L}=n)^2 + L\frac{r^\eta}{\rho^\beta}\hat{\mu}(Z_0^{\rho,L}=n) \right) + N'_{\rho,L}\frac{\alpha^L {a_{L-1}}^2}{r^2}(\Delta L)^{-\mathfrak{p}+1} \\
&- {N'_{\rho,L}}^2 \hat{\mu}(Z_0^{\rho,L}=n)^2. \\
&\lesssim \Delta\rho^{d_0} N_\rho \hat{\mu}(Z_0^{\rho,L}=n) + {N'_{\rho,L}}^2 L \frac{r^\eta}{\rho^\beta}\hat{\mu}(Z_0^{\rho,L}=n) + N'_{\rho,L}\frac{\alpha^L {a_{L-1}}^2}{r^2}(\Delta L)^{-\mathfrak{p}+1} \\
&\overset{(\star)}{\lesssim} {N_\rho}^\alpha L\rho^{d_0} + N_\rho \frac{r^\eta}{\rho^\beta} + \alpha^L {a_{L-1}}^2 L^{-\mathfrak{p}} \frac{N_\rho {N_\rho}^{\alpha(-\mathfrak{p}+1)}}{r^2} \\
&\overset{(\star\star)}{\lesssim} \frac{t^\alpha}{\hat{\mu}(\Gamma_\rho)^\alpha} L\rho^{d_0} + \frac{t}{\hat{\mu}(\Gamma_\rho)}\rho^{w\eta-\beta} + \alpha^L {a_{L-1}}^2 L^{-\mathfrak{p}} \frac{t}{\hat{\mu}(\Gamma_\rho)} \frac{t^{\alpha(-\mathfrak{p}+1)}}{\hat{\mu}(\Gamma_\rho)^{\alpha(-\mathfrak{p}+1)}} \rho^{-2w} \\
&\overset{(\mathrm{H6.1})}{\lesssim} L\rho^{d_0-\alpha d_1} + \rho^{w\eta-\beta-d_1} + \alpha^L {a_{L-1}}^2 L^{-\mathfrak{p}} \rho^{\alpha d_0(\mathfrak{p}-1)-2w-d_1} \\
&\overset{(\star\star\star)}{\lesssim} \rho^{d_0-\alpha d_1} + \rho^{w\eta-\beta-d_1} + \rho^{\alpha d_0(\mathfrak{p}-1)-2w-d_1},
\end{aligned}
$$

where $(\star)$ uses $N'_{\rho,L}\hat{\mu}(Z_0^{\rho,L}=n) \leqslant N_\rho L^{-1}\hat{\mu}(Z_0^{\rho,L}\geqslant 1) \leqslant N_\rho L^{-1} L\hat{\mu}(\Gamma_\rho) \lesssim t$ and t is incorporated into the $\lesssim$ sign; $(\star\star)$ uses the choice $r := \rho^w$ for a given $w > 1$; and $(\star\star\star)$ incorporates L dependent quantities on $\lesssim$. Notice that t and L dependent constants being incorporated inside $\lesssim$ is associated to the use of a constant $C_{t,L}$ in the statement.

Finally, we need to choose $(\alpha, w) \in (0,1)\times(1,\infty)$ so that

$$
\begin{cases} d_0 > \alpha d_1 \\ w\eta > \beta + d_1 \\ \alpha d_0(\mathfrak{p}-1) > 2w + d_1 \end{cases} \quad \text{i.e.} \quad \begin{cases} \alpha < \frac{d_0}{d_1} \wedge 1 = \frac{d_0}{d_1} \\ w > \frac{\beta+d_1}{\eta} \vee 1 \\ w < \frac{\alpha d_0(\mathfrak{p}-1)-d_1}{2} \end{cases},
$$

which admits a solution if, and only if,

$$
\frac{\beta+d_1}{\eta} \vee 1 = \frac{\frac{d_0}{d_1} d_0(\mathfrak{p}-1) - d_1}{2} \Leftrightarrow d_0(\mathfrak{p}-1) > \frac{2\left(\frac{\beta+d_1}{\eta} \vee 1\right) + d_1}{d_0/d_1}.
$$

This is guaranteed by the parametric constraint (H10.1), so there exists some solution (α_*, w_*) to the system. Actually, the space of solutions forms a triangle and one can select (α_*, w_*) as its incenter, a function of d_0, d_1, η, β and $\mathfrak{p}$, whereas the strictly positive margin this choice opens in the inequalities of the original system is denoted by $q(d_0, d_1, \eta, \beta, \mathfrak{p})$. With such a choice, we obtain that

$$
\operatorname{var}_{\mathbb{P}}(\mathfrak{W}_\rho^{L,n}) \leqslant C_{t,L}\cdot\rho^{q(d_0,d_1,\eta,\beta,\mathfrak{p})} \leqslant C_{t,L}\cdot\rho^q \ , \forall q \in (0, q(d_0,d_1,\eta,\beta,\mathfrak{p})).
$$

∎

Lemma 4.4.3. *Let $(\theta,\mathbb{P},T_\omega,\mu_\omega,\Gamma)$ be a system satisfying the hypotheses (H3.1), (H4.1), (H6.1), (H6.2), (H7.1), (H7.2) and (H9') with the parametric constraint (H10.1).*

Then: $\forall t>0, \forall n\geqslant 1, \forall(\rho_m)_{m\geqslant 1}\searrow 0$ with $\sum_{m\geqslant 1}{\rho_m}^q<\infty$ (for some $0<q<q(d_0,d_1,\eta,\beta,\mathfrak{p})$), denoting $N_\rho := \lfloor \frac{t}{\hat{\mu}(\Gamma_\rho)}\rfloor$ and $N'_{\rho,L}=\frac{N_\rho}{L}$[19]*, one has:*

1)

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}}\sum_{j=0}^{N'_{\rho_m,L}-1}\mu_\omega(Z_j^{\omega,\rho_m,L}=n)=t\alpha_1\lambda_n,\ \mathbb{P}\text{-a.s.}$$

2)

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}}\sum_{j=0}^{N'_{\rho_m,L}-1}\mu_\omega(Z_j^{\omega,\rho_m,L}\geqslant 1)=t\alpha_1,\ \mathbb{P}\text{-a.s.}$$

3)[20]

$$\lim_{m\to\infty}\sum_{j=0}^{N_{\rho_m}-1}\mu_{\theta^j\omega}(\Gamma_{\rho_m}(\theta^j\omega))=t,\ \mathbb{P}\text{-a.s.},$$

Proof. Let t,n and $(\rho_m)_{m\geqslant 1}$ be as in the statement. Consider $L\geqslant 1$ and m large enough so that $\rho_m\leqslant\rho_{\mathrm{var}}(L)$, $N_{\rho_m}\geqslant 3$ and $N'_{\rho_m,L}\geqslant 3$. Denote also $\mathfrak{W}_\rho^{L,n}(\omega)=\sum_{j=0}^{N'_{\rho,L}-1}\mu_\omega(Z_j^{\omega,\rho,L}=n)$.

Using Chebycheff's inequality combined with lemma 4.4.2, we get that

$$\mathbb{P}\Big(\Big|\mathfrak{W}_\rho^{L,n}-\mathbb{E}_{\mathbb{P}}(\mathfrak{W}_\rho^{L,n})\Big|>a\Big)\leqslant\frac{\mathrm{var}_{\mathbb{P}}(\mathfrak{W}_\rho^{L,n})}{a^2}\leqslant\frac{C_{t,L}}{a^2}\rho^q,$$

and therefore, since $\sum_{m\geqslant 1}{\rho_m}^q<\infty$, Borel-Cantelli lemma let us conclude that

$$\lim_{m\to\infty}\Big|\mathfrak{W}_{\rho_m}^{L,n}-\mathbb{E}_{\mathbb{P}}(\mathfrak{W}_{\rho_m}^{L,n})\Big|=0,\ \mathbb{P}\text{-a.s.}$$

On the other hand,

$$\mathbb{E}_{\mathbb{P}}(\mathfrak{W}_{\rho_m}^{L,n})=\frac{1}{L}\frac{t}{\hat{\mu}(\Gamma_{\rho_m})}\hat{\mu}(Z_0^{\rho_m,L}=n)=t\frac{\hat{\mu}(Z_{\Gamma_{\rho_m}}^L\geqslant 1)}{L\hat{\mu}(\Gamma_{\rho_m})}\frac{\hat{\mu}(Z_{\Gamma_{\rho_m}}^L=n)}{\hat{\mu}(Z_{\Gamma_{\rho_m}}^L\geqslant 1)},$$

so, by lemma 4.4.1, the definition of λ_n, we have that

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}}\ \mathbb{E}_{\mathbb{P}}(\mathfrak{W}_{\rho_m}^{L,n})=t\alpha_1\lambda_n$$

and therefore, combining the previous two centered limits, conclusion (1) follows:

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}}\ \mathfrak{W}_{\rho_m}^{L,n}=t\alpha_1\lambda_n,\ \mathbb{P}\text{-a.s.}$$

[19] See footnote 11.
[20] See footnote 12.

For (2), it suffices to repeat the argument noticing that the new expectation will be driven by $t\frac{\hat{\mu}(Z^L_{\Gamma_{\rho_m}}\geqslant 1)}{L\hat{\mu}(\Gamma_{\rho_m})}$, whose double limit is $t\alpha_1$.

For (3), it suffices to fix $L=1$ and $n=1$ in the above argument, and after the Borel-Cantelli step, notice that

$$\mathbb{E}_{\mathbb{P}}(\mathfrak{W}^{1,1}_{\rho_m}) = t\frac{\hat{\mu}(\Gamma_{\rho_m})}{\hat{\mu}(\Gamma_{\rho_m})} \overset{m\to\infty}{\to} t.$$

■

4.5 Proof of theorem 4.1.4

4.5.1 Applying the abstract approximation theorem

Let $t>0$, $n\geqslant 1$ ($n=0$ is the leftover case) and $\omega\in\Omega$ be any. Actually, at finitely many instances of the argument, we will restrict ω to be taken in a set of full measure. To be seen in due time.

Fix, once and for all, $(\rho_m)_{m\geqslant 1}\searrow 0$ fast enough so that $\sum_{m\geqslant 1}(\rho_m)^q<\infty$, for some $0<q<q(d_0,d_1,\eta,\beta,\mathfrak{p})$. For example, $\rho_m=m^{-2/q}$ is adapted to q (but not $q/2$) while $\rho_m=e^{-m}$ is adapted to any positive q.

Fix $L\geqslant n$. We will not choose it as a function of other variables, i.e., it will consist of a new free variable.

Define $N_m:=\lfloor\frac{t}{\hat{\mu}(\Gamma_{\rho_m})}\rfloor$. Let $v\in(0,d_0)$ and set $\Delta_m:={\rho_m}^{-v}$. we will consider m large enough (depending on L) so that $N_m\geqslant 3$, $\Delta_m\geqslant 2$, $\rho_m\leqslant\rho_{\mathrm{var}}(L)$, $L\leqslant\lfloor\frac{N_m}{3}\rfloor$ and $\Delta_m<N'_{m,L}$. Lastly, define $N'_{m,L}:=\frac{N_m}{L}\in\mathbb{N}_{\geqslant 3}$[21].

We want to study

$$\mu_\omega(Z^{\omega,N_m}_{\Gamma_{\rho_m}}=n)=\mu_\omega(\textstyle\sum_{i=0}^{N_m-1}I^{\omega,m}_i=n)=\mu_\omega\left(\displaystyle\sum_{j=0}^{N'_{m,L}-1}\sum_{i=jL}^{(j+1)L-1}I^{\omega,m}_i\right),$$

where $I^{\omega,m}_i=\mathbb{1}_{\Gamma_{\rho_m}(\theta^i\omega)}\circ T^i_\omega$.

To harmonize with the notation of theorem 4.3.1, we write

$$I^{\omega,m}_i=:X^{\omega,m}_i:(M,\mathscr{B}_M,\mu_\omega)\to\{0,1\}\ (i\in[0,N_m-1]\cap\mathbb{N}_{\geqslant 0})\text{ and}$$

$$\textstyle\sum_{i=jL}^{(j+1)L-1}I^{\omega,m}_i=\sum_{i=jL}^{(j+1)L-1}X^{\omega,m}_i=:Z^{\omega,m,L}_j:(M,\mathscr{B}_M,\mu_\omega)\to\mathbb{N}_{\geqslant 0}\ (j\in[0,N'_{m,L}-1]\cap\mathbb{N}_{\geqslant 0}).$$

Then one can plug in the variables here to those of the theorem 4.3.1, namely

$$N:=N_m,(\mathcal{X},\mathscr{X},\mathbb{Q}):=(M,\mathscr{B}_M,\mu^\omega),X_i:=X^{\omega,m}_i,L:=L,\Delta:=\Delta_m,N'_L:=N'_{m,L},Z^L_j:=Z^{\omega,m,L}_j,$$

to obtain that

$$\left|\mu_\omega\left(Z^{\omega,N_m}_{\Gamma_{\rho_m}}=n\right)-\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}^{\omega,m,L}_j=n\right)\right|$$
$$\leqslant 4\left(\tilde{\mathcal{R}}^1_{\omega,m}(N_m,L)+\mathcal{R}^1_{\omega,m}(N_m,L,\Delta_m)+\mathcal{R}^2_{\omega,m}(N_m,L,\Delta_m)+\mathcal{R}^3_{\omega,m}(N_m,L)\right),$$

[21] See footnote 11.

where objects being invoked are presented in theorem 4.3.1.

For the next sections, sections 4.5.2 to 4.5.7, it is enough to consider ω restricted to a $\mathbb{P}$-full measure set.

