Universidade Federal da Bahia - UFBA
Instituto de Matemática e Estatística - IME
Programa de Pós-Graduação em Matemática - PGMAT
Tese de Doutorado

Asymptotic probabilistic properties of orbits: return times and shortest distance

Adriana Coutinho dos Santos

Salvador-Bahia
15 de Fevereiro de 2019

Asymptotic probabilistic properties of orbits: return times and shortest distance

Adriana Coutinho dos Santos

Tese apresentada ao Colegiado do Curso de Pós-graduação em Matemática da Universidade Federal da Bahia, como requisito parcial para obtenção do Título de Doutora em Matemática.

Orientador: Prof. Dr. Jérôme François Alain Jean Rousseau
Coorientador: Prof. Dr. Benoît Saussol

Salvador-Bahia
15 de Fevereiro de 2019

Coutinho dos Santos, Adriana

Asymptotic probabilistic properties of orbits: return times and shortest distance / Adriana Coutinho dos Santos. - Salvador, 2019.

81 f. : il

Orientador: Prof. Dr. Jérôme François Alain Jean Rousseau.

Coorientador: Prof. Dr. Benoît Saussol .

Tese (Doutorado - Matemática) - Universidade Federal da Bahia, Instituto de Matemática e Estatística, Programa de Pós-graduação em Matemática, 2019.

1. Recorrência de Poincaré. 2. Taxa exponencial. 3. Repulsor conforme. 3. Grandes desvios. 4. Correspondência de sequências. 5. Teoria de códigos. 6. Entropia de Rényi. I. François Alain Jean Rousseau, Jérôme . II. Saussol, Benoît III. Título.

CDU : 517.938

: 519.723

Asymptotic probabilistic properties of orbits: return times and shortest distance

Adriana Coutinho dos Santos

Tese apresentada ao Colegiado do Curso de Pós-graduação em Matemática da Universidade Federal da Bahia, como requisito parcial para obtenção do Título de Doutora em Matemática.

Banca examinadora:

Prof. Dr. Jérôme François Alain Jean Rousseau
(Orientador - PGMAT/UFBA)

Prof. Dr. Paulo César Rodrigues Pinto Varandas
(PGMAT/UFBA)

Prof. Dr. Benoît Saussol
(Coorientador - Université de Bretagne Occidentale)

Prof. Dr. Miguel Natalio Abadi
(USP)

Prof. Dr. Rodrigo Lambert
(UFU)

À minha família,
pelo carinho e apoio.

Acknowledgement

Em primeiro lugar, agradeço a Deus pelo dom da vida e por ter me sustentado até aqui, sem Ele eu nada posso. Como sempre Deus me mostrou que vale a pena acreditar e que os seus planos são maiores do que imaginei. Obrigada Deus pelo seu amor incondicional.

Agradeço a minha família, principalmente meus pais, Giselia e Edvaldo, e meu irmão Alisson. Obrigada por tudo que vocês fizeram por mim. Obrigada, pelo apoio em cada decisão que tomei ao longo da minha vida. Se eu cheguei onde estou hoje foi por causa desse apoio. Agradeço por compreenderem minha ausência nos muitos momentos de encontros em família. Amo vocês!

Agradeço a Jérôme pela orientação ao longo destes 4 anos de doutorado. No final do mestrado quando ele me apresentou as possibilidades para o doutorado, confesso que fiquei receosa. Foi sugerido o doutorado sanduíche, eu deveria ir para Brest trabalhar com Benoît. Essa foi mais uma das escolhas acertadas que fiz na vida. E quem diria que aquela menina que há 7 anos dizia que o mais distante que iria pra estudar seria Salvador, iria mais longe ainda? Hoje pra mim o longe é um lugar que não existe. O doutorado me proporcionou experiências que jamais passou em minha cabeça vivenciar tão cedo. Fui a Brest, conheci Benoît e a teoria de grandes desvios. Fiz amigos com os quais tenho contato até hoje. A vida me reservou muitas surpresas. Quando voltei a Brest pela segunda vez, conheci meu amor. Agradeço a Jérôme mais uma vez por ter me incentivado a viajar. Agradeço pelas oportunidades a mim apresentadas. Agradeço pela atenção, pelo cuidado e por me ensinar sobre pesquisa. Eu aprendi muito com ele.

Je dresse mes remerciements les plus distingués au professeur Benoît, qui m'a accueilli tour à tour à l'Université de Bretagne Occidentale (UBO). Votre disponibilité constante et votre attention, même quand j'avais des multiples doutes sur certains problèmes de mon sujet de recherche par mails m'ont encouragé à poursuivre dans ce chemin pénible. Je peux dire que j'ai eu de la chance d'avoir Jérôme et Benoît comme directeurs de thèse. Grâce à eux j'ai pu acquérir beaucoup d'expériences à travers de nombreux conférences et séminaires sans oublier toutes ces expériences de recherches mathématiques que j'ai pu avoir. Je tiens très sincèrement lui dire encore merci.

Aos outros membros da banca meu muito obrigada. Agradeço a Paulo, pelos conselhos mesmo que indiretamente. Ele sempre foi muito sábio nas suas colocações. Durante as aulas de formalismo termodinâmico e nas visitas, com alguma pausa na sala 282, deixou palavras de incentivo para nós que muitas vezes estávamos desanimados com a pesquisa. Agradeço a Rodrigo, por ter aceitado participar da banca e pelo conhecimento partilhado nesses meses de trabalho em conjunto. Aproveito para me desculpar pelas minhas chatices. Agradeço a Miguel, pela participação na banca e também por ter aceitado ser meu supervisor de pós-doutorado. Agradeço também aos professores suplentes, Edgar e Davide, por terem participado da defesa, pelos questionamentos e pelas contribuições.

Je remercie mon fiancé Clé. Il était très important pendant la période du doctorat à Brest. J'étais loin de ma famille et de mes amis et il a comblé cette distance. Je crois que c'est Dieu qui a fait qu'on se croise dans les couloirs de l'université. Merci pour être toujours avec moi. Merci pour tes conseils et ton soutien constant. Loin de mes yeux, mais prêt de mon coeur. Notre histoire c'est juste au début. Je t'aime!

Obrigada aos meus amigos, aos de perto e aos distantes. Obrigada Carol, pelo companheirismo ao longo dos quase 6 anos em Salvador. Mais que uma amiga, é como uma irmã. Agradeço a Mille, que iniciou essa caminhada conosco no mestrado. Nossos laços se estreitaram ao longo dos anos, prezo muito por sua amizade. Não posso esquecer de Verônica e Carol Morais. Obrigada meninas, pelo incentivo, pelas orações e por se fazerem presentes nas minhas conquistas. Obrigada Moacyr, pela amizade sincera e pela prestatividade de sempre. Agradeço também a Mari, pela amizade, pelo carinho e disponibilidade em ajudar. Obrigada a todos que continuaram na minha vida, mesmo depois dos cursos técnicos, da graduação e do mestrado. Obrigada Marcia, que me ensinou francês. Você me encorajou a não ter medo.

Obrigada a todos os colegas de doutorado. Agradeço a Junilson, Heides e Edvan pela luta no início do doutorado, pelos estudos em conjunto para o exame de qualificação e por estarem sempre disponíveis a ajudar nos momentos de dúvidas. Agradeço aos mais assíduos da sala 282: Alejandra, Diego, Diogo, Elaine, Fabíola e Vinícius. Vocês alegraram muitos dos meus dias tristes e solitários. Obrigada pela companhia diária, pela ajuda nos momentos de dúvidas e pelo riso nos muitos momentos de descontração. Eu torço pelo sucesso e realização de cada um de vocês.

Agradeço a todos os funcionários do IME-UFBA, pelo profissionalismo e pela disposição em ajudar.

Agradeço também aos professores com os quais tive a honra de cursar disciplinas. Obrigada pela acolhida, pelos ensinamentos e pelo exemplo.

Por fim, agradeço a CAPES pelo apoio financeiro.

"C'est le temps que tu as perdu pour ta rose
qui fait ta rose si importante"

Antoine de Saint-Exupéry

Abstract

This work provides some original contributions to the study of large deviation for return times and the asymptotic behavior of the shortest distance between observed orbits. In the first part, we prove a large deviation result for return time of the orbits of a dynamical system in a r-neighbourhood of an initial point x. Our first result may be seen as a differentiable version of the work by Jain and Bansal, who considered the return time of a stationary and ergodic process defined in the space of infinite sequences. We obtain large deviation estimates for dynamical systems in general and in the case of conformal repellers we compute the rate functions in terms of HP-spectrum for dimensions of multifractal analysis.

In the second part of this work, we investigate the shortest distance between two observed orbits and its asymptotic behavior. The main result is a strong law of large numbers for a re-scaled version of this quantity, which presents an explicit relation with the correlation dimension of the pushforward measure. We apply this result to study the shortest distance between orbits of a random dynamical system. In the symbolic case, this problem corresponds to the longest common substring problem for encoded sequences and its limiting relationship with the Rényi entropy. We apply this results to the zero-inflated contamination model and to the stochastic scrabble.

Keywords: Poincaré recurrence, exponential rate, conformal repeller, large deviation, string-matching, coding theory, Rényi entropy.

Resumo

Este trabalho fornece algumas contribuições originais para o estudo de grandes desvios para tempo de retorno e comportamento assintótico da menor distância entre duas órbitas transformadas. Na primeira parte, provamos um resultado de grandes desvios para o tempo de retorno de uma órbita de um sistema dinâmico numa r-vizinhança de seu ponto inicial x. Nosso primeiro resultado pode ser visto como uma versão diferenciável do trabalho de Jain e Bansal, que consideraram o tempo de retorno de um processo estacionário e ergódico definido no espaço das sequências finitas. Obtemos estimativas de grandes desvios para sistemas dinâmicos gerais, e no caso de repulsor conforme calculamos as funções taxas em termos do HP-espectro para dimensão da análise multifractal.

Na segunda parte deste trabalho, investigamos a menor distância entre duas órbitas transformadas e seu comportamento assintótico. O principal resultado é uma lei forte dos grandes números para uma versão reescalonada desta quantidade. A quantidade limite apresenta uma relação explícita com a dimensão de correlação da medida *pushforward*. Aplicamos este resultado ao estudo da menor distância entre órbitas para um sistema dinâmico aleatório. No caso simbólico, este problema corresponde ao problema da maior subsequência comum entre sequências codificadas, e o seu limitante está relacionado com a entropia de Rényi do processo. Aplicamos este resultado aos modelos de contaminação inflada por zeros, e sequências de caracteres com pesos.

Palavras-chave: Recorrência de Poincaré, taxa exponencial, repulsor conforme, grandes desvios, correspondência de sequências, teoria de códigos, entropia de Rényi.

Contents

Introduction

Consider a dynamical system $(M, \mathcal{A}, g, \mu)$ where M is a compact metric space, $\mathcal{A}$ is a σ-algebra on M, $g : M \to M$ is a measurable map and μ an invariant probability measure on $(M, \mathcal{A})$. An essential result in ergodic theory is Poincaré's recurrence theorem. It states that any probability measure preserving map has almost everywhere recurrence. It is natural to ask for more quantitative results of recurrence. In [35], Kac has proven that, when the system is ergodic, the mean of the return time in a measurable set is equal to the inverse of the measure of this set. This subject has been further studied by many authors. In particular, Boshernitzan [19] established a link between the Hausdorff dimension of M and the time needed by an orbit to approach its initial point. To review results on quantitative recurrence see, for example [29, 52]. In the present work we are interested in large deviation for return times for a class of systems with exact dimensional measures.

Moreover, finding patterns on symbolic strings has been a widely studied subject matter on Genetics, Probability and Information Theory over the years. The investigations about how many information a n-string have on the whole realization of the process are naturally linked with the concept of redundancy and compression algorithms. On the other hand, the overlap between (some proportion of) two different strings can give us some knowledge about the similarity of the sources that generate those processes. In addition, repetition and similarity are two well-exploited concepts in the study of DNA sequences. In this direction, we will focus on the longest common substring problem and its dynamical correspondent, the shortest distance between observed orbits.

The remainder of this introduction will be devoted to discuss about these individual topics.

Results on large deviation theory

Several works already addressed large deviations for return time. Abadi and Vaienti in [6] proved large deviation properties of $\tau(C_n)/n$, where $\tau(C_n)$ is the first return of a n-cylinder to itself. More precisely, if the system is ψ-mixing, if $\psi(0) < 1$ and the Rényi entropies exist for all integers β, then for $\delta \in (0,1]$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \mu \{x : \tau(C_n) \leq [\delta n]\} := M(\delta)$$

exists. In addition, they provide an explicit expression for $M(\delta)$. Generalizations were later shown in [2, 31].

A large deviation result for the n-th return times $\tau_A^n(x)$ into a fixed set A was also considered by Chazottes and Leplaideur [21] (see also [37]). Birkhoff's theorem gives that for μ-almost every point x

$$\lim_{n\to\infty} \frac{\tau_A^n(x)}{n} = \frac{1}{\mu(A)}.$$

For Axiom A diffeomorphisms and equilibrium states μ, they proved the existence of a rate function Φ_A, such that for every $u \geq \frac{1}{\mu(A)}$,

$$\lim_{n\to\infty} \frac{1}{n} \log \mu \left\{ \frac{\tau_A^n}{n} \geq u \right\} = \Phi_A(u),$$

with the appropriate change in the definition when $0 \leq u \leq \frac{1}{\mu(A)}$.

Our first result concerns a different notion of large deviation for return time, and may be seen as a differentiable version of a recent work by Jain and Bansal [33]. They studied a large deviation property for repetition times under ϕ-mixing conditions. Let H denote the entropy rate of a finite-valued process $X = (X_n)$ and x a particular realization of X. Define the first return time of x_1^n as

$$R_n(x) = \min \left\{ j \geq 1 : x_1^n = x_{-j+1}^{-j+n} \right\}.$$

We say that X has exponential rates for entropy if for every $\epsilon > 0$, we have

$$\mathbb{P}\left(\left\{ x_1^n : 2^{-n(H+\epsilon)} \leq \mathbb{P}(x_1^n) \leq 2^{-n(H-\epsilon)} \right\} \right) \leq 1 - r(n,\epsilon),$$

where $r(\epsilon, n) = e^{-k(\epsilon)n}$, with $k(\epsilon)$ a real valued positive function of ϵ. They proved that for an exponentially ϕ-mixing process with exponential rates for entropy,

$$\mathbb{P}\left(\left| \frac{\log R_n(X)}{n} - H \right| > \epsilon \right) \leq 2e^{-I(\epsilon)n} \text{ for any } n \text{ sufficiently large,}$$

where $I(\epsilon)$ is a real positive valued function for all $\epsilon > 0$ and $I(0) = 0$.

Here, we will study the return time $\tau_r(x)$ of a point $x \in M$ under the map g in its r-neighborhood, defined as follows:

$$\tau_r(x) = \tau_{B(x,r)}(x) = \min\{n \geq 1 : d(g^n x, x) < r\}.$$

It was proved by Barreira and Saussol [17] that

$$\underline{R}(x) \leq \underline{d}_\mu(x) \quad \text{and} \quad \overline{R}(x) \leq \overline{d}_\mu(x),$$

for μ-almost every $x \in M$, where $\underline{R}(x)$, $\overline{R}(x)$, $\underline{d}_\mu(x)$ and $\overline{d}_\mu(x)$ are the lower and upper recurrence rates and the lower and upper pointwise dimensions of the measure μ at the point $x \in M$, respectively. If the system has a super-polynomial decay of correlations, Saussol in [51] showed that equalities will hold for the expressions above.

In the first part of this work, for a measure μ exact dimensional, we are interested in studying the limiting behavior as r goes to zero of $\mu\left(\tau_r \geq r^{-d_\mu-\epsilon}\right)$ and $\mu\left(\tau_r \leq r^{-d_\mu+\epsilon}\right)$. This characterization is made via asymptotic exponential bound. We consider the limits

$$\lim_{r\to 0}\frac{1}{\log r}\log\mu\left(\tau_r \geq r^{-d_\mu-\epsilon}\right) \quad \text{and} \quad \lim_{r\to 0}\frac{1}{\log r}\log\mu\left(\tau_r \leq r^{-d_\mu+\epsilon}\right).$$

The choice of the normalization $\log r$ is suggested by the discrete case [33]. Moreover, this choice is strengthen by the large deviation principle for the pointwise dimension (see Corollary 2.2.7) where the normalization factor comes directly form Gartner-Ellis Theorem.

We apply our first result to conformal repellers. More precisely, given $J \subset M$ an invariant and compact set, if (J, g) is a conformal repeller and μ is an equilibrium state for a Hölder potential, we estimate large deviation rate functions which are related to HP-spectrum for dimensions.

Large deviations results are often related to multifractal analysis [47]. It turns out that in the case of conformal repellers, the multifractal spectra is degenerated [53, 27], that is

$$\dim_H\left\{x \in M : \lim_{r\to 0}\frac{\log \tau_r(x)}{-\log r} = \alpha\right\} = \dim_H M$$

for any $0 \leq \alpha \leq \infty$. It is not clear if this fact does influence large deviations for return time.

String matching problem

Along the second part of this work we will adopt the following terminology about searching and finding patterns. When the search occurs on the same string, we are talking about *repetition*. Otherwise we treat this as a *coincidence*. In what follows we present a brief discussion about this concepts and present some previous results in the literature.

Let $Y = Y_0^\infty$ be a stochastic process taking values on $\Omega = \chi^{\mathbb{N}}$, where χ is an alphabet. Consider a string $x_1^n \in \chi^n$ and a realization $y = y_1^\infty$. In view of repetition, one of the earliest studied quantities is the (Ornstein-Weiss) return time, defined in the previous subsection. Let $R_n(y)$ be the first return of a realization y to its own n-cylinder (or to its first n-string), that is, the first time that the string y_1^n recurs in the past of y. In [42], it was stated that $\frac{\log R_n(y)}{n} \longrightarrow h_\mu$ for μ-almost every realization y, where h_μ is the entropy of the measure μ.

An interesting and intuitive link between return times and the notion of data compression schemes can be found in [60]. In that paper the author present a quantity that essentially measures the smallest string on the process that did not appear in the n-sized past database of the realization. Formally

$$L_n(x) = \inf\left\{j \geq 1 : x_1^j \neq x_{-m+1}^{-m+j} \text{ , for some } 1 \leq m \leq n\right\} \ .$$

The authors then have used the duality $R_n > m \Longleftrightarrow L_m < n$ to prove that $\frac{\log n}{L_n} \longrightarrow h_\mu$ for μ-almost realization of the process. In the sequel, an entropy statistical estimator based on L_n and the proofs for its consitence were provided in [36].

The notion of coincidence has been exploited on the context *waiting times* (or *string matching*) concept (see [25, 59, 60]). Let x and y two realizations of the independent stochastic processes X and Y. The waiting time between x and y, defined as the first time that the string y_1^n appears in x is given by

$$W_n(x,y) = \inf\left\{j \geq 1 : x_j^{j+n-1} = y_1^n\right\} .$$

In [59], an exponential limiting distribution was proved for the waiting time (properly re-scaled), in the case that the measure is ψ-mixing with exponential decay of correlations.

Since most of the above mentioned quantities are typically exponentially large in the size of the cylinders, it becomes necessary to investigate some smaller-order quantity that gives an information about the process. In that sense, we get the first-return (or short-return) function of a cylinder, defined as

$$T_n(x) = \inf_{z : z_1^n = x_1^n} \tau_{x_1^n}(z),$$

where $\tau_{x_1^n}(z)$ is a hitting time to a string x_1^n of a realization z of the process that starts with the initial condition: $z_1^n = x_1^n$.

In [8] and [55], the authors used different techniques to state that $T_n/n \rightarrow 1$ almost surely when n diverges, which provides a linear feature of T_n as a function of n. The rate for this convergence was also investigated, and large deviation principles for T_n (and its relationship with the Rényi entropy) were presented in [1, 6, 31]. A weak convergence theorem for the fluctuations of T_n was presented on [3, 4].

In view of coincidence, a two-dimensional version of the short-return function was presented on [5]. It is the shortest path between two observables, and is given as follows: for two realizations x and y

$$T_n^{(2)}(x,y) = \inf_{z : z_1^n = y_1^n} W_n(x,z).$$

For independent sources, the authors proved a linear increasing of T_n^2 with respect to n, a large deviation principle for $T_n^{(2)}$ and a weak convergence for a re-scalled version of $T_n^{(2)}$.

The string matching problem is essentially motivated by biomolecular sequence comparison. The alignments of DNA and protein sequences, for example, consists of identifying common subsequences to understand evolutionary relationships. On the scenario of Erdös-Rényi laws, a remarkable matching quantity has been studied in [10]: $M_n(x,y)$, the length of the *longest matching consecutive subsequence* (or longest common substring) between two sequences. More precisely, if x and y are two realizations of the stochastic processes $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$,

$$M_n(x,y) = \max\left\{k \ : x_i^{i+k-1} = y_j^{j+k-1} \text{ for some } 0 \leq i,j \leq n-k\right\},$$

where x_i^{i+k-1} (respectively y_j^{j+k-1}) denotes the substring $x_i x_{i+1} \cdots x_{i+k-1}$ (respectively $y_j y_{j+1} \cdots y_{j+k-1}$).

If the two processes are independent and identically distributed, and generated by the same source $\mathbb{P}$, the authors proved that $M_n/(\log_{1/p} n) \longrightarrow 2$ for almost every realization (x,y), where $p = \mathbb{P}(X_0 = Y_0)$ [10]. Furthermore, if $\mathbb{P}$ defines a Markov chain, p is the largest eigenvalue of the matrix $[(p_{ij})^2]$, where $[p_{ij}]$ is its matrix. This result was recently generalized in [18] for ψ-mixing processes with polynomial decay of correlations. For another works related to matching sequences, see for example [24, 41].

Following the direction of the pattern investigation between strings, one can ask if the above mentioned results hold if we put a perturbation on the orbits. In other words: what happens if we consider encoded sequences as our interest objects of investigation?

In view of this, we study a version of the longest matching substring problem when the orbits are encoded by a measurable function (which we call *code* or *observation*, depending on the context). We call it the *longest common substring between encoded strings*. More precisely, let χ (respectively $\tilde{\chi}$) be an alphabet, $\Omega = \chi^{\mathbb{N}}$ (respectively $\tilde{\Omega} = \tilde{\chi}^{\mathbb{N}}$) the space of all sequences with symbols in χ (respectively $\tilde{\chi}$) and let $f : \Omega \to \tilde{\Omega}$ be a code. Given two sequences $x, y \in \Omega$, we define the n-length of the longest common substring for the encoded pair $(f(x), f(y))$ by

$$M_n^f(x,y) = \max\left\{k \ : f(x)_i^{i+k-1} = f(y)_j^{j+k-1} \text{ for some } 0 \leq i,j \leq n-k\right\},$$

where $f\,(x)_i^{i+k-1}$ and $f\,(y)_j^{j+k-1}$ denotes the substrings (of the encoded sequences $f(x)$ and $f(y)$) of length k beginning in $f(x)_i$ and $f(x)_j$ respectively.

Our theorem generalizes the results from the stochastic scrabble given by [9], from a Markov chain to a general α-mixing process with exponential decay. Another application deals with the zero-inflated contamination model defined in [22, 30]. In dynamical system, the correspondent of the longest common substring for the encoded pair is the shortest distance between observed orbits. If we consider a dynamical system $(M, \mathcal{A}, g, \mu)$ and an observation $f : M \to Y$, we investigate the asymptotic behavior of

$$m_n^f(x, y) = \min_{i,j=0,\dots,n-1} \left(d(f(g^i x), f(g^j y)) \right),$$

proving that its limiting behavior is related to the correlation dimension of the pushfoward measure $f_*\mu$. An application of this result is given for the shortest distance between random orbits.

Structure of the work

In the first part of this work we are interested in studying the asymptotic behavior of return times in dynamical systems. In view of this, in Chapter 1 we recall basic concepts in ergodic theory and thermodynamic formalism. We also present a construction of Markov partitions for repellers and a few useful inequalities from probability theory. Some classic results of multifractal analysis theory will also appear. We try to make the reader familiar with some concepts in large deviation theory and also present fundamental results such as Gartner-Ellis Theorem. These results will be used in Chapter 2 to obtain large deviations estimates for return times. Section 2.1 has essential definitions of rate functions in order to give a precise statement of Theorem 2.1.3 whose proof is presented in the sequel. An application for conformal repellers is given in Section 2.2: the rate functions are related with the HP-spectrum for dimensions of multifractal analysis.

In Chapter 3, for general dynamical systems, we study the shortest distance between two observed orbits (see Definition 3.1.1), that is, the orbits are encoded by a measurable function. For this case, we states a strong law of large numbers in which the limiting rate is given by the correlation dimension of the pushforward measure $f_*\mu$ (see Section 3.1). We also investigate this distance in the case of random dynamical systems, in Section 3.2, proving a similar result. In Section 3.3, we study the longest matching substring problem for encoded orbits. Under suitable mixing conditions on the source we prove a strong convergence for this quantity, and concludes that it grows logarithmically fast in n. This is in fact a law of large numbers which has Rényi's entropy as limiting-rate in the symbolic case. The rest of the chapter is dedicate to present particular examples: the zero-inflated contamination and the matching string with scores models.

Finally, in Chapter 4, we discuss about future perspectives for further scientific investigations of this subject on the context of Dynamical System and Stochastic Process.

Chapter 1

Preliminary results

In this chapter we recall some notions from ergodic theory, Markov partitions, repeller and large deviation. We also present some useful results that will be used in the proofs of our theorems.

We recall that a triple $(M, \mathcal{A}, \mu)$ is said to be a measure space if M is a space, $\mathcal{A}$ is a σ-algebra on M and μ is a measure on $(M, \mathcal{A})$.

1.1 Ergodic theory

Definition 1.1.1. *Let $(M, \mathcal{A}, \mu)$ be a measure space and let $g : M \to M$ be a measurable map. We say that μ is g-invariant or that g preserves μ if*

$$\mu(g^{-1}(A)) = \mu(A)$$

for every $A \in \mathcal{A}$.

Proposition 1.1.2. *Let $g : M \to M$ be a measurable map and μ a measure on $(M, \mathcal{A})$. Then g preserves μ if, and only if,*

$$\int \phi \; d\mu = \int \phi \circ g \; d\mu.$$

for any μ-integrable $\phi : M \to \mathbb{R}$.

For a proof we refer the reader to [16, Proposition 2.1] or [58, Proposition 1.1.1].

We remark that if M is a metric space, a version of this result is true for any continuous and limited function $\phi : M \to \mathbb{R}$.

Let us give some examples: let $M = [0, 1]$ and consider μ the Lebesgue measure. Let $g : M \to M$ be the map $x \mapsto 2x \mod 1$, called doubling map, and $g : M \to M$ defined by $g(x) = x + \alpha,\ \alpha \in M$, the rotation of angle α on the circle. All these maps preserves μ.

Definition 1.1.3. *A system $(M, \mathcal{A}, g, \mu)$ is called a measure preserving system if M is a space, $\mathcal{A}$ is a σ-algebra on M, $g : M \to M$ is a measurable map and μ is a g-invariant probability measure.*

Theorem 1.1.4 (Poincaré Recurrence Theorem)**.** *Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system. Let $A \subset M$ be a measurable set with $\mu(A) > 0$. Then, for μ-almost every $x \in A$, for infinitely many n's, $g^n(x) \in A$.*

Proof. Let A be a fixed set with $\mu(A) > 0$. Let A_∞ be the set of points of A which never come back to A. Namely,

$$A_\infty = \{x \in A : g^n(x) \notin A, \ \forall n \geq 1\}.$$

We first show that A_∞ has zero measure. We observe that $g^{-n}(A_\infty) \cap g^{-m}(A_\infty) = \emptyset$, for every $m \neq n$. Indeed, suppose that there exist $m > n \geq 1$ such that $x \in g^{-n}(A_\infty) \cap g^{-m}(A_\infty)$, thus $y = g^n(x) \in A_\infty$ and $g^{m-n}(y) = g^m(x) \in A_\infty \subset A$. This means that y come back to A, contradicting the definition of A_∞. So, we proved that these two preimages of g are disjoint. Since g is measure preserving, we have

$$\mu\left(\bigcup_{n=1}^{\infty} g^{-n}(A_\infty)\right) = \sum_{n=1}^{\infty} \mu(g^{-n}(A_\infty)) = \sum_{n=1}^{\infty} \mu(A_\infty).$$

Since μ is finite, we should have

$$\mu\left(\bigcup_{n=1}^{\infty} g^{-n}(A_\infty)\right) < \infty, \text{ then } \sum_{n=1}^{\infty} \mu(A_\infty) < \infty.$$

This last expression is an infinite sum of identic terms, thus, $\mu(A_\infty) = 0$ and the claim is proved.

