# Interactions of geometry, probability theory and dynamics

*From infinite entropy to fractals*

By

Felipe Pérez

School of Mathematics
University of Bristol

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Science.

May 2020

# Abstract

In this thesis we approach three different problems concerning statistical properties of dynamical systems. In the first problem we study the dimension theoretical properties of infinite entropy ergodic invariant measures for a certain family of one dimensional maps in the unit interval. The main difficulty in our setting is that the infinite entropy condition makes several of the standard tools in ergodic theory unavailable. We bypass this difficulty by using more refined covering that allow us to see the asymptotic interaction between the measure we are interested in and the Lebesgue measure. Our main result provides an almost sure value for the lower and upper local dimensions of the measure.
In the second problem, we study the limit laws of non-stationary dynamical systems comprised of intermittent (non-uniformly hyperbolic) maps in the unit interval. In particular, give large deviations probabilities bounds for sequential and random systems, as well as establish a central limit theorems and determine the role that random centering plays on them. The techniques we use differ from the usual spectral methods used for uniformly hyperbolic systems, and are based on concentration inequalities and new results in the area.
In the final problem, we study statistical properties of self-affine systems. In particular, we are interested in the asymptotic properties of equilibrium measures for the system: we prove that such measures are not mixing, and we prove the existence of zero-one laws for shrinking target problems.

# Dedication and acknowledgements

I would like to thank my advisor Thomas Jordan. Without his guidance and encouragement to carry on, I would have not arrived to this point. I could have not wished for a better supervisor for my doctorate studies.
I would like to give special thanks to Matthew Nicol and Vaughn Climenhaga who made my visit to the United States possible during my third year. I also extend my thanks to the University of Houston and their Faculty of Mathematics, who welcomed me and treated me as one of them since the moment I arrived. I reiterate my thanks to Matt, who invited me to do research with him as if I were a student of his own.
I would also like to thank my family, both my partner and my daughter, whose support was fundamental to complete this monumental challenge. They also drove me to continue with this task even in the times where everything seemed impossible.
Finally, I would like to thank everyone who helped me in any way to get here. Naming everyone would be impossible, but I would like to at least give special thanks to my former supervisor, Godofredo Iommi and to my friends Juan Pablo, Gastón, Betzabé, Chris and Joe who where always there for me and always had something relevant to say.

# Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .................................................... DATE: ..........................................

# TABLE OF CONTENTS

**Page**

CHAPTER 1

# INTRODUCTION

The theory of dynamical systems has a long history and has developed in many different forms. The common philosophy is to consider *abstract models of systems which evolve in time according to certain laws*. Throughout its history, this theory has found applications and motivations from diverse fields, including modeling of traffic, weather, kinematics, biological systems, among others. While the theory started with systems evolving in continuous time, nowadays it also deals with systems evolving in discrete time steps.

One of the main and historically first objects of study in dynamical systems are *deterministic systems*. In this context laws of evolution of the system are known and given the state of the system at time $n$, we can determine the state of the system at time $n+1$. Deterministic systems have the advantage that, given an initial condition, we know the state at any time. However, there are still many challenges. For example: if we have limited precision for our measurement of the current state of the system, then as we iterate the evolution law of our system, this error may grow in time, thus our predictions of the future states of the system will have a much larger uncertainty than we would wish. This phenomenon is known as *sensitivity to initial conditions* and is one of the defining characteristics of *chaotic systems*. It is particularly prevalent in systems with some degree of *hyperbolicity*, a property which for us will mean that the evolution law is such that it *pushes away* points in a certain direction, and *brings together* points in a different direction. Some of the mathematical tools used in order to build a theoretical framework for this theory include measure and probability theory, functional analysis, complex analysis, geometric measure theory, among others.