4.5.2 Estimating the error $\mathcal{R}^1$

Recall that

$$\mathcal{R}^1_{\omega,m}(N_m,L,\Delta_m) =$$
$$\sum_{j=0}^{N'_{m,L}-1} \max_{q\in[1,n]} \sum_{u=1}^{q} \left| \mu_\omega \left(Z_j^{\omega,m,L}=u, \sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q-u \right) - \mu_\omega \left(Z_j^{\omega,m,L}=u \right) \mu_\omega \left(\sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q-u \right) \right|.$$

Recycling the construction and notation used in the proof of lemma 4.4.2 to control the term (II): for a given $j \in [0, N'_{m,L}-1]$, writing $\omega' = \theta^{jL}\omega$ and considering $r \in (0, \rho_m/2)$, $v \in [0, L-1]$, we once again have the objects: $U_{v,\omega'}, \overset{-}{U}_{v,r,\omega'}, \overset{+}{U}_{v,r,\omega'}, \mathcal{U}_{\omega'}, \overset{-}{\mathcal{U}}_{r,\omega'}, \overset{+}{\mathcal{U}}_{r,\omega'}, \overset{-}{\phi}{}_r^{\omega'}$ and $\overset{+}{\phi}{}_r^{\omega'}$. Then:

$$\left| \mu_\omega \left(Z_j^{\omega,m,L}=u, \sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q-u \right) - \mu_\omega \left(Z_j^{\omega,m,L}=u \right) \mu_\omega \left(\sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q-u \right) \right|$$

$$= \left| \mu_\omega \left(\sum_{i=jL}^{(j+1)L-1} I_i^{\omega,m} = u, \sum_{i=(j+\Delta_m)L}^{N_m-1} I_i^{\omega,m} = q-u \right) \right.$$
$$\left. - \mu_\omega \left(\sum_{i=jL}^{(j+1)L-1} I_i^{\omega,m} = u \right) \mu_\omega \left(\sum_{i=(j+\Delta_m)L}^{N_m-1} I_i^{\omega,m} = q-u \right) \right|$$

$$= \left| \mu_{\omega'} \left(\sum_{i=0}^{L-1} I_i^{\omega',m} = u, \sum_{i=\Delta_m L}^{(N_m-1)-jL} I_i^{\omega',m} = q-u \right) \right.$$
$$\left. - \mu_{\omega'} \left(\sum_{i=0}^{L-1} I_i^{\omega',m} = u \right) \mu_{\theta^{\Delta_m L}\omega'} \left(\sum_{i=0}^{(N_m-1)-(j+\Delta_m)L} I_i^{\theta^{\Delta_m L}\omega',m} = q-u \right) \right|.$$

$$= \left| \mu_{\omega'} \left(\mathbb{1}_{\mathcal{U}_{\omega'}} \mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}} \circ T_{\omega'}^{\Delta_m L} \right) - \mu_{\omega'} \left(\mathbb{1}_{\mathcal{U}_{\omega'}} \right) \mu_{\theta^{\Delta_m L}\omega'} \left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}} \right) \right|$$

where we used that $V_j^{\omega,m,L,\Delta_m} := \sum_{i=0}^{(N_m-1)-(j+\Delta_m)L} I_i^{\theta^{\Delta_m L}\omega',m}$, and thus $\sum_{i=\Delta_m L}^{(N_m-1)-jL} I_i^{\omega',m} = V_j^{\omega,m,L,\Delta_m} \circ T_{\omega'}^{\Delta_m L}$,

$$\leqslant \left| \mu_{\omega'} \left(\overset{\pm}{\phi}{}_r^{\omega'} \mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}} \circ T_{\omega'}^{\Delta_m L} \right) - \mu_{\omega'} \left(\mathbb{1}_{\mathcal{U}_{\omega'}} \right) \mu_{\theta^{\Delta_m L}\omega'} \left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}} \right) \right|,$$

where $\overset{\pm}{\phi}{}^{\omega'}_r$ means that either $\overset{+}{\phi}{}^{\omega'}_r$ or $\overset{-}{\phi}{}^{\omega'}_r$ will make the inequality true,

$$\begin{aligned}
\leqslant\ & \left|\mu_{\omega'}\left(\overset{\pm}{\phi}{}^{\omega'}_r \mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}}\circ T^{\Delta_m L}_{\omega'}\right)-\mu_{\omega'}\left(\overset{\pm}{\phi}{}^{\omega'}_r\right)\mu_{\theta^{\Delta_m L}\omega'}\left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}}\right)\right| \\
+\ & \left|\left[\mu_{\omega'}\left(\overset{\pm}{\phi}{}^{\omega'}_r\right)-\mu_{\omega'}\left(\mathbb{1}_{\mathcal{U}_{\omega'}}\right)\right]\mu_{\theta^{\Delta_m L}\omega'}\left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q-u\}}\right)\right| \\
=:\ & (A)+(B).
\end{aligned}$$

Now notice that

$$(A)\lesssim(\Delta_m L)^{-\mathfrak{p}}\|\overset{\pm}{\phi}{}^{\omega'}_r\|_{\mathrm{Lip}_{d_M}}1\lesssim(\Delta_m L)^{-\mathfrak{p}}{a_{L-1}}^2/r^2,$$

where the first estimate used (H7.1) while the later used (H3.1), (H4.1) and (H3.2), as in the quenched argument in the proof of lemma 4.4.2[22].

Moreover,

$$(B)\leqslant\mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m}=q-u\right)\mu_{\omega'}(\overset{+}{\mathcal{U}}_{r,\omega'}\setminus\overset{-}{\mathcal{U}}_{r,\omega'})\overset{\text{(H6.2)}}{\lesssim}\mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m}=q-u\right)L\frac{r^\eta}{{\rho_m}^\beta}.$$

Therefore

$$\mathcal{R}^1_{\omega,m}(N_m,L,\Delta_m)\lesssim\sum_{j=0}^{N'_{m,L}-1}\max_{q\in[1,n]}\sum_{u=1}^{q}\left[(\Delta_m L)^{-\mathfrak{p}}\frac{{a_{L-1}}^2}{r^2}+\mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m}=q-u\right)L\frac{r^\eta}{{\rho_m}^\beta}\right]$$

$$\leqslant\sum_{j=0}^{N'_{m,L}-1}\sum_{u=1}^{L}(\Delta_m L)^{-\mathfrak{p}}\frac{{a_{L-1}}^2}{r^2}+L\frac{r^\eta}{{\rho_m}^\beta}\sum_{j=0}^{N'_{m,L}-1}\max_{q\in[1,N_m]}\mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m}\in[0,q]\right)$$

$$\lesssim N_m(\Delta_m L)^{-\mathfrak{p}}\frac{{a_{L-1}}^2}{r^2}+N_m\frac{r^\eta}{{\rho_m}^\beta}\leqslant N_m(\Delta_m L)^{-\mathfrak{p}}\frac{{a_{L-1}}^2}{r^2}+LN_m\frac{r^\eta}{{\rho_m}^\beta},$$

where $V_j^{\omega,m,L,\Delta_m}$ takes values between 0 and $N_m-(j+\Delta_m)L\leqslant N_m$.

4.5.3 Estimating the error $\tilde{\mathcal{R}}^1$

This section is going to follow the lines of the previous one, with minor modifications.

Recall that

$$\tilde{\mathcal{R}}^1_{\omega,m}(N_m,L,\Delta_m)=$$
$$\sum_{j=0}^{N'_{m,L}-1}\max_{q\in[0,n]}\left|\mu_\omega\left(Z_j^{\omega,m,L}\geqslant1,\sum_{k=j+\Delta_m}^{N'_{m,L}-1}Z_k^{\omega,m,L}=q\right)-\mu_\omega\left(Z_j^{\omega,m,L}\geqslant1\right)\mu_\omega\left(\sum_{k=j+\Delta_m}^{N'_{m,L}-1}Z_k^{\omega,m,L}=q\right)\right|.$$

For a given $j\in[0,N'_{m,L}-1]$, writing $\omega'=\theta^{jL}$ and considering $r\in(\rho_m/2)$, $v\in[0,L]$, recalling the objects introduced in the proof of lemma 4.4.2, we reuse $U_{v,\omega'},\overset{-}{U}_{v,r,\omega'}$ and $\overset{+}{U}_{v,r,\omega'}$, whereas $\mathcal{U}_{\omega'},\overset{-}{\mathcal{U}}_{r,\omega'}$ and $\overset{+}{\mathcal{U}}_{r,\omega'}$ are modified by including a union $\bigcup_{n=1}^{L}$ before the original definitions therein (in

[22]In the present passage, the a.s. validity of $\|\overset{+}{\phi}{}^{\omega'}_r\|_{\mathrm{Lip}_{d_M}}\leqslant{a_{L-1}}^2/r^2$ would be enough, but, after recalling the argument of lemma 4.4.2 we see that it actually holds for every ω. The validity for every ω was important back then, but not here.

particular, $\{Z_0^{\omega',\rho_m,L} \geqslant 1\} = \mathcal{U}_{\omega'}$), while $\overset{-}{\phi}{}_r^{\omega'}$ and $\overset{+}{\phi}{}_r^{\omega'}$ are kept the same (but considering the previous modification).

Following the same steps and notation from the previous section, we get that

$$\begin{aligned}
&\left|\mu_\omega\left(Z_j^{\omega,m,L}\geqslant 1, \sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q\right) - \mu_\omega\left(Z_j^{\omega,m,L}\geqslant 1\right)\mu_\omega\left(\sum_{k=j+\Delta_m}^{N'_{m,L}-1} Z_k^{\omega,m,L}=q\right)\right| \\
&\leqslant \left|\mu_{\omega'}\left(\overset{\pm}{\phi}{}_r^{\omega'} \mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q\}} \circ T_{\omega'}^{\Delta_m L}\right) - \mu_{\omega'}\left(\overset{\pm}{\phi}{}_r^{\omega'}\right)\mu_{\theta^{\Delta_m L}\omega'}\left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q\}}\right)\right| \\
&\quad + \left|\left[\mu_{\omega'}\left(\overset{\pm}{\phi}{}_r^{\omega'}\right) - \mu_{\omega'}\left(\mathbb{1}_{\mathcal{U}_{\omega'}}\right)\right]\mu_{\theta^{\Delta_m L}\omega'}\left(\mathbb{1}_{\{V_j^{\omega,m,L,\Delta_m}=q\}}\right)\right| \\
&=: (A) + (B).
\end{aligned}$$

As before,

$$(A) \lesssim (\Delta_m L)^{-\mathfrak{p}} \left\|\overset{\pm}{\phi}{}_r^{\omega'}\right\|_{\mathrm{Lip}_{d_M}} 1 \lesssim (\Delta_m L)^{-\mathfrak{p}} {a_{L-1}}^2/r^2.$$

For the first inequality we use (H7.1). For the second, we adapt the previous reasoning as follows. Intuitively, the Lipschitz constant of, say, the modified function $\overset{-}{\phi}{}_r^{\omega'}$ is bounded by the inverse of $d\big(\overset{-}{\mathcal{U}}_{r,\omega'}, \mathcal{U}^c_{\omega'}\big)$. For a point to $x \in \mathcal{U}^c_{\omega'}$, with no hits, to be minimally displaced to $\overset{-}{\mathcal{U}}_{r,\omega'}$, among x itself being displaced or the consequently-displaced points in its orbit, a) at least one r-stringent hit has to be created while b) the other instances should turned into r-stringent non-hits (if they are not already). The situation where this would occur with minimal displacement is one where (b) starts already fulfilled and only (a) has to be accomplished by displacing x in such that its $L-1$ iterate changes from a non-hit to a r-stringent hit. This can be made with a minimum displacement of r/a_{L-1}, where again we use (H3.1), (H3.2) and (H4.1).

Moreover,

$$(B) \leqslant \mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m} = q\right)\mu_{\omega'}(\overset{+}{\mathcal{U}}_{r,\omega'} \setminus \overset{-}{\mathcal{U}}_{r,\omega'}) \overset{\text{(H6.2)}}{\lesssim} \mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m} = q\right) L^2 \frac{r^\eta}{{\rho_m}^\beta}.$$

Therefore

$$\begin{aligned}
\tilde{\mathcal{R}}^1_{\omega,m}(N_m, L, \Delta_m) &\lesssim \sum_{j=0}^{N'_{m,L}-1} \max_{q\in[0,n]} \left[(\Delta_m L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2} + \mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m} = q\right) L^2 \frac{r^\eta}{{\rho_m}^\beta}\right] \\
&\leqslant \sum_{j=0}^{N'_{m,L}-1} (\Delta_m L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2} + L^2 \frac{r^\eta}{{\rho_m}^\beta} \sum_{j=0}^{N'_{m,L}-1} \max_{q\in[0,N_m]} \mu_{\theta^{\Delta_m L}\omega'}\left(V_j^{\omega,m,L,\Delta_m} \in [0,q]\right) \\
&\lesssim N'_{m,L}(\Delta_m L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2} + LN_m \frac{r^\eta}{{\rho_m}^\beta} \leqslant N_m (\Delta_m L)^{-\mathfrak{p}} \frac{{a_{L-1}}^2}{r^2} + LN_m \frac{r^\eta}{{\rho_m}^\beta},
\end{aligned}$$

where $V_j^{\omega,m,L,\Delta_m}$ takes values between 0 and $N_m - (j+1)L \leqslant N_m$.

4.5.4 Estimating the error $\mathcal{R}^2$

To start

$$\mathcal{R}^2_{\omega,m}(N_m,L,\Delta_m) = \sum_{j=0}^{N'_{m,L}-1} \mu_\omega(Z_j^{\omega,m,L} \geqslant 1, \sum_{k=j+1}^{j+\Delta_m-1} Z_k^{\omega,m,L} \geqslant 1)$$
$$\leqslant \sum_{j=0}^{N'_{m,L}-1} \sum_{k=j+1}^{j+\Delta_m-1} \mu_\omega(Z_j^{\omega,m,L} \geqslant 1, Z_k^{\omega,m,L} \geqslant 1)$$

where we reverse the double sum and single out the $k = j+1$ terms

$$= \sum_{k=1}^{N'_{m,L}+\Delta_m-2} \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)} \mu_\omega(Z_j^{\omega,m,L} \geqslant 1, Z_k^{\omega,m,L} \geqslant 1) + \sum_{k=1}^{N'_{m,L}} \mu_\omega(Z_{k-1}^{\omega,m,L} \geqslant 1, Z_k^{\omega,m,L} \geqslant 1)$$
$$=: (I) + (II)$$

To estimate (I) we notice that:

$$(I) \leqslant \sum_{k=1}^{N'_{m,L}+\Delta_m-2} \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)} \sum_{i=jL}^{(j+1)L-1} \sum_{l=kL}^{(k+1)L-1} \mu_\omega\left((T^i_\omega)^{-1}\Gamma_{\rho_m}(\theta^i\omega) \cap (T^l_\omega)^{-1}\Gamma_{\rho_m}(\theta^l\omega)\right) \ (l > i)$$
$$\leqslant \sum_{k=1}^{N'_{m,L}+\Delta_m-2} \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)} \sum_{i=jL}^{(j+1)L-1} \sum_{l=kL}^{(k+1)L-1} \mu_{\omega'}\left(\Gamma_{\rho_m}(\omega') \cap (T^{l-i}_{\omega'})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega') \cap \overset{++}{\mathcal{C}}{}^{\omega'}_{l-i}\right)$$
$$+ \sum_{k=1}^{N'_{m,L}+\Delta_m-2} \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)} \sum_{i=jL}^{(j+1)L-1} \sum_{l=kL}^{(k+1)L-1} \mu_{\omega'}\left(\Gamma_{\rho_m}(\omega') \cap (T^{l-i}_{\omega'})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega') \cap \left[\overset{+-}{\mathcal{C}}{}^{\omega'}_{l-i} \cup \overset{-}{\mathcal{C}}{}^{\omega'}_{l-i}\right]\right)$$
$$=: (I_{\text{good}}) + (I_{\text{bad}})$$

where $\omega' := \theta^i\omega$.