Now, let F be the set of $x \in A$ that come back to A only finitely many times, formally

$$F = \{x \in A : \exists k \in \mathbb{N} \ g^n(x) \notin A, \ \forall n > k\}.$$

So, we have that every point $x \in F$ has some iterated $g^k(x)$ in A_∞. That is,

$$F \subset \bigcup_{k=0}^{\infty} g^{-k}(A_\infty).$$

Since $\mu(A_\infty) = 0$ and μ is invariant, we get:

$$\mu(F) \leq \mu\left(\bigcup_{k=0}^{\infty} g^{-k}(A_\infty)\right) \leq \sum_{k=0}^{\infty} \mu(g^{-k}(A_\infty)) = \sum_{k=0}^{\infty} \mu(A_\infty) = 0.$$

Therefore, $\mu(F) = 0$. This proves the theorem. $\square$

Definition 1.1.5. *Let $g : M \to M$ be a measurable map and μ a g-invariant probability measure. We say that μ is ergodic if for all measurable invariant set A, (i.e., $g^{-1}A = A$), either $\mu(A) = 0$ or $\mu(A) = 1$.*

The doubling map and the rotation of angle $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ on the circle are ergodic with respect to the Lebesgue measure μ.

Definition 1.1.6. *Let $g : M \to M$ be a measurable map and μ a finite g-invariant measure on M. Consider $A \subset M$ a measurable set with $\mu(A) > 0$ and a point $x \in A$. The first return time of the orbit of x to the set A is defined by*

$$\tau_A(x) = \min \{n \geq 1 : g^n x \in A\}.$$

Poincaré's recurrence theorem states that under a measure preserving system, almost every point of a measurable set A returns infinitely many times to A. However, it does not give us an estimate of the expected time for an orbit to returns to A. The following result shows that, for an ergodic measure, the mean of the return time to A is $1/\mu(A)$.

Theorem 1.1.7 (Kac's Lemma [35])**.** *Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system such that μ is ergodic. Let $A \subset M$ be a measurable set with $\mu(A) > 0$. Then,*

$$\int_A \tau_A \, d\mu = 1.$$

Equivalently, $\frac{1}{\mu(A)} \int_A \tau_A \, d\mu = \frac{1}{\mu(A)}$, i.e. the mean of the return time is inversely proportional to the measure of A.

Proof. Consider the set

$$A^*_\infty = \{x \in X : g^n(x) \notin A, \ \forall n \geq 0\}.$$

For each $n \geq 1$ we define

$$A_n = \{x \in A : \ g(x) \notin A, \ldots, g^{n-1}(x) \notin A, \text{ but } g^n(x) \in A\} \text{ and}$$

$$A^*_n = \{x \in M : \ x \notin A, g(x) \notin A, \ldots, g^{n-1}(x) \notin A, \text{ but } g^n(x) \in A\}.$$

That is, A_n is the set of points of A that return to A for the first time exactly at moment n,

$$A_n = \{x \in A : \tau_A(x) = n\}$$

and A^*_n is the set of points that are not in A which enter into A for the first time exactly at time n. These sets are measurable and then τ_A is measurable. Moreover, for each $n \geq 0$ these sets are disjoint and their union gives the space M. Hence,

$$1 = \mu(M) = \sum_{n=1}^{\infty} (\mu(A_n) + \mu(A^*_n)) + \mu(A_\infty) + \mu(A^*_\infty). \tag{1.1}$$

Since g is ergodic, almost every point of M enters in A and then $\mu(A^*_\infty) = 0$. Moreover, by the proof of Theorem 1.1.4, $\mu(A_\infty) = 0$. It is a straightforward calculation to verify that $g^{-1}(A^*_n) = A^*_{n+1} \cup A_{n+1}$ for all $n \geq 1$. Then by invariance on μ,

$$\mu(g^{-1}(A^*_n)) = \mu(A^*_n) = \mu(A^*_{n+1}) + \mu(A_{n+1}) \text{ for all } n \geq 1.$$

By applying this successively, we get

$$\mu(A^*_n) = \mu(A^*_m) + \sum_{i=n+1}^{m} \mu(A_i) \text{ for all } m > n.$$

The expression (1.1) implies that $\mu(A^*_m) \to 0$ when $m \to \infty$. Therefore, taking the limit when $m \to \infty$ in last equality we obtain

$$\mu(A^*_n) = \sum_{i=n+1}^{\infty} \mu(A_i). \tag{1.2}$$

By replacing (1.2) in (1.1) it follows that

$$1 = \mu(M) = \sum_{n=1}^{\infty}\left(\sum_{i=n}^{\infty} \mu(A_i)\right) = \sum_{n=1}^{\infty} n\mu(A_n) = \int_A \tau_A \, d\mu,$$

and this complete the proof. $\square$

In the first part of this work we will focus on studying the behavior of return time of a point to the ball. Thus, to use results that relate return times and dimension we need some conditions of asymptotic independence that are stronger than ergodicity.

Definition 1.1.8. *Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system. The correlation function for measurable observables $\psi, \phi : M \to \mathbb{R}$ is defined by*

$$C_n(\psi, \phi) = \int (\psi \circ g^n)\phi \, d\mu - \int \psi \, d\mu \int \phi \, d\mu.$$

Definition 1.1.9. *Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system. The system is mixing if we have for all $A, B \in \mathcal{A}$,*

$$\lim_{n\to\infty} \mu(g^{-n}(A) \cap B) - \mu(A)\mu(B) = 0.$$

Roughly speaking: if g is mixing, the events $g^{-n}(A)$ and B become independent as n diverge.

Notice that by changing the observables by characteristic functions in the formula of correlation function we get that the mixing definition is equivalent to $\lim_{n\to\infty} C_n(\chi_A, \chi_B) = 0$, for all $A, B \in \mathcal{A}$.

Remark 1.1.10. *Mixing implies ergodicity. In fact, suppose that there exists an invariant set $A \subset M$, with $0 < \mu(A) < 1$. Taking $B = A^c$ we obtain that $g^{-n}(A) \cap B = \emptyset$ for all n. Then, $\mu(g^{-n}(A) \cap B) = 0$ for all n. Since $\mu(A)\mu(B) > 0$ by definition we get a contradiction with*

$$\lim_{n\to\infty} \mu(g^{-n}(A) \cap B) - \mu(A)\mu(B) = 0.$$

Ergodicity is a weaker property. We observe that the doubling map is mixing but the rotation of angle $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is not mixing, and both are ergodic.

Definition 1.1.11. *Let $\Phi : \mathbb{N} \to \mathbb{R}$ be a function such that $\Phi(n) \underset{n\to\infty}{\longrightarrow} 0$. Consider V a normed vector space. The system $(M, \mathcal{A}, g, \mu)$ has decay of correlations with speed Φ if for all $\psi, \phi \in V$,*

$$\left| \int (\psi \circ g^n)\phi \, d\mu - \int \psi \, d\mu \int \phi \, d\mu \right| \leq \|\phi\|_V \|\psi\|_V \Phi(n),$$

where $\|\cdot\|_V$ is a norm on V.

We can also define a decay of correlations where $\Phi \to 0$ with some rate and V is a space of functions. Let $\mathcal{H}^\alpha(M, \mathbb{R})$ be the space of real Hölder functions on M, for some $\alpha > 0$. In Chapter 3 we will consider the rate of decay of correlations for observables $\psi, \phi \in \mathcal{H}^\alpha(M, \mathbb{R})$. If Φ has a form $\Phi(n) = a^n$ with $0 \leq a < 1$, we say that the system has an exponential decay of correlations.

We present another notion of rapid mixing.

Definition 1.1.12. *We say that $(M, \mathcal{A}, g, \mu)$ has super-polynomial decay of correlations if for all $\psi, \phi \in \mathcal{H}^\alpha(M, \mathbb{R})$, the speed Φ satisfies,*

$$\lim_{n\to\infty} \Phi(n) n^q = 0$$

for all $q > 0$.

There exists a wide class of systems that satisfy the condition of super-polynomial decay of correlations. For more details and examples about this notion we refer the reader to Section 1.2 in [49].

We introduce now briefly some notions of dimension theory.

Let (M, d) be a metric space. We define the diameter of the set $U \subset M$ by

$$\text{diam } U = \sup \{d(x, y) : x, y \in U\}.$$

Let $\mathcal{U}$ denote the collection of subsets of M. The diameter of $\mathcal{U}$ is defined by

$$\text{diam } \mathcal{U} = \sup \{\text{diam } U : U \in \mathcal{U}\}.$$

Given $Z \subset M$ and $s \in \mathbb{R}$, the s-dimensional Hausdorff measure of Z is defined by

$$m(Z, s) = \lim_{\epsilon \to 0} \inf_{\mathcal{U}} \sum_{U \in \mathcal{U}} (\text{diam } U)^s,$$

where the infimum is taken over all finite or countable covers $\mathcal{U}$ of the set Z with diam $\mathcal{U} \leq \epsilon$.

Thus, we can present the notion of Hausdorff dimension.

Definition 1.1.13. *The Hausdorff dimension of a set $Z \subset M$ is defined by*

$$\dim_H Z = \inf\{s : m(Z, s) = 0\} = \sup\{s : m(Z, s) = \infty\}.$$

Definition 1.1.14. *The Hausdorff dimension of a measure μ is defined by*

$$\dim_H \mu = \inf\{\dim_H Z : \mu(M \setminus Z) = 0\}.$$

In what follows we present another notion of dimension.

Definition 1.1.15. *The lower and upper pointwise dimensions of the measure μ at the point $x \in M$are defined by*

$$\underline{d}_\mu(x) = \varliminf_{r \to 0} \frac{\log \mu(B(x, r))}{\log r} \quad \textit{and} \quad \overline{d}_\mu(x) = \varlimsup_{r \to 0} \frac{\log \mu(B(x, r))}{\log r},$$

where $B(x, r)$ is the ball of radius r centered at x.

If there exists a constant d_μ such that

$$\underline{d}_\mu(x) = \overline{d}_\mu(x) = d_\mu \quad \text{for} \quad \mu\text{-almost every} \quad x \in M,$$

we call μ exact dimensional. And d_μ is called pointwise dimensions of the measure μ.

For an exact dimensional measure, the Hausdorff dimension and the local dimension coincide. Young established the following criterion, which we start without proof:

Proposition 1.1.16 ([61])**.** *If μ is exact dimensional, then*

$$d_\mu = \dim_H \mu.$$

We now present results that relate quantitative recurrence and dimension. Firstly let us state a key concept.

The first return time of a point $x \in M$ to the ball $B(x, r)$ is given by

$$\tau_r(x) = \min\{n \geq 1 : d(g^n x, x) < r\}.$$

Definition 1.1.17. *The lower and upper recurrence rates of x are defined by*

$$\underline{R}(x) = \varliminf_{r \to 0} \frac{\log \tau_r(x)}{-\log r} \quad \textit{and} \quad \overline{R}(x) = \varlimsup_{r \to 0} \frac{\log \tau_r(x)}{-\log r}.$$

When $\underline{R}(x) = \overline{R}(x)$ we denote the common value by $R(x)$.

Barreira and Saussol showed in [17] a result that relates these quantities with lower and upper pointwise dimensions.

Theorem 1.1.18 ([17])**.** *Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system. Set $M \subset \mathbb{R}^d$ for some $d \in \mathbb{N}$. Then,*

$$\underline{R}(x) \leq \underline{d}(x) \quad \text{and} \quad \overline{R}(x) \leq \overline{d}(x)$$

for μ-almost every $x \in M$.

The authors also showed that these inequalities becomes equalities when the measure μ has a condition called long return time.

Saussol in [51] extends the previous theorem for a class of systems such that the map g is Lipschitz, with positive entropy and super-polynomial decay of correlation. Namely,

Theorem 1.1.19 ([51])**.** *Let (M, g, μ) be a measure preserving system. If the entropy $h_\mu(g) > 0$, g is Lipschitz and the decay of correlation is super-polynomial then*

$$\underline{R}(x) = \underline{d}(x) \quad \text{and} \quad \overline{R}(x) = \overline{d}(x)$$

for μ-almost every $x \in M$.

One notice that in the case that μ is exact dimensional this theorem implies that

$$\log \tau_r(x) \underset{r \to 0}{\sim} \log \left(r^{-d_\mu(x)} \right).$$

The remainder of this section is dedicate to present the definitions of entropy for a continuous map of a compact metric space, and pressure.

Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving system. Let $\mathcal{P}$ be a measurable partition of M, that is, a collection of pairwise disjoint measurable sets whose union is M. Denote by $\mathcal{P}(x)$ the partition element that contains a point x.

We define the entropy of $\mathcal{P}$ as

$$H_\mu(\mathcal{P}) = -\sum_{P \in \mathcal{P}} \mu(P) \log \mu(P).$$

Given a partition $\mathcal{P}$ of X with finite entropy, we denote

$$\mathcal{P}^n = \bigvee_{i=0}^{n-1} g^{-i}(\mathcal{P}) \text{ for any } n \geq 1.$$

The element $\mathcal{P}^n(x)$ that contains x is given by

$$\mathcal{P}^n(x) = \mathcal{P}(x) \cap g^{-1}(\mathcal{P}(g(x))) \cap \cdots \cap g^{-n+1}(\mathcal{P}(g^{n-1}(x))).$$

We define the entropy of g with respect to μ and the partition $\mathcal{P}$ as

$$h_\mu(g,\mathcal{P}) = \lim_n \frac{1}{n} H_\mu(\mathcal{P}^n) = \inf_n \frac{1}{n} H_\mu(\mathcal{P}^n).$$

Finally the entropy of the system (g,μ) is defined by

$$h_\mu(g) = \sup_{\mathcal{P}} h_\mu(g,\mathcal{P}),$$

where the supremum is taken over all partitions with finite entropy.

The notion of pressure was established by Ruelle and extended by Walters. The variational principle says that for all continuous function φ,

$$P(\varphi) = \sup_\mu \left(h_\mu(g) + \int \varphi \, d\mu \right),$$

where the supremum is taken over all g-invariant probability measures μ in M. A g-invariant probability measure μ is called an equilibrium measure for φ if

$$P(\varphi) = h_\mu(g) + \int \varphi \, d\mu.$$

Now we give a notion of cohomology in dynamical systems.

Let $S : M \to M$ be a continuous map of a topological space M. Two continuous functions $\varphi_1 : M \to \mathbb{R}$ and $\varphi_2 : M \to \mathbb{R}$ are said to be cohomologous to a constant if there exists a continuous function $\phi : M \to \mathbb{R}$ and a constant $c \in \mathbb{R}$ such that

$$\varphi_1 - \varphi_2 = \phi - \phi \circ S + c.$$

1.2 Markov partition and repeller

Let $g : M \to M$ be a $C^{1+\alpha}$ map of a smooth manifold and consider a g-invariant compact set $J \subset M$. The map g is said to be expanding on J if there exist constants $c > 0$ and $\rho > 1$ such that

$$\|d_x g^n v\| \geq c\rho^n \|v\|$$

for every $n \in \mathbb{N}$, $x \in J$ and $v \in T_x M$. In addition, we call J a repeller if there exists an open neighborhood V of J such that

$$J = \bigcap_{n \geq 0} g^{-n} V.$$

The map g is said to be conformal on J if

$$d_x g = a(x) Isom_x,$$

where $Isom_x$ denotes an isometry of the tangent space $T_x M$.

Definition 1.2.1. *We say that (J, g) is expansive when there exists $\delta > 0$ such that for any $x, y \in J$,*

$$\text{if for all } n \geq 0 \text{ we have } d(g^n(x), g^n(y)) < \delta \text{ then } x = y.$$

We call δ an expansiveness constant of g.

Remark 1.2.2. *It is possible to show that all repellers are expansive.*

Definition 1.2.3. *Given $\alpha > 0$, a sequence $(x_n)_{n\geq 0}$ is called α-pseudo-orbit of (J, g) if*

$$d(g(x_n), x_{n+1}) < \alpha, \text{ for all } n \geq 0.$$

We call a sequence $x_0, x_1, \ldots, x_{m-1}, x_m = x_0$ an α-periodic orbit if $d(g(x_n), x_{n+1}) < \alpha$. A particular case of an α-periodic orbit is provided by $x_0, g(x_0), \ldots, g^{m-1}(x_0)$ such that $d(g^m(x_0), x_0) < \alpha$.

We now present the shadowing property. The proof due to Saussol [54].

Proposition 1.2.4 (Shadowing lemma). *If (J, g) is a repeller then for every $\beta > 0$ there exists $\alpha > 0$ such that given an α-pseudo-orbit $(x_n)_{n\geq 0}$ in J there exists $z \in J$ such that its orbit β-shadows $(x_n)_{n\geq 0}$, that is, $d(g^n(z), x_n) < \beta$ for all $n \geq 0$. If β is less than half of an expansive constant of g then the point z is unique. Moreover, if the pseudo-orbit is periodic, then the orbit of z is periodic.*

Proof. Since g is C^1 and expanding on a neighborhood V of J, it is a local diffeomorphism. By compacity there exists $\epsilon > 0$ such that for all $x \in J$, $g : B(x, 2\epsilon) \to g(B(x, 2\epsilon))$ is an expanding diffeomorphism. In particular $g(B(x, 2\epsilon)) \supset B(g(x), 2\epsilon)$ and for all $x \in J$, the branch of the inverse $g_x^{-1} : B(g(x), 2\epsilon) \to B(x, 2\epsilon)$ is well defined. Without loss of generality we will assume that $B(x, 2\epsilon) \subset J$ for all $x \in J$ and that $\beta < \epsilon$. Let $\alpha \in (0, \epsilon)$ be such that $\beta = \frac{\alpha}{1-\rho}$.

If the pseudo-orbit is infinite then for all $p > 0$ we can make the following construction that gives a z^p which is β-shadowed by $x_0, \ldots, x_p$. Let us put $z_p = x_p$. We will define by induction $(z_j)_{j\leq p}$. Put $r_j = d(z_j, x_j)$. We have $r_p = 0 < \epsilon$. Suppose we have defined $z_p, \ldots, z_{j+1}$ and that $r_{j+1} < \epsilon$. Then

$$d(g(x_j), z_{j+1}) \leq d(g(x_j), x_{j+1}) + d(x_{j+1}, z_{j+1}) \leq \alpha + r_{j+1} < 2\epsilon.$$

Therefore the preimage $z_j := g_{x_j}^{-1} z_{j+1}$ is well defined. Moreover,

$$r_j = d(z_j, x_j) \leq \rho d(g(z_j), g(x_j)) \leq \rho(\alpha + r_{j+1}).$$

By an immediate recurrence we get

$$r_j < (\rho + \rho^2 + \cdots + \rho^{p-j})\alpha = \frac{\alpha}{1-\rho} < \epsilon$$

for all $j \leq p$ and hence the sequence $(z_j)_j$ is well defined. The point

$$z^p := z_0 = g_{x_0}^{-1} \circ g_{x_1}^{-1} \circ \cdots \circ g_{x_{p-1}}^{-1}(x_p)$$

verifies the conditions presented.

Let z be an accumulation point of z^p, which exists because the ball $\overline{B(x_0, \epsilon)}$ is compact. Let $n \leq 0$. For all $p \leq n$ we have

$$d(g^n(z), x_n) \leq d(g^n(z), g^n(z^p)) + d(g^n(z^p), x_n) \leq \beta + d(g^n(z), g^n(z^p)).$$

By continuity of g^n one obtain, taking the limit $p \to \infty$, that $d(g^n(z), x_n) \leq \beta$. So the orbit of z is shadowed by the infinite orbit $x_0, x_1, \ldots$ and since $\beta < \epsilon$ we have $g^n(z) \in V$ for all n, i.e. $z \in J$. If the pseudo-orbit is finite, it is enough to apply the previous part to the infinite pseudo-orbit. The remaining statements are simple consequences. If z' is another point satisfying the conclusion of the proposition then

$$d(g^n(z), g^n(z')) \leq d(g^n(z), x_n) + d(x_n, g^n(z')) < 2\beta \text{ for all } \; n \geq 0.$$

By expansiveness, it follows that $z = z'$. Finally, if the pseudo-orbit is periodic, with period $k \geq 1$, we also have

$$d(g^n(g^k(z)), x_n) \leq d(g^{n+k}(z), x_{n+k}) < \beta \text{ for all } \; n \geq 0.$$

By uniqueness, we obtain that $g^k(z) = z$. □

It is important to note that the proof of Proposition 1.2.4 shows us that we can take $\alpha = c_1\beta$, where $c_1 > 0$ depends only of ρ.

Theorem 1.2.5 (Closing lemma)**.** *If (J, g) is a repeller then for all r, k, x such that $d(g^k(x), x) < r$ there exists a point z with $g^k(z) = z$ and $d(x, z) < c_1 r, c_1 > 0$.*

Proof. The proof follows immediately from Proposition 1.2.4. □

Assume that g is topologically mixing in J, that is, for all A, B open sets of M there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0, \; A \cap g^{-n}(B) \neq \emptyset$.

Definition 1.2.6. *Let J be a repeller of the map g. A collection of closed sets $\mathcal{J} = \{J_1, \ldots, J_k\}$ is called a Markov partition of J (with respect to g) if:*

1. $J = \cup_1^k J_i$ and $J_i = \overline{\text{int } J_i}$ for each i;

2. $\text{int } J_i \cap \text{int } J_j = \emptyset$ whenever $i \neq j$;

3. $g(J_i) \cap \text{int } J_j \neq \emptyset$, then $g(J_i) \supset J_j$.

Example 1.2.7. *The collection* $\mathcal{J} = \left\{\left[0, \frac{1}{2}\right], \left[\frac{1}{2}, 1\right]\right\}$ *is a natural Markov partition of the doubling map.*

Fix a Markov partition $\mathcal{J}$ and consider the $k \times k$ matrix $A = (a_{ij})$ with entries

$$a_{ij} = \begin{cases} 1 & if\ g(J_i) \cap \operatorname{int} J_j \neq \emptyset, \\ 0 & otherwise. \end{cases}$$

Let $\mathcal{A} = \{1, \ldots, k\}$ and $\Sigma_A^+ \subset \mathcal{A}^{\mathbb{N}}$ the set of sequences defined by

$$\Sigma_A^+ = \left\{\omega = (\omega_i)_{i \geq 0} : a_{\omega_i \omega_{i+1}} = 1 \text{ for every } i \in \mathbb{N}\right\}.$$

Consider $\sigma : \Sigma_A^+ \to \Sigma_A^+$ the shift map defined by $\sigma(\omega)_i = \omega_{i+1}$ for every $i \in \mathbb{N}$. This define the symbolic coding $\chi : \Sigma_A^+ \to J$ such that

$$\chi(\omega) = \bigcap_{i=0}^{\infty} g^{-i} J_{\omega_i}$$

and

$$\chi \circ \sigma = g \circ \chi.$$

The map χ is Hölder continuous and injective except on the set $S = \bigcup_{n=0}^{\infty} g^{-n} \partial \mathcal{J}$, where $\partial \mathcal{J} = \bigcup_i \partial J_i$.

For $\omega \in \Sigma_A^+$ we denote by $C_n(\omega) = \left\{\omega' \in \Sigma_A^+ : \omega_i' = \omega_i \text{ for all } 0 \leq i \leq n-1\right\}$ the n-cylinder containing ω. We set $\mathcal{J}_n(x) = \chi(C_n(\omega))$ when $x = \chi(\omega) \notin S$.

We can now define the class of Gibbs measure.

Recall that $S_n(\varphi) = \sum_{k=0}^{n-1} \varphi\left(g^k(x)\right)$.

Definition 1.2.8. *Let* φ *be a Hölder function and* μ *be a* g*-invariant probability measure. We say that* μ *is a Gibbs measure for the potential* φ *if there exists a constant* $P(\varphi) \in \mathbb{R}$ *such that for some* $\kappa_\varphi \geq 1$, *for any* x *and* n, *the following holds:*

$$\frac{1}{\kappa_\varphi} \leq \frac{\mu(\mathcal{J}_n(x))}{exp(S_n\varphi(x) - nP(\varphi))} \leq \kappa_\varphi.$$

Let ζ be a Hölder continuous function on J and $\mu = \mu_\zeta$ be the equilibrium measure for (g, ζ). Let $\nu = \nu_\varphi$ be the Gibbs measure of the Hölder potential $\varphi = \zeta \circ \chi$ on Σ_A^+. Note that $\mu = \chi_* \nu$. Finally, consider the function ψ such that $\log \psi = \varphi - P(\varphi)$. ψ is a Hölder continuous function on M such that $P(\log \psi) = 0$ and ν is a unique equilibrium measure for $\log \psi$.

We collect some facts about a notion of dimension denominated HP-spectrum for dimensions, that was introduced by Hentschel and Procaccia in [32].

The following result was proved by Pesin and Weiss in [47] (see Theorem 1.(2) and Lemma 5).

Proposition 1.2.9 ([47]). *For all $q \in \mathbb{R}$, the following limit exists*

$$T(q) = \lim_{r \to 0} \frac{\log \int_J \mu(B(x,r))^{q-1} d\mu(x)}{-\log r}. \tag{1.3}$$

In addition, the function $T(q)$ is real analytic for all $q \in \mathbb{R}$, $T(0) = \dim_H J$, $T(1) = 0$, $T'(q) \leq 0$ and $T''(q) \geq 0$. And $T''(q) > 0$ if and only if the function $\log \psi - T'(q) \log |a(\chi(w))|$ is not cohomologous to a constant, if and only if μ is not a measure of maximal dimension.

Remark 1.2.10. *Given $q \in (-\infty, \infty)$, define ϕ_q on Σ_A^+ the one parameter family of functions by*

$$\phi_q(w) = -T(q) \log |a(\chi(w))| + q \log \psi(w).$$

The function $T(q)$ is chosen such that $P(\phi_q) = 0$. Moreover, for any $q > 1$,

$$\frac{T(q)}{1-q} = HP_\mu(q). \tag{1.4}$$

Note that μ is exact dimensional, see for instance [46, Theorem 9].

Theorem 1.2.11 (Dimension of repellers of conformal maps)**.** *If (J, g) is a conformal repeller then*

$$dim_H J = s,$$

where s is the unique real number such that $P(s\varphi) = 0$, for the function $\varphi : J \to \mathbb{R}$ defined by $\varphi(x) = -\log \|d_x g\|$.

Proof. See e.g. Section 4.1 in [15]. □

Remark 1.2.12. *Ruelle in [50] showed that a conformal repeller (J, g) such that g is topologically mixing satisfies $\dim_H J = \dim_H \mu$. In addition, the equilibrium measure μ of $s\varphi$ is equivalent to the s-dimensional Hausdorff measure m. The equilibrium measure μ is called the measure of maximal dimension.*

We will now introduce another notion of dimension on dynamical system that is related to invariant ergodic measures.

For $q = 2$, the formula (1.4) coincides with the *correlation dimension* of the measure μ (see Section 17 in [44]). For simplicity of notation, we write C_μ instead of $HP_\mu(2)$, that is,

$$C_\mu = \lim_{r \to 0} \frac{\log \int_M \mu(B(x,r)) \; d\mu(x)}{\log r}. \tag{1.5}$$

If μ is ergodic, Pesin and Tempelman [45] showed that for all $q > 1$ this limit exists.

Note that the limit (1.5) depends on the metric on M and on the invariant measure but does not depend on the map.

The *lower and upper correlation dimension* of μ are denoted, respectively as

$$\underline{C}_\mu = \varliminf_{r \to 0} \frac{\log \int_M \mu(B(x,r)) \; d\mu(x)}{\log r} \quad \text{and} \quad \overline{C}_\mu = \varlimsup_{r \to 0} \frac{\log \int_M \mu(B(x,r)) \; d\mu(x)}{\log r}.$$

1.3 Probability results

In what follows we present inequalities in probability theory that will provide bounding quantities throughout this work.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is the sample space, $\mathcal{F}$ is the event space and $\mathbb{P} : \mathcal{F} \to [0,1]$ is a function that assigns probabilities to events.

Definition 1.3.1. *A real valued function X defined on Ω is said to be a random variable if for every Borel set $B \subset \mathbb{R}$ we have $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$.*

Theorem 1.3.2 (Markov's inequality)**.** *Let X be a non-negative random variable and suppose that $\mathbb{E}[X]$ exists. Therefore, for any $t > 0$,*

$$\mathbb{P}(X > t) < \frac{\mathbb{E}[X]}{t}.$$

Theorem 1.3.3 (Chebyshev's inequality)**.** *Let $\mu = \mathbb{E}[X]$ and $\sigma^2 = Var[X]$. Then,*

$$\mathbb{P}(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2} \quad \textit{and} \quad \mathbb{P}(|Z| \geq k) \leq \frac{1}{k^2}$$

where $Z = (X - \mu)/\sigma$.

The next theorem is a classical result that establishes if certain events occur infinitely often or only finitely often.

Theorem 1.3.4 (Borel-Cantelli's lemma)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and consider a sequence $A_n \in \mathcal{F},\ n \geq 1$.*

(i) *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then*

$$\mathbb{P}(\{x : x \in A_n \ \ \textit{for infinitely many } n\}) = 0.$$

(ii) *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$ and the A_n's are independent, then*

$$\mathbb{P}(\{x : x \in A_n \ \ \textit{for infinitely many } n\}) = 1.$$

Proof. A proof can be found in [34]. □

1.4 The large deviation principle

In this section we introduce a large deviation lower and upper bound that characterizes the limiting behavior of a family of probability measures in terms of the logarithmic moment generating function. This approach is due to Dembo and Zeitouni [26] who concerns this to the study of rare events and its relation with large deviation theory.