The theory also considers *random dynamical systems*, in which the laws of evolution change over time according to random laws, or are known up to a certain degree of uncertainty. In this case, we cannot predict the state of the system once we let it evolve, but we can only make statistical predictions about it. The results in this context can be classified in three different categories:

1. Sequential: here the results are proved for *all* realizations of the random laws of evolution,

2. Annealed: here the results are proved for *the average* realization of the random laws of evolution,

3. Quenched: here the results are proved for *almost all* the realizations of the random laws of evolution.

Note that in the last two categories, a notion of a probability distribution of the evolution laws is needed. This problem is addressed by introducing the notion of *noise space*, which drives the evolution of the system. The idea is to consider all the different possible evolution laws as outcomes of a random selection process.
Writing a comprehensive history of the area is far beyond the scope of this thesis, so we mention three key milestones in the theory which are fundamental for the investigations developed in this thesis.

1. In [Poi90], Poincare formulated what today is known as the Poincare's recurrence theorem (see theorem 2.1.12), which states that under certain conditions, most trajectories of our system visit infinitely often all subsets of the underlaying space of the system which have positive measure.

2. In [Bir31], Birkhoff gives a quantitative version of Poincare's result: his Ergodic theorem shows that the frequency at which trajectories return to a given set is proportional to the measure of such set.

3. Inspired by the work of Shannon, [Sha48], Kolmogorov and Sinai introduced (see [Kol58], [Kol59] and [Sin59]) the concept of Entropy in the context of dynamical systems. This notion tries to capture the idea of *growth of complexity* of the systems as they evolve in time, and represents an invariant which enables us to distinguish systems which are not equivalent.

Despite the nuances in the different approaches to dynamical systems, there are three objects at the heart of the theory: *phase space*, *composition of transformations* and *orbits*

of points in the phase space. The phase space represents a set which we will let evolve in time. The evolution of the phase space is represented by the successive composition of a family of endomorphisms of the phase space. The orbits represent, for each point of the phase space, the set consisting of all the images of such point under the compositions of the family of endomorphisms. One of the central problems in dynamical systems is understanding the behavior of orbits for a given family of endomorphisms.
When studying the behavior of orbits, there are two main paradigms to choose such orbits (paradigms which are not mutually exclusive): the topological one, in which we are interested in generic points in the sense of all points belonging to an open dense set exhibiting similar orbit behavior, and the measure theoretic one, in which we are interested in generic points in the sense of all points belonging to a set of total measure exhibiting similar orbit behavior. In our investigation, we focus on the second point of view. How we choose to measure the phase space is a non-trivial issue, which will be addressed when pertained.
The dynamical systems considered can be of a physical nature, where the phase space represents the set of configuration of a mechanical system, the transformations are the laws of evolution of such set of configurations, and the orbits are the physical trajectories of the objects. Examples can also arise in a mathematical setting. For instance, it is common to let the phase space represent a set of numbers and the endomorphisms of the space a set of transformation of such numbers.
In this thesis, we will focus on situations arising from considering mathematical objects. In particular, our results pertain to three different settings:

1. Iterations of uniformly expanding maps of the unit interval,

2. Composition of sequences, taken either deterministically or at random, of non-uniformly expanding maps of the unit interval,

3. Composition of sequences of contractions on the real plane.

In each of the settings we investigate different problems, however in each case we study the properties of the measures with which we look at our phase spaces with: statistical properties, geometric properties and combinations of both.
In the first problem, we study maps belonging to a family that generalizes the Gauss map. We are interested in understanding some geometric properties of probability measures which are invariant and ergodic with respect to such maps, and posses some asymptotic independence properties (Gibbs/Bernoulli measures) as well as infinite entropy with

respect to the map. The kind of properties we are interested in are related to fractal geometry: we would like to understand the concentration of mass at local scale around a typical point with respect to the measure (the local dimension of the measure). In the context of finite entropy, the local dimension can be calculated almost everywhere in terms of the entropy and the average exponent of the derivative of the map (Lyapunov exponent), using standard techniques of fractal geometry: coverings by cylinders and metric estimates of their diameters, as well as measure estimates of their mass. When the entropy is infinite, these methods no longer work, as they rely on asymptotic results such as the Ergodic Theorem and the Shannon-McMillan-Breiman theorem, which are no longer valid in our context. To tackle this difficulty, we use finer coverings, as well as some tools of infinite ergodic theory.