To estimate (I_{good}) we begin evaluating the following:

$$\mu_{\omega'}\left(\overset{++}{\mathcal{C}}{}^{\omega'}_{l-i} \cap \Gamma_{\rho_m}(\omega') \cap (T^{l-i}_{\omega'})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega')\right)$$

$$\leqslant \sum_{\substack{\xi=\varphi(\text{dom}(\varphi))\in \overset{++}{\mathcal{C}}{}^{\omega'}_{l-i}: \\ \xi\cap\Gamma_{\rho_m}(\omega')\neq\varnothing}} \frac{\mu_{\omega'}|_\xi\left(\xi \cap (T^{l-i}_{\omega'})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega')\right)}{\mu_{\omega'}|_\xi(\xi)} \mu_{\omega'}(\xi),$$

where, from (H3.3), $\varphi \in \text{IB}(T^{l-i}_{\omega'})$ implies $\mu_{\omega'}|_{\varphi(\text{dom}(\varphi))} = J_\varphi{}^{-1}\left[\varphi_*(\mu_{\theta^{l-i}\omega'}|_{\text{dom}(\varphi)})\right]$, and so

$$\leqslant \sum_{\xi \text{ as above}} \frac{\left[J_\varphi{}^{-1}\left[\varphi_*(\mu_{\theta^{l-i}\omega'}|_{\text{dom}(\varphi)})\right]\right]\left(\varphi(\text{dom}(\varphi)) \cap (T^{l-i}_{\omega'})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega')\right)}{\left[J_\varphi{}^{-1}\left[\varphi_*(\mu_{\theta^{l-i}\omega'}|_{\text{dom}(\varphi)})\right]\right](\varphi(\text{dom}(\varphi)))} \mu_{\omega'}(\xi)$$

$$\leqslant \sum_{\xi \text{ as above}} \frac{\sup_{x\in\xi} J_\varphi{}^{-1}(x)}{\inf_{x\in\xi} J_\varphi{}^{-1}(x)} \frac{\mu_{\theta^{l-i}\omega'}|_{\mathrm{dom}(\varphi)}\left(\mathrm{dom}(\varphi)\cap\varphi^{-1}(T_{\omega'}^{l-i})^{-1}\Gamma_{\rho_m}(\theta^{l-i}\omega')\right)}{\mu_{\theta^{l-i}\omega'}|_{\mathrm{dom}(\varphi)}(\mathrm{dom}(\varphi))}\mu_{\omega'}(\xi)$$

$$\underset{\text{(H2.6)}}{\overset{\text{(H3.3)}}{\lesssim}} (l-i)^{\mathfrak{d}}\iota^{-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\sum_{\xi \text{ as above}}\mu_{\omega'}(\xi) \overset{\text{(H2.2)}}{\leqslant} (l-i)^{\mathfrak{d}}\iota^{-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\mathcal{N}\mu_{\omega'}\left(\bigcup_{\xi \text{ as above}}\xi\right)$$

$$\overset{\text{(H3.4)}}{\leqslant} (l-i)^{\mathfrak{d}}\iota^{-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\mathcal{N}\mu_{\omega'}\left(B_{D(l-i)^{-\kappa}}(\Gamma_{\rho_m}(\omega'))\right)$$

$$\overset{\text{(H6.1)}}{\leqslant} (l-i)^{\mathfrak{d}}\iota^{-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\mathcal{N}C_0(\rho_m+D(l-i)^{-\kappa})^{d_0} \lesssim \mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))(l-i)^{\mathfrak{d}}\left[\rho_m{}^{d_0}+(l-i)^{-\kappa d_0}\right].$$

Then

$$(I_{\text{good}}) \leqslant \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\ \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)}\ \sum_{i=jL}^{(j+1)L-1}\ \sum_{l=kL}^{(k+1)L-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))(l-i)^{\mathfrak{d}}\left[\rho_m{}^{d_0}+(l-i)^{-\kappa d_0}\right]$$
$$= \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\ \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)}\ \sum_{l=kL}^{(k+1)L-1}\left(\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\sum_{i=jL}^{(j+1)L-1}(l-i)^{\mathfrak{d}}\left[\rho_m{}^{d_0}+(l-i)^{-\kappa d_0}\right]\right)$$

where, for each l fixed, as i runs, we have $l-i\in[kL-jL-L+1, kL-jL+L-1]$, so

$$\leqslant \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\ \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)}\ \sum_{l=kL}^{(k+1)L-1}\left(\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\sum_{s=kL-jL-L+1}^{kL-jL+L-1}s^{\mathfrak{d}}\left[\rho_m{}^{d_0}+s^{-\kappa d_0}\right]\right)$$

$$= \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\ \sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)}\left[\left(\sum_{l=kL}^{(k+1)L-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\right)\left(\sum_{s=kL-jL-L+1}^{kL-jL+L-1}s^{\mathfrak{d}}\left[\rho_m{}^{d_0}+s^{-\kappa d_0}\right]\right)\right]$$
$$\leqslant \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\left(\sum_{l=kL}^{(k+1)L-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\right)\left(\sum_{j=(k-\Delta_m+1)\vee 0}^{(k-2)\wedge(N'_{m,L}-1)}\ \sum_{s=kL-jL-L+1}^{kL-jL+L-1}s^{\mathfrak{d}}\left[\rho_m{}^{d_0}+s^{-\kappa d_0}\right]\right)$$

where $s\in[L+1, 3\Delta_m L]$[23], so

$$\lesssim \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\left(\sum_{l=kL}^{(k+1)L-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\right)\left(\sum_{u=L+1}^{3\Delta_m L}u^{\mathfrak{d}}\left[u^{-\kappa d_0}+\rho_m{}^{d_0}\right]\right)$$

$$\lesssim \sum_{k=1}^{N'_{m,L}+\Delta_m-2}\left(\sum_{l=kL}^{(k+1)L-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\right)\left(L^{\mathfrak{d}-\kappa d_0+1}+(\Delta_m L)^{\mathfrak{d}+1}\rho_m{}^{d_0}\right),$$

where for the first term in the square bracket we have used that, for $\alpha>1$, $\sum_{n=m}^{\infty}n^{-\alpha}\lesssim m^{-\alpha+1}$ together with $\mathfrak{d}-\kappa d_0<-1$, which is guaranteed by (H10.3), whereas for the second we have used that $u^{\mathfrak{d}}$ is increasing and the summation interval is bounded above by $3\Delta_m L$.

we will leave (I_{bad}) to the end.

[23] The interval where s ranges basically has length $2L$ and it is translated by L when j moves one unit, therefore the original and the new interval overlap by half, so eventual repetitions are more than compensated by a factor of two.

For (II), we consider $L' < L$ and proceed as follows

$$
\begin{aligned}
&\sum_{k=1}^{N'_{m,L}} \mu_\omega(Z_{k-1}^{\omega,m,L} \geqslant 1, Z_k^{\omega,m,L} \geqslant 1) \\
&= \sum_{k=1}^{N'_{m,L}} \mu_\omega \left(\sum_{i=(k-1)L}^{kL-1} I_i^{\omega,m} \geqslant 1, \sum_{l=kL}^{kL+L'-1} I_l^{\omega,m} + \sum_{l=kL+L'}^{(k+1)L-1} I_l^{\omega,m} \geqslant 1 \right) \\
&\leqslant \sum_{k=1}^{N'_{m,L}} \mu_\omega \left(\sum_{i=(k-1)L}^{kL-1} I_i^{\omega,m} \geqslant 1, \sum_{l=kL}^{kL+L'-1} I_l^{\omega,m} \geqslant 1 \right) + \mu_\omega \left(\sum_{i=(k-1)L}^{kL-1} I_i^{\omega,m} \geqslant 1, \sum_{l=kL+L'}^{(k+1)L-1} I_l^{\omega,m} \geqslant 1 \right)
\end{aligned}
$$

and, denoting $\omega' = \theta^i \omega$,

$$
\begin{aligned}
&\leqslant \sum_{k=1}^{N'_{m,L}} \sum_{l=kL}^{kL+L'-1} \mu_{\theta^l \omega} \left(\Gamma_{\rho_m}(\theta^l \omega) \right) \\
&+ \sum_{k=1}^{N'_{m,L}+\Delta_m-2} \sum_{i=(k-1)L}^{kL-1} \sum_{l=kL+L'}^{(k+1)L-1} \mu_{\omega'} \left(\Gamma_{\rho_m}(\omega') \cap (T_{\omega'}^{l-i})^{-1} \Gamma_{\rho_m}(\theta^{l-i}\omega') \cap \overset{++}{\mathcal{C}}{}^{\omega'}_{l-i} \right) \\
&+ \sum_{k=1}^{N'_{m,L}} \sum_{i=(k-1)L}^{kL-1} \sum_{l=kL+L'}^{(k+1)L-1} \mu_{\omega'} \left(\Gamma_{\rho_m}(\omega') \cap (T_{\omega'}^{l-i})^{-1} \Gamma_{\rho_m}(\theta^{l-i}\omega') \cap \left[\overset{+-}{\mathcal{C}}{}^{\omega'}_{l-i} \cup \overset{-}{\mathcal{C}}{}^{\omega'}_{l-i} \right] \right) \\
&=: (II_{\text{rest}}) + (II_{\text{good}}) + (II_{\text{bad}}).
\end{aligned}
$$

The term (II_{rest}) will not be improved, whereas the term (II_{good}) is approached just like (I_{good}), as follows:

$$
(II_{\text{good}}) \lesssim \sum_{k=1}^{N'_{m,L}} \sum_{l=kL+L'}^{(k+1)L-1} \left(\mu_{\theta^l \omega}(\Gamma_{\rho_m}(\theta^l \omega)) \sum_{i=(k-1)L}^{kL-1} (l-i)^{\mathfrak{d}} \left[{\rho_m}^{d_0} + (l-i)^{-\kappa d_0} \right] \right)
$$

where, for each l fixed, as i runs, we have $l - i \in [L'+1, 2L-1]$, so

$$
\begin{aligned}
&\leqslant \sum_{k=1}^{N'_{m,L}} \left(\sum_{l=kL+L'}^{(k+1)L-1} \mu_{\theta^l \omega}(\Gamma_{\rho_m}(\theta^l \omega)) \right) \left(\sum_{u=L'+1}^{2L-1} u^{\mathfrak{d}} \left[{\rho_m}^{d_0} + u^{-\kappa d_0} \right] \right) \\
&\lesssim \sum_{k=1}^{N'_{m,L}} \left(\sum_{l=kL}^{(k+1)L-1} \mu_{\theta^l \omega}(\Gamma_{\rho_m}(\theta^l \omega)) \right) \left({L'}^{\mathfrak{d}-\kappa d_0+1} + L^{\mathfrak{d}+1} {\rho_m}^{d_0} \right).
\end{aligned}
$$

Now we combine (I_{bad}) and (II_{bad}) and their domain of summation[24] to see that

$$
\begin{aligned}
(I_{\text{bad}})+(II_{\text{bad}}) &\lesssim \sum_{i=0}^{N_m-1}\sum_{l=i+L+L'}^{i+\Delta_m L}\mu_{\theta^i\omega}\left(\left[\overset{+}{\overline{\mathcal{C}}}{}^{\theta^i\omega}_{l-i}\cup\overline{\mathcal{C}}^{\theta^i\omega}_{l-i}\right]\cap\Gamma_{\rho_m}(\theta^i\omega)\right)\\
&=\sum_{i=0}^{N_m-1}\sum_{s=L+L'}^{\Delta_m L}\mu_{\theta^i\omega}\left(\left[\overset{+}{\overline{\mathcal{C}}}{}^{\theta^i\omega}_{s}\cup\overline{\mathcal{C}}^{\theta^i\omega}_{s}\right]\cap\Gamma_{\rho_m}(\theta^i\omega)\right)\\
&\leqslant\sum_{s=L'}^{\Delta_m L}\sum_{i=0}^{N_m-1}\mu_{\theta^i\omega}\left(\left[\overset{+}{\overline{\mathcal{C}}}{}^{\theta^i\omega}_{s}\cup\overline{\mathcal{C}}^{\theta^i\omega}_{s}\right]\cap\Gamma_{\rho_m}(\theta^i\omega)\right).
\end{aligned}
$$

Combining the bounds of (I_{good}) and (II_{good}), we conclude that

$$
\begin{aligned}
\mathcal{R}^2_{\omega,m}(N_m,L,\Delta_m) \lesssim &\sum_{l=0}^{5N_m-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\left(L'^{\eth-\kappa d_0+1}+(\Delta_m L)^{\eth+1}{\rho_m}^{d_0}\right)\\
&+\sum_{k=1}^{N'_{m,L}}\sum_{l=kL}^{kL+L'-1}\mu_{\theta^l\omega}\left(\Gamma_{\rho_m}(\theta^l\omega)\right)+\sum_{s=L'}^{\Delta_m L}\sum_{i=0}^{N_m-1}\mu_{\theta^i\omega}\left(\left[\overset{+}{\overline{\mathcal{C}}}{}^{\theta^i\omega}_{s}\cup\overline{\mathcal{C}}^{\theta^i\omega}_{s}\right]\cap\Gamma_{\rho_m}(\theta^i\omega)\right).
\end{aligned}
$$

4.5.5 Estimating the error $\mathcal{R}^3$

Here we use (H6.1) to see that

$$
\begin{aligned}
\mathcal{R}^3_{\omega,m}(N_m,L,\Delta_m) &= \sum_{i=0}^{N_m-1}\sum_{\ell=0\vee(i-\Delta_m L)}^{i}\mu_{\theta^i\omega}(\Gamma_{\rho_m}(\theta^i\omega))\mu_{\theta^\ell\omega}(\Gamma_{\rho_m}(\theta^\ell\omega)))\\
&\lesssim \Delta_m L\,{\rho_m}^{d_0}\sum_{i=0}^{N_m-1}\mu_{\theta^i\omega}(\Gamma_{\rho_m}(\theta^i\omega)),
\end{aligned}
$$

which, noticing that $\Delta_m L\leqslant(\Delta_m L)^{\eth+1}$, reveals to be bounded above by $\mathcal{R}^2_{\omega,m}(N_m,L,\Delta_m)$.

4.5.6 Controlling the total error

Put $r={\rho_m}^w$ $(w>1)$ and $L'=L^\alpha$ $(0<\alpha<1)$. Then

$$
\left|\mu_\omega\left(Z^{\omega,N_m}_{\Gamma_{\rho_m}}=n\right)-\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}^{\omega,m,L}_j=n\right)\right|
$$

$$
\begin{aligned}
&\lesssim a_{L-1}{}^2{\rho_m}^{\mathfrak{p}v-2w-d_1}+L{\rho_m}^{w\eta-\beta-d_1}\\
&+\sum_{l=0}^{5N_m-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega))\left(L'^{\eth-\kappa d_0+1}+L^{\eth+1}{\rho_m}^{d_0-v(\eth+1)}\right)\\
&+\sum_{k=1}^{N'_{m,L}}\sum_{l=kL}^{kL+L'-1}\mu_{\theta^l\omega}\left(\Gamma_{\rho_m}(\theta^l\omega)\right)+\sum_{s=L'}^{\Delta_m L}\sum_{i=0}^{N_m-1}\mu_{\theta^i\omega}\left(\left[\overset{+}{\overline{\mathcal{C}}}{}^{\theta^i\omega}_{s}\cup\overline{\mathcal{C}}^{\theta^i\omega}_{s}\right]\cap\Gamma_{\rho_m}(\theta^i\omega)\right),
\end{aligned}
$$

where in the first line of the RHS accounts for both $\mathcal{R}^1$ and $\tilde{\mathcal{R}}^1$.

---

[24] Notice that the initial L'-strip of the first component of the original summation has already been singled out inside (II_{rest}).

Until this point, parameters v (accompanying Δ_m, see section 4.5), w (accompanying r), and α (accompanying L'), which are local to the proof, were not fine-tuned.

In the last equation, we need the exponents accompanying ρ to be strictly positive. In particular, we need

$$w > \frac{\beta+d_1}{\eta} \vee 1,\ pv - 2w - d_1 > 0 \text{ and } d_0 - v(\mathfrak{d}+1) > 0.$$

The space of solutions $(w,v) \in (1,\infty)\times(0,d_0)$ to those inequalities is non-empty if $\mathfrak{p} > \frac{2\left(\frac{\beta+d_1}{\eta}\vee 1\right)+d_1}{d_0/(\mathfrak{d}+1)}$, which is guaranteed by (H10.2).

We will take double limits of the type $\overline{\lim}_{L\to\infty}\overline{\lim}_{m\to\infty}$ on the RHS. Initially, taking the $\overline{\lim}_{m\to\infty}$, we use that, by lemma 4.4.3,

$$\lim_{m\to\infty}\sum_{l=0}^{5N_m-1}\mu_{\theta^l\omega}(\Gamma_{\rho_m}(\theta^l\omega)) = 5t,\ \mathbb{P}\text{-a.s.}$$

and, by similar arguments[25],

$$\lim_{m\to\infty}\sum_{k=1}^{N'_{m,L}}\sum_{l=kL}^{kL+L'-1}\mu_{\theta^l\omega}\left(\Gamma_{\rho_m}(\theta^l\omega)\right) = tL^{\alpha-1},\ \mathbb{P}\text{-a.s..}$$

Finally, using hypothesis (H4.2) and noticing that $\mathfrak{d}-\kappa d_0+1<0$ (by (H10.3)) and $\alpha-1<0$ (by design), we conclude that the RHS under the double limit $\overline{\lim}_{L\to\infty}\overline{\lim}_{m\to\infty}$ goes to 0. The same thing occurs if we adopt the double limits $\underline{\lim}_{L\to\infty}\overline{\lim}_{m\to\infty}$, $\overline{\lim}_{L\to\infty}\underline{\lim}_{m\to\infty}$ and $\underline{\lim}_{L\to\infty}\underline{\lim}_{m\to\infty}$. Therefore

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}}\left|\mu_\omega\left(Z^{\omega,N_m}_{\Gamma_{\rho_m}}=n\right)-\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}^{\omega,m,L}_j=n\right)\right| = 0,\ \mathbb{P}\text{-a.s..}$$

4.5.7 Convergence of the leading term to the compound Poisson distribution

It remains to show that $\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}^{\omega,m,L}_j=n\right)$ to $\mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})$.