1.4.1 Basic definitions and properties

Throughout this section M denotes a topological space and $\mathcal{B}$ the Borel σ-algebra on M.

Definition 1.4.1. *The function $I : M \to [0, \infty)$ is a rate function if it is lower continuous (i.e., for all $\alpha \in [0, \infty)$ the level set $\{x : I(x) \leq \alpha\}$ is a closed subset of M). If the level sets are compacts subsets of M, I is called a good rate function.*

Definition 1.4.2. *Let $(\mu_n)_{n\geq 0}$ be a family of probability measures on $(M, \mathcal{B})$. W say that $(\mu_n)_{n\geq 0}$ satisfies the large deviation principle (LDP) with a rate function I if:*

1. *for any closed set $F \subset M$,*
$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} I(x);$$
2. *for any open set $G \subset M$,*
$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x\in G} I(x).$$

Consider a sequence $(X_j)_{j\geq 0}$ of d-dimensional random vectors independent and identically distributed (i.i.d.) according to the probability law $\mu \in \mathcal{M}_1(\mathbb{R}^d)$ and let the sequence of empirical means $S_n := \frac{1}{n}\sum_{j=1}^{n} X_j$.

Denote by μ_n the law of S_n and $\overline{x} = \mathbb{E}[X_1]$ and assume that $\overline{x}$ exists and is finite. From the classical theory of probability we have two results: the law of large numbers and the central limit theorem. The law of large numbers states that S_n converges to $\overline{x}$ almost surely when n goes to infinity. If in addition $\sigma^2 = Var[X_1]$ is finite, the central limit theorem states that $\frac{\sqrt{n}(S_n - \overline{x})}{\sigma}$ converges to the normal distribution. With large deviations techniques it is possible to estimate the rate at which $\mu_n(S_n > a)$ converges to zero for $a > \overline{x}$.

Definition 1.4.3. *The logarithmic moment generating function associated with the law μ is defined as*

$$\Lambda(\lambda) = \log M_X(\lambda) = \log \mathbb{E}\left[e^{\langle \lambda, X_1\rangle}\right], \tag{1.6}$$

where $\langle \lambda, x\rangle = \sum_{j=1}^{d} \lambda_j x_j$ is the usual scalar product in $\mathbb{R}^d$.

Definition 1.4.4. *The Fenchel-Legendre transform of Λ is defined by*

$$\Lambda^*(x) := \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, x\rangle - \Lambda(\lambda)\}.$$

Let us define $D_\lambda = \{\lambda : \Lambda(\lambda) < \infty\}$ and $D_{\Lambda^*} = \{x : \Lambda^*(x) < \infty\}$ the domain of Λ and Λ^*, respectively.

We consider the first result for random variables taking values in $\mathbb{R}$.

Theorem 1.4.5 (Cramér)**.** *The sequence of measures $(\mu_n)_{n\geq 0}$ satisfies the LDP with the convex rate function $\Lambda^*(\cdot)$, that is:*

(a) *For any closed set $F \subset \mathbb{R}$,*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \Lambda^*(x).$$

(b) *For any open set $G \subset \mathbb{R}$,*

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x\in G} \Lambda^*(x).$$

Note that this theorem is a LDP with convex rate function $\Lambda^*(\cdot)$. Moreover it is a result limited to the i.i.d. case.

The following lemma presents some properties of $\Lambda^*(\cdot)$ and $\Lambda(\cdot)$.

Lemma 1.4.6. (a) *Λ is a convex function and Λ^* is a convex rate function.*

(b) *$\Lambda(\cdot)$ is differentiable on D_Λ^o with*

$$\Lambda^{'}(\eta) = \frac{1}{M_X(\eta)} \mathbb{E}\left[X_1 e^{\eta X_1}\right] \tag{1.7}$$

and

$$\Lambda^{'}(\eta) = y \Rightarrow \Lambda^*(y) = \eta y - \Lambda(\eta). \tag{1.8}$$

Proof. **(a)** Given $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\theta \in [0,1]$, applying Hölder's inequality for the conjugate exponents $\dfrac{1}{\theta}$ and $\dfrac{1}{1-\theta}$ we get,

$$\begin{aligned}
\Lambda(\theta\lambda_1 + (1-\theta)\lambda_2) &= \log \mathbb{E}\left[e^{[\theta\lambda_1+(1-\theta)\lambda_2]X_1}\right] \\
&= \log \mathbb{E}\left[\left(e^{\lambda_1 X_1}\right)^\theta \left(e^{\lambda_2 X_1}\right)^{(1-\theta)}\right] \\
&\leq \log\left\{\left[\mathbb{E}\left[e^{\lambda_1 X_1}\right]\right]^\theta \left[\mathbb{E}\left[e^{\lambda_2 X_1}\right]\right]^{(1-\theta)}\right\} \\
&= \log\left[\mathbb{E}\left[e^{\lambda_1 X_1}\right]\right]^\theta + \log\left[\mathbb{E}\left[e^{\lambda_2 X_1}\right]\right]^{(1-\theta)} \\
&= \theta\Lambda(\lambda_1) + (1-\theta)\Lambda(\lambda_2),
\end{aligned}$$

which proves convexity.

The convexity of Λ^* follows by definition,

$$\begin{aligned}
\theta\Lambda^*(x_1) + (1-\theta)\Lambda^*(x_2) &= \theta \sup_{\lambda\in\mathbb{R}}\{\lambda x_1 - \Lambda(\lambda)\} + (1-\theta)\sup_{\lambda\in\mathbb{R}}\{\lambda x_2 - (1-\theta)\Lambda(\lambda))\} \\
&= \sup_{\lambda\in\mathbb{R}}\{\theta\lambda x_1 - \theta\Lambda(\lambda)\} + \sup_{\lambda\in\mathbb{R}}\{(1-\theta)\lambda x_2 - \Lambda(\lambda))\} \\
&\geq \sup_{\lambda\in\mathbb{R}}\{(\theta x_1 + (1-\theta)x_2)\lambda + \Lambda(\lambda)\} \\
&= \Lambda^*(\theta x_1 + (1-\theta)x_2).
\end{aligned}$$

By definition $\Lambda(0) = 0$, so $\Lambda^*(x) = \sup_{\lambda\leq 0}[\lambda x - \Lambda(\lambda)] \geq 0x - \Lambda(0) = 0$ is nonnegative. Now, fix a sequence $x_n \to x$. Then, for every $\lambda \in \mathbb{R}$,

$$\underline{\lim}_{x_n\to x} \Lambda^*(x_n) \geq \underline{\lim}_{x_n\to x}[\lambda x_n - \Lambda(\lambda)] = \lambda x - \Lambda(\lambda).$$

Thus,

$$\underline{\lim}_{x_n\to x} \Lambda^*(x_n) \geq \sup_{\lambda\in\mathbb{R}}[\lambda x_n - \Lambda(\lambda)] = \Lambda^*(x).$$

And this proves that Λ^* is lower semicontinuous.

(b) The identity (1.7) follows using the dominated convergence theorem, since $f_\epsilon(x) = \left(e^{(\eta+\epsilon x)} - e^{\eta x}\right)/\epsilon$ converges pointwise to $xe^{\eta x}$ as $\epsilon \to 0$, and $|f_\epsilon(x)| \leq e^{\eta x}\left(e^{\delta|x|-1}\right)/\delta := h(x)$ for every $\epsilon \in (-\delta,\delta)$, while $\mathbb{E}[|h(X_1)|] < \infty$ for $\delta > 0$ small enough. By convexity of $\Lambda(\lambda)$,

$$\Lambda'(\eta)(\lambda - \eta) + \lambda(\eta) \leq \Lambda(\lambda)$$

implies

$$y(\lambda - \eta) + \Lambda(\eta) \leq \Lambda(\lambda).$$

Therefore (1.8) is established.

□

Lemma 1.4.7. *Let $\{a_1(n),\ldots,a_N(n)\}$ be a collection of N sequences. Then, for every $a_i(n) \geq 0$,*

$$\overline{\lim}_{n\to\infty}\frac{1}{n}\log\left(\sum_{i=1}^N a_i(n)\right) = \max_{i=1,\ldots,N}\overline{\lim}_{n\to\infty}\frac{1}{n}\log a_i(n). \tag{1.9}$$

Proof. We observe that

$$\max_{i=1,\ldots,N} a_i(n) \leq \sum_{i=1}^N a_i(n) \leq N\max_{i=1,\ldots,N} a_i(n).$$

Since the max is being taken over finitely many terms, $\frac{1}{n}\log N \to 0$ as $n \to \infty$ and

$$\overline{\lim}_{n\to\infty}\frac{1}{n}\log\left(\max_{i=1,\ldots,N} a_i(n)\right) = \max_{i=1,\ldots,N}\overline{\lim}_{n\to\infty}\frac{1}{n}\log a_i(n).$$

This concludes the proof.

□

1.4.2 Gartner-Ellis theorem

Consider a family of random vectors $(Z_n)_n$ in $\mathbb{R}^d$, which will play a role as the empirical mean S_n in the i.i.d. case. Consider the logarithmic moment generating function

$$\Lambda_n(\lambda) = \log \mathbb{E}\left[e^{\lambda Z_n}\right]. \tag{1.10}$$

The family $(\mu_n)_{n\geq 0}$ may satisfy the large deviation property if there exists a limit of properly scaled logarithmic moment generating functions.

Assumption 1.4.8. *For each $\lambda \in \mathbb{R}^d$, the logarithmic moment generating function defined as the limit*

$$\Lambda(\lambda) := \lim_{n\to\infty} \frac{1}{n}\Lambda_n(n\lambda)$$

exists in $\overline{\mathbb{R}}$. Furthermore, the origin belongs to the interior of the set $D_\lambda := \{\lambda \in \mathbb{R}^d : \Lambda(\lambda) < \infty\}$, Λ is C^2 and a strictly convex function.

In particular, if μ_n is the law of S_n, then for all $n \in \mathbb{Z}_+$,

$$\frac{1}{n}\Lambda_n(n\lambda) = \Lambda(\lambda) := \log \mathbb{E}\left[e^{\langle\lambda, X_1\rangle}\right],$$

and the assumption above holds when $0 \in D_\Lambda^o$.

More general, one can prove that $\Lambda(\cdot)$ is strictly convex and then $\Lambda^*(\cdot)$ is also strictly convex.

In fact, computing the second derivative we get

$$\Lambda''(\lambda) = \frac{\mathbb{E}\left[X_1^2 e^{\eta X_1}\right]}{M(\eta)} - \left(\frac{\mathbb{E}\left[X_1^2 e^{\eta X_1}\right]}{M(\eta)}\right)^2 > 0.$$

Note that $\Lambda(0) = \mathrm{var}(X_1)$. Assume that $\mathrm{var}(X_1) > 0$.

Lemma 1.4.9. *Λ^* is strictly convex function and C^1 on its support.*

Proof. For simplicity of the proof we will consider $d = 1$.

By (1.8) we have that $\Lambda^*(\Lambda'(\eta)) = \eta\Lambda'(\eta) - \Lambda(\eta)$. Thus

$$(\Lambda^*)'(\Lambda'(\eta)) = (\Lambda^*)'(\Lambda'(\eta))\Lambda''(\eta) = \Lambda'(\eta) + \eta\Lambda''(\eta) - \Lambda''(\eta) = \Lambda''(\eta).$$

Then,

$$(\Lambda^*)'(\Lambda'(\eta)) = \eta.$$

Now,

$$(\Lambda^*)''(\Lambda'(\eta)) = (\Lambda^*)'(\Lambda'(\eta))\Lambda''(\eta) = 1.$$

Therefore,

$$(\Lambda^*)'(\Lambda'(\eta)) = \frac{1}{\Lambda''(\eta)} > 0,$$

since Λ is strictly convex. Thus, $\Lambda^*(y)$ is also strictly convex. $\square$

Definition 1.4.10. *Suppose that all compact subsets of M belong to $\mathcal{B}$. A family of probability measures (μ_n) on M is exponentially tight if for every $\alpha < \infty$, there exists a compact set $K_\alpha \subset M$ such that*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K_\alpha^c) < -\alpha.$$

Definition 1.4.11. *We say that $y \in \mathbb{R}^d$ is an exposed points of Λ^* if for some $\Lambda \in \mathbb{R}^d$ and for all $x \neq y$,*

$$\langle \lambda, y \rangle - \Lambda^*(y) > \langle \lambda, x \rangle - \Lambda^*(x). \tag{1.11}$$

The vector λ is called an exposing hyperplane.

Definition 1.4.12. *A convex function $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ is essentially smooth if*

1. *D_Λ^o is non empty;*
2. *$\Lambda(\cdot)$ is differentiable on D_Λ^o;*
3. *$\Lambda(\cdot)$ is steep, i.e., if (λ_n) is a sequence on D_Λ^o converging to a boundary point of D_Λ^o then $\lim_{n\to\infty} |\nabla\Lambda(\lambda_n)| = \infty$.*

We also need two auxiliary lemmas that presents the elementary properties of Λ and Λ^*.

Lemma 1.4.13. *Let Assumption 1.4.8 hold. Then,*

(a) *$\Lambda(\lambda)$ is a convex function, $\Lambda(\lambda) > -\infty$ everywhere, and $\Lambda^*(x)$ is a good convex rate function.*

(b) *Suppose that $y = \nabla\Lambda(\eta)$ for some $\eta \in D_\Lambda^o$. Then $\lambda^*(y) = \langle \eta, y \rangle - \Lambda(\eta)$.*

Moreover, $y \in \mathcal{F}$, with η being the exposing hyperplane for y.

For every non empty convex set C, the *relative interior* of C, denoted by ri C, is defined as the set

$$\mathrm{ri} C = \{y \in C : x \in C \Rightarrow y - \epsilon(x - y) \in C \text{ for some } \epsilon > 0\}.$$

Lemma 1.4.14 (Rockafellar)**.** *If $\Lambda : \mathbb{R}^d \to (-\infty, \infty]$ is an essentially smooth, lower semicontinuous, convex function, then ri $D_{\Lambda^*} \subseteq \mathcal{F}$.*

Theorem 1.4.15 (Gartner-Ellis)**.** *Suppose that the Assumption 1.4.8 holds. Then,*

(a) *For any closed set F,*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \mu_n(F) \leq -\inf_{x\in F} \Lambda^*(x). \tag{1.12}$$

(b) *For any open set G,*

$$\varliminf_{n\to\infty} \frac{1}{n}\mu_n(G) \geq -\inf_{x\in G\cap\mathcal{F}} \Lambda^*(x) \tag{1.13}$$

where $\mathcal{F}$ is the set of exposing points of Λ^* whose exposing hyperplane belongs to D_Λ^o.

(c) *If Λ is an essentially smooth, lower semicontinuous function, then the LDP holds with the good rate function $\Lambda^*(\cdot)$.*

Proof. **(a)** The upper bound for compact sets is established by the same argument from the proof of the theorem of Cramér-$\mathbb{R}^d$ (see Section 2.2 in [26] for details). The extension to all closed sets follows by proving that the sequence of measures (μ_n) is exponentially tight. For that, let μ_j denote the j-th unique vector in $\mathbb{R}^d$ for $j = 1, \ldots, d$. Since $0 \in D_\Lambda^o$, there exist $\theta_j > 0$, $\eta_j > 0$ such that $\Lambda(\theta_j u_j) < \infty$ and $\Lambda(-\eta_j u_j) < \infty$ for $j = 1, \ldots, d$. Then, by Chebycheff's inequality,

$$\mu_n^j((-\infty, -\rho]) \leq e^{-n\eta_j\rho+\Lambda_n(-n\eta_j u_j)} \text{ and}$$

$$\mu_n^j([\rho, \infty)) \leq e^{-n\theta_j\rho+\Lambda_n(n\theta_j u_j)},$$

$j = 1, \ldots, d$, where μ_n^j are the laws of the coordinates of the random vector Z_n. Thus, for $j = 1, \ldots, d$,

$$\lim_{\rho\to\infty} \varlimsup_{n\to\infty} \frac{1}{n} \log \mu((-\infty, -\rho]) = -\infty,$$

$$\lim_{\rho\to\infty} \varlimsup_{n\to\infty} \frac{1}{n} \log \mu([\rho, \infty)) = -\infty.$$

Consequently, combining these limits with Lemma1.4.7, we get

$$\lim_{\rho\to\infty} \varlimsup_{n\to\infty} \frac{1}{n} \log \mu(([-\rho, \rho]^d)^c) = -\infty$$

i.e., (μ_n) is an exponentially tight sequence of probability measures.

(b) To establish the lower bound for any open set, it is sufficient to prove that for every $y \in \mathcal{F}$,

$$\lim_{\delta\to\infty} \varliminf_{n\to\infty} \log \mu_n(B(y, \delta)) \geq -\Lambda^*(y). \tag{1.14}$$

Fix $y \in \mathcal{F}$ and let $\eta \in D_\Lambda^o$ an exposing hyperplane for y. Then for n large enough, $\Lambda_n(n\eta) < \infty$ and the probability measures $\widetilde{\mu}_n$ are well defined via,

$$\frac{d\widetilde{\mu}_n}{d\mu_n}(z) = e^{n\langle\eta,z\rangle-\Lambda_n(n\eta)}.$$

Thus,

$$\begin{aligned}
\frac{1}{n}\log\mu_n(B(y,\delta)) &= \frac{1}{n}\Lambda_n(n\eta) - \langle\eta, z\rangle + \frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)) \\
&= \frac{1}{n}\Lambda_n(n\eta) - \langle\eta, y\rangle + \langle\eta, y-z\rangle + \frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)) \\
&\geq \frac{1}{n}\Lambda_n(n\eta) - \langle\eta, y\rangle - |\eta|\delta + \frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\lim_{\delta\to 0}\underline{\lim}_{n\to\infty}\frac{1}{n}\log\mu_n(B(y,\delta)) &\geq \Lambda(\eta) - \langle\eta, y\rangle + \lim_{\delta\to 0}\underline{\lim}_{n\to\infty}\frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)) \quad (1.15)\\
&\geq -\Lambda^*(y) + \lim_{\delta\to 0}\underline{\lim}_{n\to\infty}\frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)).
\end{aligned}$$

Since the weak law of large numbers can not be applied, the strategy now is to use the upper bound proved in (a). At first we verify that $\widetilde{\mu}_n$ satisfies Assumption 1.4.8 with the limiting logarithmic moment generating function $\widetilde{\Lambda}(\cdot) := \Lambda(\cdot+\eta) - \Lambda(\eta)$. In fact, $\widetilde{\Lambda}(0) = \Lambda(\eta) - \Lambda(\eta) = 0$ and since $\eta \in D^o_\lambda$ it follows that $\widetilde{\Lambda}(\lambda) < \infty$ for every $|\lambda|$ sufficiently small. Let $\widetilde{\Lambda}_n(\cdot)$ denote the logarithmic moment generating function corresponding to the law $\widetilde{\mu}_n$. Then for every $\lambda \in \mathbb{R}^d$,

$$\begin{aligned}
\frac{1}{n}\widetilde{\lambda}_n(\eta\lambda) &:= \frac{1}{n}\log\left[\int_{\mathbb{R}^d} e^{n\langle\lambda, z\rangle} d\widetilde{\mu}_n(z)\right] \\
&= \frac{1}{n}\log\left[\int_{\mathbb{R}^d} e^{n\langle\lambda+\eta, z\rangle - \Lambda_n(n\eta)} d\mu_n(z)\right] \\
&= -\frac{1}{n}\Lambda_n(n\eta) + \frac{1}{n}\log\left[\int_{\mathbb{R}^d} e^{\langle n(\lambda+\eta), z\rangle} d\mu_n(z)\right] \\
&= \frac{1}{n}\Lambda_n(n(\lambda+\eta)) - \frac{1}{n}\Lambda_n(n\eta) \to \widetilde{\Lambda}(\lambda)
\end{aligned}$$

because $\Lambda_n(n\eta) < \infty$ for n large enough. Let us define

$$\widetilde{\Lambda}^*(x) := \sup_{\lambda\in\mathbb{R}^d}\{\langle\lambda, x\rangle - \widetilde{\Lambda}(\lambda)\} = \Lambda^*(x) - \langle\eta, x\rangle + \Lambda(\eta). \quad (1.16)$$

Since Assumption 1.4.8 also holds for $\widetilde{\mu}_n$, it follows, applying Lemma 1.4.13 to $\widetilde{\Lambda}$, that $\widetilde{\Lambda}^*$ is a good rate function. Moreover, by part (a), a large deviations upper bound of the form of (2.8) holds for the sequence of measures $\widetilde{\mu}_n$ with the good rate function $\widetilde{\Lambda}^*$. In particular, for the closed set $B(y,\delta)^c$ it holds

$$\overline{\lim_{n\to\infty}}\frac{1}{n}\log\widetilde{\mu}_n(B(y,\delta)^c) \leq -\inf_{x\in B(y,\delta)^c}\widetilde{\Lambda}^*(x) = -\widetilde{\Lambda}^*(x_0)$$

for some $x_0 \neq y$. Since y is an exposed point of λ^* with η being the exposing hyperplane, we get that when $\Lambda^*(y) \geq [\langle\eta, y\rangle - \Lambda(\eta)]$ e $x_0 \neq y$, follows

$$\widetilde{\Lambda}^*(x_0) \geq [\Lambda^*(x_0) - \langle\eta, x_0\rangle] - [\Lambda^*(y) - \langle\eta, y\rangle] > 0.$$

Hence, for every $\delta > 0$,

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \widetilde{\mu}_n(B(y,\delta)^c) \leq 0.$$

This inequality implies that $\widetilde{\mu}_n(B(y,\delta)^c) \to 0$ and then , $\widetilde{\mu}_n(B(y,\delta)) \to 1$ for every $\delta > 0$. In particular,

$$\lim_{\delta\to 0} \underline{\lim_{n\to\infty}} \frac{1}{n} \log \widetilde{\mu}_n(B(y,\delta)) = 0,$$

and (1.14) follows by (1.15).

(c) In view of parts (a) and (b) and Lemma 1.4.14, it is sufficient to show that for any open set G,

$$\inf_{x\in G\cap \mathrm{ri} D_{\Lambda^*}} \Lambda^*(x) \leq \inf_{x\in G} \Lambda^*(x).$$

If $G \cap D_{\Lambda^*} = \emptyset$, there is nothing to prove. Then, assume that $D_{\Lambda^*} \neq \emptyset$. This implies that there exists some $z \in \mathrm{ri}\ D_{\Lambda^*}$. Fix $y \in G \cap D_{\Lambda^*}$. Hence, for all $\alpha > 0$ sufficiently small,

$$\alpha z + (1-\alpha)y \in G \cap riD_{\Lambda^*}.$$

Therefore,

$$\inf_{x\in G\cap \mathrm{ri} D_{\Lambda^*}} \Lambda^*(x) \leq \lim_{\alpha\searrow 0} \Lambda^*(\alpha z + (1-\alpha)y) \leq \Lambda^*(y).$$

The arbitrariness of y completes the proof.

□

Remark 1.4.16. *In $\mathbb{R}$, a point y is exposed if the curve $\Lambda^*(y)$ lies strictly above the line of slope λ through the point $(x, \Lambda^*(x))$. It was proved that $\Lambda^*(\cdot)$ is convex and differentiable, then we can make this the tangent hyperplane (see lecture notes [56, Chapter 33]). So, $\mathcal{F}$ is the interval where the tangent is well defined. And then the Theorem 2.2.5 can be reduced to a simpler version since in $\mathbb{R}$ we have $G \cap \mathcal{F} = G$.*

Chapter 2

Large deviation estimates for return times

In this chapter we focus on large deviation results for a dynamical systems with an exact dimensional measure. In the first section we present a generalization of [33] for return time of the orbit of x to the ball $B(x, r)$. We establish a link between return time and rate functions for dimension and for fast return times. We prove that when a dynamical system has an exact dimensional measure, the large deviation rate function that is given in terms of the rate functions mentioned above. As an application, a large deviation result for repellers is proven in Section 2.2.

This chapter is based on article [23], Large deviation for return times, written with Benoît Saussol and Jérôme Rousseau and published in Nonlinearity.

2.1 Large deviation estimates for return times in a general setting

Throughout this section we consider $g : M \to M$ a measurable map and μ an ergodic invariant probability measure on $(M, \mathcal{A})$.

2.1.1 Definitions and statements

We define the rate functions which will appear in our large deviations estimates. The first one is related to the deviations in the pointwise dimension; it has been computed in [47] in the case of conformal repellers.

Definition 2.1.1. *The exponential rate for dimension is defined for $\epsilon > 0$ by:*

$$\underline{\psi}(\pm\epsilon) = \varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\left\{\frac{\log \mu(B(x,r))}{-\log r} \in I_{\pm\epsilon}\right\}\right), \tag{2.1}$$

where $I_\epsilon = (-\infty, -d_\mu - \epsilon)$ *and* $I_{-\epsilon} = (-d_\mu + \epsilon, +\infty)$.

If we denote by

$$\mu_{B(x,r)}(A) = \frac{\mu(A \cap B(x,r))}{\mu(B(x,r))}$$

the conditional measure on $B(x,r)$, where A is a measurable set, then we can present the second rate function that quantifies the probability of quick returns near the origin.

Definition 2.1.2. *The exponential rate for fast return times is defined for* $\epsilon, a > 0$ *by:*

$$\underline{\varphi}(a,\epsilon) = \varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\left\{x_0 : \mu_{B(x_0,2r)}\left(\tau_{B(x_0,2r)} \le r^{-d_\mu+\epsilon}\right) \ge Cr^a\right\}\right), \tag{2.2}$$

for some constant $C > 0$.

We may now state our main result. We emphasize that the value of C in (2.2) is irrelevant in the theorem.

Theorem 2.1.3. *Let* $(M, \mathcal{A}, g, \mu)$ *be a measure preserving system. Suppose that* μ *is an exact dimensional measure. Given* $\epsilon > 0$, *we have:*

$$\varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \ge r^{-d_\mu-\epsilon}\right) \ge \max_{\gamma\in(0,1)} \min\left\{(1-\gamma)\epsilon, \underline{\psi}(\gamma\epsilon)\right\} \tag{2.3}$$

$$\varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \le r^{-d_\mu+\epsilon}\right) \ge \max_{\substack{\gamma\in(0,1)\\ a,\epsilon''>0}} \min\left\{-\gamma\epsilon - \epsilon'' + a, \underline{\psi}(\gamma\epsilon), \underline{\varphi}(a,\epsilon), \underline{\psi}(-\epsilon'')\right\}. \tag{2.4}$$

This result is satisfactory in the sense that it can be applied to a broad class of dynamical systems, provided one can estimate the rate functions $\underline{\psi}$ and $\underline{\varphi}$.

The rate function for dimension $\underline{\psi}$ is rather classical. We can observe that in (2.3) if the rate function for dimension $\underline{\psi}$ is positive in some interval $(0,\epsilon)$, it readily implies that $\mu\left(\tau_r \ge r^{-d_\mu-\epsilon}\right)$ has a fast decay.

The rate function $\underline{\varphi}$ is not so well known. However, for several dynamical systems an estimation of the error in the approximation to the exponential law for return time has been computed. In many cases, including Hénon maps [20, Theorem 3.1], it is possible to show that for some $a, b > 0$, and any sufficiently small $r > 0$,

E1 there exists a set $\Omega_r \subset M$ such that $\mu(\Omega_r^c) < r^b$;

E2 for all $x \in \Omega_r$,

$$\sup_{t\ge 0}\left|\mu_{B(x,r)}\left(\tau_{B(x,r)} > \frac{t}{\mu(B(x,r))}\right) - e^{-t}\right| \le r^a.$$

The conditions E1-E2 imply that $\underline{\varphi}(a,\epsilon) \ge \min\{\underline{\psi}(a-\epsilon), b\}$ (see Proposition 2.1.5 in Subsection 2.1.2).

2.1.2 Proof of the general result

In this section we prove the Theorem 2.1.3 using the method developed in [52]. We start by the result that is an elementary property related to the lower bound that will be used throughout the chapter.