In the second problem, we consider both deterministic and random compositions of maps belonging to a certain family of non-uniformly hyperbolic maps: the Liverani-Saussol-Vaienti (LSV) maps. The composition can be done according to a fixed sequence of choices of the maps, or according to a probability distribution defined in a finite collection of maps from this family. The elements of the LSV family are endomorphisms of $[0,1]$ having derivative strictly bigger than one in all points but $x = 0$, where the maps and all of their compositions have derivative equal to one. We are interested in studying statistical laws of the considered compositions from the point of view of the Lebesgue measure, laws such as central limit theorems and large deviations. The methods that have been historically used to derive these results in the context of uniformly-hyperbolic dynamics rely heavily on the quasi-compactness of an operator (the transfer operator) associated to the dynamics of the system. When the maps are no longer uniformly-hyperbolic, the transfer operator of the system is not quasi-compact, so the methods no longer work. To overcome this difficulty, we use a *martingale approximation* method, in which the sums which are being studied can be written as the sum of a martingale difference and an error term which can be controlled. Therefore, by keeping track of the induced error, we can use probabilistic techniques from the theory of martingales in the dynamical setting.

In the third problem, we investigate statistical laws associated to systems consisting of iterations of different affine contractions of the two-dimensional Euclidean space. These systems give rise to fractal sets called self-affine sets. The problem of determining the Hausdorff dimension of such sets has been historically difficult, and the tools and results developed by Falconer and Käenmäki have substantially pushed forward the subject. In particular, the construction of Käenmäki measures on the self-affine sets makes it possible to find the dimension of such sets through equilibrium measures. We investigate the mixing and zero-one laws for such measures, using recent results by Fraser, Jordan & Jurga.

## 1.1 Structure of the thesis

The organization of the thesis, as well as the main results of each section is the following:
*Chapter 2: Preliminaries.* In this chapter we present the basic terminology and results needed to establish the results of the subsequent chapters. We introduce standard tools from ergodic theory that will be used throughout the thesis. We also introduce the relevant notions of dimension theory, which will be fundamental for chapter 3. Finally, we also introduce the language of iterated function systems and random and sequential dynamical systems, which will be the objects of study of chapters 5 and 4 respectively.
*Chapter 3: Dimension of measures with infinite entropy.* In this chapter, we consider a class of maps of the unit interval $[0,1]$, and for each of them, a class of invariant, ergodic probability measures (see section 2.1 for the corresponding definitions). The class of maps is characterized by having coding by a countable alphabet, a symbolic shift and metric properties which are similar to the ones of the Gauss map, and the measures are defined in terms of their asymptotic behavior (Gibbs measures), as well as the fact that they have infinite entropy with respect to the map. Our main result consists of finding the lower and upper local dimensions of the measure, yielding in particular the values of the Hausdorff and the packing dimensions of the measure. These results have been accepted for publication by *Nonlinearity*, and can also be found in the pre-print [PP18]. After that, we construct an ergodic invariant probability measure which has infinite entropy with respect to an EMR and positive Hausdorff dimension.
*Chapter 4: Limit laws for sequential and random dynamical systems.* In this chapter we consider deterministic and random compositions of maps belonging to a class of non-uniformly hyperbolic maps of the unit interval (Liverani-Saussol-Vaienti maps). For this class of maps, there is no common invariant probability measure, hence it is natural to consider the Lebesgue measure as a reference measure. The main results of this chapter are a large deviations bounds for sequential and random systems, as well as an upgrade to a central limit theorem proved previously by Nicol, Török, and Vaienti, and Hella and Leppänen. The results of this chapter are joint work with Matthew Nicol and Andrew Török from the University of Houston, and have been accepted for publication in *Ergodic Theory and Dynamical Systems*. The pre-print version of the article can be found in [NPPT19].