Due to the independence and distributional properties of the $\tilde{Z}^{\omega,m,L}_j$'s (see theorem 4.3.1):

$$\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}^{\omega,m,L}_j=n\right)$$

$$=\sum_{l=1}^{n}\sum_{0\leqslant j_1<\ldots<j_l\leqslant N'_{m,L}-1}\left(\prod_{\substack{j\in[0,N'_{m,L}-1]\\ \setminus\{j_i:i=1,\ldots,l\}}}\mu_\omega(Z^{\omega,m,L}_j=0)\cdot\sum_{\substack{(n_1,\ldots,n_l)\in\mathbb{N}^l_{\geqslant 1}\\ n_1+\ldots+n_l=n}}\prod_{i=1}^{l}\mu_\omega(Z^{\omega,m,L}_{j_i}=n_i)\right)$$

[25] Adapting the argument of lemma 4.4.3 item (III) to the new term, we see that the new $\mathbb{P}$-expectation is $tL^{\alpha-1}$, but the variance lemma used therein, lemma 4.4.2, would need to be adapted as well, what we omitted.

$$\overset{(\star)}{=} (1+o(1)) \prod_{j=0}^{N'_{m,L}-1} \mu_\omega(Z_j^{\omega,m,L}=0) \sum_{l=1}^{n} \frac{1}{l!} \sum_{\substack{j_i\in[0,N'_{m,L}-1]\\ i=1,\dots,l}} \sum_{\substack{(n_1,\dots,n_l)\in\mathbb{N}^l_{\geqslant 1}\\ n_1+\dots+n_l=n}} \prod_{i=1}^{l} \mu_\omega(Z_{j_i}^{\omega,m,L}=n_i)$$

$$\overset{(\star\star)}{=} (1+o(1)) \prod_{j=0}^{N'_{m,L}-1} \mu_\omega(Z_j^{\omega,m,L}=0) \sum_{l=1}^{n} \frac{1}{l!} \sum_{\substack{(n_1,\dots,n_l)\in\mathbb{N}^l_{\geqslant 1}\\ n_1+\dots+n_l=n}} \prod_{i=1}^{l} \left(\sum_{j=0}^{N'_{m,L}-1} \mu_\omega(Z_j^{\omega,m,L}=n_i) \right),$$

where i) $o(1)$ refers to a function $g(\omega,m,L)$ so that $\overline{\lim}_{L\to\infty}\overline{\lim}_{m\to\infty}|g(\omega,m,L)|=0$, $\mathbb{P}$-a.s.; ii) equality $(\star)$ included $1/l!$ to account for j_i's not being anymore increasing and used that the error terms that come from different j_i's being equal are small, as one can see in the case when two j_i agree; and iii) equality $(\star\star)$ uses that a product of sums distributes as a sum of products.

We then notice that, by lemma 4.4.3,

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}} \sum_{j=0}^{N'_{\rho_m,L}-1} \mu^\omega(Z_j^{\omega,\rho_m,L}=n_i) = t\alpha_1\lambda_{n_i}, \mathbb{P}\text{-a.s.}$$

and

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}} \prod_{j=0}^{N'_{m,L}-1} \mu_\omega(Z_j^{\omega,m,L}=0) = \lim_{L\to\infty}\overline{\lim_{m\to\infty}} \exp\left(\sum_{j=0}^{N'_{m,L}-1} \ln\Big(1-\mu_\omega(Z_j^{\omega,m,L}\geqslant 1)\Big)\right)$$

$$= \lim_{L\to\infty}\overline{\lim_{m\to\infty}} \exp\left(\sum_{j=0}^{N'_{m,L}-1} -\mu_\omega(Z_j^{\omega,m,L}\geqslant 1)+o(1)\right) = e^{-t\alpha_1}, \mathbb{P}\text{-a.s..}$$

Therefore

$$\lim_{L\to\infty}\overline{\lim_{m\to\infty}} \left| \mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1} \tilde{Z}_j^{\omega,m,L}=n\right) - e^{-t\alpha_1}\sum_{l=1}^{n}\frac{(t\alpha_1)^l}{l!} \sum_{\substack{(n_1,\dots,n_l)\in\mathbb{N}^l_{\geqslant 1}\\ n_1+\dots+n_l=n}} \prod_{i=1}^{l}\lambda_{n_i} \right| = 0, \mathbb{P}\text{-a.s.}$$

$$\Leftrightarrow \lim_{L\to\infty}\overline{\lim_{m\to\infty}} \left| \mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1} \tilde{Z}_j^{\omega,m,L}=n\right) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\}) \right| = 0, \mathbb{P}\text{-a.s.,}$$

where the equivalence is because the former term is precisely the density of such a compound Poisson distribution (see equation (2.1)).

As a consequence,

$$\left|\mu_\omega(Z_{\Gamma_{\rho_m}}^{\omega,N_m}=n) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})\right| \overset{\forall L\geqslant n}{\leqslant} \left|\mu_\omega(Z_{\Gamma_{\rho_m}}^{\omega,N_m}=n) - \mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}_j^{\omega,m,L}=n\right)\right|$$

$$+ \left|\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1}\tilde{Z}_j^{\omega,m,L}=n\right) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})\right|$$

$$
\begin{aligned}
\Rightarrow \overline{\lim_{m\to\infty}} \left|\mu_\omega(Z^{\omega,N_m}_{\Gamma_{\rho_m}} = n) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})\right| & \\
\leqslant & \lim_{L\to\infty} \overline{\lim_{m\to\infty}} \left|\mu_\omega(Z^{\omega,N_m}_{\Gamma_{\rho_m}} = n) - \mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1} \tilde{Z}^{\omega,m,L}_j = n\right)\right| \\
+ & \lim_{L\to\infty} \overline{\lim_{m\to\infty}} \left|\mu_\omega\left(\sum_{j=0}^{N'_{m,L}-1} \tilde{Z}^{\omega,m,L}_j = n\right) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})\right| \\
= & \ 0,\ \mathbb{P}\text{-a.s.}
\end{aligned}
$$

We then conclude that $\lim_{m\to\infty} \left|\mu_\omega(Z^{\omega,N_m}_{\Gamma_{\rho_m}} = n) - \mathrm{CPD}_{t\alpha_1,(\lambda_\ell)_\ell}(\{n\})\right| = 0$, $\mathbb{P}$-a.s., as desired.

Chapter 5

Application: random piecewise expanding one-dimensional systems

We consider a class of random piecewise expanding one-dimensional systems $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$ prescribed by the following conditions. Elements in this class immediately comprise a system as in the general setup of section 4.1.1 and will check that they also comprise a system as in the working setup of section 4.1.2 (i.e., satisfying hypotheses (H1-H10)).

C1. Consider finitely many maps of the unit interval (or circle), $T_v : M \to M$, for $v \in \{0, \dots, u-1\}$. For ease of exposition, say that $u = 2$. They carry a family of open intervals $A_v = (\zeta_{v,i})_{i=1}^{I_v}$ $(I_v < \infty)$ so that $M \backslash \bigcup_{i=1}^{I_v} \zeta_{v,i}$ is finite and $T_v|_{\zeta_{v,i}}$ is surjective and C^2-differentiable with

$$
\begin{aligned}
1 < d_{\min} \leqslant\ & \inf\{|T_v{}'(x)| : x \in \zeta_{v,i}, v = 1, \dots, I_v, i = 0, \dots, u-1\}, \\
& \sup\{|T_v{}''(x)| : x \in \zeta_{v,i}, v = 1, \dots, I_v, i = 0, \dots, u-1\} \leqslant c_{\max} < \infty.
\end{aligned}
$$

For $n \geqslant 1$, let $A_n^\omega = \bigvee_{j=0}^{n-1} (T_\omega^j)^{-1} A_{\pi_j(\omega)}$. For $n = 0$, we adopt the convention $A_0^\omega = \{(0,1)\}$ $(\forall \omega \in \Omega)$. Write $\mathcal{A}_n^\omega = \bigcup_{\zeta \in A_n^\omega} \zeta$ (co-finite) and, for $x \in \mathcal{A}_n^\omega$, denote by $A_n^\omega(x)$ the element of A_n^ω containing x. In particular, $x \in \mathcal{A}_n^\omega$ implies that x is a point of differentiability for T_ω^n.

C2. Let $\Omega = \{0,1\}^{\mathbb{Z}}$. Set $T_\omega := T_{\pi_0(\omega)}$, where $\pi_j(\omega) = \omega_j$ $(j \in \mathbb{Z})$. Consider $\theta : \Omega \to \Omega$ to be the bilateral shift map.

C3. Consider $\mathbb{P} \in \mathcal{P}_\theta(\Omega)$ an equilibrium state associated to a Lipschitz potential. Usual instances are Bernoulli and Markov measures.

C4. Consider $\Gamma(\omega) = \{x(\omega)\}$ $(\omega \in \Omega)$, where $x : \Omega \to M$ is a random variable taking values either x_0 or x_1 (possibly coincident) in the form $x(\omega) = x_{\pi_0(\omega)}$, with $\{x_0, x_1\} \subset \bigcap_{\omega \in \Omega} \bigcap_{l=1}^{\infty} \mathcal{A}_l^\omega$ [1] (which needs to be a non-empty set).

[1]The intersection $\bigcap_{l=1}^{\infty} \mathcal{A}_l^\omega$ is a co-countable set $(\forall \omega \in \Omega)$.

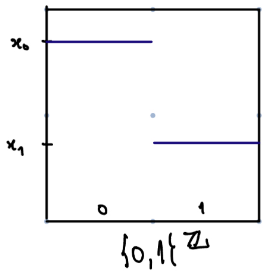


Figure 5.1 A representation of $x(\omega) = x_{\pi_0(\omega)}$

Moreover, for each $\omega \in \Omega$, with the minimal period

$$m(\omega) := \min\{m \geqslant 1 : T_\omega^{m(\omega)} x(\omega) = x(\theta^{m(\omega)}\omega)\} \in \mathbb{N}_{\geqslant 1} \cup \{\infty\},$$

one defines the number of finite-periods occurring along the ω fiber ($K(\omega) \in \mathbb{N}_{\geqslant 0} \cup \{\infty\}$) and the associated sequence of such periods ($(m_j(\omega))_{j=0}^{K(\omega)-1} \subset \mathbb{N}_{\geqslant 1}$), using the conventions $m_{-1}(\omega) :\equiv 0$ and $\max \varnothing := 0$, letting

$$K(\omega) := \max \left\{ k \geqslant 1 : \begin{array}{ll} m_0(\omega) := m(\omega) & \in \mathbb{N}_{\geqslant 1} \\ m_1(\omega) := m(\theta^{m_0(\omega)}\omega) & \in \mathbb{N}_{\geqslant 1} \\ m_2(\omega) := m(\theta^{m_1(\omega)+m_0(\omega)}\omega) & \in \mathbb{N}_{\geqslant 1} \\ \dots & \\ m_{k-1}(\omega) := m(\theta^{m_{k-2}(\omega)+\dots+m_0(\omega)}\omega) & \in \mathbb{N}_{\geqslant 1} \end{array} \right\} \in \mathbb{N}_{\geqslant 0} \cup \{\infty\}.$$

In particular, writing $M_j(\omega) := \sum_{k=0}^{j-1} m_k(\omega)$ for $1 \leqslant j \leqslant K(\omega)$ (with $M_0(\omega) :\equiv 0$), one has:

$$x(\omega) \xmapsto{T_\omega^{m_0(\omega)}} x(\theta^{M_1(\omega)}\omega) \xmapsto{T_{\theta^{M_1(\omega)}\omega}^{m_1(\omega)}} x(\theta^{M_2(\omega)}\omega) \xmapsto{T_{\theta^{M_2(\omega)}\omega}^{m_2(\omega)}} x(\theta^{M_3(\omega)}\omega) \; \dots \; .$$

We conclude (C4) assuming that the target satisfies the dynamical condition that

$$\sup\{m_j(\omega) : \omega \in \Omega, j = 0, \dots, K(\omega) - 1\} =: M_\Gamma < \infty,$$

where the convention $\max \varnothing := 0$ is adopted.

C5. Consider that there exists $r > 0$, $K, Q > 1$ and $\beta \in (0,1]$ so that $\mu_\omega = h_\omega \mathrm{Leb}$ forms a quasi-invariant family satisfying: i) $(\omega, x) \mapsto h_\omega(x)$ is measurable, ii) $K^{-1} \leqslant h_\omega|_{B_r(x(\omega))} \leqslant K$ a.s., and iii) $h_\omega|_{B_r(x(\omega))} \in \mathrm{Hol}_\beta(M)$ with $H_\beta(h_\omega|_{B_r(x(\omega))})$ a.s.. See remark 8.

The following result says that theorem 4.1.4 applies to systems in the class (C1-C5) and, in particular, they have quenched limit entry distributions in the compound Poisson class with the needed statistical quantities presented explicitly.

Theorem 5.0.1. *Let $(\theta, \mathbb{P}, T_\omega, \mu_\omega, \Gamma)$ be a system satisfying conditions (C1-C5). Then the hypotheses of theorem 4.1.4 are satisfied with*

$$\alpha_\ell = \int_\Omega \begin{cases} \dfrac{h_\omega(x(\omega))}{\int_\Omega h_\omega(x(\omega))d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} - \left(JT_\omega^{M_\ell(\omega)}(x(\omega)) \right)^{-1} \right], & \text{if } \ell \leqslant K(\omega) \\ \dfrac{h_\omega(x(\omega))}{\int_\Omega h_\omega(x(\omega))d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} \right] & \text{, if } \ell = K(\omega)+1 \\ 0 & \text{, if } \ell \geqslant K(\omega)+2 \end{cases} d\mathbb{P}(\omega).$$

The quantities α_ℓ comply with (H9) and theorem 4.1.3, allowing for $\lambda_\ell = (\alpha_\ell - \alpha_{\ell+1})/\alpha_1$ to hold.

In particular: $\forall t > 0, \forall n \geqslant 0, \forall (\rho_m)_{m \geqslant 1} \searrow 0$ with $\sum_{m \geqslant 1} {\rho_m}^q < \infty$ (for some $0 < q < 1$) one has

$$\mu_\omega(Z_{\Gamma_{\rho_m}}^{\omega, \lfloor t/\hat{\mu}(\Gamma_{\rho_m}) \rfloor} = n) \xrightarrow[m \to \infty]{\mathbb{P}\text{-a.s.}} \mathrm{CPD}_{t\alpha_1, (\lambda_\ell)_\ell}(\{n\}).$$

We will prove the theorem after a few remarks on relevant subclasses within (C1-C5) and examples.

Remark 5. When the maps T_v are piecewise expanding linear maps, they preserve Lebesgue and conditions (C1)-(C3),(C5) are immediately satisfied.

To illustrate condition (C4), or, better said, condition $M_\Gamma < \infty$, we can look at deterministic targets $x(\omega) \equiv x$. Two noticeable cases occur:

i) *Pure periodic points x*: when there is some $m_* = m_*(x) \geqslant 1$ so that x is (minimally) fixed by any concatenations of m_* maps in $(T_v)_{v=0}^{u-1}$. In this case, $m(\omega) \equiv m_*, K(\omega) \equiv \infty, m_j(\omega) \equiv m_*$ and $M_\Gamma = m_*$.

It is convenient to represent these types of examples with diagrams (that can neglect topological information), where the deterministic target x is highlighted with a green ball, each arrow indicates how each map T_v acts, blue cycles indicate cycles that avoid the target, purple paths indicate paths between the blue cycles and the target and yellow cycles indicate cycles that include the target (but are not obtained composing blue cycles with purple paths).

Figure 5.2 (a) Pure one-periodic.

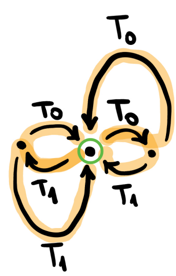


Figure 5.3 (b) Pure two-periodic.