Lemma 2.1.4. *Let* $\{a_1(r), \ldots, a_p(r)\}$ *be a collection of* p *sequences with* $a_i(r) > 0$. *If*

$$\forall i \leq p, \quad \gamma_i = \varliminf_{r\to 0} \frac{1}{\log r} \log a_i(r) > 0.$$

Then

$$\varliminf_{r\to 0} \frac{1}{\log r} \log \left(\sum_{i=1}^{p} a_i(r) \right) \geq \min_{i=1,\ldots,p} \gamma_i.$$

Proof. For all $\epsilon > 0$ there exists $r_i > 0$ such that $r < r_i$ implies $a_i \leq r^{\gamma_i - \epsilon}$. Let $\epsilon > 0$ sufficiently small such that $\gamma_i - \epsilon > 0$. We have,

$$\sum_{i=1}^{p} a_i(r) \leq \sum_{i=1}^{p} r^{\gamma_i - \epsilon} \leq p r^{\min\{\gamma_i\} - \epsilon}$$

and this implies

$$\frac{1}{\log r} \log \left(\sum_{i=1}^{p} a_i(r) \right) \geq \min_{i=1,\ldots,p} \{\gamma_i\} - \epsilon + \frac{\log p}{\log r}.$$

Finally,

$$\varliminf_{r\to 0} \frac{1}{\log r} \log \left(\sum_{i=1}^{p} a_i(r) \right) \geq \min_{i=1,\ldots,p} \{\gamma_i\} - \epsilon.$$

The result is proved since ϵ can be chosen arbitrarily small. □

Denote

$$\underline{f}(\epsilon) = \varliminf_{r\to 0} \frac{1}{\log r} \log \mu \left(\tau_r \geq r^{-d_\mu - \epsilon} \right)$$

and

$$\underline{f}(-\epsilon) = \varliminf_{r\to 0} \frac{1}{\log r} \log \mu \left(\tau_r \leq r^{-d_\mu + \epsilon} \right).$$

Given $\epsilon, \xi > 0$, define

$$A_\epsilon(r) = \left\{ x \in M : \mu(B(x,r)) \geq r^{d_\mu + \epsilon} \right\} \tag{2.5}$$

and

$$A_{-\xi}(r) = \left\{ x \in M : \mu(B(x,r)) \leq r^{d_\mu - \xi} \right\}. \tag{2.6}$$

Proof of the Theorem 2.1.3. Let $\gamma \in (0,1)$. We have

$$\begin{aligned} \mu(\{x : \tau_r(x) \geq r^{-d_\mu - \epsilon}\}) \;&\leq\; \mu \left(\left\{ x \in A_{\gamma\epsilon}\left(\frac{r}{4}\right) : \tau_r(x) \geq r^{-d_\mu - \epsilon} \right\} \right) \\ &+\; \mu \left(\left\{ x \in A^c_{\gamma\epsilon}\left(\frac{r}{4}\right) : \tau_r(x) \geq r^{-d_\mu - \epsilon} \right\} \right). \end{aligned}$$

Let us define the set

$$M_r = \left\{ x \in A_{\gamma\epsilon}\left(\frac{r}{4}\right) : \tau_r(x) \geq r^{-d_\mu - \epsilon} \right\}.$$

Let $\left\{B\left(x_i, \frac{r}{2}\right)\right\}_i$ be a family of balls of radius $r/2$ centered at points of $A_{\gamma\epsilon}(\frac{r}{4})$ that covers M_r and such that $B\left(x_i, \frac{r}{4}\right) \cap B\left(x_j, \frac{r}{4}\right) = \emptyset\ if\ x_i \neq x_j$. We have

$$\begin{aligned}
\mu\left(\left\{x : \tau_r(x) \geq r^{-d_\mu-\epsilon}\right\}\right) &\leq \mu(\cup_i B_i \cap M_r) + \mu\left(\left\{x \in A^c_{\gamma\epsilon}\left(\frac{r}{4}\right) : \tau_r(x) \geq r^{-d_\mu-\epsilon}\right\}\right) \\
&\leq \sum_i \mu(B_i \cap M_r) + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right).
\end{aligned}$$

Using first the triangle inequality and then Kac's lemma and Markov inequality, we obtain

$$\mu(B_i \cap M_r) \leq \mu\left(B_i \cap \left\{\tau_{B_i} \geq r^{-d_\mu-\epsilon}\right\}\right) \leq r^{d_\mu+\epsilon} \int_{B_i} \tau_{B_i} d\mu = r^{d_\mu+\epsilon}.$$

Observe that $\displaystyle\sum_i \left(\frac{r}{4}\right)^{d_\mu+\gamma\epsilon} \leq \sum_i \mu\left(B\left(x_i, \frac{r}{4}\right)\right) \leq 1$. Thus, since the balls are disjoint it follows that the number of balls is bounded by $\left(\frac{1}{4}r\right)^{-d_\mu-\gamma\epsilon}$. Therefore,

$$\begin{aligned}
\mu\left(\left\{x : \tau_r(x) \geq r^{-d_\mu-\epsilon}\right\}\right) &\leq \sum_i r^{d_\mu+\epsilon} + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right) \\
&\leq \left(\frac{1}{4}r\right)^{-d_\mu-\gamma\epsilon} r^{d_\mu+\epsilon} + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right) \\
&\leq 4^{d_\mu+\gamma\epsilon} r^{(1-\gamma)\epsilon} + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right).
\end{aligned}$$

Thus,

$$\frac{1}{\log r} \log \mu\left(\left\{x : \tau_r(x) \geq r^{-d\mu-\epsilon}\right\}\right) \geq \frac{1}{\log r} \log\left(4^{d_\mu+\gamma\epsilon} r^{(1-\gamma)\epsilon} + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right)\right).$$

Hence, by Lemma 2.1.4, we get

$$\begin{aligned}
\underline{f}(\epsilon) &\geq \underline{\lim}_{r\to 0} \frac{1}{\log r} \log\left(4^{d_\mu+\gamma\epsilon} r^{(1-\gamma)\epsilon} + \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right)\right) \\
&\geq \min\left\{\underline{\lim}_{r\to 0} \frac{1}{\log r} \log\left(4^{d_\mu+\gamma\epsilon} r^{(1-\gamma)\epsilon}\right), \underline{\lim}_{r\to 0} \frac{1}{\log r} \log \mu\left(A^c_{\gamma\epsilon}\left(\frac{r}{4}\right)\right)\right\} \\
&= \min\left\{(1-\gamma)\epsilon, \underline{\psi}(\gamma\epsilon)\right\}.
\end{aligned}$$

This proves the first statement.

Now, let $\epsilon'' > 0$. We define

$$\Gamma_r = \left\{x \in A_{\gamma\epsilon}(2r) \cap A_{(-\epsilon'')}(2r) : \tau_r(x) \leq r^{-d_\mu+\epsilon}\right\}$$

and

$$D_r = \left\{x_0 : \mu_{B(x_0,2r)}(\tau_{B(x_0,2r)} \leq r^{-d_\mu+\epsilon}) \leq Cr^a\right\},\ C > 0.$$

Let $\{B(x_i,2r\}_i$ be a family of balls of radius $2r$ centered at points of $A_{\gamma\epsilon}(2r)\cap D_r\cap A_{(-\epsilon'')}(2r)$ that covers $\Gamma_r\cap D_r$ and such that $B(x_i,r)\cap B(x_j,r)=\emptyset \; if \; x_i\neq x_j$. We have

$$\begin{aligned}
& \mu\left(\left\{x:\tau_r(x)\le r^{-d_\mu+\epsilon}\right\}\right)\\
\le\; & \mu\left(\left\{x\in A_{\gamma\epsilon}(2r)\cap D_r\cap A_{(-\epsilon'')}(2r):\tau_r(x)\le r^{-d_\mu+\epsilon}\right\}\right)\\
& +\mu\left(\left\{x\in (A_{\gamma\epsilon}(2r)\cap D_r\cap A_{(-\epsilon'')}(2r))^c:\tau_r(x)\le r^{-d_\mu+\epsilon}\right\}\right)\\
\le\; & \mu\left(\cup_i B(x_i,2r)\cap\Gamma_r\cap D_r\right)+\mu\left(A^c_{\gamma\epsilon}(2r)\right)+\mu\left(D^c_r\right)+\mu\left(A^c_{(-\epsilon'')}(2r)\right).
\end{aligned}$$

We remark that

$$\begin{aligned}
& \mu\left(\cup_i B(x_i,2r)\cap\Gamma_r\cap D_r\right)\\
\le\; & \sum_i \mu(B(x_i,2r)\cap\Gamma_r\cap D_r)\\
\le\; & \sum_i \mu(B(x_i,2r))\frac{1}{\mu(B(x_i,2r))}\mu\left(B(x_i,2r)\cap\left\{\tau_{B(x_i,2r)}\le r^{-d_\mu+\epsilon}\right\}\right)
\end{aligned}$$

where the last inequality follows from $\left\{\tau_{B(x_i,r)}\le r^{-d_\mu+\epsilon}\right\}\subset\left\{\tau_{B(x_i,2r)}\le r^{-d_\mu+\epsilon}\right\}$. Therefore, by definition of D_r,

$$\begin{aligned}
& \mu\left(\left\{x:\tau_r(x)\le r^{-d_\mu+\epsilon}\right\}\right)\\
\le\; & \sum_i \mu(B(x_i,2r))\mu_{(B(x_i,2r))}\left(\tau_{B(x_i,2r)}\le r^{-d_\mu+\epsilon}\right)+\mu\left(A^c_{\gamma\epsilon}(2r)\right)+\mu\left(D^c_r\right)+\mu\left(A^c_{(-\epsilon'')}(2r)\right)\\
\le\; & \sum_i \mu(B(x_i,2r))Cr^a+\mu\left(A^c_{\gamma\epsilon}(2r)\right)+\mu\left(D^c_r\right)+\mu\left(A^c_{(-\epsilon'')}(2r)\right).
\end{aligned}$$

Observe that $\sum_i r^{d_\mu+\gamma\epsilon}\le\sum_i\mu(x_i,r)\le 1$. Thus, since balls are disjoint it follows that the number of balls is bounded by $r^{-d_\mu-\gamma\epsilon}$ and

$$\begin{aligned}
\sum_i \mu(B(x_i,2r)) &\le \sum_i (2r)^{d_\mu-\epsilon''}\\
&\le r^{-d_\mu-\gamma\epsilon}(2r)^{d_\mu-\epsilon''}\\
&\le 2^{d_\mu-\epsilon''}r^{-\gamma\epsilon-\epsilon''}.
\end{aligned}$$

Then, we obtain that

$$\mu\left(\left\{x:\tau_r(x)\le r^{-d_\mu+\epsilon}\right\}\right)\le C2^{d_\mu-\epsilon''}r^{-\gamma\epsilon-\epsilon''+a}+\mu\left(A^c_{\gamma\epsilon}(2r)\right)+\mu\left(D^c_r\right)+\mu\left(A^c_{(-\epsilon'')}(2r)\right).$$

Hence,

$$\underline{f}(-\epsilon)\ge\underline{\lim}_{r\to 0}\frac{1}{\log r}\log\left(C2^{d_\mu-\epsilon''}r^{-\gamma\epsilon-\epsilon''+a}+\mu\left(A^c_{\gamma\epsilon}(2r)\right)+\mu\left(D^c_r\right)+\mu\left(A^c_{(-\epsilon'')}(2r)\right)\right).$$

Finally, using the definitions of $\underline{\psi}$ and $\underline{\varphi}$ we get by Lemma 2.1.4 that

$$\underline{f}(-\epsilon)\ge\min\left\{-\gamma\epsilon-\epsilon''+a,\underline{\psi}(\gamma\epsilon),\underline{\varphi}(a,\epsilon),\underline{\psi}(-\epsilon'')\right\}.$$

This concludes the proof of the theorem. $\square$

We finish with a brief result that may help to estimate the rate function for fast returns.

Proposition 2.1.5. *If there exist constants $a, b > 0$ such that for all $r \in (0,1)$:*

- *there exists a set Ω_r such that*
$$\mu(\Omega_r^c) < r^b;$$
- *for all $x \in \Omega_r$,*
$$\left| \mu_{B(x,r)} \left(\tau_{B(x,r)} > \frac{t}{\mu(B(x,r))} \right) - e^{-t} \right| \leq r^a,$$
for every $t > 0$.

Then, $\underline{\varphi}(a, \epsilon) \geq \min\{\underline{\psi}(a - \epsilon), b\}$.

Proof. Take $t = Cr^a, C > 0$. Making the first order expansion of e^{-t}, we have for $x \in \Omega_r$
$$\left| \mu_{B(x,r)} \left(\tau_{B(x,r)} > \frac{Cr^a}{\mu(B(x,r))} \right) - 1 + Cr^a + o(r^{2a}) \right| \leq r^a,$$
which implies
$$\left| \mu_{B(x,r)} \left(\tau_{B(x,r)} < \frac{Cr^a}{\mu(B(x,r))} \right) + Cr^a + o(r^{2a}) \right| \leq r^a.$$
So, it follows that
$$\mu_{B(x,r)} \left(\tau_{B(x,r)} < \frac{Cr^a}{\mu(B(x,r))} \right) < r^a.$$
Let N_r be a set defined by $N_r = \left\{ x : \mu(B(x,r)) \geq r^{d_\mu + a - \epsilon} \right\}$. For $x \in N_r \cap \Omega_r$ we obtain
$$\mu_{B(x,r)} \left(\tau_{B(x,r)} < Cr^{-d_\mu + \epsilon} \right) < r^a.$$

Thus,
$$\mu \left(\left\{ x : \mu_{B(x,2r)} \left(\tau_{B(x,2r)} \leq r^{-d_\mu + \epsilon} \right) > 2^a r^a \right\} \right) \leq \mu((N_{2r} \cap \Omega_{2r})^c) \leq \mu(N_{2r}^c) + \mu(\Omega_{2r}^c).$$

Finally, by Lemma 2.1.4, we get
$$\underline{\varphi}(a, \epsilon) \geq \min\{\underline{\psi}(a - \epsilon), b\}.$$

□

2.2 Large deviation estimates for return times for conformal repeller

From now on, let (J, g) be a conformal repeller.

If we consider a conformal repeller and an equilibrium state of a Hölder potential ζ we obtain a somewhat more concrete version of our principal result:

1. in this setting we can compute the exponential rate for the dimension $\underline{\psi}$, using thermodynamic formalism;
2. we can also estimate the exponential rate for fast return times $\underline{\varphi}$, using a technique similar to the one used to prove exponential return time statistics.

Thus, applying Theorem 2.1 to this setting, we obtain a large deviation result with a rate that is given in terms of Legendre transform of the convex function $T(\cdot)$ defined in (1.3).

Theorem 2.2.1. *Let (J, g) be a conformal repeller and μ an equilibrium state for a Hölder potential ζ. For any $\epsilon > 0$, we have:*

$$\begin{aligned} \varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \geq r^{-d_\mu - \epsilon}\right) &\geq g_1(\epsilon) \\ \varliminf_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \leq r^{-d_\mu + \epsilon}\right) &\geq g_2(\epsilon), \end{aligned}$$

where

$$g_1(\epsilon) = \max_{\gamma\in(0,1)} \min\left\{(1-\gamma)\epsilon, \Lambda^*(-d_\mu - \gamma\epsilon)\right\} > 0$$

and

$$g_2(\epsilon) = \max_{\substack{\gamma\in(0,1)\\ \epsilon'>0 \\ \epsilon''>0}} \min\left\{a', \Lambda^*(-d_\mu - \gamma\epsilon), \min\{a_0, \Lambda^*(-d_\mu + \epsilon')\}, \Lambda^*(-d_\mu + \epsilon'')\right\} > 0$$

with $a' = -\gamma\epsilon - \epsilon'' + \min\{d_2, \epsilon - \epsilon'\}$, $\Lambda^*(x) = -x + T^*(x) = -x + \sup_{\lambda\in\mathbb{R}}\{\lambda x - T(\lambda)\}$ and a_0, d_2 are some constants.

Lemma 2.2.2. *If μ is the measure of maximal dimension then $HP_\mu(q)$ is constant and equal to d_μ.*

Proof. See for instant the proof of [15, Lemma 6.1.7]. □

Remark 2.2.3. *If μ is the measure of maximal dimension the above theorem remains valid. However, since $HP_\mu(q)$ is constant and equal to d_μ, it follows that $\Lambda^*(x) = +\infty$ for $x \neq 0$, and thus $g_1(\epsilon) = \max_{\gamma\in(0,1)}(1-\gamma)\epsilon$ and $g_2(\epsilon) = \max_{\substack{\gamma\in(0,1)\\ \epsilon'>0 \\ \epsilon''>0}} \min\{-\gamma\epsilon - \epsilon'' + \min\{d_2, \epsilon - \epsilon'\}, a_0\}$.*

Proposition 2.2.4. *Suppose that μ is not the measure of maximal dimension. Then for any $\kappa < 1$ and ϵ sufficiently small, $g_1 \geq \kappa c \epsilon^2$ and $g_2 \geq \kappa c \left(\frac{\epsilon}{3}\right)^2$, with $c = \frac{1}{2}(\Lambda^*)''(-d_\mu) = \frac{\lambda_\mu}{\sigma_\mu^2}$, where λ_μ is the Lyapunov exponent of g and σ_μ^2 the variance of $\log\psi + \log|a|$ with respect to μ.*

Proof. By definition $T^*(x) = \sup_{q\in\mathbb{R}}\{qx - T(q)\}$. The supremum is achieved for q such that $\frac{d}{dq}(qx - T(q)) = x - T'(q) = 0$, that is, $T'(q) = x$. Thus, for any $q \in \mathbb{R}$

$$T^*(T'(q)) = qT'(q) - T(q).$$

So, it follows that

$$(T^*)'(T'(q))T''(q) = qT''(q),$$

and hence, $(T^*)'(T'(q)) = q$ and, differentiating, we obtain

$$(T^*)''(T'(q))T''(q) = 1,$$

that is

$$(T^*)''(T'(q)) = \frac{1}{T''(q)} \text{ for every } q \in \mathbb{R} \text{ such that } T'(q) = x.$$

Since $\Lambda(\lambda) = T(\lambda + 1)$ and $\Lambda^*(x) = -x + T^*(x)$ we conclude that

$$(\Lambda^*)''(x) = \frac{1}{T''(q)}, \text{ where } T'(\lambda + 1) = x.$$

Moreover, by Proposition 1.2.9 this is non negative.

For $x = -d_\mu$ we have $\lambda = 0$. Then, by Lemma 5 in [47]

$$(\Lambda^*)''(-d_\mu) = \frac{1}{T''(1)} = \frac{\lambda_\mu}{\sigma_\mu^2}.$$

Finally, for ϵ sufficiently small, $g_1 \geq \kappa c \epsilon^2$ and taking $\epsilon' = \epsilon'' = \gamma\epsilon = \frac{\epsilon}{\eta}$, with $\eta > 3$, we have $g_2 \geq \kappa c \left(\frac{\epsilon}{3}\right)^2$.

□

To obtain Theorem 2.2.1, we need a fundamental theorem of large deviation theory, the Gartner-Ellis Theorem, that was presented in Subsection 1.4.2.

In this context, for $r \in (0, 1)$ let $(\mu_r)_r$ be a family of probability measures and consider a family of random variables $(Z_r)_r$ in $\mathbb{R}$.

Hence, for any $\lambda \in \mathbb{R}$ and taking $n = -\log r$, the logarithmic moment generating function defined in Assumption 1.4.8 can be reformulated as

$$\Lambda(\lambda) := \lim_{r\to 0} \frac{1}{-\log r}\Lambda_r(-\lambda \log r). \tag{2.7}$$

Thanks to Remark 1.4.16, in $\mathbb{R}$ we can enunciate Gartner-Ellis theorem as follows.

Theorem 2.2.5. *If Assumption 1.4.8 holds, then*

(a) for any closed set F,

$$\overline{\lim_{r\to 0}}\frac{1}{-\log r}\log \mu_r(F) \leq -\inf_{x\in F}\Lambda^*(x); \tag{2.8}$$

(b) for any open set G,

$$\underline{\lim_{r\to 0}}\frac{1}{-\log r}\log \mu_r(G) \geq -\inf_{x\in G}\Lambda^*(x). \tag{2.9}$$

We will apply this Theorem 2.2.5 to the family Z_r defined by

$$Z_r = \frac{\log \mu(B(x,r))}{-\log r}.$$

Replacing Z_r in (1.10) we get that

$$\Lambda_r(\lambda) = \log \int e^{\lambda \frac{\log \mu(B(x,r))}{-\log r}} d\mu(x). \tag{2.10}$$

Thus, substituting (2.10) into (2.7) we obtain

$$\begin{aligned}\Lambda(\lambda) &= \lim_{r\to 0}\frac{1}{-\log r}\log \int e^{\lambda \log \mu(B(x,r))} d\mu(x) \\ &= \lim_{r\to 0}\frac{1}{-\log r}\log \int \mu(B(x,r))^{\lambda} d\mu(x). \end{aligned} \tag{2.11}$$

The next result is basically a consequence of Proposition 1.2.9 combined with (2.11).

Proposition 2.2.6. *Let (J,g) be a conformal repeller and μ an equilibrium state for the Hölder potential ζ. Then, for $\lambda > 0$, the following limit exists*

$$\Lambda(\lambda) = \lim_{r\to 0}\frac{1}{-\log r}\log \int \mu(B(x,r))^{\lambda} d\mu(x) = T(\lambda+1).$$

Proof. The result follows taking $\lambda = q-1$ in (1.3). $\square$

Applying Gartner-Ellis Theorem, we obtain that the quantity $\mu\left(\left\{\frac{\log \mu(B(x,r))}{-\log r} \in I\right\}\right)$, where I is an interval, decrease exponentially when r goes to zero. Namely,

Corollary 2.2.7. *Under the same conditions of Proposition 2.2.6 we have that for all interval I,*

$$\lim_{r\to 0}\frac{1}{-\log r}\log \mu\left(\left\{\frac{\log \mu(B(x,r))}{-\log r} \in I\right\}\right) = -\inf_{x\in I}\Lambda^*(x),$$

where $\Lambda^*(x) = -x + T^*(x)$ is continuous on its domain.

Proof. This equality is a direct consequence of Theorem 2.2.5. Since the logarithmic moment generating function is defined by $\Lambda(\lambda) = T(\lambda+1)$, the Fenchel-Legendre transform of $\Lambda(\lambda)$ is

$$\begin{aligned}
\Lambda^*(x) &= \sup_{\lambda\in\mathbb{R}}\{\lambda x - \Lambda(\lambda)\} \\
&= \sup_{\lambda\in\mathbb{R}}\{\lambda x - T(\lambda+1)\} \\
&= \sup_{\nu\in\mathbb{R}}\{(\nu-1)x - T(\nu)\} \\
&= -x + \sup_{\nu\in\mathbb{R}}\{\nu x - T(\nu)\} \\
&= -x + T^*(x).
\end{aligned}$$

The continuity of $\Lambda^*(x)$ follows from its convexity. □

In Figure 2.1, one can see a graph of the Fenchel-Legendre transform of Λ, where the interval $I \subset (-\infty, -d_\mu - \epsilon) \cup (-d_\mu + \epsilon, +\infty) = I_\epsilon \cup I_{-\epsilon}$. Λ^* is strictly convex and its minimum is reached at $-d_\mu$.

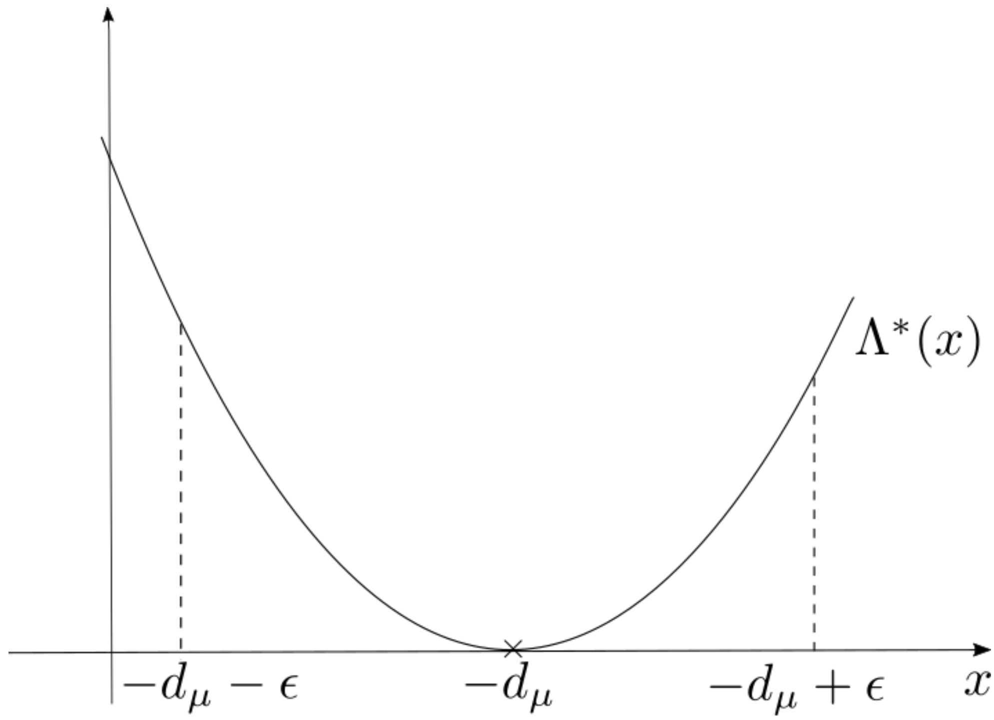

Figure 2.1: Graph of Λ^*.

Recall we defined the exponential rate for the dimension

$$\underline{\psi}(\pm\epsilon) = \underline{\lim}_{r\to 0}\frac{1}{\log r}\log\mu\left(\left\{\frac{\log\mu(B(x,r))}{-\log r} \in I_{\pm\epsilon}\right\}\right)$$

(see Section 2.1). Using Corollary 2.2.7, we get the rate function for the dimension for conformal repellers.

Proposition 2.2.8. *For any $\epsilon > 0$, the exponential rate for the dimension is given by:*

$$\underline{\psi}(\epsilon) = \inf_{x \in (-\infty, -d_\mu - \epsilon)} \Lambda^*(x) = \Lambda^*(-d_\mu - \epsilon) > 0$$

and

$$\underline{\psi}(-\epsilon) = \inf_{x \in (-d_\mu + \epsilon, +\infty)} \Lambda^*(x) = \Lambda^*(-d_\mu + \epsilon) > 0.$$

Proof. By Corollary 2.2.7 and from the convexity of Λ^*, we deduce that

$$\begin{aligned} \underline{\psi}(\pm\epsilon) &= \inf_{x \in I_{\pm\epsilon}} \Lambda^*(x) \\ &= \Lambda^*(-d_\mu \mp \epsilon) \end{aligned}$$

which proves the proposition. □

From now on, assume that ζ is an Hölder potential such that $P(\zeta) = 0$. To obtain an exponential rate for the fast return times (Proposition 2.2.9), we will need the Closing lemma (recall Theorem 1.2.5).

We will use these properties to obtain information about the rate function for the fast return times.

Proposition 2.2.9. *There exist constants $a_0, d_2 > 0$ such that for $\epsilon, \epsilon' > 0$, the exponential rate for fast return times satisfies:*

$$\underline{\varphi}(\min\{d_2, \epsilon - \epsilon'\}, \epsilon) \geq \min\{a_0, \underline{\psi}(-\epsilon')\} > 0.$$

This proposition is a consequence of the following lemma.

Lemma 2.2.10. *For any $d_0 \in (0, d_\mu)$ there exist constants $a_0, c_3, c_4, c_5, d_1, D > 0$ and a set Ω_r such that*

$$\mu(\Omega_r^c) < D r^{a_0}$$

and for all $x_0 \in \Omega_r$, one has

$$\mu_{B(x_0,2r)} \left(\tau_{B(x_0,2r)} \leq r^{-d_0} \right) \leq (c_4 - c_5 \log r)\, \mu(B(x_0, c_3 r^{d_1})) + r^{-d_0} \mu(B(x_0, 3r)).$$

Proof. We first claim that there exists Ω_r with $\mu(\Omega_r^c) \leq D r^{a_0}$ such that for all $x_0 \in \Omega_r$ and for all $k \leq c_0 \log \frac{1}{2r}$ we have $B(x_0, 2r) \cap g^{-k}(B(x_0, 2r)) = \emptyset$.

Indeed, let $c_0 = \frac{d_\mu}{2 \log m}$, where m is the degree of the map g. If x_0 is such that $B(x_0, 2r) \cap g^{-k}(B(x_0, 2r)) \neq \emptyset$, there exists x such that $d(x, x_0) < 2r$ and $d(g^k(x), x_0) < 2r$, thus $d(x, g^k(x)) < 4r$. By the Closing lemma, there exists a point z such that $g^k(z) = z$ and $d(z, x) < 4c_1 r$.

Define $\mathcal{P}_k = \{z : g^k(z) = z\}$ that is, the set of points that are arbitrarily close to x such that are also periodic. Thus $d(x, \mathcal{P}_k) < 4c_1 r$. Observe that

$$\{x_0 : B(x_0, 2r) \cap g^{-k}(B(x_0, 2r)) \neq \emptyset\} \subset B(\mathcal{P}_k, (4c_1+2)r) = \bigcup_{y\in\mathcal{P}_k} B(y, (4c_1+2)r). \quad (2.12)$$

Moreover, let $\xi = \frac{d_\mu}{4\log m}$. Using (2.6), we have the inequality

$$\begin{aligned} \mu\left(A_{(-\xi)}((8c_1+4)r) \cap B(\mathcal{P}_k, (4c_1+2)r)\right) &\leq \#\mathcal{P}_k \sup_{x\in A_{(-\xi)}((8c_1+4)r)} \mu(B(x, (8c_1+4)r)) \\ &\leq m^k((8c_1+4)r)^{d_\mu-\xi}. \end{aligned}$$

Now, take $K = c_0 \log \frac{1}{2r}$ and define

$$\Omega_r = A_{(-\xi)}((8c_1+4)r) \cap \bigcap_{k\leq K} B(\mathcal{P}_k, (4c_1+2)r)^c.$$

We proceed to compute a upper bound for the quantity $\mu(\Omega_r^c)$. In order to get this, it will be necessary to combine the previous inequality with Corollary 2.2.7. Hence,

$$\begin{aligned} \mu(\Omega_r^c) &\leq \sum_{k=1}^{K} \mu\left(A_{(-\xi)}((8c_1+4)r) \cap B(\mathcal{P}_k, (4c_1+2)r)\right) + \mu\left(A^c_{(-\xi)}((8c_1+4)r)\right) \\ &\leq ((8c_1+4)r)^{d_\mu-\xi} \sum_{k=1}^{K} m^k + ((8c_1+4)r)^{\underline{\psi}(-\xi)-\delta}, \\ &\leq ((8c_1+4)r)^{d_\mu-\xi} m^{K+1} + ((8c_1+4)r)^{\underline{\psi}(-\xi)-\delta} \\ &\leq Dr^{a_0}, \end{aligned}$$

for $\delta > 0$ sufficiently small and $a_0 = \min\{d_\mu - \xi - c_0\log m, \underline{\psi}(-\xi) - \delta\}$.