*Chapter 5: Statistical properties of Käenmäki measures.* In this chapter we present the first results of an ongoing project at the time of writing of this thesis. We investigate some statistical properties of Käenmäki measures associated to two dimensional iterated function systems consisting of affine transformations whose linear parts can be represented using diagonal or anti-diagonal matrices. Our main results show that such measures are not mixing, and that they satisfy a zero-one law in the context of shrinking targets.

All logarithms considered in this thesis are in base $e$.

# Chapter 2

# Preliminaries

## 2.1 Measure preserving transformations

The main objects of study of this thesis are dynamical systems and the statistical properties of measures on the underlying phase space. We provide a general definition, which will include the systems treated in all chapters of this thesis.
Let $(X_i, \mathscr{B}_i)$ for $i = 1, 2$ be measurable spaces, that is, $X_i$ is a set and $\mathscr{B}_i$ a sigma-algebra on $X_i$. A function $T: X_1 \to X_2$ is *measurable* if $T^{-1}(B) \in \mathscr{B}_1$ for all sets $B \in \mathscr{B}_2$. In general we will take $X_2 = X_1$ and $\mathscr{B}_2 = \mathscr{B}_1$, or $X_2 = \mathbb{R}$ and $\mathscr{B}_2$ the Borel sigma-algebra of $\mathbb{R}$. We will refer to measurable functions $f: X \to \mathbb{R}$ indistinctly as *random variables*, *observables* or *potentials*.

**Definition 2.1.1.** *A* (deterministic) dynamical system *is an action $\zeta$ of a semigroup $S$ on a measurable space $(X, \mathscr{B})$ such that $\zeta(s, \cdot): X \to X$ is measurable for all $s \in S$. We call $X$ the* phase space.

We think of $X$ evolving according to $\zeta(s, \cdot): X \to X$ for different elements of $s$. For instance, if $s_1, s_2 \in S$, we can think of the succesive evolved states of $X$ given by $\zeta(s_1, \cdot): X \to X$ and then $\zeta(s_1 s_2, \cdot): X \to X$.
The main examples of deterministic dynamical systems that we will work with are the following:

**Example 2.1.2.** *Suppose $T: X \to X$ is a measurable function, and consider the semigroup $\{T^i\}_{i \geq 0}$, with composition as operation. Then $(X, \{T^i\}_{i \geq 0})$ is a dynamical system. If $f$ is* *invertible, then we can also consider the action of the group* $\{T^i\}_{i\in\mathbb{Z}}$ *on* $X$*, and this also defines a dynamical system.*

**Example 2.1.3.** *Suppose* $S = \{T_\lambda\}_{\lambda\in\Lambda}$ *is a family of measurable transformations on* $X$*, and consider the free semigroup* $F(S) = \{T_{i_1} \circ \cdots \circ T_{i_n} : i_k \in \Lambda\}$*. The operation of* $F(S)$ *is given by composition of the transformations in* $S$*. Here we do not include the empty product as an identity element of* $F(S)$*.*

A dynamical system represents the basic structure that models evolution of measurable spaces over time. In order to make measurements of the space and its subsequent evolved states, we need a measure on the phase space. In general, the phase space can have a very pathological image after the action of dynamics. To avoid this, we will make certain assumptions which ensure compatibility between the transformations and the measures on the phase space.

**Definition 2.1.4.** *Let* $m$ *be a measure on* $(X,\mathscr{B})$*, and a measurable transformation* $T\colon X \to X$*. We say that* $T$ *is* non-singular *with respect to the measure* $m$ *if* $m(T^{-1}(A)) = 0$ *if and only if* $m(A) = 0$ *for every measurable set* $A \in \mathscr{B}$*. We say that the dynamics* $\zeta$ *is* non-singular *if all the transformations* $\zeta(s,\cdot)$ *are non-singular.*

A more rigid notion than non-singularity occurs when the transformation do not alter the structure of the phase space as seen by the measure.