Considering remark 5, we can easily present explicit examples of systems complying with cases (a) and (b) above. In both examples, $x(\omega) \equiv 1/2$ and all maps preserve Lebesgue. Constructions of this kind are possible for any $m_* \geqslant 1$ and $u \geqslant 1$.

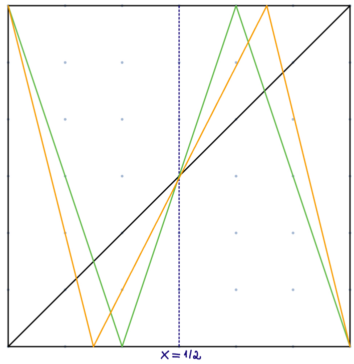


Figure 5.4 (a) A pure one-periodic system.[2]

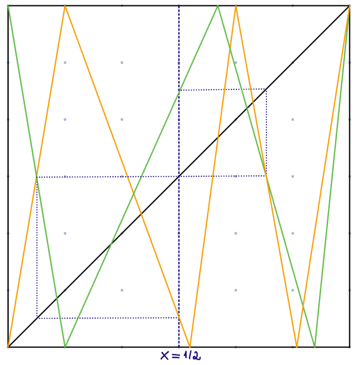
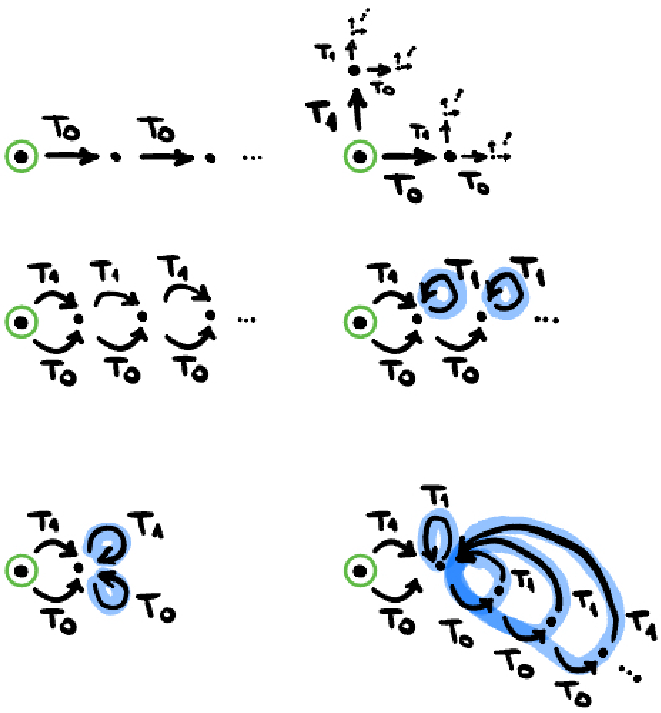


Figure 5.5 (b) A pure two-periodic system.

ii) *Pure aperiodic points* x: when x is not fixed by any finite concatenation of maps in $(T_\nu)_{\nu=0}^{u-1}$. In this case, $m(\omega) \equiv \infty, K(\omega) \equiv 0$ and $M_\Gamma = 0$.

Here are some compatible diagrams in this case:



Explicit examples realizing these structures (or exhibiting these sorts of behaviors) can be tricky to construct[3], especially when the diagram is infinite and one has to control the behavior of

[2]More precisely, take two piecewise expanding linear maps T_ν $(\nu = 0,1)$ of the unit interval, with three surjective branches: T_0 fixes the midpoint $1/2$ in its center branch, with slope 2, whereas its left branch maps 0 to 1 and its right branch maps 1 to 0; T_1 is just the same, but having slope 3 in the fixed midpoint. Let $x(\omega) \equiv 1/2$. This choice satisfies the inclusion in (C4), because regardless of the random seed ω and order of iteration l, the midpoint is always a differentiable fixed point right in the middle of the center branch of T_ω^l. Condition $M_\Gamma < \infty$ is verified as in (i) with $m_* = 1$.

[3]We are not claiming that every (possible) diagram compatible with (ii) can be realized by examples in the class (C1-C5).

infinitely many iterates of the system[4]. Notice, however, that, once the maps are fixed, the set of pure aperiodic x's is generic, because it is given by

$$M \setminus \bigcup_{p \geqslant 1} \bigcup_{(v_0,\ldots,v_{p-1}) \in \{0,\ldots,u-1\}^p} \mathrm{Fix}(T_{v_{p-1}} \circ \ldots \circ T_{v_0}),$$

which is co-countable.

For a finite diagram such as the last one in the first column, we can consider the following explicit example:

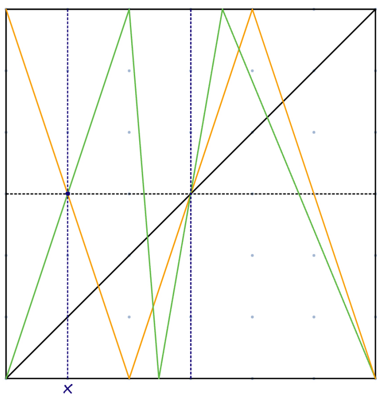
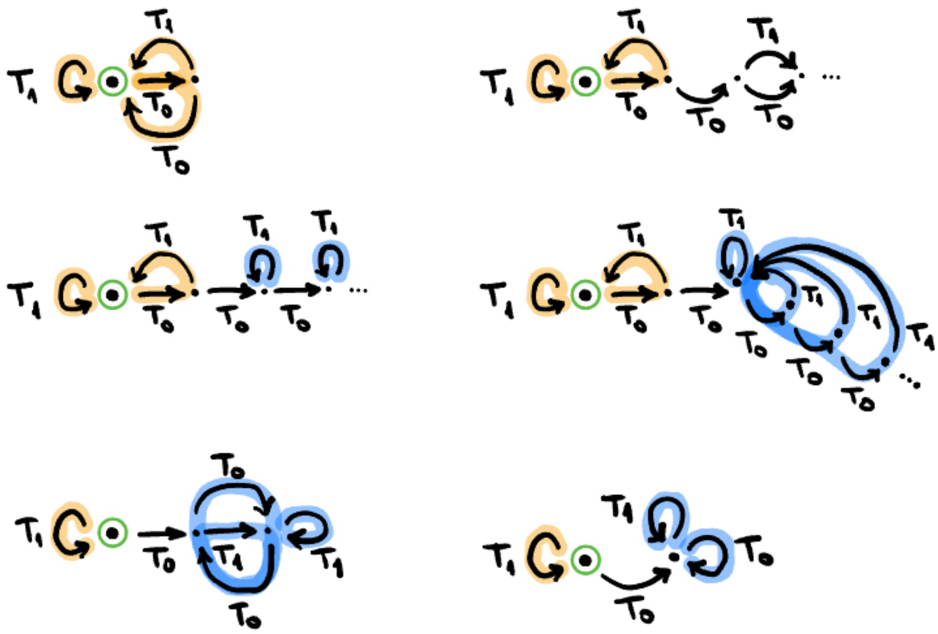


Figure 5.6 A pure aperiodic system.

iii) *Hybrid*. This is the general case. They can combine the behavior in (i) and (ii) while still verifying $M_\Gamma < \infty$. Here are some possible diagrams in this case:

[4]In this direction, beta maps with irrational translation and rational (random) targets were studied in [8]. They do not fit exactly in the class (C1-C5) because they do not have subjective branches. However, they can be dealt with here by considering their action on S^1 rather than on $[0,1]$. See remark 7.



For a finite diagram such as the last one, we can consider the following explicit example:

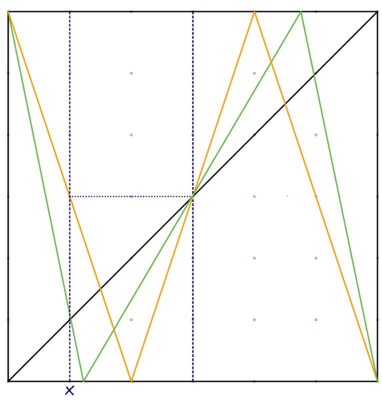
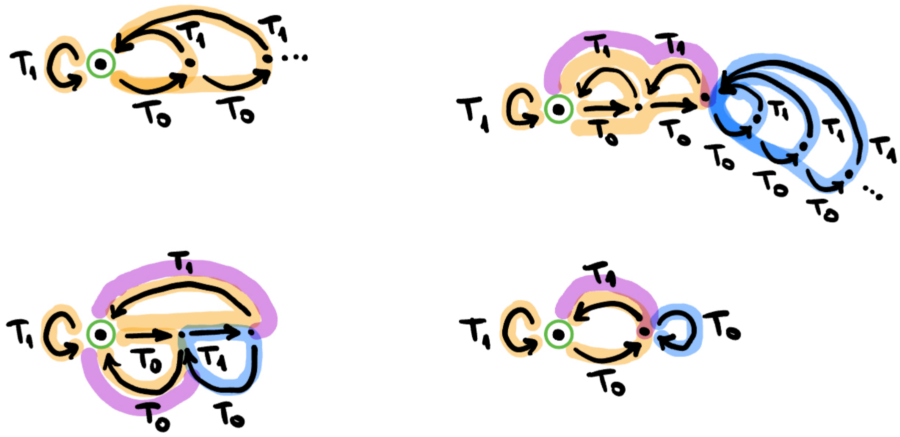


Figure 5.7 A hybrid system.

iv) *Non-examples*. Here are some diagrams which do not satisfy $M_\Gamma < \infty$.



Notice that whenever a purple path occurs arbitrarily large periods can be formed. But this can occur without purple paths as well, as in the first diagram. Moreover, this can occur both with infinite diagrams (the first two) and with finite diagrams (the last two).

Remark 6. It is not being claimed that systems as in (iv) are not covered by the theory in chapter 4. It is just being said that systems as in (iv) are not treated with the techniques used in this section (to calculate underlying α_ℓ's).

Proof of theorem 5.0.1. It is enough to check that conditions (C1)-(C5) imply the hypotheses (H1-H7, H9-H10) of section 4.

(H1). Immediate, since $M = [0,1]$ and $\Omega = \{0,1\}^{\mathbb{Z}}$.

(H2). (H2.1) holds because the finitely many branches are injective and $M\backslash\bigcup_{i=1}^{I_v}\zeta_{v,i}$ is finite. (H2.2) holds with $R := 1/2$, $\mathcal{N} := 1$, and $(y_k^{\omega,n})_{k\in K_{\omega,n}}$ assigned the singleton $1/2$ $(\forall\omega,n)$, in which case, for all ω and n, the family of balls is the single one $B_{1/2}(1/2) = (0,1)$, that leaves only two points uncovered, and trivially has at most 1 overlaps. (H2.3) holds because every power T_ω^n also satisfies the property in the second sentence of (C1). (H2.5) holds because $M\backslash\bigcup_{i=1}^{I_v}\zeta_{v,i}$ and the measures have no singleton (see (H6.1) below). Finally, (H2.6) holds with $\iota := K^{-1}$ (that of (C5)), because one always has $\mu_{\theta^n\omega}(B_R(y_k^{\omega,n})) \geqslant K^{-1}$.

This choices imply that $C_n^\omega = A_n^\omega$ and $\mathcal{C}_n^\omega = \mathcal{A}_n^\omega$.

(H3). (H3.1) is immediate and we let $\overset{+}{C}{}_n^{\omega} := C_n^{\omega}$ and $\overset{++}{C}{}_n^{\omega} := \mathcal{C}_n^{\omega}$. Then (H3.2) holds immediately, whereas (H3.3) holds with $\mathfrak{d} := 0$ and $C := K^4 e^{c_{\max}\frac{d_{\min}}{d_{\min}-1}}$, because of the usual distortion control based on the second derivative bounded as in (C1) (see, e.g., [70], section 3.3): for any $\omega \in \Omega, n \geqslant 1, \varphi \in \overset{+}{\mathrm{IB}}(T_\omega^n)$ and $x, y \in \xi = \varphi(\mathrm{dom}(\varphi))$:

$$K^{-2}DT_\omega^n = K^{-2}\frac{d\varphi_*\,\mathrm{Leb}}{d\,\mathrm{Leb}} \leqslant \frac{d\varphi_*\left[\mu_{\theta^n\omega}|_{\mathrm{dom}(\varphi)}\right]}{d\mu_\omega|_{\varphi(\mathrm{dom}(\varphi))}} \leqslant K^2\frac{d\varphi_*\,\mathrm{Leb}}{d\,\mathrm{Leb}} = K^2 DT_\omega^n \Rightarrow \frac{J_\varphi(x)}{J_\varphi(y)} \leqslant K^4\frac{DT_\omega^n(x)}{DT_\omega^n(y)}$$

but

$$\ln\frac{DT_\omega^n(x)}{DT_\omega^n(y)} = \sum_{i=0}^{n-1}\ln\frac{DT_{\theta^i\omega}(T_\omega^i x)}{DT_{\theta^i\omega}(T_\omega^i y)} = \sum_{i=0}^{n-1}\ln\left(\frac{DT_{\theta^i\omega}(T_\omega^i x) - DT_{\theta^i\omega}(T_\omega^i y)}{DT_{\theta^i\omega}(T_\omega^i y)} + 1\right)$$

$$\overset{\ln(1+x)<x}{<} \sum_{i=0}^{n-1} \frac{DT_{\theta^i\omega}(T_\omega^i x) - DT_{\theta^i\omega}(T_\omega^i y)}{DT_{\theta^i\omega}(T_\omega^i y)} \leqslant c_{\max} \sum_{i=0}^{n-1} |T_\omega^i x - T_\omega^i y|$$

$$\leqslant c_{\max} \sum_{i=0}^{n-1} {d_{\min}}^{-(n-i)} |T_\omega^n(x) - T_\omega^n(y)| \leqslant c_{\max} \frac{1}{1 - {d_{\min}}^{-1}}. \tag{5.1}$$

On the other hand, (H3.4) follows with $D := 1$ and $\kappa > 1$ arbitrary since

$${d_{\min}}^n \leqslant DT_\omega^n(x) \leqslant {d_{\max}}^n, \forall \omega \in \Omega, n \geqslant 1, x \in \mathcal{A}_n^\omega$$

then

$${d_{\max}}^{-n} \leqslant D\varphi(z) \leqslant {d_{\min}}^{-n} \lesssim n^{-\kappa}, \forall \kappa > 1, \omega \in \Omega, n \geqslant 1, \varphi \in \mathrm{IB}(T_\omega^n), z \in \mathrm{dom}(\varphi).$$

(H4). (H4.2) is immediate because $\overset{-}{\mathcal{C}}_n^\omega, \overset{+-}{\mathcal{C}}_n^\omega = \varnothing$. On the other hand, (H4.1) follows from the finitely many maps in (C1), the dependence of T_ω simply on the first coordinate of ω as in (C2) and the inclusion in (C4).

(H5). Holds immediately because maps T_ω and the target $\Gamma(\omega)$ depend only on the first coordinate of ω, thus their Lipschitz constant is 0. Also, $\mathrm{Lip}(\theta) < \infty$ is immediate.

(H6). (H6.1) holds with $d_0, d_1 := 1$, $C_0 := K$ and $C_1 := K^{-1}$. (H6.2) holds with $\eta, \beta := 1$ and $E = 2K^2$.

(H7). Items (H7.1) and (H7.2) hold with any $\mathfrak{p} > 1$. This is because a) (C3) implies that $(\theta, \mathbb{P})$ satisfies exponential decay of the type described in (H7), with test functions in $\mathrm{Lip}(\Omega)$ and $L^\infty(\Omega)$ (see [17]) and b) (C5) considers quenched ACIPs, whose quenched decay is exponential of the type described in (H7), with test functions in $\mathrm{Lip}(M)$ and $L^\infty(M)$ (see [32] page 1130 and equation (19)[5]).

(H10). Since $\mathfrak{p}$ arbitrarily large is available, simply $\kappa d_0 > 1$ has to be checked. But once $d_0 = 1$ and κ arbitrarily large was possible, the desired parametric constraint is immediately satisfied.