We observe that $x_0 \in \Omega_r$ implies that $x_0 \notin B(\mathcal{P}_k, (4c_1+2)r)$ for all $k \leq c_0 \log \frac{1}{2r}$. Therefore, from (2.12) we obtain $B(x_0, 2r) \cap g^{-k}(B(x_0, 2r)) = \emptyset$ which proves our initial claim.

We now turn to estimate the quantity $\mu_{B(x_0,2r)}\left(g^{-k}B(x_0, 2r)\right)$ for large values of k.

Recall that ζ is a Hölder potential such that $P(\zeta) = 0$. We also recall that the Ruelle-Perron-Frobenius operator $\mathcal{L}_\zeta : C(M) \to C(M)$ defined on the space $C(M)$ of continuous function by

$$\mathcal{L}_\zeta(f)(x) = \sum_{y\in g^{-1}(x)} e^{\zeta(y)} f(y).$$

By induction, for every $n \geq 1$,

$$\mathcal{L}^n_\zeta(f)(x) = \sum_{y\in g^{-n}(x)} e^{S_n\zeta(y)} f(y), \quad (2.13)$$

where $S_n\zeta = \sum_{k=0}^{n-1} \zeta \circ g^k$. Now we have that

$$\begin{aligned}
\mu\left(B(x_0,2r)\cap g^{-k}B(x_0,2r)\right) &= \int \mathbb{1}_{B(x_0,2r)}\mathbb{1}_{B(x_0,2r)}\circ g^k \; d\mu \\
&= \int \mathcal{L}^k(\mathbb{1}_{B(x_0,2r)})\mathbb{1}_{B(x_0,2r)} \; d\mu \\
&\leq \mu(B(x_0,2r)) \left\|\mathcal{L}^k(\mathbb{1}_{B(x_0,2r)})\right\|_\infty .
\end{aligned}$$

Hence, the conditional measure is limited to

$$\mu_{B(x_0,2r)}\left(g^{-k}B(x_0,2r)\right) \leq \left\|\mathcal{L}^k(\mathbb{1}_{B(x_0,2r)})\right\|_\infty . \tag{2.14}$$

Now let $f = \mathbb{1}_R$ be the characteristic function of $R \in \mathcal{J}_k$. Applying (2.13) we have

$$\begin{aligned}
\mathcal{L}^k(\mathbb{1}_R)(x) &= \sum_{y\in g^{-k}x} e^{S_k\zeta(y)}\mathbb{1}_R(y) \\
&\leq \sum_{y\in g^{-k}x,\ y\in R} k_\zeta\mu(R)
\end{aligned}$$

where the last inequality follows from the Gibbs property since $P(\zeta)=0$. In addition, the preimage of x under g^k has just one element in R, thus

$$\mathcal{L}^k(\mathbb{1}_R)(x) \leq k_\zeta\mu(R). \tag{2.15}$$

By (2.15) we have

$$\begin{aligned}
\mathcal{L}^k(\mathbb{1}_{B(x_0,2r)}) &= \sum_{R\in\mathcal{J}_k, R\cap B(x_0,2r)\neq\emptyset} \mathcal{L}^k(\mathbb{1}_R) \\
&\leq \sum_{R\in\mathcal{J}_k, R\subset B(x_0,2r+\operatorname{diam}(\mathcal{J}_k))} k_\zeta\mu(R) \\
&\leq k_\zeta\mu(B(x_0,2r+\operatorname{diam}(\mathcal{J}_k))).
\end{aligned} \tag{2.16}$$

Substituting (2.16) into (2.14) see that

$$\mu_{B(x_0,2r)}\left(g^{-k}B(x_0,2r)\right) \leq k_\zeta\mu(B(x_0,2r+\operatorname{diam}(\mathcal{J}_k))).$$

Let $k > c_0\log\frac{1}{2r}$. We have that $\operatorname{diam}(\mathcal{J}_k) < c_2\beta^{-k}$. Then, for k such that $c_2\beta^{-k} > r$, we have

$$\beta^{-k} < \beta^{-c_0\log\frac{1}{2r}} = (2r)^{c_0\log\beta},$$

which implies

$$\begin{aligned}
\mu(B(x_0,2r+\operatorname{diam}(\mathcal{J}_k))) &\leq \mu(B(x_0,3c_2\beta^{-k})) \\
&\leq \mu(B(x_0,c_3r^{c_0\log\beta})).
\end{aligned}$$

When k satisfies $c_2\beta^{-k} \leq r$, we obtain

$$\mu(B(x_0, 2r + \operatorname{diam}(\mathcal{J}_k))) \leq \mu(B(x_0, 3r)).$$

Recall that for all $x_0 \in \Omega_r$ and for all $k \leq c_0 \log \frac{1}{2r}$, $B(x_0, 2r) \cap g^{-k}(B(x_0, 2r)) = \emptyset$. We may now combining these informations to conclude that

$$\begin{aligned}
&\mu_{B(x_0,2r)}\left(\tau_{B(x_0,2r)} \leq r^{-d_0}\right) \\
\leq\ & \sum_{k=1}^{r^{-d_0}} \mu_{B(x_0,2r)}\left(g^{-k}B(x_0, 2r)\right) \\
=\ & \sum_{k=c_0 \log \frac{1}{2r}}^{\lfloor \frac{\log c_2 - \log r}{\log \beta} \rfloor} \mu_{B(x_0,2r)}\left(g^{-k}B(x_0, 2r)\right) + \sum_{k=\lfloor \frac{\log c_2 - \log r}{\log \beta} \rfloor + 1}^{r^{-d_0}} \mu_{B(x_0,2r)}\left(g^{-k}B(x_0, 2r)\right) \\
\leq\ & \sum_{k=c_0 \log \frac{1}{2r}}^{\lfloor \frac{\log c_2 - \log r}{\log \beta} \rfloor} \mu(B(x_0, c_3 r^{d_1})) + \sum_{k=\lfloor \frac{\log c_2 - \log r}{\log \beta} \rfloor + 1}^{r^{-d_0}} \mu(B(x_0, 3r)) \\
\leq\ & (c_4 - c_5 \log r)\, \mu(B(x_0, c_3 r^{d_1})) + r^{-d_0} \mu(B(x_0, 3r))
\end{aligned}$$

with $d_1 = c_0 \log \beta$, which ends the proof. □

Remark 2.2.11. *Given a subset $A \subset \mathbb{R}^n$ by the usual metric on $\mathbb{R}^n$ we have that $d(a, b) = |a - b|$, for all $a, b \in A$. However, we can also consider the intrinsic metric on $\mathbb{R}^n$, defined as the infimum of the lengths of curves that connect a to b in A. Thus, the intrinsic diameter is the longest of all shortest paths on the surface between pairs of points. Following this notion, let us denote by int diam $\mathcal{J}_k(x)$ the intrinsic diameter of the cylinder $\mathcal{J}_k(x)$.*

From the theory of conformal repellers we obtain the lemma that gives a uniform bound for the measure.

Lemma 2.2.12. *There exists $d_3 > 0$ such that for all $x \in J$ and $r > 0$, one has $\mu(B(x, r)) \leq r^{d_3}$.*

Proof. Let $r > 0$. Given a point $x \in J$ by [46, Proposition 2], there exist positive constants c_8 and c_9 such that, for every x,

$$\operatorname{diam}\mathcal{J}_k(x) \leq c_8 \prod_{i=0}^{k-1} \left|a\left(g^i(x)\right)\right|^{-1} \quad \text{and} \quad \operatorname{int\,diam}\mathcal{J}_k(x) \geq c_9 \prod_{i=0}^{k-1} \left|a\left(g^i(x)\right)\right|^{-1}.$$

Let k be the minimum value such that $\operatorname{diam}\mathcal{J}_k(x) < r$. We claim that there exists $c_{10} > 0$ such that int $\operatorname{diam}\mathcal{J}_k(x) > c_{10} r$. Indeed,

$$\operatorname{diam}\mathcal{J}_{k-1}(x) \geq r \text{ implies that } c_8 \prod_{i=0}^{k-2} \left|a\left(g^i(x)\right)\right|^{-1} \geq r.$$

And then,

$$\text{int diam}\mathcal{J}_k(x) \geq c_9 \prod_{i=0}^{k-1} |a(g^i(x))|^{-1} \geq \frac{c_9}{c_8} \left|a\left(g^{k-1}(x)\right)\right|^{-1} r > c_{10} r,$$

where c_{10} is independent of x and r (see proof of Theorem 4.1.7 in [15]). Thus the claim is proved.

Therefore, it follows that there exists y such that $B(y, c_{10}r) \subset \mathcal{J}_k(x)$. On the other hand, since diam$\mathcal{J}_k(x) < r$ we have $B(x, r) \subset B(y, 2r)$. Then, since μ is diametrically regular (see [44, Proposition 21.4]), that is, there exist constants $\gamma_0 > 1$ and $C_0 > 0$ such that for any point z and any $r > 0$

$$\mu(B(z, \gamma_0 r)) \leq C_0 \mu((B(z, r)),$$

we obtain

$$\begin{aligned} \mu(B(x,r)) &\leq \mu(B(y, 2r)) \\ &\leq C_0^{k+1} \mu(B(y, 2^{-k}r)) \\ &\leq C_0^{k+1} \mu(B(y, c_{10}r)), \end{aligned} \tag{2.17}$$

for k large enough such that $2^{-k} < c_{10}$.

Now, we claim that there exist positive constants b and C_1 such that

$$\mu(\mathcal{J}_k(x)) \leq C_1 e^{-bk}. \tag{2.18}$$

If h is the density of μ with respect to the conformal measure, i.e. $\mathcal{L}_\zeta h = h$ then, the potential $\overline{\zeta} = \zeta - \log(h \circ g) + \log h$ has the same equilibrium measure μ as ζ. Moreover, $\mathcal{L}_{\overline{\zeta}} 1 = 1$, implies that $\overline{\zeta}(y) < 0$ for all y (if we suppose that each x has at least two preimages, otherwise we should to normalise with g^N for N such that each x has at least two preimages for g^N). Thus, there exists $b > 0$ such that

$$\max_y \{\overline{\zeta}(y)\} \leq -b.$$

And then,

$$\max_y \{\exp S_k \overline{\zeta}(y)\} \leq e^{-bk}.$$

By Gibbs property, we obtain the claim.

Therefore, by (2.17) and (2.18), it follows that $\mu(B(x,r)) < C_1 C_0^{k+1} e^{-bk}$.

Let $k = -\log r$, thus $\mu(B(x,r)) < C_{11} r^b$. Choosing d_3 such that $r^{d_3} > C_{11} r^b$, we conclude the proof of the Lemma. $\square$

Proof of Proposition 2.2.9. Using the above lemma we have that there exist constants $c_6, d_2 > 0$ such that $(c_4 - c_5 \log r)\, \mu(B(x_0, c_3 r^{d_1})) \leq c_6 r^{d_2}$, for all x_0.

Let $0 < \epsilon' < \epsilon$, for $x \in \Omega_r \cap A_{(-\epsilon')}(3r)$ and using Lemma 2.2.10 we obtain

$$\begin{aligned} \mu_{B(x_0,2r)}\left(\tau_{B(x_0,2r)} \leq r^{-d_0}\right) &\leq c_6 r^{d_2} + r^{-d_0}(3r)^{d\mu-\epsilon'} \\ &\leq c_7 r^{\min\{d_2,-d_0+d_\mu-\epsilon'\}}. \end{aligned}$$

Consequently

$$\begin{aligned} \mu\left(\left\{x : \mu_{B(x_0,2r)}\left(\tau_{B(x_0,2r)} \leq r^{-d_0}\right) > c_7 r^{\min\{d_2,-d_0+d_\mu-\epsilon'\}}\right\}\right) &\leq \mu\left(\left(\Omega_r \cap A_{(-\epsilon')}(3r)\right)^c\right) \\ &\leq \mu(\Omega_r^c) + \mu\left(A^c_{(-\epsilon')}(3r)\right). \end{aligned}$$

Finally, taking $d_0 = d_\mu - \epsilon$ and using Lemma 2.1.4, we get

$$\underline{\varphi}(\min\{d_2, \epsilon - \epsilon'\}, \epsilon) \geq \min\{a_0, \underline{\psi}(-\epsilon')\}$$

and the proposition is proved.

□

We are now able to prove Theorem 2.2.1.

Proof of the Theorem 2.2.1. For $\gamma \in (0,1)$, by Proposition 2.2.8

$$\underline{\psi}(\gamma\epsilon) = \Lambda^*(-d_\mu - \gamma\epsilon) > 0$$

and

$$\underline{\psi}(-\epsilon'') = \Lambda^*(-d_\mu + \epsilon'') > 0.$$

Moreover, by Proposition 2.2.9,

$$\underline{\varphi}(\min\{d_2, \epsilon - \epsilon'\}, \epsilon) \geq \min\{a_0, \underline{\psi}(-\epsilon')\} = \min\{a_0, \Lambda^*(-d_\mu + \epsilon')\} > 0.$$

Thus, it follows from Theorem 2.1.3, that

$$\underline{\lim}_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \geq r^{-d_\mu - \epsilon}\right) \geq \max_{\gamma\in(0,1)} \min\{(1-\gamma)\epsilon, \Lambda^*(-d_\mu - \gamma\epsilon)\} > 0$$

and

$$\begin{aligned} &\underline{\lim}_{r\to 0} \frac{1}{\log r} \log \mu\left(\tau_r \leq r^{-d_\mu + \epsilon}\right) \\ \geq\ &\max_{\substack{\gamma\in(0,1)\\ \epsilon'>0 \\ \epsilon''>0}} \min\left\{a', \Lambda^*(-d_\mu - \gamma\epsilon), \min\{a_0, \Lambda^*(d_\mu + \epsilon')\}, \Lambda^*(-d_\mu + \epsilon'')\right\} \\ >\ &0, \end{aligned}$$

with $a' = -\gamma\epsilon - \epsilon'' + \min\{d_2, \epsilon - \epsilon'\}$. Thus the theorem is proved. □

Chapter 3

Shortest distance between observed orbits and matching strings in encoded sequences

In this chapter we present our results related to behaviour of the shortest distance between observed orbits. For a dynamical system $(M, \mathcal{A}, g, \mu)$ and an observation f from M to a metric space (Y, d), in Section 3.1, we study the shortest distance between two observed orbits, proving that the limiting rate is related to the correlation dimension of the pushfoward measure $f_*\mu$. In Section 3.2, we present a result in the case of random dynamical systems and give some examples for which the theorem applies. Finally, in Section 3.3, under mixing conditions, we present the symbolic theorem which establishes a relation between the longest common substring between encoded strings and the Rényi entropy. This chapter is a joint work with Rodrigo Lambert and Jérôme Rousseau.

3.1 Shortest distance between observed orbits

Let $(M, \mathcal{A}, g, \mu)$ be a measure preserving dynamical system. In what follows, we present one of the main quantities of this chapter.

Definition 3.1.1. *Let $f : M \to Y \subset \mathbb{R}^n$ be a measurable function, called the observation. We define the shortest distance between two observed orbits as follows*

$$m_n^f(x,y) = \min_{i,j=0,\ldots,n-1} \left(d(f(g^i x), f(g^j y)) \right).$$

We will show that the shortest distance between two observed orbits is related with the correlation dimension of the pushforward measure $f_*\mu$. Recall that the pushforward measure is given by $f_*\mu(\cdot) := \mu(f^{-1}(\cdot))$.

We also recall that the lower and upper correlation dimension of $f_*\mu$ are denoted by $\underline{C}_{f_*\mu}$ and $\overline{C}_{f_*\mu}$.

Theorem 3.1.2. *Let $(M, \mathcal{A}, g, \mu)$ be a dynamical system. Consider an observation $f : M \to Y$ such that $\underline{C}_{f_*\mu} > 0$. Then for $\mu \otimes \mu$-almost every $(x, y) \in M \times M$*

$$\overline{\lim_{n\to\infty}} \frac{\log m_n^f(x,y)}{-\log n} \leq \frac{2}{\underline{C}_{f_*\mu}} \ . \tag{3.1}$$

We recall that the condition $\underline{C}_{f_*\mu} = 0$ can lead to unknown values for the above limit. However, one can observe that if $m_n^f = 0$ on a set of positive measure, our result implies immediately that $\underline{C}_{f_*\mu} = 0$. In the following example we present a measure μ and an observation f for which $C_{f_*\mu}$ is zero. Moreover, $m_n^f(x, y)$ is also zero for the system $(M, \mathcal{A}, g, \mu)$ such that μ is invariant by g.

Example 3.1.3. *Let $M \subset \mathbb{R}$ and $\mu = Leb$ the Lebesgue measure on M. Given $A \subset M$ with $\mu(A) > 0$ we define a function $f : M \to M$ by*

$$f(x) = \begin{cases} x, & \textit{if } x \in A^c \\ c, & \textit{if } x \in A \end{cases}$$

where $c \in \mathring{A}$ is a constant and $\mathring{A}$ is the interior of A.

Now, take any transformation g which is μ-invariant. By Poincaré's recurrence Theorem, we obtain that, for some finite n, the pair $(g^i x, g^j y)$ visits $A \times A$ for some $0 \leq i, j \leq n$. Therefore, for n sufficiently large $m_n^f(x, y) = 0$, and then $\underline{C}_{f_*\mu} = 0$.

Note that $Im(f) = A^c \cup \{c\}$. Then for $r < d(c, A^c)$, if $f(x) = c$ we get that $f_*\mu(B(f(x), r)) = \mu(A)$. On the other hand, if $f(x) \in A^c$, $f_*\mu(B(f(x), r)) \leq 2r$. Thus

$$\begin{aligned} \int_M f_*\mu(B(f(x),r)) \ d\mu(x) &= \int_A f_*\mu(B(f(x),r)) \ d\mu(x) + \int_{A^c} f_*\mu(B(f(x),r)) \ d\mu(x) \\ &\leq \int_A \mu(A) \ d\mu(x) + \int_{A^c} 2r \ d\mu(x) \\ &\leq \mu^2(A) + 2r\mu(A^c). \end{aligned}$$

It is easy to see that $C_{f_*\mu} = 0$. As an illustration, take: $M = [0, 1]$, $c = 3/4$, $A = [1/2, 1]$ and $g(x) = 2x \mod 1$.

Now let us present some technical notation as well as some tools that will be used to proof our results. For $\epsilon > 0$ we define

$$k_n = \frac{2 \log n + \log \log n}{\underline{C}_{f_*\mu} - \epsilon}.$$

We also define

$$A_{ij}^f(y) = g^{-i}\left[f^{-1} B\left(f(g^j y), e^{-k_n}\right)\right]$$

and

$$S_n^f(x,y) = \sum_{i,j=1,\ldots,n} \mathbb{1}_{A_{ij}^f(y)}(x).$$

Lemma 3.1.4. *Under the same conditions of Theorem 3.1.2 we have*

$$m_n^f(x,y) \leq e^{-k_n} \iff S_n^f(x,y) > 0 \ .$$

Proof. $S_n^f(x,y) > 0$ if and only if there exists at least a pair $(i,j) \in \{1,\ldots,n\} \times \{1,\ldots,n\}$ such that $x \in A_{ij}^f(y)$, i.e. $f(g^i x) \in B(f(g^j y), e^{-k_n})$, thus $d\left(f(g^i x), f(g^j y)\right) \leq e^{-k_n}$. This occurs if and only if, for all $1 \leq i,j \leq n$

$$m_n^f(x,y) = \min\left\{d(f(g^i x), f(g^j y))\right\} \leq e^{-k_n}.$$

□

In general the pushforward measure is not invariant. Nonetheless, since μ is g-invariant, given $y \in M$ and taking $\phi(y) = f_*\mu(B(f(y), e^{-k_n}))$ in Proposition 1.1.2 we have:

Lemma 3.1.5. *By invariance of μ follows the equality*

$$\int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right)\ d\mu(y) = \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y).$$

Proof of the Theorem 3.1.2. We first show that the event $\left\{m_n^f < e^{-k_n}\right\}$ occurs only finitely many times. By Lemma 3.1.4 and Markov inequality, we get that

$$\mu \otimes \mu\left(\left\{(x,y) : m_n^f(x,y) < e^{-k_n}\right\}\right) \leq \mathbb{E}\left(S_n^f\right).$$

By definition, the expected value of S_n^f is given by

$$\begin{aligned}
\mathbb{E}\left(S_n^f\right) &= \int\int \sum_{i,j=1,\ldots,n} \mathbb{1}_{A_{ij}^f(y)}(x)\ d\mu(x)\ d\mu(y) \\
&= \sum_{i,j=1,\ldots,n} \int \left(\int \mathbb{1}_{A_{ij}^f(y)}(x)\ d\mu(x)\right)\ d\mu(y) \\
&= \sum_{i,j=1,\ldots,n} \int \mu\left(f^{-1}B\left(f(g^j y), e^{-k_n}\right)\right)\ d\mu(y) \\
&= \sum_{i,j=1,\ldots,n} \int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right)\ d\mu(y) \\
&= n^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y),
\end{aligned}$$

where the last inequality follows is according to Lemma 3.1.5. Thus,

$$\mu \otimes \mu\left(\left\{(x,y) : m_n^f(x,y) < e^{-k_n}\right\}\right) \leq n^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y).$$

For large values of n, by definition of $\underline{C}_{f_*\mu}$ it holds

$$\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y) \leq e^{-k_n(\underline{C}_{f_*\mu}-\epsilon)}.$$

Moreover by definition of k_n,

$$\mu \otimes \mu\left(\left\{(x,y) : m_n^f(x,y) < e^{-k_n}\right\}\right) \leq n^2 e^{-k_n(\underline{C}_{f_*\mu}-\epsilon)} = \frac{1}{\log n}\ .$$

Choosing a subsequence $(n_\kappa)_{\kappa\in\mathbb{N}}$ such that $n_\kappa = \left\lceil e^{\kappa^2}\right\rceil$, we have that

$$\mu \otimes \mu\left(\left\{(x,y) : m_{n_\kappa}^f(x,y) < e^{-k_{n_\kappa}}\right\}\right) \leq \frac{1}{\kappa^2}\ .$$

Since the last quantity is summable in κ, the Borel-Cantelli lemma gives that if κ is large enough, then for $\mu\otimes\mu$-almost every pair (x,y) it holds

$$m_{n_\kappa}^f(x,y) \geq e^{-k_{n_\kappa}}$$

and then

$$\frac{\log m_{n_\kappa}^f(x,y)}{-\log n_\kappa} \leq \frac{1}{\underline{C}_{f_*\mu}-\epsilon}\left(2 + \frac{\log\log n_\kappa}{\log n_\kappa}\right). \tag{3.2}$$

We observe that for all n, there exists κ such that $e^\kappa \leq n \leq e^{\kappa+1}$. In addition, since $\left(m_n^f\right)_{n\in\mathbb{N}}$ is a decreasing sequence and $\log z$ is a monotone function, it follows that

$$\frac{\log m_{n_\kappa}^f(x,y)}{-\log n_{\kappa+1}} \leq \frac{\log m_n^f(x,y)}{-\log n} \leq \frac{\log m_{n_{\kappa+1}}^f(x,y)}{-\log n_\kappa}.$$

Taking the limit superior in the above inequalities and observing that $\lim_{\kappa\to\infty}\frac{\log n_\kappa}{\log n_{\kappa+1}} = 1$, we get

$$\overline{\lim_{n\to\infty}}\frac{\log m_n^f(x,y)}{-\log n} = \overline{\lim_{\kappa\to\infty}}\frac{\log m_{n_\kappa}^f(x,y)}{-\log n_\kappa}.$$

Thus, by (3.2) we obtain

$$\overline{\lim_{n\to\infty}}\frac{\log m_n^f(x,y)}{-\log n} \leq \frac{2}{\underline{C}_{f_*\mu}-\epsilon}.$$

Since ϵ can be arbitrarily small, the prove is complete. □

As in [18], to obtain an equality in (3.1), we will need more assumptions on the system.

(H1) Let $\mathcal{H}^\alpha(M,\mathbb{R})$ be the space of Hölder observables. For all $\psi,\phi\in\mathcal{H}^\alpha(M,\mathbb{R})$ and for all $n\in\mathbb{N}^*$, we have:

$$\left|\int_M \psi\circ f(g^n x)\phi\circ f(x)\ d\mu(x) - \int_X \psi\circ f\ d\mu\int_M \phi\circ f\ d\mu\right| \leq \|\psi\circ f\|_\alpha\|\phi\circ f\|_\alpha\theta_n$$

with $\theta_n = a^n$ and $a\in[0,1)$.

(HA) There exist $r_0 > 0$, $\xi \geq 0$ and $\beta > 0$ such that for $f_*\mu$-almost every $y \in \mathbb{R}^n$ and any $r_0 > r > \rho > 0$,

$$f_*\mu(B(y, r+\rho)\backslash B(y, r-\rho)) \leq r^{-\xi}\rho^{\beta}.$$

One can observe that, if f is Lipschitz, assuming hypothesis (H1) is weaker than assuming a exponential decay of correlations (for Hölder observables) for the system $(M, \mathcal{A}, g, \mu)$. Indeed, note that if f is Lipschitz then $\psi \circ f$ is Hölder for every Hölder function ψ.

Theorem 3.1.6. *Let $(M, \mathcal{A}, T, \mu)$ be a dynamical system and consider a Lipschitz observation $f : M \to Y$ such that $\underline{C}_{f_*\mu} > 0$. If the system satisfies (H1) and (HA), then for $\mu \otimes \mu$-almost every $(x, y) \in M \times M$*

$$\varliminf_{n\to\infty} \frac{\log m_n^f(x,y)}{-\log n} \geq \frac{2}{\overline{C}_{f_*\mu}} \ .$$

Furthermore, in the case that $\underline{C}_{f_*\mu} = \overline{C}_{f_*\mu} = C_{f_*\mu}$ we get

$$\lim_{n\to\infty} \frac{\log m_n^f(x,y)}{-\log n} = \frac{2}{C_{f_*\mu}} \ .$$

To prove Theorem 3.1.6, the main difficulty is that we cannot apply mixing as simply. In particular, we can only apply mixing to Hölder observables and indicator functions are not even continuous. To overthrow this difficulty, we will first prove that a particular function is Hölder in order to apply the mixing property, then, in the proof of Theorem 3.1.6, we will also approximate characteristic functions by Lipschitz functions in order to apply mixing again.

Lemma 3.1.7. *Let $(M, \mathcal{A}, g, \mu)$ be a dynamical system with observation f. If it satisfies (HA), then there exist $0 < r_0 < 1$, $c \geq 0$ and $\zeta \geq 0$ such that for any $0 < r < r_0$, the function $\psi_1 : x \mapsto f_*\mu(B(x, r))$ belongs to $\mathcal{H}^{\alpha}(M, \mathbb{R})$ and*

$$||\psi_1||_{\alpha} \leq 2r^{-\zeta} \ .$$

Proof. Let $x, y \in M$ and $0 < r < r_0$, if $||x - y|| < r$ we have

(i) $B(x, r - ||x-y||) \subset B(x, r)$ and $B(x, r - ||x-y||) \subset B(y, r)$;

(ii) $B(x, r) \subset B(x, r + ||x-y||)$ and $B(y, r) \subset B(x, r + ||x-y||)$.

Then it holds

$$f_*\mu\left(B(x, r - ||x-y||)\right) \leq f_*\mu(B(x,r)) \text{ and } f_*\mu(B(y,r)) \leq f_*\mu(B(x, r + ||x-y||)) \ .$$

Hence

$$||f_*\mu(B(y,r)) - f_*\mu(B(x,r))|| \leq f_*\mu(B(x, r + ||x-y||)) - f_*\mu(B(x, r - ||x-y||)).$$

Thus, by (HA),

$$||f_*\mu(B(y,r)) - f_*\mu(B(x,r))|| \leq r^{-\xi}||x-y||^{\beta}.$$

On the other hand, if $||x-y|| \geq r$ then

$$||f_*\mu(B(y,r)) - f_*\mu(B(x,r))|| \leq 2 \leq \frac{2}{r}||x-y|| \ .$$

Thus, ψ_1 is Hölder and $||\psi_1||_\alpha \leq 2r^{-\zeta}$ with $\zeta = \max\{1, \xi\}$.

□

In the sequel, we present the proof of Theorem 3.1.6. This proof mainly follows the ideas of the proof of [18, Theorem 5].