**Definition 2.1.5.** *If the transformation* $T\colon X \to X$ *is such that* $m(T^{-1}(A)) = m(A)$ *for all measurable sets* $A \in \mathscr{B}$*, we say that* $f$ *is* measure preserving *(or that the measure is* $T$*-invariant). We say that the dynamics* $\zeta$ *is* measure preserving *if all the transformations* $\zeta(s,\cdot)$ *are measure preserving.*

It is easy to see that a composition of non-singular/measure preserving transformations is non-singular/measure preserving, thus we can form a category whose objects are measure spaces and its arrows are non-singular/measure preserving transformations. We will refer to this category as the category of non-singular/measure preserving transformations. This means in particular that when we have dynamics given by the action of the free semigroup generated by a set $S$, it suffices to check that each of the elements of $S$ is non-singular/measure preserving to check that the whole action has the same property. If the sigma-algebra of the phase space is generated by a semi-algebra, then we do not need to check the equality for all the sigma-algebra but just for the generators:

**Lemma 2.1.6.** *Suppose that $\mathscr{S}$ is a semi-algebra on $X$ such that $\sigma(\mathscr{S}) = \mathscr{B}$, and suppose that the measure $\mu$ is such that $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \mathscr{S}$. Then $\mu(T^{-1}(B)) = \mu(B)$ for all $B \in \mathscr{B}$.*

**Proof.** This is theorem 1.1 in [Wal00]. ∎

The definition of measure preserving transformations can be characterized in terms of integrals of functions:

**Proposition 2.1.7.** *A transformation $T$ preserves $m$ if and only if*

$$\int_X \varphi \circ T dm = \int_X \varphi dm$$

*for every $\varphi \in L^1(m)$.*

**Proof.** This is proposition 1.1.1 [VO16]. ∎

We list now some of the main examples that we are going to work with throughout this thesis.

**Example 2.1.8.** *Suppose $\mathscr{A}$ is a finite/countable alphabet, and consider the set of infinite sequences with elements in $\mathscr{A}$, denoted by $X := \mathscr{A}^{\mathbb{N}}$. The topology of $X$ is the product topology of the discrete spaces $\mathscr{A}$, and the sigma-algebra of $X$ is the Borel sigma-algebra generated by this topology. We define a transformation $T$ on $X$ given by $T(\omega)_n = \omega_{n+1}$ for $\omega = (\omega_1, \omega_2, \ldots) \in X$, or equivalently, $T(\omega) = (\omega_2, \omega_3, \ldots)$. We call $X$ a* symbolic space *with finitely/countably many symbols, and $T$ the* left shift operator *on the symbolic space.*
*The structure of $X$ is best described by the* cylinder sets*: for every finite sequence $a = (a_1, \ldots, a_k) \in \mathscr{A}^k$, define the* cylinder *associated to the sequence $a$ as the set $C(a) = [a_1, \ldots, a_k] = \{\omega \in X : \omega_1 = a_1, \ldots, \omega_k = a_k\}$. In this case, we say that the* length *of the cylinder is $k$. Note that the collection of all cylinders of length $k$ forms a partition of the phase space $X$. The topology of $X$ can also be described as the topology generated by all cylinders.*
*Let $p = (p_1, \ldots) \in (0,1)^{|\mathscr{A}|}$ be a probability vector. Define a measure $\mu$ on $X$ by setting*

$$\mu([a_1, \ldots, a_k]) = \prod_{i=1}^{k} p_{a_i}$$

*for all words $a = (a_1, \ldots, a_k) \in \mathscr{A}^k$, for all $k \geq 1$. The existence of such measure can be proved using Kolmogorov's extension theorem (see theorem 2.1.5 in [Oks13]). We call this* *kind of measure a* Bernoulli measure. *Note that by construction, Bernoulli measures leave the measure of cylinders invariant under the application of the left shift map. Since the cylinders generate the sigma-algebra of measurable sets, by lemma 2.1.6 we conclude that the measure is indeed invariant under $T$.*