(H9). We start calculating α_ℓ's. Consider $\ell \geqslant 1$ and $\omega \in \Omega$ (eventually taken in a set of full measure).

Consider

$$L \geqslant M_{\ell \wedge K(\omega)}(\omega). \tag{5.2}$$

Then take $\rho_0(\omega, L) = \rho_0(\pi_0(\omega), \ldots, \pi_L(\omega))$ small enough so that $\rho \leqslant \rho_0(\omega, L)$ implies

$$T_\omega^i B_\rho(x(\omega)) \cap B_\rho(x(\theta^i\omega)) = \varnothing, \forall i \in [1, L] \backslash \{M_k(\omega) : k \in [1, K(\omega)]\}, \tag{5.3}$$

which can be guaranteed noticing that

a) returns occur precisely in the instants $\{M_k(\omega) : k \in [1, K(\omega)]\}$ and not in between (by minimality),

[5]From here one has that BV is the good Banach space and that decay appears in terms of BV against L^∞ (constants are uniform). In particular, for any Lipschitz function, we can apply this result to get decay of type the Lip against L^∞, as we need. Related comments can be found in the second paragraph of remark 9. Moreover, noticing that the cone of continuous functions is preserved (see, e.g., [1] page 5), that the eigenfunctions $h_\omega \in$ BV upgrade to Holder ones, showing that condition (C5) is adequate. Related comments can be found in remark 8.

b) T_{ω}^{i} is continuous on $x(\omega)$ ($\forall i \geqslant 1$), a.s., because, by (C4), one has

$$x(\omega) \in \{x_0, x_1\} \subset \bigcap_{l=1}^{\infty} \mathcal{A}_l^{\omega} \subset \mathcal{A}_i^{\omega}, \text{ a.s.}$$

Because of the previous constraint, one could have started with L's of the form $L = M_{q_L \wedge K(\omega)}(\omega)$, $q_L \geqslant \ell$ (so still satisfying equation (5.2)), in the sense that other choices of L are superfluous from the viewpoint of the quantity we will study, $Z_{*\Gamma_\rho}^{\omega,L}$. Then one could restrict $\rho_0(\omega, L)$ further so that $\rho \leqslant \rho_0(\omega, L)$ implies:

$$T_{\theta^{M_{k'}(\omega)}\omega}^{\overbrace{M_k(\omega)-M_{k'}(\omega)}^{M_{k-k'}\left(\theta^{M_{k'}(\omega)}\omega\right)}} B_\rho(x(\theta^{M_{k'}(\omega)}\omega)) \subset A_{M_{q_L \wedge K(\omega)}(\omega)-M_k(\omega)}^{\theta^{M_k(\omega)}\omega}\left(x(\theta^{M_k(\omega)}\omega)\right), \; \forall k', k \in [0, q_L \wedge K(\omega)], k' \leqslant k, \tag{5.4}$$

which can be guaranteed noticing that

a) $T_{\theta^{M_{k'}(\omega)}\omega}^{M_{k-k'}\left(\theta^{M_{k'}(\omega)}\omega\right)} x(\theta^{M_{k'}(\omega)}\omega) = x(\theta^{M_{k-k'}\left(\theta^{M_{k'}(\omega)}\omega\right)}\theta^{M_{k'}(\omega)}\omega) = x(\theta^{M_k(\omega)}\omega)$, with the later in $\{x_0, x_1\} \overset{(C4)}{\subset} \bigcap_{l=1}^{\infty} \mathcal{A}_l^{\theta^{M_k(\omega)}\omega} \subset \mathcal{A}_{M_{q_L \wedge K(\omega)}(\omega)-M_k(\omega)}^{\theta^{M_k(\omega)}\omega}$,

b) $T_{\theta^{M_{k'}(\omega)}\omega}^{M_{k-k'}\left(\theta^{M_{k'}(\omega)}\omega\right)}$ is continuous at $x(\theta^{M_{k'}(\omega)}\omega)$, because, again by (C4), one has $x(\theta^{M_{k'}(\omega)}\omega) \in \mathcal{A}_{M_{k-k'}\left(\theta^{M_{k'}(\omega)}\omega\right)}^{\theta^{M_{k'}(\omega)}\omega}$.

The point with condition (5.4) is to say that, ρ is so small that, starting from any pre-intermediary time $M_{k'}(\omega)$ and going to any post-intermediary step $M_k(\omega)$, the initial ρ-sized ball grows under iteration up to time $M_k(\omega)$ but still fitting inside a partition domain (thus an injectivity domain) of the map evolving from time $M_k(\omega)$ until the end, $M_{q_L \wedge K(\omega)}$. In particular, the image balls won't break injectivity (or wrap around). Most importantly, it is implied that for any $z \in B_\rho(x(\omega))$:

$$\left(I_0^{\omega,\rho}(z), I_{M_1(\omega)}^{\omega,\rho}(z), \ldots, I_{M_{q_L \wedge K(\omega)}(\omega)}^{\omega,\rho}(z)\right)$$

is a binary sequence starting with a batch of 1's followed by a (possibly degenerate) batch of 0's (e.g. 11100, 1111 or 11000).

Then, for ω, L and ρ as above, one has:

$$\hat{\alpha}_\ell^{\omega}(L,\rho)\mu_\omega(\Gamma_\rho(\omega)) = \mu_\omega(Z_{*\Gamma_\rho}^{\omega,L} \geqslant \ell - 1, I_0^{\omega,\rho} = 1)$$

$$\overset{(5.3)}{=} \mu_\omega\left(\sum_{j \in \{M_k(\omega): k \in [1, q_L \wedge K(\omega)]\}} I_j^{\omega,\rho} \geqslant \ell - 1, I_0^{\omega,\rho} = 1\right)$$

$$\overset{(5.4)}{=} \begin{cases} \mu_\omega\left(I_0^{\omega,\rho} = 1, I_{M_1(\omega)}^{\omega,\rho} = 1, \ldots, I_{M_{\ell-1}(\omega)}^{\omega,\rho} = 1\right), & \text{if } \ell - 1 \leqslant K(\omega) \\ 0 & \text{, otherwise} \end{cases}$$

$$\overset{(5.4)}{=}\begin{cases}\mu_\omega\left(I^{\omega,\rho}_{M_{\ell-1}(\omega)}=1\right), & \text{if } \ell-1\leqslant K(\omega)\\ 0 & \text{, otherwise}\end{cases},$$

so that

$$\alpha^\omega_\ell(L,\rho)\frac{\mu_\omega(\Gamma_\rho(\omega))}{\hat{\mu}(\Gamma_\rho)}\overset{(4.5)}{=}\begin{cases}\dfrac{\mu_\omega\left((T^{M_{\ell-1}(\omega)}_\omega)^{-1}\Gamma_\rho(\theta^{M_{\ell-1}(\omega)}\omega)\right)}{\hat{\mu}(\Gamma_\rho)}-\dfrac{\mu_\omega\left((T^{M_\ell(\omega)}_\omega)^{-1}\Gamma_\rho(\theta^{M_\ell(\omega)}\omega)\right)}{\hat{\mu}(\Gamma_\rho)}, & \text{if } \ell\leqslant K(\omega),\\ \dfrac{\mu_\omega\left((T^{M_{\ell-1}(\omega)}_\omega)^{-1}\Gamma_\rho(\theta^{M_{\ell-1}(\omega)}\omega)\right)}{\hat{\mu}(\Gamma_\rho)} & \text{, if } \ell=K(\omega)+1,\\ 0 & \text{, if } \ell\geqslant K(\omega)+2.\end{cases}$$

Notice that

$$\frac{\mu_\omega\left((T^{M_{\ell-1}(\omega)}_\omega)^{-1}\Gamma_\rho(\theta^{M_{\ell-1}(\omega)}\omega)\right)}{\hat{\mu}(\Gamma_\rho)}=\frac{\mathrm{Leb}\left(h_\omega\mathbb{1}_{(T^{M_\ell(\omega)}_\omega)^{-1}\Gamma_\rho(\theta^{M_\ell(\omega)}\omega)}\right)}{\int_\Omega\mathrm{Leb}(h_\omega\mathbb{1}_{\Gamma_\rho(\omega)})d\mathbb{P}(\omega)}$$

$$=\frac{[h_\omega(x(\omega))+\mathcal{O}(\varepsilon)]\,\mathrm{Leb}\left((T^{M_\ell(\omega)}_\omega)^{-1}B_\rho(x(\theta^{M_\ell(\omega)}\omega))\right)}{\int_\Omega[h_\omega(x(\omega))+\mathcal{O}(\varepsilon)]\mathrm{Leb}(B_\rho(x(\omega)))d\mathbb{P}(\omega)}$$

$$=\frac{[h_\omega(x(\omega))+\mathcal{O}(\varepsilon)]\left[\left(JT^{M_\ell(\omega)}_\omega(x(\omega))\right)^{-1}+\mathcal{O}(\varepsilon)\right]\mathrm{Leb}\left(B_\rho(x(\theta^{M_\ell(\omega)}\omega))\right)}{\int_\Omega[h_\omega(x(\omega))+\mathcal{O}(\varepsilon)]\mathrm{Leb}(B_\rho(x(\omega)))d\mathbb{P}(\omega)}$$

$$=\frac{h_\omega(x(\omega))+\mathcal{O}(\varepsilon)}{\int_\Omega h_\omega(x(\omega))+\mathcal{O}(\varepsilon)d\mathbb{P}(\omega)}\left[\left(JT^{M_\ell(\omega)}_\omega(x(\omega))\right)^{-1}+\mathcal{O}(\varepsilon)\right] \tag{5.5}$$

where, given $\varepsilon>0$ (for ω and L chosen as above), we've considered $\rho\leqslant\rho_1(\omega,\varepsilon)<r$ (see (C5)), with $\rho_1(\omega,\varepsilon)$ small enough so that for any $\rho\leqslant\rho_1(\omega,\varepsilon)$:

$$h_\omega(z)=h_\omega(x(\omega))+\mathcal{O}(\varepsilon),\ \forall z\in B_\rho(x(\omega))$$

and

$$\left(JT^{M_\ell(\omega)}_\omega(z)\right)^{-1}=\left(JT^{M_\ell(\omega)}_\omega(x(\omega))\right)^{-1}+\mathcal{O}(\varepsilon),\ \forall z\in B_\rho(x(\omega)).$$

We can write

$$\rho_1(\omega,\varepsilon)=\left(\varepsilon/H_\beta(h_\omega|_{B_r(x(\omega))})\right)^{1/\beta}\wedge\left(\varepsilon/H_\beta\left([JT^{M_\ell(\omega)}_\omega]^{-1}|_{B_r(x(\omega))}\right)\right)^{1/\beta}\wedge 1.$$

We can use (C1) (finitely many maps and uniformly bounded second derivatives), (C4) (uniformly bounded finite-periods) and (C5) (uniform Holder constants for the densities) to pass to controls that are uniform on ω and then integrate: for any $\varepsilon>0$, $L\geqslant L_*:=\ell M_\Gamma$ and

$$\rho\leqslant\rho_*(L,\varepsilon):=\min_{\substack{(v_0,\ldots,v_L)\\ \in\ \{0,1\}^{L+1}}}\rho_1(v_0,\ldots,v_L)\wedge\operatorname*{ess\,inf}_\omega\rho_1(\omega,\varepsilon)\in(0,1],$$

one has

$$\alpha_\ell(L,\rho) = \int_\Omega \begin{cases} \dfrac{h_\omega(x(\omega)) + \mathcal{O}(\varepsilon)}{\displaystyle\int_\Omega h_\omega(x(\omega)) + \mathcal{O}(\varepsilon) d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} + \mathcal{O}(\varepsilon) - \left(JT_\omega^{M_\ell(\omega)}(x(\omega)) \right)^{-1} - \mathcal{O}(\varepsilon) \right], & \text{if } \ell \leqslant K(\omega) \\ \dfrac{h_\omega(x(\omega)) + \mathcal{O}(\varepsilon)}{\displaystyle\int_\Omega h_\omega(x(\omega)) + \mathcal{O}(\varepsilon) d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} + \mathcal{O}(\varepsilon) \right] & , \text{if } \ell = K(\omega)+1 \\ 0 & , \text{if } \ell \geqslant K(\omega)+2 \end{cases} d\mathbb{P}(\omega),$$

then taking iterated limits of the type $\lim_\varepsilon \lim_L \overline{\lim}_\rho$ one finds that

$$\alpha_\ell = \int_\Omega \begin{cases} \dfrac{h_\omega(x(\omega))}{\displaystyle\int_\Omega h_\omega(x(\omega)) d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} - \left(JT_\omega^{M_\ell(\omega)}(x(\omega)) \right)^{-1} \right], & \text{if } \ell \leqslant K(\omega) \\ \dfrac{h_\omega(x(\omega))}{\displaystyle\int_\Omega h_\omega(x(\omega)) d\mathbb{P}(\omega)} \left[\left(JT_\omega^{M_{\ell-1}(\omega)}(x(\omega)) \right)^{-1} \right] & , \text{if } \ell = K(\omega)+1 \\ 0 & , \text{if } \ell \geqslant K(\omega)+2 \end{cases} d\mathbb{P}(\omega). \tag{5.6}$$

The following diagram helps one to visualize how the integrand in equation (5.6), with the factor $\frac{h_\omega(x(\omega))}{\int_\Omega h_\omega(x(\omega)) d\mathbb{P}(\omega)}$ suppressed, changes

a) for ω's with varying amount of periodicity (read the different lines),

b) as ℓ grows (read the different columns).

$$\begin{array}{rl} & \qquad\qquad \ell=1 \qquad\qquad\qquad\qquad \ell=2 \qquad\qquad\qquad\qquad\qquad \ell=3 \\ K(\omega)=\infty: & \left(1 - 1/JT_\omega^{m_0(\omega)}(x(\omega)), \dfrac{1 - 1/JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\omega))}{JT_\omega^{m_0(\omega)}(x(\omega))}, \dfrac{1 - 1/JT_{\theta^{m_0(\omega)+m_1(\omega)}\omega}^{m_2(\omega)}(x(\omega))}{JT_\omega^{m_0(\omega)}(x(\omega)) JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\omega))}, \dots \right) \\ K(\omega)=0: & (\qquad 1 \qquad , \qquad 0 \qquad , \qquad 0 \qquad , \overset{\bar{0}}{\dots}) \\ K(\omega)=1: & \left(1 - 1/JT_\omega^{m_0(\omega)}(x(\omega)), \dfrac{1}{JT_\omega^{m_0(\omega)}(x(\omega))}, \qquad 0 \qquad , \overset{\bar{0}}{\dots} \right) \\ K(\omega)=2: & \left(1 - 1/JT_\omega^{m_0(\omega)}(x(\omega)), \dfrac{1 - 1/JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\omega))}{JT_\omega^{m_0(\omega)}(x(\omega))}, \dfrac{1}{JT_\omega^{m_0(\omega)}(x(\omega)) JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\omega))}, \overset{\bar{0}}{\dots} \right). \end{array} \tag{5.7}$$

Having found that α_ℓ's exist and have explicit representation, it remains to check that $\alpha_1 > 0$ and $\sum_{\ell=1}^\infty \ell^2 \alpha_\ell < \infty$.

It holds that $\alpha_1 > 0$ because the quantity found in the first column of diagram (5.7) is bounded below by $1 - 1/d_{\min} > 0$.

Moreover, considering the integrand of equation (5.6), we see that α_ℓ is at most $(1/d_{\min})^{\ell-1}$, therefore

$$\sum_{\ell=1}^\infty \ell^2 \hat{\alpha}_\ell \leqslant \sum_{\ell=1}^\infty \ell^2 ({}^1\!/_{d_{\min}})^{\ell-1} < \infty,$$

since $d_{\min} > 1$.

This concludes that conditions (C1)-(C4) imply the hypotheses of theorem 4.1.4 and that the associated α_ℓ's satisfy (H9) and the hypotheses of theorem 4.1.3.