Proof of Theorem 3.1.6. Without loss of generality we will assume that $\theta_\ell = e^{-\ell}$. Let $b < -4$. Given $\epsilon > 0$, we define

$$k_n = \frac{2\log n + b \log\log n}{\overline{C}_{f_*\mu} + \epsilon}.$$

Remember that in the proof of Theorem 3.1.2 we had

$$\mathbb{E}\left(S_n^f\right) = n^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\, d\mu(y).$$

In addition, by Lemma 3.1.4

$$\begin{aligned}\mu\otimes\mu\left(\left\{(x,y) : m_n^f(x,y) \geq e^{-k_n}\right\}\right) &\leq \mu\otimes\mu\left(\left\{(x,y) : S_n^f(x,y) = 0\right\}\right) \\ &\leq \mu\otimes\mu\left(\left\{(x,y) : \left|S_n^f(x,y) - \mathbb{E}\left(S_n^f\right)\right| \geq \left|\mathbb{E}\left(S_n^f\right)\right|\right\}\right).\end{aligned}$$

By Chebyshev's inequality we get that this last quantity is limited by $\frac{\operatorname{var}\left(S_n^f\right)}{\mathbb{E}\left(S_n^f\right)^2}$. And thus,

$$\mu\otimes\mu\left(\left\{(x,y) : m_n^f(x,y) \geq e^{-k_n}\right\}\right) \leq \frac{\operatorname{var}\left(S_n^f\right)}{\mathbb{E}(S_n^f)^2}.$$

We now proceed to estimate the variance of S_n^f.

We see at once that

$$\begin{aligned}\operatorname{var}\left(S_n^f\right) &= \sum_{1\leq i,i',j,j'\leq n} \operatorname{cov}\left(\mathbb{1}_{A_{ij}^f}, \mathbb{1}_{A_{i'j'}^f}\right) \\ &= \sum_{1\leq i,i',j,j'\leq n} \int\int \mathbb{1}_{A_{ij}^f(y)}\mathbb{1}_{A_{i'j'}^f(y)} - \int\int \mathbb{1}_{A_{ij}^f(y)} \int\int \mathbb{1}_{A_{i'j'}^f(y)} \\ &= \sum_{1\leq i,i',j,j'\leq n} \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y), e^{-k_n}\right)}(g^i x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y), e^{-k_n}\right)}(g^{i'} x) \\ &- n^4\left(\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\, d\mu(y)\right)^2. \end{aligned} \tag{3.3}$$

Since we would like to apply the mixing property to estimate the previous sum, we will present an approximation for characteristic functions by Lipschitz functions following the construction of the proof of Lemma 9 in [51].

Let $\rho > 0$ (to de defined properly later). Let $\eta_{e^{-k_n}} : [0,\infty) \to \mathbb{R}$ be the $\frac{1}{\rho e^{-k_n}}$-Lipschitz function such that $\mathbb{1}_{[0,e^{-k_n}]} \leq \eta_{e^{-k_n}} \leq \mathbb{1}_{[0,(1+\rho)e^{-k_n}]}$ and set $\varphi_{f(y),e^{-k_n}}(x) = \eta_{e^{-k_n}}(d(f(y),x))$. Since f is L-Lipschitz it follows that $\varphi_{f(y),e^{-k_n}} \circ f$ is $\frac{L}{\rho e^{-k_n}}$-Lipschitz. Moreover, we have

$$\begin{aligned}
\mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(x) &= \mathbb{1}_{B\left(f(g^j y),e^{-k_n}\right)}(f(x)) \\
&= \mathbb{1}_{[0,e^{-k_n}]}(d(f(g^j y),f(x))) \\
&\leq \eta_{e^{-k_n}}(d(f(g^j y),f(x))) \\
&= \varphi_{f(g^j y),e^{-k_n}}(f(x)). \qquad (3.4)
\end{aligned}$$

We are now able to apply the mixing property. We will consider four different cases. Let us fix $\ell = \ell(n) = \log(n^\gamma)$ for some $\gamma > 0$ to be defined later.

Case 1: $|i - i'| > \ell$. By (H1) and (3.4) we obtain

$$\begin{aligned}
&\int\int \mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(g^i x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y),e^{-k_n}\right)}(g^{i'} x)\; d\mu(x)\; d\mu(y) \\
=\;& \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(g^{i-i'} x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y),e^{-k_n}\right)}(x)\; d\mu(x)\; d\mu(y) \\
\leq\;& \int\left(\int \varphi_{f(g^j y),e^{-k_n}}(f(g^{i-i'}x))\; d\mu(x) \int \varphi_{f(g^{j'} y),e^{-k_n}}(f(x))\; d\mu(x)\right)\; d\mu(y) \\
+\;& \theta_g \left\|\varphi_{f(g^j y),e^{-k_n}}\right\| \left\|\varphi_{f(g^{j'} y),e^{-k_n}}\right\| \\
\leq\;& \frac{L^2}{\rho^2 e^{-2k_n}}\theta_\ell + \int f_*\mu\left(B\left(f(g^j y),(1+\rho)e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y),(1+\rho)e^{-k_n}\right)\right)\; d\mu(y).
\end{aligned}$$

To estimate the second part of the last inequality, we can observe that using (HA) we obtain

$$\begin{aligned}
&\int f_*\mu\left(B\left(f(g^j y),(1+\rho)e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y),(1+\rho)e^{-k_n}\right)\right)\; d\mu(y) \\
-\;& \int f_*\mu\left(B\left(f(g^j y),e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y),e^{-k_n}\right)\right)\; d\mu(y) \\
\leq\;& \int f_*\mu\left(B\left(f(g^j y),(1+\rho)e^{-k_n}\right)\right)\left(f_*\mu\left(B\left(f(g^{j'} y),(1+\rho)e^{-k_n}\right)\right) - f_*\mu\left(B\left(f(g^{j'} y),e^{-k_n}\right)\right)\right)\; d\mu(y) \\
+\;& \int f_*\mu\left(B\left(f(g^{j'} y),e^{-k_n}\right)\right)\left(f_*\mu\left(B\left(f(g^j y),(1+\rho)e^{-k_n}\right)\right) - f_*\mu\left(B\left(f(g^j y),e^{-k_n}\right)\right)\right)\; d\mu(y) \\
\leq\;& \int f_*\mu\left(B\left(f(g^j y),(1+\rho)e^{-k_n}\right)\right) e^{\xi k_n}\rho^\beta\; d\mu(y) + \int f_*\mu\left(B\left(f(g^{j'} y),e^{-k_n}\right)\right) e^{\xi k_n}\rho^\beta\; d\mu(y).
\end{aligned}$$

Therefore, choosing $\rho = n^{-\delta}$ for some $\delta > 0$ to be defined later, we have for n large enough

$$\begin{aligned}
&\int\int \mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(g^i x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y),e^{-k_n}\right)}(g^{i'} x)\ d\mu(x)\ d\mu(y)\\
\leq\ & \frac{L^2}{\rho^2 e^{-2k_n}}\theta_\ell + 2e^{\xi k_n}\rho^\beta \int f_*\mu\left(B\left(f(g^j y), 2e^{-k_n}\right)\right)\ d\mu(y)\\
&+\int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y).
\end{aligned}$$

To deal with the third term of the last inequality we need to consider two different cases.

Case 1.1: $|j-j^{'}|>\ell$. Using Lemma 3.1.7, we can use the mixing property (H1)

$$\begin{aligned}
&\int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y)\\
\leq\ & \int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right)\ d\mu(y) \int f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y)\\
&+\theta_\ell \left\|f_*\mu\left(B\left(\cdot, e^{-k_n}\right)\right)\right\| \left\|f_*\mu\left(B\left(\cdot, e^{-k_n}\right)\right)\right\|\\
\leq\ & 4\theta_\ell e^{2\zeta k_n} + \left(\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y)\right)^2.
\end{aligned}$$

Case 1.2: $|j-j^{'}|\leq\ell$. Using Holder's inequality together with Lemma 3.1.5 we have

$$\begin{aligned}
&\int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y)\\
\leq\ & \left(\int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right)^2\ d\mu(y)\right)^{1/2} \left(\int f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)^2\ d\mu(y)\right)^{1/2}\\
=\ & \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2\ d\mu(y).
\end{aligned}$$

Combining these cases we can write

$$\begin{aligned}
&\sum_{|i-i'|>\ell,|j-j'|>\ell} \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(g^i x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y),e^{-k_n}\right)}(g^{i'} x)\ d\mu(x)\ d\mu(y)\\
+\ & \sum_{|i-i'|>\ell,|j-j'|\leq\ell} \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y),e^{-k_n}\right)}(g^i x)\mathbb{1}_{f^{-1}B\left(f(g^{j'} y),e^{-k_n}\right)}(g^{i'} x)\ d\mu(x)\ d\mu(y)\\
\leq\ & n^4 L^2\rho^{-2}e^{2k_n}\theta_\ell + 2n^2 e^{\xi k_n}\rho^\beta \sum_{|j-j'|>\ell}\int f_*\mu\left(B\left(f(g^j y), 2e^{-k_n}\right)\right)\ d\mu(y)\\
+\ & 4n^4\theta_\ell e^{2\zeta k_n} + n^2\sum_{|j-j'|>\ell}\left(\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y)\right)^2\\
+\ & n^2\sum_{|j-j'|\leq\ell}\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2\ d\mu(y)\\
\leq\ & n^4 L^2\rho^{-2}e^{2k_n}\theta_\ell + 2n^4 e^{\xi k_n}\rho^\beta \int f_*\mu\left(B\left(f(g^j y), 2e^{-k_n}\right)\right)\ d\mu(y) + 4n^4\theta_\ell e^{2\zeta k_n}\\
+\ & n^4\left(\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right) d\mu(y)\right)^2 + 2n^3\ell\int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2 d\mu(y). \quad (3.5)
\end{aligned}$$

Case 2: $|i-i'| \leq \ell$. At first we observe that for all $l, m \in \mathbb{N}$ it holds

$$\mathbb{1}_{f^{-1}B\left(f(g^l y), e^{-k_n}\right)}(g^m x) = \mathbb{1}_{f^{-1}B\left(f(g^m x), e^{-k_n}\right)}(g^l y).$$

And this implies that

$$\begin{aligned} & \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y), e^{-k_n}\right)}(g^i x) \mathbb{1}_{f^{-1}B\left(f(g^{j'} y), e^{-k_n}\right)}(g^{i'} x)\ d\mu(x)\ d\mu(y) \\ = & \int\int \mathbb{1}_{f^{-1}B\left(f(g^i x), e^{-k_n}\right)}(g^j y) \mathbb{1}_{f^{-1}B\left(f(g^{i'} x), e^{-k_n}\right)}(g^{j'} y)\ d\mu(y)\ d\mu(x). \end{aligned} \tag{3.6}$$

Case 2.1: $|j-j^{'}| > \ell$. In this case, we can proceed in the same way as in the second sum using the above symmetry.

Case 2.2: $|j-j^{'}| \leq \ell$. For this, the boundedness of the indicator function and Lemma 3.1.5 give that,

$$\begin{aligned} & \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y), e^{-k_n}\right)}(g^i x) \mathbb{1}_{f^{-1}B\left(f(g^{j'} y), e^{-k_n}\right)}(g^{i'} x)\ d\mu(x)\ d\mu(y) \\ \leq & \int\int \mathbb{1}_{f^{-1}B\left(f(g^j y), e^{-k_n}\right)}(g^i x)\ d\mu(x)\ d\mu(y) \\ \leq & \int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right)\ d\mu(y) \\ = & \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y). \end{aligned} \tag{3.7}$$

Finally, for these cases we have

$$\begin{aligned} & \sum_{|i-i'|\leq\ell, |j-j'|>\ell} \int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y) \\ & + \sum_{|i-i'|\leq\ell, |j-j'|\leq\ell} \int f_*\mu\left(B\left(f(g^j y), e^{-k_n}\right)\right) f_*\mu\left(B\left(f(g^{j'} y), e^{-k_n}\right)\right)\ d\mu(y) \\ \leq\ & n^4 L^2 e^{2k_n}\theta_\ell + 2n\ell \sum_{|j-j'|>\ell} \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2\ d\mu(y) \\ +\ & 2n\ell \sum_{|j-j'|\leq\ell} \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y) \\ \leq\ & n^4 L^2 \rho^{-2} e^{2k_n}\theta_\ell + 2n^3\ell \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2\ d\mu(y) \\ +\ & 4n^2\ell^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y). \end{aligned} \tag{3.8}$$

Putting all the previous estimates in (3.3) we obtain

$$\begin{aligned} \frac{\mathrm{var}\left(S_n^f\right)}{\mathbb{E}\left(S_n^f\right)^2} \leq & \frac{2n^4 L^2 \rho^{-2} e^{2k_n}\theta_\ell + 4n^4\theta_\ell e^{2\zeta k_n} + 2n^4 e^{\xi k_n}\rho^\beta \int f_*\mu\left(B\left(f(g^j y), 2e^{-k_n}\right)\right)\ d\mu(y)}{\left(n^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y)\right)^2} \\ + & \frac{4n^2\ell^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y) + 4n^3\ell \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)^2\ d\mu(y)}{\left(n^2 \int f_*\mu\left(B\left(f(y), e^{-k_n}\right)\right)\ d\mu(y)\right)^2}. \end{aligned} \tag{3.9}$$

By definition of $\overline{C}_{f_*\mu}$ we have for n large enough

$$\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\geq e^{-k_n(\overline{C}_{f_*\mu}+\epsilon)}.$$

Recalling that $\rho=n^{-\delta}$, we can observe that we can choose δ large enough (depending on $\xi,\beta,\overline{C}_{f_*\mu},\underline{C}_{f_*\mu},b$ and ϵ) so that

$$\frac{2n^4e^{\xi k_n}\rho^\beta\int f_*\mu\left(B\left(f(y),2e^{-k_n}\right)\right)\ d\mu(y)}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2}\leq\frac{1}{n}. \tag{3.10}$$

Recalling that $\ell=\log(n^\gamma)$, we can observe that we can choose γ large enough (depending on $\delta,\overline{C}_{f_*\mu},\zeta,b$ and ϵ) so that

$$\frac{2n^4L^2\rho^{-2}e^{2k_n}\theta_\ell}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2}\leq\frac{1}{n} \tag{3.11}$$

and so that

$$\frac{4n^4\theta_\ell e^{\zeta k_n}}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2}\leq\frac{1}{n}. \tag{3.12}$$

For the fourth term we have

$$\begin{aligned}\frac{4n^2\ell^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2} &\leq \frac{4\ell^2}{n^2e^{-k_n(\overline{C}_{f_*\mu}+\epsilon)}}\\ &\leq 4\gamma^2(\log n)^{2+b}. \end{aligned}\tag{3.13}$$

To estimate the last term, we will use the following lemma.

Lemma 3.1.8 (Lemma 14 [18])**.** *Let $Z\subset\mathbb{R}^n$ and let ν be a probability measure on Z. There exists a constant $K>0$ depending only on n such that for every r small enough*

$$\int_Z\mu\left(B(y,r)\right)^2d\nu(y)\leq K\left(\int_Z\mu\left(B(y,r)\right)d\nu(y)\right)^{3/2}.$$

Applying the previous lemma with $Z=Y$ and $\nu=f_*\mu$ we obtain

$$\begin{aligned}\frac{4n^3\ell\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)^2\ d\mu(y)}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2} &\leq \frac{4n^3\ell K\left(\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^{3/2}}{\left(n^2\int f_*\mu\left(B\left(f(y),e^{-k_n}\right)\right)\ d\mu(y)\right)^2}\\ &\leq \frac{4\ell K}{n\left(\int f_*\mu(B(f(y),e^{-k_n}))\ d\mu(y)\right)^{1/2}}\\ &\leq \frac{4\ell K}{n}e^{\frac{k_n(\overline{C}_{f_*\mu}+\epsilon)}{2}}\\ &\leq 4\ell K(\log n)^{b/2}\\ &\leq 4K\gamma(\log n)^{1+\frac{b}{2}}. \end{aligned}\tag{3.14}$$

Taking $b < -4$ and substituting (3.10), (3.11), (3.12), (3.13) and (3.14) into (3.9) we get

$$\begin{aligned} \mu \otimes \mu \left(\left\{ (x,y) : m_n^f(x,y) \geq e^{-k_n} \right\} \right) &\leq \frac{\operatorname{var}\left(S_n^f\right)}{\mathbb{E}\left(S_n^f\right)^2} \\ &\leq \mathcal{O}((\log n)^{-1}). \end{aligned} \tag{3.15}$$

Thus, taking a subsequence $n_\kappa = \left\lceil e^{\kappa^2} \right\rceil$. As in the proof of Theorem 3.1.2, by the Borel-Cantelli Lemma we obtain

$$\varliminf_{n\to\infty} \frac{\log m_n^f(x,y)}{-\log n} = \varliminf_{\kappa\to\infty} \frac{\log m_{n_\kappa}^f(x,y)}{-\log n_\kappa} \geq \frac{2}{\underline{C}_{f_*\mu} + \epsilon}.$$

Since ϵ can be arbitrarily small, the theorem follows. □

3.2 Shortest distance between orbits for random dynamical systems

Let $M \subset \mathbb{R}^n$ and let $(\Omega, \theta, \mathbb{P})$ be a probability measure preserving system, where Ω is a metric space and $B(\Omega)$ its Borelian σ-algebra. We first introduce the notion of random dynamical system.

Definition 3.2.1. *A random dynamical system $\mathcal{G} = (g_\omega)_{\omega\in\Omega}$ on M over $(\Omega, B(\Omega), \mathbb{P}, \theta)$ is generated by maps g_ω such that $(\omega, x) \mapsto g_\omega(x)$ is measurable and satisfies:*

$$g_\omega^0 = Id \text{ for all } \omega \in \Omega,$$

$$g_\omega^n = g_{\theta^{n-1}(\omega)} \circ \cdots \circ g_{\theta(\omega)} \circ g_\omega \text{ for all } n \geq 1.$$

The map $S : \Omega \times M \to \Omega \times M$ defined by $S(\omega, x) = (\theta(\omega), g_\omega(x))$ is the dynamics of the random dynamical systems generated by $\mathcal{G}$ and is called skew-product.

Definition 3.2.2. *A probability measure μ is said to be an invariant measure for the random dynamical system $\mathcal{G}$ if it satisfies*

1. *μ is S-invariant*
2. *$\pi_*\mu = \mathbb{P}$*

where $\pi : \Omega \times M \to \Omega$ is the canonical projection.

Let $(\mu_\omega)_\omega$ denote the decomposition of μ on M, that is, $d\mu(\omega, x) = d\mu_\omega(x)d\mathbb{P}(\omega)$. We denote by $\nu = \int \mu_\omega d\mathbb{P}$ the marginal of ν on M.

To obtain a result that links the shortest distance between orbits and random dynamical systems we need to assume a hypothesis for the measure and an (annealed) decay of correlations for the random dynamical system. Namely,

(a) There exist $r_0 > 0$, $\xi \leq 0$ and $\beta > 0$ such that for almost every $y \in M$ and any $r_0 > r > \rho > 0$,

$$\nu(B(y, r+\rho) \backslash B(y, r-\rho)) \leq r^{-\xi} \rho^{\beta}.$$

(b) (Annealed decay of correlations) $\forall n \in \mathbb{N}^*$, ψ and ϕ Hölder observables from M to $\mathbb{R}$,

$$\left| \int_{\Omega \times M} \psi(g_\omega^n(x)) \phi(x) \, d\mu(\omega, x) - \int_{\Omega \times M} \psi \, d\mu \int_{\Omega \times M} \phi \, d\mu \right| \leq \|\psi\|_\alpha \|\phi\|_\alpha \theta_n$$

with $\theta_n = e^{-n}$.

Definition 3.2.3. *We define the shortest distance between two random orbits as follows*

$$m_n^{\omega,\tilde{\omega}}(x, \tilde{x}) = \min_{i,j=0,\dots,n-1} \left(d\left(g_\omega^i(x), g_{\tilde{\omega}}^j(\tilde{x}) \right) \right).$$

Theorem 3.2.4. *Let $\mathcal{G}$ be a random dynamical system on M over $(\Omega, B(\Omega), \mathbb{P}, \theta)$ with an invariant measure μ such that $\underline{C}_\nu > 0$. Then for $\mu \otimes \mu$-almost every $(\omega, x, \tilde{\omega}, \tilde{x}) \in \Omega \times M \times \Omega \times M$,*

$$\overline{\lim_{n\to\infty}} \frac{\log m_n^{\omega,\tilde{\omega}}(x, \tilde{x})}{-\log n} \leq \frac{2}{\underline{C}_\nu}.$$

Moreover, if the random dynamical system satisfies assumptions (a) and (b), then

$$\underline{\lim_{n\to\infty}} \frac{\log m_n^{\omega,\tilde{\omega}}(x, \tilde{x})}{-\log n} \geq \frac{2}{\overline{C}_\nu}$$

and if $\underline{C}_\nu = \overline{C}_\nu$, then

$$\lim_{n\to\infty} \frac{\log m_n^{\omega,\tilde{\omega}}(x, \tilde{x})}{-\log n} = \frac{2}{C_\nu}.$$

Proof. This proof will follow the idea given in [49].

This theorem is proved using Theorem 3.1.2 and Theorem 3.1.6 applied to the dynamical system $(\Omega \times M, B(\Omega \times M), \mu, S)$ with the observation f defined by

$$f : \Omega \times M \to M$$
$$(\omega, x) \mapsto x.$$

With this observation, for all z and $t \in \Omega \times M$ we can link the shortest distance between two observed orbits and the shortest distance between two random orbits. Set $z = (\omega, x)$ and $t = (\tilde{\omega}, \tilde{x})$ then

$$\begin{aligned} m_n^f(z,t) &= \min_{i,j=0,\dots,n-1} \left(d\left(f\left(S^i(\omega, x)\right), f\left(S^j(\tilde{\omega}, \tilde{x})\right)\right)\right) \\ &= \min_{i,j=0,\dots,n-1} \left(d\left(g_\omega^i(x), g_{\tilde{\omega}}^j(\tilde{x}) \right)\right) \\ &= m_n^{\omega,\tilde{\omega}}(x, \tilde{x}). \end{aligned}$$

Moreover, we can identify the pushforward measure: $f_*\mu = \nu$. Therefore, in view of the lower and upper correlation dimensions, the following statement finishes the proof

$$\underline{C}_{f_*\mu} = \underline{C}_\nu \text{ and } \overline{C}_{f_*\mu} = \overline{C}_\nu.$$

□

3.2.1 Examples

We will present some examples of random dynamical systems where for which we can apply the last statement.

Non-i.i.d. random dynamical system

The first example is a non-i.i.d. random dynamical system for which it was computed recurrence rates in [40] and hitting times statistics in [49].

Consider the two linear maps which preserve Lebesgue measure Leb on $M = \mathbb{T}^1$, the one-dimensional torus:

$$\begin{aligned} g_1 : M &\to M & \text{and} \quad g_2 : M &\to M \\ x &\mapsto 2x & x &\mapsto 3x. \end{aligned}$$

The following skew product gives the dynamics of the random dynamical system:

$$\begin{aligned} S : \Omega \times M &\to \Omega \times M \\ (\omega, x) &\mapsto (\theta(\omega), g_\omega(x)) \end{aligned}$$

with $\Omega = [0,1]$, $g_\omega = g_1$ if $\omega \in [0, 2/5)$ and $g_\omega = g_2$ if $\omega \in [2/5, 1]$ where ω is the following piecewise linear map:

$$\theta(\omega) = \begin{cases} 2\omega & \text{if } \omega \in [0, 1/5) \\ 3\omega - 1/5 & \text{if } \omega \in [1/5, 2/5) \\ 2\omega - 4/5 & \text{if } \omega \in [2/5, 3/5) \\ 3\omega/2 - 1/2 & \text{if } \omega \in [3/5, 1]. \end{cases}$$

Note that the random orbit is constructed by choosing one of these two maps following a Markov process with the stochastic matrix

$$A = \begin{pmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{pmatrix}.$$

The associated skew-product S is Leb $\otimes$ Leb-invariant. It is easy to check that Lebesgue measure satisfies (a). Moreover, by [13] the skew product S has an exponential decay of correlations. Since in this example $C_\nu = 1$ with $\nu = \text{Leb}$, Theorem 3.2.4 implies that for Leb $\otimes$ Leb $\otimes$ Leb $\otimes$ Leb-almost every $(\omega, x, \tilde{\omega}, \tilde{x}) \in [0,1] \times \mathbb{T}^1 \times [0,1] \times \mathbb{T}^1$,

$$\lim_{n\to\infty} \frac{\log m_n^{\omega,\tilde{\omega}}(x, \tilde{x})}{-\log n} = 2.$$

Randomly perturbed dynamical systems

Consider a deterministic dynamical system (M, g, μ) where M is a compact Riemannian manifold, g is a map and μ is a g-invariant probability measure. We will present a random dynamical system constructed by perturbing the map g with a random additive noise. For $\epsilon > 0$, set $\Lambda_\epsilon = B(0, \epsilon)$ and let $\mathcal{P}_\epsilon$ be a probability measure on Λ_ϵ. For each $\omega \in \Lambda_\epsilon$, we denote the family of transformations $\{g_\omega\}_\omega$ where the map $g_\omega : M \to M$ are given by

$$g_\omega(x) = g(x) + \omega.$$

Denote $\mathcal{G}$ the i.i.d dynamical system on M over $(\Lambda_\epsilon^{\mathbb{N}}, \mathcal{P}_\epsilon^{\mathbb{N}}, \sigma)$. In the case where $M = \mathbb{T}^d$, for some expanding and piecewise expanding maps, if ϵ is sufficiently small, it was proved (see e.g. [11, 14, 57]) that the random dynamical system has a stationary measure μ_ϵ absolutely continuous with respect to Lebesgue measure with density h_ϵ such that $0 < \underline{h}_\epsilon \leq h_\epsilon \leq \overline{h}_\epsilon < \infty$ and the system has a superpolynomial decay of correlation. Thus, since the assumptions (a) and (b) are satisfied one can apply Theorem 3.2.4 and obtain behavior of the shortest distance $m_n^{\omega,\tilde{\omega}}$.

Random hyperbolic toral automorphisms

A linear toral automorphism is a map $g : \mathbb{T}^2 \to \mathbb{T}^2$ defined by the matrix action $x \mapsto Ax$, where the matrix A has integer entries and $\det A = \pm 1$. We say that g is hyperbolic if A has eigenvalues with modulus different from 1. For more simplicity, we will use the notation A for both the matrix and the associated automorphism.

For an hyperbolic toral automorphism A, we denote E_u^A the subspace spanned by e_u^A, the eigenvector associated to the eigenvalue whose absolute value is greater than 1 and we denote E_s^A the subspace spanned by e_s^A, the eigenvector associated to the eigenvalue whose absolute value is less than 1.

Following the definition from [12], we say that a pair (A_0, A_1) of hyperbolic toral automorphisms has the cone property if there exists an expansion cone $\mathcal{E}$ such that

1. $A_i\mathcal{E} \subset \mathcal{E}$,
2. there exists $\lambda_\mathcal{E} > 1$ such that $|A_i x| \geq \lambda_\mathcal{E}|x|$ for $x \in \mathcal{E}$,
3. $E_u^{A_i} \cap \partial\mathcal{E} = 0$, where $\partial\mathcal{E}$ denote the boundary of $\mathcal{E}$,

and there exists a contraction cone $\mathcal{C}$ such that $\mathcal{C} \cap \mathcal{E} = 0$ and

1. $A_i^{-1}\mathcal{C} \subset \mathcal{C}$,
2. there exists $\lambda_\mathcal{C} < 1$ such that $|A_i^{-1}x| \geq \lambda_\mathcal{C}^{-1}|x|$ for $x \in \mathcal{C}$,

3. $E_s^{A_i} \cap \partial\mathcal{C} = 0$.

One can observe that for example a pair of hyperbolic toral automorphisms with positive entries, or a pair of hyperbolic toral automorphisms with negative entries, has the cone property.

Let $\Lambda = \{0, 1\}$ and $\theta = \sigma$ be the left shift on $\Lambda^{\mathbb{N}}$. Let A_0, A_1 two hyperbolic automorphisms satisfying the cone property. Let A_0 be chosen with a probability q and A_1 with a probability $1-q$, i.e. $\mathbb{P} = \mathcal{P}^{\mathbb{N}}$ with $\mathcal{P}(0) = q$ and $\mathcal{P}(1) = 1-q$.

Then, for the i.i.d. random dynamical system on $\mathbb{T}^2$ over $(\Lambda^{\mathbb{N}}, \mathcal{P}^{\mathbb{N}}, \sigma)$, the Lebesgue measure is stationary (and thus hypothesis (a) is satisfied) and the system has an exponential decay of correlations (see [12]).