Let us finally notice that in this case, where $d_0, d_1, \eta, \beta = 1$ and $\mathfrak{p} = \infty$ (i.e., can be taken arbitrarily large), $q(d_0, d_1, \eta, \beta, \mathfrak{p})$, reduces to 1. This is because the system of inequalities appearing at end of proof of lemma 4.4.2 reduces to only two ($1 > \alpha$ and $w > 2$ for $(\alpha, w) \in (0,1) \times (1, \infty)$) which admit a solution that opens a margin of (at least) 1 in both equations. ■

Remark 7. As it comes to $M = [0,1]$, the use of surjective branches in (C1) was to facilitate as much as possible the presentation of covers and cylinders in (H2) below. But these can be still presented without surjective branches. For example, one could present them for the beta maps $T_0(x) = 1/2 + 2x$ (mod 1) and $T_1(x) = 1/2 + 3x$ (mod 1). On the other hand, to have the type of decay against Lipschitz test functions we will be after in (H7), the interval maps ought to have subjective branches (otherwise the good functional space becomes bounded variation instead of Lipschitz), which is not the case of the previous beta maps. In this situation, one has to resort to seeing these beta maps as acting smoothly in $M = S^1$, and cylinders will not anymore mark regions of continuity/differentiability, but will still mark injective regions, so to speak.

Remark 8. Condition (C5) was included to make transparent what is really used in the argument above. But one should be aware that conditions (C1-C3) suffice to conclude that densities are a.s. bounded away from 0 and ∞ and a.s. admit a uniform Holder constant (on the entire manifold M). See [77] Example 21. This is stronger than (C5), which then can, technically, be omitted from the list of conditions.

Remark 9. Some points have to be carefully evaluated in case one wants to generalize (C5) in such a way as to accommodate general quenched equilibrium states where $\mu_\omega = h_\omega \nu_\omega$ where h_ω is an eigenvector for the RPF operator in a suitable Banach space $\mathscr{B}$ (i.e., $\mathcal{L}_\omega h_\omega = \lambda_\omega h_{\theta\omega}$) and $\nu_\omega \in \mathcal{P}(M)$ is a conformal measure (i.e., $\mathcal{L}^*_\omega \nu_{\theta\omega} = \lambda_\omega \nu_\omega$). In this direction, useful existential results with decay of correlations are found, for example, in [63] [16], [65], [64], [72], [15], [7] [6], [9], [10], [82] and [83].

The first thing to notice is that the hypotheses of theorem 4.1.4 don't touch on which suitable Banach space $\mathscr{B}$ is used to reveal a spectral. Even if $\mathscr{B}$ is taken to be the BV space and decay appears in terms of BV against L^1, one can still recover what is required in terms of Lipschitz against L^∞. Despite decay being fine, and nowhere else in the hypotheses of theorem 4.1.4 properties of functional spaces and the associated density appearing to intervene, the reader will notice that we've just used knowledge of this kind to get α_ℓ's and guarantee (H9) in the above application.

In particular, we used that h_ω is continuous at $x(\omega)$, around which (in a vicinity of uniform size) h_ω is a continuous function with uniform moduli of continuity (in our case, uniform Holder constant). For example, when BV is the good Banach space and $h_\omega \in$ BV, it should still be required that h_ω is continuous on $x(\omega)$ (actually, in the qualified way we just described). This is not a cost-less requirement: BV functions are continuous Lebesgue-a.e., but when one wants compound limiting behavior, the target points generally aren't generic (or Lebesgue-generic) points — as in the periodic target case. So even mere continuity can not be guaranteed for free (not to say the qualified type of continuity we used).

Moreover, when $\nu_\omega \equiv \text{Leb}$, we have used that $\text{Leb}(B_\rho(x)) \overset{\forall x \in M}{=} Cr^{\upsilon}$ to do cancellations and get to equation (5.5). In general, one needs that: $\exists \rho_* > 0$, ω a.s., $\forall \rho \leqslant \rho_*$, $\forall k \geqslant 0$

$$\nu_{\theta^k\omega}(B_\rho(x(\theta^k\omega)))/\nu_\omega(B_\rho(x(\omega))) = 1$$

or at least: ω a.s., $\forall k \geqslant 0$

$$\lim_{\rho \to 0} \nu_{\theta^k\omega}(B_\rho(x(\theta^k\omega)))/\nu_\omega(B_\rho(x(\omega))) = 1,$$

with uniform-on-(ω, k) control on how the function approaches 1 as ρ shrinks to 0. Even if $\lim_{\rho \to 0} \frac{\log \nu_\omega(B_\rho(x))}{\log r} = C_\omega$ for ν_ω-a.e $x \in M$ (ω a.s.) and even if $C_\omega \equiv C$ a.s., it is not enough — because, once again, points x of interest aren't usually typical ones, but of periodic type. This is emphasized in contrast to hypothesis (H6.1), where dimensions were controlled roughly, utilizing a pair of inequalities. However, if one wants to assess the existence of the α_ℓ's and calculate them following the previous approach, exact dimensional control is needed. The needed fine dimensional control on quenched conformal measures is not available in the literature (except when $\nu_\omega \equiv \text{Leb}$).

It is also needed that the potential $\phi : \Omega \times M \to \mathbb{R}$ is so that $\operatorname{ess\,sup}_\omega H_\beta(\phi_\omega) < \infty$.

Provided that all the conditions needed to mimic the argument developed in the above proof are met, the associated α_ℓ's will read as in equation (5.6) but with $\prod_{j=0}^{M_\ell(\omega)-1} \lambda^{-1}_{\theta^i\omega} e^{\sum_{j=0}^{M_\ell(\omega)-1} \phi_{\theta^j\omega} \circ T^j_\omega(x(\omega))}$ instead of $\left(JT^{M_\ell(\omega)}_\omega(x(\omega))\right)^{-1}$.

End of remark.

Now we concentrate on analyzing the these conclusions of theorem 5.0.1 refine (or how α_ℓ's in equation (5.6) simplify) when additional conditions are considered.

Corollary 5.0.2. *Consider the assumptions of theorem 5.0.1 and assume further that $K(\omega) = 0$ a.s.. Then*

$$\alpha_\ell = \begin{cases} 1, \text{ if } \ell = 1 \\ 0, \text{ if } \ell \geqslant 2 \end{cases}, \tag{5.8}$$

and CPD in the limit theorem boils down to a standard Poisson.

Proof. Immediate. ∎

Corollary 5.0.3. *Consider the assumptions of theorem 5.0.1 and assume further that $\mathbb{P}$ is Bernoulli, $K(\omega) = \infty$ a.s. and*

$$h_\omega(x(\omega)) \perp \left(JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\right)_j.$$[6]

Then

$$\alpha_\ell = (D-1)D^{-\ell}, \text{ with } D^{-1} := \int_\Omega [JT^{m_0(\omega)}_\omega x(\omega)]^{-1} d\mathbb{P}(\omega),$$

and the CPD in the limit theorem boils down to a Polya-Aeppli (or geometric) one.

[6] This occurs when, for example, when $h_\omega \equiv 1$ a.s., or much more generally, when h_ω depends only on the past entries of ω (see, e.g., [68] prop. 1.2.3 and [65] prop. 3.3.2).

Proof. Notice that $K(\omega)=\infty$ a.s. and the independence of $h_\omega(x(\omega))$ from the rest implies

$$\alpha_\ell=\int_\Omega\prod_{j=0}^{\ell-2}\Big[JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\Big]^{-1}d\mathbb{P}(\omega)-\int_\Omega\prod_{j=0}^{\ell-1}\Big[JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\Big]^{-1}d\mathbb{P}(\omega),$$

then, after we make the point in I) that $\Big(\omega\mapsto JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\Big)_j$ is independent under $\mathbb{P}$, we will find that

$$\alpha_\ell=\prod_{j=0}^{\ell-2}\int_\Omega\Big[JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\Big]^{-1}d\mathbb{P}(\omega)-\prod_{j=0}^{\ell-1}\int_\Omega\Big[JT^{m_j(\omega)}_{\theta^{M_j(\omega)}\omega}(x(\theta^{M_j(\omega)}\omega))\Big]^{-1}d\mathbb{P}(\omega),$$

which, we will argue in II), equals

$$\alpha_\ell=\prod_{j=0}^{\ell-2}\int_\Omega\Big[JT^{m_0(\omega)}_{\omega}x(\omega)\Big]^{-1}d\mathbb{P}(\omega)-\prod_{j=0}^{\ell-1}\int_\Omega\Big[JT^{m_0(\omega)}_{\omega}x(\omega)\Big]^{-1}d\mathbb{P}(\omega)=(D-1)D^{-\ell},$$

where $D^{-1}:=\int_\Omega[JT^{m_0(\omega)}_\omega x(\omega)]^{-1}d\mathbb{P}(\omega)$, as desired.

Let us make the points that are missing.

I) Notice first that

$$\mathbb{P}(m_0(\omega)=i_0,m_1(\omega)=i_1)=\mathbb{P}(m_0(\omega)=i_0,m_0(\theta^{i_0}\omega)=i_1)=\mathbb{P}(\mathbb{1}_{\mathrm{Per}_{i_0}(\Gamma)}\mathbb{1}_{\theta^{-i_0}\mathrm{Per}_{i_1}(\Gamma)})$$

$$=\mathbb{P}(\mathbb{1}_{\mathrm{Per}_{i_0}(\Gamma)})\mathbb{P}(\mathbb{1}_{\theta^{-i_0}\mathrm{Per}_{i_1}(\Gamma)})=\mathbb{P}(m_0(\omega)=i_0)\mathbb{P}(m_0(\omega)=i_1),$$

where the first equality in the second line is because (π_j)'s are independent under $\mathbb{P}$ and the indicator functions can be expressed in terms of disjoint blocks of (π_j)'s, namely $\pi_0,\dots,\pi_{i_0-1}$ and $\pi_{i_0},\dots,\pi_{i_0+i_1-1}$. On the other hand

$$\mathbb{P}(m_1(\omega)=i_1)=\sum_{i_0}\mathbb{P}(m_0(\omega)=i_0,m_1(\omega)=i_1)$$

$$=\sum_{i_0}\mathbb{P}(m_0(\omega)=i_0)\mathbb{P}(m_0(\omega)=i_1)=\mathbb{P}(m_0(\omega)=i_1).$$

So combining the two previous chains of equality, we find that m_0 and m_1 are independent, i.e., $m_0\perp m_1$.

Once again, since $(\pi_j)_j$ is an independency under $\mathbb{P}$, whenever two random variables X and Y can be expressed as $X=\phi\circ(\pi_0,\dots,\pi_{i_0-1})$ and $Y=\psi\circ(\pi_{i_0},\dots,\pi_{i_0+i_1-1})$, then $X\perp Y$. Similarly for π instead of π. This is the case for $(JT^{i_0}_{\cdot}(x(\cdot)),\mathbb{1}_{m_0(\cdot)=i_0})\perp(JT^{i_1}_{\theta^{i_0}\cdot}(x\circ\theta^{i_0}(\cdot)),\mathbb{1}_{m_1(\cdot)=i_1})$.

Therefore

$$
\begin{aligned}
&\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = a, \left[JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\theta^{m_0(\omega)}\omega))\right]^{-1} = b\right\}\right) \\
&= \sum_{i_0}\sum_{i_1}\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{i_0}(x(\omega))\right]^{-1} = a, \left[JT_{\theta^{i_0}\omega}^{i_1}(x(\theta^{i_0}\omega))\right]^{-1} = b, m_0(\omega) = i_0, m_0(\theta^{i_0}\omega) = i_1\right\}\right) \\
&= \sum_{i_0}\sum_{i_1}\left[\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{i_0}(x(\omega))\right]^{-1} = a, m_0(\omega) = i_0\right\}\right)\mathbb{P}\left(\left\{\omega : \left[JT_{\theta^{i_0}\omega}^{i_1}(x(\theta^{i_0}\omega))\right]^{-1} = b, m_0(\theta^{i_0}\omega) = i_1\right\}\right)\right] \\
&= \left[\sum_{i_0}\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{i_0}(x(\omega))\right]^{-1} = a, m_0(\omega) = i_0\right\}\right)\right]\left[\sum_{i_1}\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{i_1}(x(\omega))\right]^{-1} = b, m_0(\omega) = i_1\right\}\right)\right] \\
&= \mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = a\right\}\right)\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = b\right\}\right).
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
&\mathbb{P}\left(\left\{\omega : \left[JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\theta^{m_0(\omega)}\omega))\right]^{-1} = b\right\}\right) \\
&= \sum_{a}\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = a, \left[JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}(x(\theta^{m_0(\omega)}\omega))\right]^{-1} = b\right\}\right) \\
&= \sum_{a}\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = a\right\}\right)\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = b\right\}\right) \\
&= \mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = b\right\}\right).
\end{aligned}
$$

So combining the two previous chains of equality, we find that

$$
JT_{\cdot}^{m_0(\cdot)}(x(\cdot)) \perp JT_{\theta^{m_0(\cdot)}\cdot}^{m_1(\cdot)}(x(\theta^{m_0(\cdot)}\cdot)),
$$

as desired.

II) Notice that

$$
\begin{aligned}
&\int_{\Omega}\left[JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}x(\theta^{m_0(\omega)}\omega)\right]^{-1}d\mathbb{P}(\omega) = \sum_{b} b\mathbb{P}\left(\left\{\omega : JT_{\theta^{m_0(\omega)}\omega}^{m_1(\omega)}x(\theta^{m_0(\omega)}\omega)\right]^{-1} = b\right\}\right) \\
&= \sum_{b} b\mathbb{P}\left(\left\{\omega : \left[JT_{\omega}^{m_0(\omega)}(x(\omega))\right]^{-1} = b\right\}\right) = \int_{\Omega}\left[JT_{\omega}^{m_0(\omega)}x(\omega)\right]^{-1}d\mathbb{P}(\omega),
\end{aligned}
$$

where we have used the last equality in I). ■

Chapter 6

Future research

The theory developed in this thesis has an inclination to cover non-uniformly expanding behavior with possibly unbounded derivatives, in the sense of hypotheses (H3.1), (H7) and (H4.2), but the examples in section 5 were uniformly expanding ones.

On-going research efforts are then directed to produce and accommodate examples of bonafide non-uniformly expanding random systems with associated targets which randomly approach their neutral points (unlikely and slowly).

Another relevant extension of the theory is to get rid of the subsequence need, $\rho_m \searrow 0$ as $m \to \infty$, entailed by lemma 4.4.3, and to upgrade theorem 4.1.4 to work under the plain limit $\rho \to 0$. This is undergoing investigation by Jiakang Wang and Nicolai Haydn.

Investigation can consider targets that contain indifferent points (or approaching them likely and/or fastly), in which case special time-normalization has to be considered.

Other research directions include adapting the theory to maps with Jacobians in $\mathbb{R}_{\geqslant 0}$, maps with infinitely many branches and applications drawing infinitely many maps.

Finally, it could also be interesting to explore how this approach adapts to handle quenched time-normalization. Ideally, assuming an ergodic driving would be enough, but a reformulation of the strategy to control almost surely the ω-dependent errors and leading term would probably be needed.

Bibliography

[1] Aaronson, J. and Denker, M. (1997). *Local limit theorems for Gibbs-Markov maps*. Mathematisches Institut der Universität Göttingen.

[2] Abadi, M., Freitas, A. C. M., and Freitas, J. M. (2019). Clustering indices and decay of correlations in non-markovian models. *Nonlinearity*, 32(12):4853.

[3] Abadi, M., Freitas, A. C. M., and Freitas, J. M. (2020). Dynamical counterexamples regarding the extremal index and the mean of the limiting cluster size distribution. *Journal of the London Mathematical Society*, 102(2):670–694.

[4] Abadi, M. and Galves, A. (2001). Inequalities for the occurrence times of rare events in mixing processes. the state of the art. *Markov Process. Related Fields*, 7(1):97–112.

[5] Arnold, L., Jones, C. K., Mischaikow, K., Raugel, G., and Arnold, L. (1995). *Random dynamical systems*. Springer.

[6] Atnip, J., Froyland, G., González-Tokman, C., and Vaienti, S. (2021a). Thermodynamic formalism for random interval maps with holes. *arXiv preprint arXiv:2103.04712*.

[7] Atnip, J., Froyland, G., González-Tokman, C., and Vaienti, S. (2021b). Thermodynamic formalism for random weighted covering systems. *Communications in Mathematical Physics*, 386(2):819–902.

[8] Atnip, J., Froyland, G., González-Tokman, C., and Vaienti, S. (2023a). Compound poisson statistics for dynamical systems via spectral perturbation. *arXiv preprint arXiv:2308.10798*.

[9] Atnip, J., Froyland, G., González-Tokman, C., and Vaienti, S. (2023b). Equilibrium states for non-transitive random open and closed dynamical systems. *Ergodic Theory and Dynamical Systems*, 43(10):3193–3215.

[10] Atnip, J., Froyland, G., González-Tokman, C., and Vaienti, S. (2023c). Thermodynamic formalism and perturbation formulae for quenched random open dynamical systems.

[11] Atnip, J., Haydn, N., and Vaienti, S. (2020). Extreme value theory with spectral techniques: application to a simple attractor. *arXiv preprint arXiv:2002.10863*.

[12] Aytaç, H., Freitas, J., and Vaienti, S. (2015). Laws of rare events for deterministic and random dynamical systems. *Transactions of the American Mathematical Society*, 367(11):8229–8278.

[13] Azevedo, D., Freitas, A. C. M., Freitas, J. M., and Rodrigues, F. B. (2016). Clustering of extreme events created by multiple correlated maxima. *Physica D: Nonlinear Phenomena*, 315:33–48.

[14] Azevedo, D., Freitas, A. C. M., Freitas, J. M., and Rodrigues, F. B. (2017). Extreme value laws for dynamical systems with countable extremal sets. *Journal of Statistical Physics*, 167(5):1244–1261.

[15] Bahsoun, W., Bose, C., and Ruziboev, M. (2019). Quenched decay of correlations for slowly mixing systems. *Transactions of the American Mathematical Society*, 372(9):6547–6587.

[16] Baladi, V. (1997). Correlation spectrum of quenched and annealed equilibrium states for random expanding maps. *Communications in mathematical physics*, 186:671–700.

[17] Baladi, V. (2000). *Positive transfer operators and decay of correlations*, volume 16. World scientific.

[18] Barbour, A. D., Chen, L. H., and Loh, W.-L. (1992). Compound poisson approximation for nonnegative random variables via stein's method. *The Annals of Probability*, 20(4):1843–1866.

[19] Basrak, B., Planinić, H., and Soulier, P. (2018). An invariance principle for sums and record times of regularly varying stationary sequences. *Probability Theory and Related Fields*, 172:869–914.

[20] Bogenschütz, T. (1992). Entropy, pressure, and a variational principle for random dynamical systems. *Random Comput. Dynam*, 1(1):99–116.

[21] Bruin, H., Saussol, B., Troubetzkoy, S., and Vaienti, S. (2003). Return time statistics via inducing. *Ergodic theory and dynamical systems*, 23(4):991–1013.

[22] Buhl, S. (2013). Modelling and estimation of extremes in space and time.

[23] Carney, M., Holland, M., and Nicol, M. (2021). Extremes and extremal indices for level set observables on hyperbolic systems. *Nonlinearity*, 34(2):1136.

[24] Carvalho, M., Freitas, A. C. M., Freitas, J. M., Holland, M., and Nicol, M. (2015). Extremal dichotomy for uniformly hyperbolic systems. *Dynamical Systems*, 30(4):383–403.

[25] Chazottes, J.-R. and Collet, P. (2013). Poisson approximation for the number of visits to balls in non-uniformly hyperbolic dynamical systems. *Ergodic Theory and Dynamical Systems*, 33(1):49–80.

[26] Coles, S., Bawa, J., Trenner, L., and Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, volume 208. Springer.

[27] Collet, P. (1996). Some ergodic properties of maps of the interval. dynamical and disordered systems. r. bamon, jm gambaudo and s. martinez ed.

[28] Crauel, H. (2002). *Random probability measures on Polish spaces*, volume 11. CRC press.

[29] Crimmins, H. (2019). Stability of hyperbolic oseledets splittings for quasi-compact operator cocycles. *arXiv preprint arXiv:1912.03008*.

[30] Crimmins, H. and Saussol, B. (2022). Quenched poisson processes for random subshifts of finite type. *Nonlinearity*, 35(6):3036.

[31] Doeblin, W. (1940). Remarques sur la théorie métrique des fractions continues. *Compositio mathematica*, 7:353–371.

[32] Dragičević, D., Froyland, G., Gonzalez-Tokman, C., and Vaienti, S. (2018). A spectral approach for quenched limit theorems for random expanding dynamical systems. *Communications in Mathematical Physics*, 360:1121–1187.

[33] Faranda, D., Ghoudi, H., Guiraud, P., and Vaienti, S. (2018). Extreme value theory for synchronization of coupled map lattices. *Nonlinearity*, 31(7):3326.

[34] Freitas, A. C., Freitas, J., and Magalhães, M. (2020a). Complete convergence and records for dynamically generated stochastic processes. *Transactions of the American Mathematical Society*, 373(1):435–478.

[35] Freitas, A. C. M., Freitas, J. M., and Magalhães, M. (2018). Convergence of marked point processes of excesses for dynamical systems. *Journal of the European Mathematical Society*, 20(9):2131–2179.

[36] Freitas, A. C. M., Freitas, J. M., Magalhães, M., and Vaienti, S. (2020b). Point processes of non stationary sequences generated by sequential and random dynamical systems. *Journal of Statistical Physics*, 181:1365–1409.

[37] Freitas, A. C. M., Freitas, J. M., Rodrigues, F. B., and Soares, J. V. (2020c). Rare events for cantor target sets. *Communications in Mathematical Physics*, 378:75–115.

[38] Freitas, A. C. M., Freitas, J. M., and Soares, J. V. (2021). Rare events for product fractal sets. *Journal of Physics A: Mathematical and Theoretical*, 54(34):345202.

[39] Freitas, A. C. M., Freitas, J. M., and Todd, M. (2010). Hitting time statistics and extreme value theory. *Probability Theory and Related Fields*, 147(3-4):675–710.

[40] Freitas, A. C. M., Freitas, J. M., and Todd, M. (2012). The extremal index, hitting time statistics and periodicity. *Advances in Mathematics*, 231(5):2626–2665.

[41] Freitas, A. C. M., Freitas, J. M., and Todd, M. (2013). The compound poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics. *Communications in Mathematical Physics*, 321(2):483–527.

[42] Freitas, A. C. M., Freitas, J. M., and Todd, M. (2015). Speed of convergence for laws of rare events and escape rates. *Stochastic Processes and their Applications*, 125(4):1653–1687.

[43] Freitas, A. C. M., Freitas, J. M., and Todd, M. (2020d). Enriched functional limit theorems for dynamical systems. *arXiv preprint arXiv:2011.10153*.

[44] Freitas, A. C. M., Freitas, J. M., Todd, M., and Vaienti, S. (2016). Rare events for the manneville–pomeau map. *Stochastic Processes and their Applications*, 126(11):3463–3479.

[45] Freitas, A. C. M., Freitas, J. M., and Vaienti, S. (2017). Extreme value laws for non stationary processes generated by sequential and random dynamical systems.

[46] Froyland, G., Lloyd, S., and Quas, A. (2010). A semi-invertible oseledets theorem with applications to transfer operator cocycles. *arXiv preprint arXiv:1001.5313*.

[47] Gallager, R. G. (2013). *Stochastic processes: theory for applications*. Cambridge University Press.

[48] Gallo, S., Haydn, N., and Vaienti, S. (2021). Number of visits in arbitrary sets for ϕ-mixing dynamics. *arXiv preprint arXiv:2107.13453*.

[49] Galves, A. and Schmitt, B. (1997). Inequalities for hitting times in mixing dynamical systems. *Random and Computational Dynamics*, 5(4):337–348.

[50] González-Tokman, C. and Quas, A. (2014). A semi-invertible operator oseledets theorem. *Ergodic Theory and Dynamical Systems*, 34(4):1230–1272.

[51] Haydn, N., Lunedei, E., and Vaienti, S. (2007). Averaged number of visits. *Chaos: An interdisciplinary journal of nonlinear science*, 17(3).

[52] Haydn, N. and Todd, M. (2016). Return times at periodic points in random dynamics. *Nonlinearity*, 30(1):73.

[53] Haydn, N. and Vaienti, S. (2020). Limiting entry and return times distribution for arbitrary null sets. *Communications in Mathematical Physics*, 378:149–184.

[54] Hirata, M. (1993). Poisson law for axiom a diffeomorphisms. *Ergodic Theory and Dynamical Systems*, 13(3):533–556.

[55] Hirata, M., Saussol, B., and Vaienti, S. (1999). Statistics of return times: A general framework and new applications. *Communications in Mathematical Physics*, 206:33–55.

[56] Holland, M., Nicol, M., and Török, A. (2012). Extreme value theory for non-uniformly expanding dynamical systems. *Transactions of the American Mathematical Society*, 364(2):661–688.

[57] Hsing, T. (1987). On the characterization of certain point processes. *Stochastic processes and their applications*, 26:297–316.

[58] Kallenberg, O. (1983). *Random measures*. De Gruyter.

[59] Kato, T. (2013). *Perturbation theory for linear operators*, volume 132. Springer Science & Business Media.

[60] Keller, G. (2012). Rare events, exponential hitting times and extremal indices via spectral perturbation. *Dynamical Systems*, 27(1):11–27.

[61] Keller, G. and Liverani, C. (1999). Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 28(1):141–152.

[62] Keller, G. and Liverani, C. (2009). Rare events, escape rates and quasistationarity: some exact formulae. *Journal of Statistical Physics*, 135:519–534.

[63] Khanin, K. and Kifer, Y. (1995). Thermodynamic formalism for random transformations and statistical mechanics. *American Mathematical Society Translations: Series 2*, pages 107–140.

[64] Kifer, Y. (2008). Thermodynamic formalism for random transformations revisited. *Stochastics and Dynamics*, 8(01):77–102.

[65] Kifer, Y. and Liu, P.-D. (2006). Random dynamics. *Handbook of dynamical systems*, 1:379–499.

[66] Last, G. and Penrose, M. (2017). *Lectures on the Poisson process*, volume 7. Cambridge University Press.

[67] Leadbetter, M. R., Lindgren, G., and Rootzén, H. (2012). *Extremes and related properties of random sequences and processes*. Springer Science & Business Media.

[68] Ledrappier, F. and Young, L.-S. (1988). Entropy formula for random transformations. *Probability theory and related fields*, 80(2):217–240.

[69] Lucarini, V., Faranda, D., de Freitas, J. M. M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, S., et al. (2016). *Extremes and recurrence in dynamical systems*. John Wiley & Sons.

[70] Luzzatto, S. (2005). Stochastic-like behaviour in nonuniformly expanding maps. *Handbook of Dynamical Systems: Volume 1B*, 2:265.

[71] Mantica, G. and Perotti, L. (2016). Extreme value laws for fractal intensity functions in dynamical systems: Minkowski analysis. *Journal of Physics A: Mathematical and Theoretical*, 49(37):374001.

[72] Mayer, V., Skorulski, B., and Urbanski, M. (2011). *Distance expanding random mappings, thermodynamical formalism, Gibbs measures and fractal geometry*, volume 2036. Springer Science & Business Media.

[73] Mori, T. (1977). Limit distributions of two-dimensional point processes generated by strong-mixing sequences.

[74] Pène, F. and Saussol, B. (2020). Spatio-temporal poisson processes for visits to small sets. *Israel Journal of Mathematics*, 240:625–665.

[75] Pitskel, B. (1991). Poisson limit law for markov chains. *Ergodic Theory and Dynamical Systems*, 11(3):501–513.

[76] Resnick, S. I. (2008). *Extreme values, regular variation, and point processes*, volume 4. Springer Science & Business Media.

[77] Rousseau, J., Saussol, B., and Varandas, P. (2014). Exponential law for random subshifts of finite type. *Stochastic Processes and their Applications*, 124(10):3260–3276.

[78] Rousseau, J. and Todd, M. (2015). Hitting times and periodicity in random dynamics. *Journal of Statistical Physics*, 161:131–150.

[79] Sabourin, A. (2021). Extremes-lecture notes master 2 'mathématiques et applications', parcours 'mathématiques de l'aléatoire', orsay.

[80] Sabourin, A. and LTCI, T. P. (2020). Extreme value analysis: Lecture.

[81] Serfozo, R. F. (1990). Point processes. *Handbooks in operations research and management science*, 2:1–93.

[82] Stadlbauer, M., Suzuki, S., and Varandas, P. (2021). Thermodynamic formalism for random non-uniformly expanding maps. *Communications in Mathematical Physics*, 385(1):369–427.

[83] Stadlbauer, M., Varandas, P., and Zhang, X. (2023). Quenched and annealed equilibrium states for random ruelle expanding maps and applications. *Ergodic Theory and Dynamical Systems*, 43(9):3150–3192.

[84] Stein, C. (1986). Approximate computation of expectations. IMS.

[85] Thieullen, P. (1987). Fibrés dynamiques asymptotiquement compacts exposants de lyapounov. entropie. dimension. In *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, volume 4, pages 49–97. Elsevier.

Lucas, AMORIM VILAS BOAS

CPT - Centre de Physique Théorique, Université de Toulon

Statistique de Poisson composée pour les systèmes dynamiques aléatoires

Résumé en français

Nous obtenons des distributions d'entrée limites quenched dans la classe composée de Poisson pour une certaine famille de systèmes dynamiques aléatoires en utilisant une approximation probabiliste par bloc pour la fonction de comptage d'entrée quenched jusqu'au temps normalisé annealed-Kac. Nous considérons des cibles aléatoires générales avec des statistiques de retour bien définies et des systèmes avec une d'ecroissance polynomiale des corrélations à la fois quenched et annealed. La théorie est rendue opérationnelle grâce à un résultat qui permet de récupérer certaines statistiques d'entrée à partir desdites statistiques de retour, qui sont calculables. Nos exemples incluent une classe de systèmes unidimensionnels à expansion aléatoire par morceaux, jetant un nouvel éclairage sur la dichotomie déterministe bien connue entre les points périodiques et apériodiques, leur formule d'indice extrême habituelle EI = 1 – $JT^p(z)$, et récupérer le cas géométrique pour les systèmes généraux pilotés par Bernoulli, mais comportement distinct dans le cas contraire. Les enquêtes futures et en cours visent à produire et à prendre en compte des exemples de véritables systèmes aléatoires à expansion non uniforme et de cibles s'approchant de leurs points neutres.

Mot clés : Systèmes Dynamiques, Perturbations aléatoires, Statistiques de Poisson composée.

Compound Poisson distributions for random dynamical systems

Résumé en anglais

We obtain quenched limiting hitting distributions in the compound Poisson class for a certain family of random dynamical systems using a probabilistic block-approximation for the quenched hit-counting function up to annealed-Kac-normalized time. We consider general random targets with well-defined return statistics and systems with both quenched and annealed polynomial decay of correlations. The theory is made operational due to a result that allows certain hitting statistics to be recovered from the said return statistics, which are computable. Our examples include a class of random piecewise expanding one-dimensional systems, casting new light on the well-known deterministic dichotomy between periodic and aperiodic points, their usual extremal index formula EI = 1 – $JT^p(z)$, and recovering the geometric case for general Bernoulli-driven systems, but distinct behavior otherwise. Future and on-going investigations aim to produce and accommodate examples of bonafide non- uniformly expanding random systems and targets approaching their neutral points.

Keywords : Dynamical systems, Random perturbations, Compound Poisson statistics.