Note that $\nu = \mathrm{Leb} \otimes \mathrm{Leb}$ implies that $C_\nu = 2$. Then, by Theorem 3.2.4 we get for $\mathbb{P} \otimes \mathrm{Leb} \otimes \mathbb{P} \otimes \mathrm{Leb}$-almost every $(\omega, x, \tilde{\omega}, \tilde{x}) \in \Omega \times \mathbb{T}^2 \times \Omega \times \mathbb{T}^2$,

$$\lim_{n\to\infty} \frac{\log m_n^{\omega,\tilde{\omega}}(x, \tilde{x})}{-\log n} = 1.$$

3.3 Matching strings in encoded sequences

The present section is dedicated to study of longest common substring of encoded sequences. We start by presenting some terminology and definitions, in order to introduce the problem.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where $\Omega = \chi^{\mathbb{N}}$ for some alphabet χ, $\mathcal{F}$ the sigma-algebra generated by the n-cylinders in Ω, and $\mathbb{P}$ is a stationary probability measure on $\mathcal{F}$. If σ is the left shift on Ω, we can see $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ as a symbolic dynamical system with $\mathbb{P}$ σ-invariant. Let $\tilde{\Omega} = \tilde{\chi}^{\mathbb{N}}$ for some alphabet $\tilde{\chi}$ and $\tilde{\mathcal{F}}$ the sigma-algebra generated by the n-cylinders in $\tilde{\Omega}$.

Definition 3.3.1. *Let $f : \Omega \to \tilde{\Omega}$ be a code. Given two sequences $x, y \in \Omega$, we define the n-length of the longest common substring for the encoded pair $(f(x), f(y))$ by*

$$M_n^f(x, y) = \max\left\{k \ : f(x)_i^{i+k-1} = f(y)_j^{j+k-1} \text{ for some } 0 \leq i, j \leq n-k\right\},$$

where $f(x)_i^{i+k-1}$ and $f(y)_j^{j+k-1}$ denote the substrings of length k beginning in $f(x)_i$ and $f(y)_j$ respectively.

For $y \in \Omega$ (respectively $\tilde{\Omega}$) we denote by $C_n(y)$ the n-cylinder containing y, that is, the set of sequences $z \in \Omega$ (respectively $\tilde{\Omega}$) such that $z_i = y_i$ for any $i = 0, \ldots, n-1$. We denote $\mathcal{F}_0^n$ (respectively $\tilde{\mathcal{F}}_0^n$) the sigma-algebra on Ω (respectively $\tilde{\Omega}$) generated by all n-cylinders.

Definition 3.3.2. *The lower and upper Rényi entropies of a measure* $\mathbb{P}$ *are defined as*

$$\underline{H}_2(\mathbb{P}) = -\varliminf_{k\to\infty} \frac{1}{k} \log \sum_{C_k} \mathbb{P}(C_k)^2 \quad \textit{and} \quad \overline{H}_2(\mathbb{P}) = -\varlimsup_{k\to\infty} \frac{1}{k} \log \sum_{C_k} \mathbb{P}(C_k)^2$$

where the sums are taken over all k-cylinders. When the limit exists we denote by $H_2(\mathbb{P})$ *the common value.*

In general, the existence of the Rényi entropy is not known. However, it was computed in some particular cases: Bernoulli shift, Markov chains and Gibbs measure of a Hölder-continuous potential [31]. The existence was also proved for ϕ-mixing measures [39], for weakly ψ-mixing processes [31] and for ψ_g-regular processes [1]. In Section 3.3.1, we will prove that for Markov chains, the Rényi entropy does not depend on the initial distribution but only on the transition matrix and that one can compute the Rényi entropy even if the measure is not stationary.

Definition 3.3.3. *Consider the dynamical system* $(\Omega, \mathbb{P}, \sigma)$*. We say that it is* α*-mixing if there exists a function* $\alpha : \mathbb{N} \to \mathbb{R}$ *where* $\alpha(\ell)$ *converges to zero when* ℓ *goes to infinity and such that*

$$\sup_{A\in\mathcal{F}_0^n \ ; \ B\in\mathcal{F}_0^m} \left|\mathbb{P}\left(A \cap \sigma^{-\ell-n}B\right) - \mathbb{P}(A)\mathbb{P}(B)\right| \le \alpha(\ell),$$

for all $m, n \in \mathbb{N}$*. Moreover, we say that the system is* ψ*-mixing if there exists a function* $\psi : \mathbb{N} \to \mathbb{R}$ *where* $\psi(\ell)$ *converges to zero when* ℓ *goes to infinity and such that*

$$\sup_{A\in\mathcal{F}_0^n \ ; \ B\in\mathcal{F}_0^m} \left|\frac{\mathbb{P}\left(A \cap \sigma^{-\ell-n}B\right) - \mathbb{P}(A)\mathbb{P}(B)}{\mathbb{P}(A)\mathbb{P}(B)}\right| \le \psi(\ell),$$

for all $m, n \in \mathbb{N}$*. In the cases that* $\alpha(\ell)$ *or* $\psi(\ell)$ *decreases exponentially fast to zero, we say that the system has an exponential decay.*

Now we are ready to present the main result of this section. It states that, under suitable conditions and large values of n, the longest common substring behaves like $\log n$, for almost all realizations.

Theorem 3.3.4. *Consider* $f : \Omega \to \tilde{\Omega}$ *a code such that* $\underline{H}_2(f_*\mathbb{P}) > 0$*. For* $\mathbb{P} \otimes \mathbb{P}$*-almost every* $(x, y) \in \Omega \times \Omega$*,*

$$\varlimsup_{n\to\infty} \frac{M_n^f(x,y)}{\log n} \le \frac{2}{\underline{H}_2(f_*\mathbb{P})}. \tag{3.16}$$

Moreover, if

(i) *the system* $(\Omega, \mathbb{P}, \sigma)$ *is* α*-mixing with an exponential decay (or* ψ*-mixing with* $\psi(\ell) = \ell^{-a}$ *for some* $a > 0$*);*

(ii) *$C_n \in \tilde{\mathcal{F}}_0^n$ implies $f^{-1}C_n \in \mathcal{F}_0^{h(n)}$, where $h(n) = o(n^\gamma)$, for some $\gamma > 0$,*

then, for $\mathbb{P} \otimes \mathbb{P}$-almost every $(x, y) \in \Omega \times \Omega$,

$$\underline{\lim}_{n\to\infty} \frac{M_n^f(x,y)}{\log n} \geq \frac{2}{\overline{H}_2(f_*\mathbb{P})} . \tag{3.17}$$

Then if the Rényi entropy exists, we get for $\mathbb{P} \otimes \mathbb{P}$-almost every $(x, y) \in \Omega \times \Omega$,

$$\lim_{n\to\infty} \frac{M_n^f(x,y)}{\log n} = \frac{2}{H_2(f_*\mathbb{P})}. \tag{3.18}$$

Proof. For simplicity we will assume that $\alpha(\ell) = e^{-\ell}$. The proof of this theorem follows the proof of the Theorem 7 in [18]. In the first part of the proof, for $\epsilon > 0$ we denote

$$k_n = \left\lceil \frac{2\log n + \log\log n}{\underline{H}_2(f_*\mathbb{P}) - \epsilon} \right\rceil .$$

Let us also denote

$$A_{i,j}^f(y) = \sigma^{-1}[f^{-1}C_{k_n}(f(\sigma^j y))]$$

and

$$S_n^f(x,y) = \sum_{i,j=1,\dots,n} \mathbb{1}_{A_{i,j}^f(y)}(x).$$

We first show that the event $\left\{M_n^f \geq k_n\right\}$ occurs only finitely many times. It follows from Lemma 3.1.4 and Markov's inequality that

$$\mathbb{P} \otimes \mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) \geq k_n\right\}\right) \leq \mathbb{P} \otimes \mathbb{P}(\{(x,y) : S_n^f(x,y) \geq 1\}) \leq \mathbb{E}\left(S_n^f\right).$$

Moreover, by computing the expected value of S_n^f we get

$$\begin{aligned}
\mathbb{E}\left(S_n^f\right) &= \int\int \sum_{i,j=1,\dots,n} \mathbb{1}_{A_{ij}^f(y)}(x)\ d\mathbb{P}(x)\ d\mathbb{P}(y) \\
&= \sum_{i,j=1,\dots,n} \int \left(\int \mathbb{1}_{A_{ij}^f(y)}(x)\ d\mathbb{P}(x)\right)\ d\mathbb{P}(y) \\
&= \sum_{i,j=1,\dots,n} \int \mathbb{P}\left(f^{-1}C_{k_n}(f(\sigma^j y))\right)\ d\mathbb{P}(y) \\
&= \sum_{i,j=1,\dots,n} \int f_*\mathbb{P}\left(C_{k_n}(f(\sigma^j y))\right)\ d\mathbb{P}(y) \\
&= n^2 \int f_*\mathbb{P}\left(C_{k_n}(f(y))\right)\ d\mathbb{P}(y),
\end{aligned}$$

where the last inequality follows as in Lemma 3.1.5.

Thus,

$$\mathbb{P} \otimes \mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) \geq k_n\right\}\right) \leq n^2 \int f_*\mathbb{P}\left(C_{k_n}(f(y))\right)\ d\mathbb{P}(y).$$

For large values of n, by definition of $\underline{H}_2(f_*\mathbb{P})$ it holds

$$\int f_*\mathbb{P}\left(C_{k_n}(f(y))\right)\ d\mathbb{P}(y) = \sum_{C_{k_n}} \mathbb{P}\left(f^{-1}\left(C_{k_n}\right)\right)^2 \leq e^{-k_n(\underline{H}_2(f_*\mathbb{P})-\epsilon)}.$$

Moreover by definition of k_n,

$$\mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) \geq k_n\right\}\right) \leq n^2 e^{-k_n(\underline{H}_2(f_*\mathbb{P})-\epsilon)} = \frac{1}{\log n}\ .$$

Choosing a subsequence $(n_\kappa)_{\kappa\in\mathbb{N}}$ such that $n_\kappa = \left\lceil e^{\kappa^2}\right\rceil$ we have that

$$\mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : M_{n_\kappa}^f(x,y) \geq k_{n_\kappa}\right\}\right) \leq \frac{1}{\kappa^2}\ .$$

Since the last quantity is summable in κ, the Borel-Cantelli lemma gives that if κ is large enough, then for almost every pair (x,y) it holds

$$M_{n_\kappa}^f(x,y) < k_{n_\kappa}$$

and then

$$\frac{M_{n_\kappa}^f(x,y)}{\log n_\kappa} \leq \frac{1}{\underline{H}_2(f_*\mathbb{P})-\epsilon}\left(2+\frac{1+\log\log n_\kappa}{\log n_\kappa}\right). \tag{3.19}$$

We observe that for all n, there exists κ such that $e^\kappa \leq n \leq e^{\kappa+1}$. In addition, since $\left(M_n^f\right)_{n\in\mathbb{N}}$ is an increasing sequence, we get

$$\frac{M_{n_\kappa}^f(x,y)}{\log n_{\kappa+1}} \leq \frac{M_n^f(x,y)}{\log n} \leq \frac{M_{n_{\kappa+1}}^f(x,y)}{\log n_\kappa}.$$

Taking the limit superior in the above inequalities and observing that $\lim_{\kappa\to\infty}\frac{\log n_\kappa}{\log n_{\kappa+1}} = 1$ by (3.19) we obtain

$$\varlimsup_{n\to\infty}\frac{M_n^f(x,y)}{\log n} = \varlimsup_{\kappa\to\infty}\frac{M_{n_\kappa}^f(x,y)}{\log n_\kappa}.$$

Thus, by (3.19) we have

$$\varlimsup_{n\to\infty}\frac{M_n^f(x,y)}{\log n} \leq \frac{2}{\underline{H}_2(f_*\mathbb{P})-\epsilon}.$$

Since ϵ can be arbitrarily small, (3.16) is proved.

We will now prove (3.17). In order to do that denote, for $\epsilon > 0$,

$$k_n = \left\lfloor \frac{2\log n + b\log\log n}{\overline{H_2}(f_*\mathbb{P})+\epsilon}\right\rfloor,$$

where b is a constant to be chosen.

Note that by Lemma 3.1.4 we have

$$\begin{aligned}
\mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) < k_n\right\}\right) &\leq \mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : S_n^f(x,y) = 0\right\}\right)\\
&\leq \mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : \left|S_n^f(x,y) - \mathbb{E}\left(S_n^f\right)\right| \geq \left|\mathbb{E}\left(S_n^f\right)\right|\right\}\right).
\end{aligned}$$

By Chebyshev's inequality we deduce that

$$\mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) < k_n\right\}\right) \leq \frac{\operatorname{var}\left(S_n^f\right)}{\mathbb{E}\left(S_n^f\right)^2}.$$

We have to estimate the variance of S_n^f.

We see at once that

$$\begin{aligned}
\operatorname{var}\left(S_n^f\right) &= \sum_{1\leq i,i',j,j'\leq n} \operatorname{cov}\left(\mathbb{1}_{A_{ij}^f}, \mathbb{1}_{A_{i'j'}^f}\right)\\
&= \sum_{1\leq i,i',j,j'\leq n} \int\int \mathbb{1}_{A_{ij}^f(y)}\mathbb{1}_{A_{i'j'}^f(y)} - \int\int \mathbb{1}_{A_{ij}^f(y)} \int\int \mathbb{1}_{A_{i'j'}^f(y)}\\
&= \sum_{1\leq i,i',j,j'\leq n} \int\int \mathbb{1}_{f^{-1}C_{k_n}(f(\sigma^j y))}(\sigma^i x)\mathbb{1}_{f^{-1}C_{k_n}(f(\sigma^{j'} y))}(\sigma^{i'} x) \qquad (3.20)\\
&- n^4\left(\int f_*\mathbb{P}\left(C_{k_n}(f(y))\right)\, d\mathbb{P}(y)\right)^2.
\end{aligned}$$

Let $\ell = \ell(n) = (\log n)^\beta$, for some $\beta > \gamma$ such that

$$(\log n)^\beta > (\log n)^\gamma. \qquad (3.21)$$

As in the proof of Theorem 3.1.6 there are four cases to consider.

Case 1: $|i - i'| > \ell + k_n$. Using the α-mixing condition we have

$$\begin{aligned}
&\int\left(\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^{(i-i')}x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(x)d\mathbb{P}(x)\right) d\mathbb{P}(y)\\
\leq\ &\int\left(\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(x)d\mathbb{P}(x)\int \mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(x)d\mathbb{P}(x)\right) d\mathbb{P}(y)\\
&+\alpha(\ell + k_n - h(k_n))\\
=\ &\int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y)\\
&+\alpha(\ell + k_n - h(k_n)). \qquad (3.22)
\end{aligned}$$

To estimate the first term of the sum above we analyse two cases.

Case 1.1: $|j-j'|>\ell+k_n$. In this case we have

$$\begin{aligned}
& \int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y) \\
= & \int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j-j'} y\right)\right)\right) f_*\mathbb{P}\left(C_{k_n}\left(f\left(y\right)\right)\right) d\mathbb{P}(y) \\
= & \sum_{C_{k_n},C'_{k_n}} \int_{f^{-1}(C_{k_n})\cap\sigma^{j-j'}\left(f^{-1}\left(C'_{k_n}\right)\right)} f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right) d\mathbb{P}(y) \\
= & \sum_{C_{k_n},C'_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right) \mathbb{P}\left(f^{-1}\left(C_{k_n}\right)\cap\sigma^{j-j'}\left(f^{-1}\left(C'_{k_n}\right)\right)\right).
\end{aligned}$$

Using the α-mixing condition in the last expression we get that

$$\begin{aligned}
& \int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y) \\
\leq & \sum_{C_{k_n},C'_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\left(f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\right) \\
& + \sum_{C_{k_n},C'_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\left(\alpha\left(\ell+k_n-h(k_n)\right)\right) \\
= & \sum_{C_{k_n},C'_{k_n}} \left[\left(f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\right)^2\right] \\
& + \sum_{C_{k_n},C'_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\left(\alpha\left(\ell+k_n-h(k_n)\right)\right) \\
\leq & \ \alpha\left(\ell+k_n-h(k_n)\right)+\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^4+\sum_{C_{k_n}\neq C'_{k_n}}\left(f_*\mathbb{P}\left(C_{k_n}\right) f_*\mathbb{P}\left(C'_{k_n}\right)\right)^2 \\
\leq & \ \alpha\left(\ell+k_n-h(k_n)\right)+\left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2.
\end{aligned} \tag{3.23}$$

Case 1.2 $|j-j'|\leq\ell+k_n$. By Hölder's inequality it follows that,

$$\begin{aligned}
& \int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y) \\
\leq & \left(\int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right)^2 d\mathbb{P}(y)\right)^{1/2}\left(\int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right)^2 d\mathbb{P}(y)\right)^{1/2} \\
= & \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^3\right)^{1/2}\left(\sum_{C'_{k_n}} f_*\mathbb{P}\left(C'_{k_n}\right)^3\right)^{1/2} \\
= & \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^3.
\end{aligned}$$

Using the subadditivity of the function $z(x) = x^{2/3}$ we obtain

$$\begin{aligned} \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^3 &\leq \left(\sum_{C_{k_n}} \left(f_*\mathbb{P}\left(C_{k_n}\right)^3\right)^{2/3}\right)^{3/2} \\ &\leq \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{3/2}. \end{aligned} \tag{3.24}$$

Using (3.23) and (3.24) we obtain

$$\begin{aligned} &\sum_{|i-i'|>\ell+k_n, |j-j'|>\ell+k_n} \int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(\sigma^{i'} x)d\mathbb{P}(x)d\mathbb{P}(y) \\ &+ \sum_{|i-i'|>\ell+k_n, |j-j'|\leq\ell+k_n} \int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(\sigma^{i'} x)d\mathbb{P}(x)d\mathbb{P}(y) \\ \leq\ & n^4\alpha\left(\ell + k_n - h(k_n)\right) + n^2 \sum_{|j-j'|>\ell+k_n} \alpha\left(\ell + k_n - h(k_n)\right) \\ &+n^2 \sum_{|j-j'|>\ell+k_n} \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2 + n^2 \sum_{|j-j'|\leq\ell+k_n} \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{3/2} \\ \leq\ & 2n^4\alpha\left(\ell + k_n - h(k_n)\right) + n^4 \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2 \\ &+2n^3(\ell + k_n) \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{3/2}. \end{aligned} \tag{3.25}$$

Case 2. $|i - i'| \leq \ell + k_n$:

Case 2.1 $|j - j'| > \ell + k_n$:

This case is analogous to the case 1.2.

Case 2.2. $|j - j'| \leq \ell + k_n$:

$$\begin{aligned} &\int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(\sigma^{i'} x)d\mathbb{P}(x)d\mathbb{P}(y) \\ \leq\ & \int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)d\mathbb{P}(x)d\mathbb{P}(y) \\ =\ & \int f_*\mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) d\mathbb{P}(y) \\ =\ & \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2. \end{aligned} \tag{3.26}$$

Using (3.24) and (3.26) we get

$$
\begin{aligned}
& \sum_{|i-i'|\le \ell+k_n, |j-j'|>\ell+k_n} \int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(\sigma^{i'} x) d\mathbb{P}(x) d\mathbb{P}(y) \\
& + \sum_{|i-i'|\le \ell+k_n, |j-j'|\le \ell+k_n} \int\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^i x)\mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(\sigma^{i'} x) d\mathbb{P}(x) d\mathbb{P}(y) \\
\le \; & n^4\alpha(\ell + k_n - h(k_n)) + 2n(\ell+k_n) \sum_{|j-j'|>\ell+k_n} \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^{3/2} \\
+ \; & 2n(\ell+k_n) \sum_{|j-j'|\le \ell+k_n} f_*\mathbb{P}\left(C_{k_n}\right)^2 \\
\le \; & n^4\alpha(\ell + k_n - h(k_n)) + 2n^3(\ell+k_n) \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^{3/2} \\
+ \; & 4n^2(\ell+k_n)^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 . \qquad (3.27)
\end{aligned}
$$

Putting these estimates together in (3.20) we get

$$
\begin{aligned}
\operatorname{var}\left(S_n^f\right) \;\le\; & 3n^4\alpha\left(\ell + k_n - h(k_n)\right) + 4n^3(\ell+k_n) \left(\sum_{C_{k_n}} \mathbb{P}\left(f^{-1}\left(C_{k_n}\right)\right)^2 \right)^{3/2} \\
& + 4n^2(\ell+k_n)^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 . \qquad (3.28)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\operatorname{var}\left(S_n^f\right)}{\mathbb{E}(S_n^f)^2} \;\le\; & \frac{3n^4\alpha\left(\ell + k_n - h(k_n)\right) + 4n^3(\ell+k_n) \left(\sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^{3/2}}{\left(n^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^2} \\
& + \frac{4n^2(\ell+k_n)^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2}{\left(n^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^2} . \qquad (3.29)
\end{aligned}
$$

We estimate each term on the right separately. Using the definition of k_n and of the Rényi entropy and the choice of ℓ, we have for the first term

$$
\frac{3n^4\alpha\left(\ell + k_n - h(k_n)\right)}{\left(n^2 \sum_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2 \right)^2} \;\le\; \frac{3n^4 e^{-(\log n)^{\beta}} n^{-2/(\overline{H}_2(f_*\mathbb{P})+\epsilon)} (\log n)^{-b/(\overline{H}_2(f_*\mathbb{P})+\epsilon)} e^{h(k_n)}}{(\log n)^{-2b}} .
$$

By hypothesis, $h(k_n) = o((\log n)^{\gamma})$ thus there exists n_0 such that for all $n \geq n_0$, $h(k_n) < (\log n)^{\gamma}$. In (3.21) take β large enough such that $(\log n)^{\beta} > (\log n)^{\gamma} + \log n^{4-2/(\overline{H}_2(f_*\mathbb{P})+\epsilon)}$. Hence,

$$\begin{aligned}
\frac{3n^4\alpha\left(\ell + k_n - h(k_n)\right)}{\left(n^2 \sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2} &\leq \frac{3(\log n)^{-b/(\overline{H}_2(f_*\mathbb{P})+\epsilon))}}{(\log n)^{-2b}} \\
&= 3(\log n)^{b(2-1/(\overline{H}_2(f_*\mathbb{P})+\epsilon))}.
\end{aligned} \tag{3.30}$$

To estimate the second term we obtain

$$\begin{aligned}
&\frac{4n^3(\ell+k_n)\left(\sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{3/2}}{\left(n^2 \sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2} \\
\leq\ & \frac{4n^3(\ell+k_n)}{n^4\left(\sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{1/2}} \\
\leq\ & 4(\ell+k_n)(\log n)^{b/2} \\
\leq\ & 4(\log n)^{\beta+b/2} + \frac{2(\log n)^{1+b/2} + b\log(\log n)(\log n)^{b/2}}{\overline{H}_2(f_*\mathbb{P})+\epsilon}.
\end{aligned} \tag{3.31}$$

Finally for the third term we get

$$\begin{aligned}
&\frac{4n^2(\ell+k_n)^2 \sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2}{\left(n^2 \sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2} \\
\leq\ & \frac{4n^2(\ell+k_n)^2}{n^4 \sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2} \\
\leq\ & 4(\ell+k_n)^2(\log n)^b \\
=\ & 4(\log n)^{2\beta+b} + 8(\log n)^{\beta+b}\frac{2\log n + b\log(\log n)}{\overline{H}_2(f_*\mathbb{P})+\epsilon} \\
+\ & \frac{4(2\log n + b\log(\log n))^2(\log n)^b}{\left(\overline{H}_2(f_*\mathbb{P})+\epsilon\right)^2}.
\end{aligned} \tag{3.32}$$

Taking $b < -2 - 2\beta$ and substituting (3.30), (3.31) and (3.32) into (3.29), we obtain

$$\mathbb{P}\otimes\mathbb{P}\left(\left\{(x,y) : M_n^f(x,y) \leq k_n\right\}\right) \leq \mathcal{O}\left((\log n)^{-1}\right). \tag{3.33}$$

Thus, taking a subsequence $(n_\kappa)_\kappa = \left\lceil e^{\kappa^2}\right\rceil$ as in the proof of (3.16) we use the Borel Cantelli Lemma to obtain (3.17).

On the other hand, if the system is ψ-mixing, for $|i-i'| > \ell + k_n$ we have the equivalent of equation (3.22):

$$\int \left(\int \mathbb{1}_{f^{-1}(C_{k_n}(f(\sigma^j y)))}(\sigma^{(i-i')}x) \mathbb{1}_{f^{-1}\left(C_{k_n}(f(\sigma^{j'} y))\right)}(x) d\mathbb{P}(x) \right) d\mathbb{P}(y)$$
$$\leq \int f_* \mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_* \mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y) \left(1 + \psi(\ell + k_n - h(k_n)\right).$$

If, moreover, $|j-j'| > \ell + k_n$, we have the equivalent of (3.23):

$$\int f_* \mathbb{P}\left(C_{k_n}\left(f\left(\sigma^j y\right)\right)\right) f_* \mathbb{P}\left(C_{k_n}\left(f\left(\sigma^{j'} y\right)\right)\right) d\mathbb{P}(y)$$
$$\leq \psi\left(\ell + k_n - h(k_n)\right) + \left(\sum_{C_{k_n}} f_* \mathbb{P}\left(C_{k_n}\right)^2 \right)^2.$$

Then,

$$\operatorname{var}\left(S_n^f\right) \leq n^4 \left[\psi\left(\ell + k_n - h(k_n)\right) + \left(\psi\left(\ell + k_n - h(k_n)\right)\right)^2\right] + 4n^2(\ell + k_n)^2 \sum_{C_{k_n}} f_* \mathbb{P}\left(C_{k_n}\right)^2$$
$$+2n^3(\ell + k_n)(2 + \psi\left(\ell + k_n - h(k_n)\right)) \left(\sum_{C_{k_n}} f_* \mathbb{P}\left(C_{k_n}\right)^2 \right)^{3/2}.$$

Using the definition of ψ we can estimate the first and the second term

$$\frac{n^4 \psi\left(\ell + k_n - h(k_n)\right)}{\left(n^2 \sum_{C_{k_n}} f_* \mathbb{P}\left(C_{k_n}\right)^2 \right)^2} \leq \frac{n^4 (\ell + k_n - h(k_n))^{-a}}{n^4 e^{-2k_n(\overline{H}_2(f_*\mathbb{P})+\epsilon)}}$$
$$= \frac{(\ell + k_n - h(k_n))^{-a}}{n^{-4}(\log n)^{-2b}}$$
$$= \frac{n^4 (\log n)^{2b}}{(\ell + k_n - h(k_n))^{a}}$$

and

$$\frac{n^4 \psi\left(\ell + k_n - h(k_n)\right)^2}{\left(n^2 \sum_{C_{k_n}} f_* \mathbb{P}\left(C_{k_n}\right)^2 \right)^2} \leq \frac{n^4 (\ell + k_n - h(k_n))^{-2a}}{n^4 e^{-2k_n(\overline{H}_2(f_*\mathbb{P})+\epsilon)}}$$
$$= \frac{(\ell + k_n - h(k_n))^{-2a}}{n^{-4}(\log n)^{-2b}}$$
$$= \frac{n^4 (\log n)^{2b}}{(\ell + k_n - h(k_n))^{2a}}.$$

The third term and the first part of the fourth term are estimated in the same way as in (3.32) and (3.31) respectively.

Finally, for the second part of the fourth term we get

$$
\begin{aligned}
& \frac{2n^3(\ell+k_n)\psi\left(\ell+k_n-h(k_n)\right)\left(\sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{3/2}}{\left(n^2\sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^2} \\
\leq\ & \frac{2n^3(\ell+k_n)\psi\left(\ell+k_n-h(k_n)\right)}{n^4\left(\sum\limits_{C_{k_n}} f_*\mathbb{P}\left(C_{k_n}\right)^2\right)^{1/2}} \\
\leq\ & \frac{2n^3(\ell+k_n)\left(\ell+k_n-h(k_n)\right)^{-a}}{n^4e^{-\frac{1}{2}k_n(\overline{H}_2(f_*\mathbb{P})+\epsilon)}} \\
=\ & \frac{2n^3(\ell+k_n)(\ell+k_n-h(k_n))^{-a}}{n^4n^{-1}(\log n)^{-b/2}} \\
=\ & \frac{2(\ell+k_n)(\ell+k_n-h(k_n))^{-a}}{(\log n)^{-b/2}} \\
=\ & \frac{2(\ell+k_n)(\log n)^{b/2}}{(\ell+k_n-h(k_n))^a}.
\end{aligned}
$$

Using the hypothesis that $h(k_n) = o((\log n)^\gamma)$, the definition of k_n and choosing $b < 2\beta$, we conclude this case as in the proof of the case α-mixing.

Finally, if the Rényi entropy exists, by (3.16) and (3.17) we conclude the proof of the theorem. □

3.3.1 Rényi entropy of Markov chains

In the sequel we present an entropy's invariance statement by change of initial distribution. In particular, we will use this result in the example of the stochastic scrabble (Subsection 3.3.2) to compute the entropy of the pushforward measure.

Theorem 3.3.5. *Let $(X_n)_{n\in\mathbb{N}}$ be a Markov chain in a finite alphabet χ, with irreducible and aperiodic transition matrix $P = [(p_{ij})]$ and stationary measure μ. For any Markov measure ν with initial distribution π and transition matrix P it holds*

$$H_2(\nu) = H_2(\mu) = -\log p$$

where p is the largest eigenvalue of the matrix $[(p_{ij})^2]$.

Proof. First of all, we observe that the second equality is a well-known result (see e.g. [31] Section 2.2). For the first equality, we will show that

$$H_2(\mu) \leq H_2(\nu) \leq H_2(\mu)\ .$$

For convenience here, we will adopt the following notation for strings of stochastic processes: $\{X_n^m = x_n^m\} = \{X_n = x_n, X_{n+1} = x_{n+1}, \cdots, X_m = x_m\}$ for every non-negative integers n, m such that $n \leq m$ and for any realization $x = x_0^\infty$.

We will use corollary (3.13) from [28], which states that there exists $\gamma \in (0,1)$ such that for all $k > 1$

$$\sup_{x_k \in \chi} |\nu(X_k = x_k) - \mu(x_k)| \leq \gamma^k \ .$$

A straightfoward computation gives that for every $n > k > 1$

$$\sup_{x_0, x_k \in \chi} |\nu(X_k = x_k \mid X_0 = x_0) - \mu(x_k)| \leq \gamma^k$$

and for every $x_k^n \in \chi^{n-k+1}$

$$|\nu(X_k^n = x_k^n) - \mu(x_k^n)| \leq c\gamma^k \mu(x_k^n)$$

with $c = (\inf_{x_0}\{\mu(x_0)\})^{-1} < +\infty$.

Let $(a_n)_{n\in\mathbb{N}}$ be a non-decreasing and unbounded sequence in n taking values on the non-negative integers and such that $n \geq a_n = o(n)$. Without loss of generality we will only consider the strings x_0^n such that $\nu(X_0^n = x_0^n) > 0$. On the one hand, we get

$$\begin{aligned} \nu(X_0^n = x_0^n) &\leq \nu(X_{a_n}^n = x_{a_n}^n) \\ &\leq \left[c\gamma^{a_n}\mu(x_{a_n}^n) + \mu(x_{a_n}^n)\right] . \end{aligned}$$

Therefore

$$\begin{aligned} \frac{1}{n}\log\sum_{x_0^n}\nu(X_0^n = x_0^n)^2 &\leq \frac{2}{n}\log(c\gamma^{a_n}+1) + \frac{1}{n}\log\sum_{x_0^n}\mu(x_{a_n}^n)^2 \\ &= \frac{2}{n}\log(c\gamma^{a_n}+1) + \frac{1}{n}\log\sum_{x_0^{a_n-1}}\sum_{x_{a_n}^n}\mu(x_{a_n}^n)^2 \\ &\leq \frac{2}{n}\log(c\gamma^{a_n}+1) + \frac{1}{n}\log|\chi|^{a_n} + \frac{1}{n}\log\sum_{x_{a_n}^n}\mu(x_{a_n}^n)^2. \end{aligned}$$

One can observe that the two first terms in the last line vanish as $n \to \infty$. Moreover, by stationarity of μ we obtain

$$\lim_{n\to\infty}\frac{1}{n}\log\sum_{x_{a_n}^n}\mu(x_{a_n}^n)^2 = \lim_{n\to\infty}\frac{1}{n}\log\sum_{x_0^{n-a_n}}\mu(x_0^{n-a_n})^2 = H_2(\mu)$$

which gives us the first inequality.

On the other hand, first notice that for strings such that $\nu(X_0^n = x_0^n) > 0$, we have for n large enough

$$\begin{aligned}
\nu(X_0^n = x_0^n) &= \pi(x_0)P_{x_0x_1}\cdots P_{x_{a_n-1}x_{a_n}}P_{x_{a_n}x_{a_n+1}}\cdots P_{x_{n-1}x_n}\\
&\geq \pi(x_0)\rho^{a_n}\frac{1}{\nu(X_{a_n}=x_{a_n})}\nu(X_{a_n}^n = x_{a_n}^n)\\
&\geq \frac{\pi(x_0)\rho^{a_n}}{\mu(x_{a_n})+\gamma^{a_n}}\left[\mu(x_{a_n}^n)(1-\gamma^{a_n})\right]\\
&\geq d\rho^{a_n}\left[\mu(x_{a_n}^n)(1-\gamma^{a_n})\right]
\end{aligned}$$

where $\rho := \min_{P_{ij}>0} P_{ij}$ and $d = \frac{1}{2}\min_{\pi(x_0)>0}\pi(x_0)$.

Now

$$\begin{aligned}
\frac{1}{n}\log\sum_{x_0^n}\nu(X_0^n = x_0^n)^2 &\geq \frac{2}{n}\log\left(d\rho^{a_n}\right) + \frac{1}{n}\log\sum_{x_0^n}\left[\mu(x_{a_n}^n)(1-\gamma^{a_n})\right]^2\\
&\geq \frac{2}{n}\log\left(d\rho^{a_n}\right) + \frac{2}{n}\log(1-\gamma^{a_n}) + \frac{1}{n}\log\sum_{x_{a_n}^n}\left[\mu(x_{a_n}^n)\right]^2.
\end{aligned}$$

As in the first part of the proof, the first two terms in the last line vanish and the third one converges to $H_2(\mu)$ as n diverges. This last statement concludes the proof.

□

3.3.2 Applications

In what follows we present some applications of the above stated theorem. They come from some well-known cases of probability's literature. The first one is a contamination code that flips to zero some symbols of the sequence, and in some sense *shrinks* the strings. The second put a weight on each symbol of χ, and has an effect of *expanding* the strings.

The zero-inflated contamination model

Example 3.3.6. *Let $(\xi_i)_{i\in\mathbb{N}}$ be a sequence of i.i.d. Bernoulli random variables taking values on $\{0,1\}$, independently of $\mathbb{P}$, governed by a Bernoulli measure given by $\mu(\xi_i = 1) = 1-\epsilon$, where ϵ is a noise parameter in $(0,1)$. Let $f_\xi : \Omega \to \Omega$ be a perturbation given by $f_\xi(z) = (\xi_i z_i)_{i\in\mathbb{N}}$. This defines the zero inflated contamination model (see [22, 30]). To shorten notation, we write f instead of f_ξ.*

Then, if $\underline{H}_2(f_*\mathbb{P}) > 0$, for $\mathbb{P}\otimes\mathbb{P}$-almost every $(z,t)\in\Omega\times\Omega$,

$$\overline{\lim_{n\to\infty}}\frac{M_n^f(z,t)}{\log n} \leq \frac{2}{\underline{H}_2(f_*\mathbb{P})}.$$

Moreover, if the system $(\Omega,\mathbb{P},\sigma)$ is α-mixing with an exponential decay, for $\mathbb{P}\otimes\mathbb{P}$-almost every $(z,t)\in\Omega\times\Omega$,

$$\underline{\lim_{n\to\infty}}\frac{M_n^f(z,t)}{\log n} \geq \frac{2}{\overline{H}_2(f_*\mathbb{P})}.$$

Indeed, for k large enough $f^{-1}C_k \in \mathcal{F}_0^{m_\epsilon(k)}$, where $m_\epsilon(k)$ is the proportion of 1's in the k-cylinder $C_k(\xi)$. Let $\mu^{\otimes\mathbb{N}}$ denote the product measure that governs the stochastic process $\{\xi_i\}_{i\in\mathbb{N}}$. One can observe that by the law of large numbers $\mu^{\otimes\mathbb{N}}$-almost every realization ξ has an ϵ-proportion of zeros, i.e.

$$\lim_{k\to\infty} \frac{m_\epsilon(k)}{k} = 1 - \epsilon.$$

Thus, for $\mu^{\otimes\mathbb{N}}$-almost every ξ, it exists $\epsilon_1 > 0$ such that $m_\epsilon(k) = o(k^{1+\epsilon_1})$ and we can apply Theorem 3.3.4.

Moreover, if $\mathbb{P}$ is a Bernoulli measure we can calculate explicitly the Rényi entropy of $f_*\mathbb{P}$. Namely, by using the binomial theorem, for k large enough we get

$$\begin{aligned}\sum_{C_k}[\mathbb{P}(f^{-1}C_k)]^2 &= \sum_{j=1}^{m_\epsilon(k)} \binom{m_\epsilon(k)}{j} p^{2j}(1-p)^{2(m_\epsilon(k)-j)} \\ &= \left[p^2 + (1-p)^2\right]^{m_\epsilon(k)} .\end{aligned}$$

Therefore the Rényi entropy is given by

$$\begin{aligned}H_2(f_*\mathbb{P}) &= -\lim_{k\to\infty} \frac{m_\epsilon(k)}{k} \log(p^2 + (1-p)^2) \\ &= -(1-\epsilon)\log\left(p^2 + (1-p)^2\right) .\end{aligned}$$

We observe that if $\chi = \{a_1, \ldots, a_n\}$ is a finite alphabet and $\mathbb{P}(X = a_i) = p_i$, by similar computations (and the multinomial theorem) we obtain

$$H_2(f_*\mathbb{P}) = -(1-\epsilon)\log\left(\sum_i p_i^2\right) = (1-\epsilon)H_2(\mathbb{P}) .$$

Therefore, in view of Theorem 3.3.4, as n diverges we get

$$\frac{M_n^f}{\log n} \longrightarrow \frac{2}{(1-\epsilon)H_2(\mathbb{P})} .$$

The case $f = Id$ is equivalent to $\epsilon = 0$ (no contamination) and if $\epsilon > 0$ we expect to observe larger values for M_n^f (in view of Theorem 3.3.4). This can be summarized with the following assertion: the more contamination, the more coincidences appear between the encoded strings. This is a rather intuitive feature of the string matching problem, which indicates that sequences which had lost much information tends to present more similarity.

The highest-scoring matching subsequence

Example 3.3.7. *In this example we will consider the case in which a shorter match can be better scored than a long one, depending on the symbols that compose the matched strings.*

For this we assume that each string is scored according to the symbols that compose it. In this sense suppose that each letter $a \in \chi$ is associated to a weight $v(a) \in \mathbb{N}^*$. We also denote the score of a string z_0^{m-1} by $V(z_0^{m-1}) = \sum_{j=0}^{m-1} v(z_j)$. If x and y are two realizations of the χ-valued stochastic processes (X_n) and (Y_n) (for short notation),

$$V_n(x,y) = \max_{0\le i,j\le n-m} \left\{V(z_0^{m-1}) : \text{there exists } 1 \le m \le n \text{ such that } z_0^{m-1} = x_i^{i+m-1} = y_j^{j+m-1}\right\}$$

is the n^{th} highest-scoring matching subsequence [9]. The authors also named it stochastic Scrabble, because of the namesake board game. For two copies independently generated by the same Markov source $\mathbb{P}$ with positive transition probabilities $[p_{ij}]$, they stated the following result:

$$\lim_{n\to\infty} \frac{V_n}{\log n} = \frac{2}{-\log p} \qquad \mathbb{P}\times\mathbb{P} - a.s.\ , \tag{3.34}$$

where $p \in (0,1)$ is the largest root of $\det(P - \lambda^V) = 0$, with $P = \left[p_{ij}^2\right]$ and $\lambda^V = \left[\delta_{ij}\lambda^{v(i)}\right]$.

One can observe that this result (3.34) can be obtained as particular case of Theorem 3.3.4. Indeed, inspired by [9], we can construct a specific code f that stretches the sequences depending on the weights of its letters. Formally

$$\begin{aligned} f\ :\ & \chi^{\mathbb{N}} \to \chi^{\mathbb{N}} \\ & x_0^\infty \mapsto \underbrace{x_0x_0\cdots x_0}_{v(x_0)}\underbrace{x_1x_1\cdots x_1}_{v(x_1)}\cdots\underbrace{x_nx_n\cdots x_n}_{v(x_n)}\cdots \end{aligned} \tag{3.35}$$

With this particular code, we get that $M_n^f(x,y) = V_n(x,y)$ and thus to get (3.34) we need to compute $H_2(f_*\mathbb{P})$ and check that conditions (i) and (ii) are satisfied.

We recall that if (X_n) is a Markov chain in $\chi = \{1,2,\ldots,d\}$, we can see $f(X_n)$ as a Markov Chain in $\tilde{\chi}$, which is a $(\sum_{i\in\chi} v(i))$-sized alphabet, given by

$$\tilde{\chi} = \left\{1_1, 1_2, \ldots, 1_{v(1)}, 2_1, 2_2, \ldots, 2_{v(2)}, \ldots, d_1, d_2, \ldots, d_{v(d)}\right\}\ .$$

In this context, we will consider that $f : \chi^{\mathbb{N}} \to \tilde{\chi}^{\mathbb{N}}$. Furthermore, if $Q = [Q_{ij}]$, $1 \le i,j \le d$ is the transition matrix for (X_n) we get that the transition matrix Q^* for the chain $f(X_n)$ on $\tilde{\chi}$ is given by

$$\begin{aligned} &Q^*_{i_\ell i_{\ell+1}} = 1 && \text{if } 1 \le \ell \le v(i)-1 \text{ and } 1 \le i,j \le d\ ; \\ &Q^*_{i_{v(i)}j_1} = Q_{ij} && \text{if } 1 \le i,j \le d\ ; \\ &Q^*_{ij} = 0 && \text{otherwise.} \end{aligned}$$

Let us give an example. Consider a i.i.d. random sequences over $\chi = \{0,1,2\}$, with $v(0) = 2$, $v(1) = 2$ and $v(2) = 1$. Note that $gdc\{v(0), v(1), v(2)\} = 1$. The transition *matrix for* $\tilde{\chi} = \{0_1, 0_2, 1_1, 1_2, 2\}$ *is*

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ Q_{00} & 0 & Q_{01} & 0 & Q_{02} \\ 0 & 0 & 0 & 1 & 0 \\ Q_{10} & 0 & Q_{11} & 0 & Q_{12} \\ Q_{20} & 0 & Q_{21} & 0 & Q_{22} \end{bmatrix}.$$

Notice that, if $v_{min} = \min_{i \in \chi}\{v(i)\}$ *is the minimum weight, we get for any cylinder* C_n*,*

$$f^{-1}C_n \in \mathcal{F}_0^{\left\lfloor \frac{n}{v_{min}} \right\rfloor},$$

and since $n/v_{min} = o(n^{1+\epsilon})$ *for all* $\epsilon > 0$*, condition (ii) of Theorem 3.3.4 is then satisfied. We recall that an irreducible and aperiodic positive recurrent Markov chain is an* α*-mixing process with exponential decay of correlation (see e.g. Theorem 4.9 in [38]) which implies condition (i).*

Finally, to obtain (3.34)*, we need to compute* $H_2(f_*\mathbb{P})$*. As in [9], to assure aperiodicity for the encoded process* $f(X_n)$ *we assume that* $gdc\{v(1), v(2), \ldots, v(d)\} = 1$*.*

Moreover, by Theorem 3.3.5 we know that the Rényi entropy of its stationary measure μ *is given by* $H_2(\mu) = -\log p$*, where* p *is the largest positive eigenvalue of the matrix* $\left[(Q^*)^2_{ij}\right]$, $1 \leq i, j \leq \sum_{i\in\chi} v(i)$ *(it was proved in [9] that this* p *is the same as the one defined in* (3.34)*). Moreover, we observe that* $f_*\mathbb{P}$ *is a Markov measure with initial distribution* π *and transition matrix* Q^**, where* π *is defined by* $\pi(i_1) = \mathbb{P}(X_0 = i)$ *and* $\pi(i_j) = 0$ *for any* $i \in \chi$ *and* $1 < j \leq v(i)$*. It is important to notice that in general,* $f_*\mathbb{P}$ *is not stationary.*

Thus, by Theorem 3.3.5, we have $H_2(\mu) = H_2(f_*\mathbb{P})$ *and we can combine it with equation* (3.18) *in Theorem 3.3.4 to conclude that, for* $\mathbb{P} \otimes \mathbb{P}$ *almost every pair of realizations, as* n *diverges it holds*

$$\frac{V_n}{\log n} \longrightarrow \frac{2}{-\log p}.$$

We remark that this example generalizes [9] to α*-mixing processes with exponential decay and* ψ*-mixing with polynomial decay, since we can apply Theorem 3.3.4 to this code* f*, and then obtain information on the highest scoring* V_n*.*

Chapter 4

Future perspectives

In view of the first part of this work, we intend to study properties of large deviation for return time in cylinders for Bernoulli shifts. More precisely, using results from [7] and techniques of [6] we expect to find estimates of large deviation for return time in cylinders in the case of shifts with a Bernoulli measure.

In Chapter 2, rates functions were estimated with elements of multifractal analysis found in the work of Pesin and Weiss [47]. Following this, we would like to estimate exponential rates for dimension and for fast return time in these above mentioned cases. Let us define:

$$\tau(C_n) = \min\{1 \leq k \leq n : C_n \cap \sigma^{-k} C_n = \emptyset\}$$

and

$$\tau_{C_n}(x) = \inf\{k \geq 1 : \sigma^k(x) \in C_n\}.$$

Then, we are interested in investigating the asymptotic behavior of $\mu\left(x : \tau_{C_n(x)}(x) \leq e^{n(h-\epsilon)}\right)$ and $\mu\left(x : \tau_{C_n(x)}(x) \geq e^{n(h+\epsilon)}\right)$, where h denote the entropy of the system.

We would like to extend these results to ϕ-mixing systems with $0 < \phi(0) < \infty$.

We recall that the dynamical system (Ω, μ, σ) is ϕ-mixing if there exists a function $\phi : \mathbb{N} \to \mathbb{R}$ where $\phi(\ell)$ converges to zero when ℓ goes to infinity and such that

$$\sup_{A \in \mathcal{F}_0^n \ ; \ B \in \mathcal{F}_0^m} \left| \frac{\mu\left(A \cap \sigma^{-\ell-n} B\right) - \mu(A)\mu(B)}{\mu(A)} \right| \leq \phi(\ell),$$

for all $m, n \in \mathbb{N}$.

As a second perspective, we propose to study the behavior of the non-aligned segment score presented by Dembo, Karlin and Zeitouni in [24]. More precisely, let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two sequences of length n, where the letters X_i and Y_i take values in a finite alphabet χ and $\mathcal{Y}$, respectively. Consider a score function $f : \chi \times \mathcal{Y} \to \mathbb{R}$ that is assigned to each pair of letters (X_i, Y_j). The non-aligned maximal segment score is given by

$$M_n(X,Y) = \max_{\substack{0\le i,j\le n-m \\ m>0}} \left\{ \sum_{k=1}^{m} f(X_{i+k}, Y_{j+k}) \right\}.$$

If we suppose that the two sequences are independent: $X_1, \ldots, X_n$ i.i.d. following the distribution law μ_X and $Y_1, \ldots, Y_n$ i.i.d following the distribution law μ_Y, where μ_X and μ_Y are probabilities measures on X and Y, respectively. Moreover, if we assume that the expected score per pair is negative and there is positive probability of attaining some positive pair score, i.e.

$$\mathbb{E}_{\mu_X \times \mu_Y}(f) < 0 \quad \text{and} \quad \mu_X \times \mu_Y(f > 0) > 0,$$

the authors proved that $M_n / \log n$ converges almost surely to a positive finite constant $\gamma(\mu_X, \mu_Y)$ defined in terms of appropriate relatives entropy.

As in Theorem 3.3.4, we would like to extend these results to ψ-mixing systems.

Bibliography

[1] M. Abadi and L. Cardeno, *Renyi entropies and large deviations for the first-match function,* IEEE Trans. Inf. Theory(61), 4 (2015), 1629-1639.

[2] M. Abadi, L. Cardeño and S. Gallo *Potential well spectrum and hitting time in renewal processes,* J. Stat. Phys. **159** (2015), no. 5, 1087-1106.

[3] M. Abadi, S. Gallo and E. Rada, *The shortest possible return time of β-mixing processes.* IEEE Trans. Inf. Theory(64), 7 (2018), 4895–4906.

[4] M. Abadi and R. Lambert, *The distribution of the short-return function,* Nolinearity (26) 5 (2013), 1143-1162.

[5] M. Abadi and R. Lambert, *From the divergence between two measures to the shortest path between two observables,* To appear on Ergod. Theory Dyn. Syst.

[6] M. Abadi and S. Vaienti, *Large deviations for short recurrence,* Discrete and Continuous Dynamical Systems-Series A,**21** (2008), no. 3, 729-747.

[7] M. Abadi and N. Vergne, *Sharp Error Terms for Return Time statistics under mixing conditions,* Journal of Theoretical Probability, **22**, no. 1, (2009), 18-37.

[8] V. Afraimovich, J.-R. Chazottes and B. Saussol, *Point-wise dimensions for Poincaré recurrence associated with maps and special flows,* Disc. Cont. Dyn. Syst. 9 (2003), no. 2, 263-280. (2003).

[9] R. Arratia, P. Morris and M. Waterman, *Stochastic scrabble: Large deviations for sequences with scores,* J. Appl. Prob. **25** no 1 (1988), 106-119.

[10] R. Arratia and M. Waterman, *An Erdös-Rényi Law with Shifts,* Adv. Math. **55** (1985), 13-23.

[11] H. Aytac, J. Freitas, and S. Vaienti, *Laws of rare events for deterministic and random dynamical system,* Trans. Amer. Math. Soc. **367** (2015), 8229-8278.

[12] A. Ayyer and M. Stenlund, *Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms,* Chaos Interdiscip. J. Nonlinear Sci. **17** 043116 (2007).

[13] V. Baladi, *Positive Transfer Operators and Decay of Correlations,* Advances Series in Nonlinear Dynamics, World Scientific Publishing Co. Inc., River Edge, NJ, **16**, (2000).

[14] V. Baladi and L.-S. Young, *On the spectra of randomly perturbed expanding maps,* Comm. Math. Phys., **156** (1993), 355-385.

[15] L. Barreira, *Dimension and Recurrence in Hyperbolic Dynamics,* (2008).

[16] L. Barreira, *Ergodic theory, hyperbolic dynamics and dimension theory,* Universitext, Springer, Heidelberg, (2012).

[17] L. Barreira and B. Saussol, *Hausdorff dimension of measures via Poincaré recurrence,* Communications in Mathematical Physics, **219** (2001), 443-463.

[18] V. Barros, L. Liao and J. Rousseau, *On the shortest distance between orbits and the longest common substring problem,* Adv. Math., **334** (2019), 311-339.

[19] M. Boshernitzan, *Quantitative recurrence results,* Invent. Math., **113** (1993), 617-631.

[20] J.-R. Chazottes and P. Collet, *Poisson approximation for the number of visits to balls in nonuniformly hyperbolic dynamical systems*, Ergodic Theory Dynam. Systems, **33** (2010), 49-80.

[21] J.-R. Chazottes and R. Leplaideur, *Fluctuations of the Nth return time for Axiom A diffeomorphisms*, Discrete Contin. Dyn. Syst., **13** (2005), 399-411.

[22] P. Colett, A. Galves and F. Leonardi, *Random perturbations of stochastic processes with unbounded variable length memory,* Electron. J. Probab. **13**, no 48 (2008), 1345-1361.

[23] A. Coutinho, J. Rousseau and B. Saussol, *Large deviation for return times*, Nonlinearity **31** no. 11 (2018), 5162-5179.

[24] A. Dembo, S. Karlin and O. Zeitouni, *Critical phenomena for sequence matching with scoring*, Ann. Probab., **22**, no. 4, (1994), 1993-2021.

[25] A. Dembo and I. Kontoyiannis, *The asymptotics of waiting otimes between stationary processes, allowing distortion,* Ann. Appl. Probab. **9**, no 2 (1999), 413-429.

[26] A. Dembo and O. Zeitoune, *Large deviations techniques and applications,* (1992).

[27] D. J. Feng and J. Wu, *The Hausdorff dimension sets in symbolic spaces*, Nonlinearity, **14** (2001), 81-85.

[28] P. Ferrari and A. Galves, *Acoplamento e Processos Estocásticos,* 21^o Colóquio Brasileiro de Matemática, IMPA-RJ (1997).

[29] A. Galves and B. Schmitt, *Inequalities for hitting times in mixing dynamical systems,* Random Comput. Dyn., **5** (1997), 337-348.

[30] N. Garcia and L. Moreira, *Stochastically perturbed chains of variable memory,* J. Stat. Phys **159**, no 5 (2015), 1107-1126.

[31] N. Haydn and S. Vaienti, *The Rényi entropy function and the large deviation of short return times,* Ergodic Theory Dynam. Systems **30** (2010), no. 1, 159-179.

[32] H. G. E. Hentschel and I. Procaccia, *The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors,* Physzca D, **8** (1983), 435-444.

[33] S. Jain and R.K. Bansal, *On large deviation property of recurrence times.* International Symposium on Information Theory Proceedings (ISIT), (2013), 2880-2884.

[34] B. J. James, *Probabilidade: um curso em nível intermediário.* IMPA, Projeto Euclides, (2004).

[35] M. Kac, *On the notion of recurrence in discrete stochastic processes,* Bull. A.M.S., **53** (1947), 1002-1010.

[36] I. Kontoyiannis, P. H. Algoet, Y. M. Suhov and A. J. Wyner, *Nonparametric Entropy Estimation for Stationary Processes and Random Fields, with Applications to English Text,* IEEE Trans. Inform. Theory **44**, no 3 (1998), 1319-1327.

[37] R. Leplaideur and B. Saussol, *Large deviations for return times in non- rectangle sets for Axiom A diffeomorphisms*, Discrete Contin. Dyn. Syst., **22** (2008), 327-344.

[38] D. A. Levin, Y. Peres and E.L. Wilmer, *Markov chains and mixing times,* American Mathematical Society, (2009).

[39] T. Luczak and W. Szpankowski, *A suboptimal lossy data compression based on approximate pattern matching*, IEEE Trans. Inform. Theory 43 (1997), 1439–1451.

[40] P. Marie and J. Rousseau, *Recurrence for random dynamical systems,* Discrete Contin. Dyn. Syst., **30** (2011), 1-16.

[41] C. Neuhauser, *A Phase Transition for the Distribution of Matching Blocks*, Combinatorics, Probability and Computing, **5**(2), (1996), 139-159.

[42] D. S. Ornstein and B. Weiss, *Entropy and data compression schemes,* IEEE Trans. Inform. Theory **39**, no. 1 (1993), 78-83.

[43] L. Peng, B. Tan and B.-W. Wang, *Quantitative Poincaré recurrence in continued fraction dynamical system,* Sci. China Math., **55** (2012), 131-140.

[44] Y. Pesin, *Dimension Theory in Dynamical Systems: Contemporary Views and Applications,* Chicago Lectures in Mathematics, University of Chicago Press, (1997).

[45] Y. Pesin and A. Tempelman, *Correlation Dimension of Measures Invariant Under Group Actions*, Random and Computational Dynamics, **3:3** (1995), 137-156.

[46] Y. Pesin and H. Weiss, *On the Dimension of Deterministic and Random Cantor-like Sets, Symbolic Dynamics, and the Eckmann-Ruelle Conjecture,* Comm. Math. Phys., **182** (1996), 105-153.

[47] Y. Pesin and H. Weiss, *A multifractal analysis of equilibrium measures for conformal expanding maps and Moran-like geometric constructions,* J. Stat. Phys., **86** (1997), 233-275.

[48] J. Rousseau, *Récurrence de Poincaré pour les observations,* PhD Thesis, University of Brest, (2010).

[49] J. Rousseau, *Hitting time statistics for observations of dynamical systems,* Nonlinearity **27** (2014), 2377-2392.

[50] D. Ruelle, *Repellers for real analytic maps,* Ergodic Theory Dynam. Systems, **2** (1982), 99-107.

[51] B. Saussol, *Recurrence rate in rapidly mixing dynamical systems,* Discrete Contin. Dyn. Sysr., **15**, no. 1 (2006), 259-267.

[52] B. Saussol, *An introduction to quantitative Poincaré recurrence in dynamical systems,* Rev. Math. Phys., **21** (2009), no. 8, 949-979.

[53] B. Saussol and J. Wu, *Recurrence spectrum in smooth dynamical system,* Nonlinearity, **16** (2003), 1991-2001.

[54] B. Saussol, *Formalisme thermodynamique, analyse multifractale et récurrence de Poincaré dans les systèmes dynamiques hyperboliques*, Brest, (2009).

[55] B. Saussol, S. Troubetzkoy and S. Vaienti, *Recurrence, dimensions and Lyapunov exponents,* J. Stat. Phys. 106 (2002), 623–634.

[56] C. Shalizi, *Advanced Probability II*, Spring, (2007).

[57] M. Viana, *Stochastic dynamics of deterministic systems,* Brazilian Math. Colloquium, IMPA, (1997).

[58] M. Viana and K. Oliveira, *Fundamentos de Teoria Ergódica* - Rio de Janeiro: SBM, (2014).

[59] A. J. Wyner, *More on recurrence and waiting otimes,* Ann. Appl. Probab. **9**, no 3 (1999), 780-796.

[60] A. Wyner and J. Ziv, *Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression,* IEEE Trans. Inform. Theory **35**, no 6 (1989), 1250-1258.

[61] L.-S. Young, *Dimension, entropy and Lyapunov exponents, Ergodic Theory Dynam. Systems* **2** (1982), 109-124.

Universidade Federal da Bahia - UFBA
Instituto de Matemática e Estatística / Colegiado da Pós-Graduação em Matemática

Av. Adhemar de Barros, s/n, Campus Universitário de Ondina, Salvador - BA
CEP: 40170 -110
<http://www.pgmat.ufba.br>