**Example 2.1.9.** *Let $X = [0,1]$ and $\mathscr{B}$ the Borel sigma-algebra. Define the* Gauss map *by*

$$T : [0,1] \to [0,1]$$
$$x \mapsto \begin{cases} \frac{1}{x} - \left[\frac{1}{x}\right] & \text{for } x \in (0,1] \\ 0 & \text{for } x = 0 \end{cases}$$

*The action of $T$ on $[0,1]$ can be visualized as in figure 2.1:*

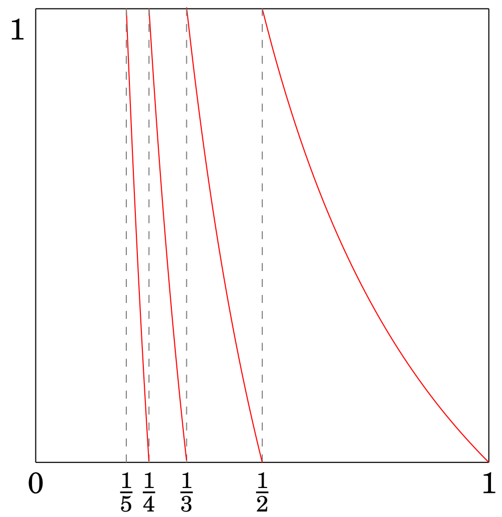


Figure 2.1: Plot of the Gauss Map $T$.

*Define the* Gauss measure *$\mu$ on $[0,1]$ by*

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{1}{1+x} dx$$

*for any measurable set $A$. We can see that this measure is invariant under $T$ by noticing that for all sets of the semi-algebra of intervals of $[0,1]$ one has*

$$T^{-1}((a,b)) = \bigcup_{n=1}^{\infty} \left( \frac{1}{n+b}, \frac{1}{n+a} \right),$$

*and consequently*

$$\mu(T^{-1}(a,b)) = \sum_{n=1}^{\infty} \mu\left( \frac{1}{n+b}, \frac{1}{n+a} \right) = \frac{1}{\log 2} \sum_{n=1}^{\infty} \log\left( \frac{\frac{1}{n+a}+1}{\frac{1}{n+b}+1} \right) = \frac{1}{\log 2} \log\left( \frac{b+1}{a+1} \right) = \mu((a,b)).$$

*By lemma 2.1.6, we concludde that the Gauss measure is invariant under the Gauss map.*

In chapter 3 we will define a class of maps of the interval which generalizes the Gauss map, yet they have similar properties to the ones of the Gauss map.

**Example 2.1.10.** *Let* $X = [0,1]$ *with the Borel sigma-algebra, and for* $\alpha \in (0,1)$*, define the* Liverani-Saussol-Vaienti *(see [LSV99]) intermittent map by*

$$T_\alpha(x) = \begin{cases} x + 2^\alpha x^{1+\alpha}, & \text{if } 0 \le x \le 1/2, \\ 2x - 1, & \text{if } 1/2 \le x \le 1. \end{cases},$$

*The issue of invariant measures for this class of maps will be addressed in chapter 4.*

We describe now the dynamics we study in chapter 5. We do not provide the full details of the constructions, as this will be provided in the corresponding chapter.

**Example 2.1.11.** *Let* $X$ *be a closed subset of* $\mathbb{R}^n$*, and let* $\{S_1, \dots, S_m : X \to X\}$*, with* $m \ge 2$ *be a family of contractions:*

$$|S_i(x) - S_i(y)| \le c_i |x - y|$$

*for* $c_i \in [0,1)$*, which we call* contraction ratios *of the maps* $S_i$*. We call the family* $\{S_i\}$ *an* iterated function system *(IFS). A non-empty compact subset* $F$ *of* $X$ *is called an* attractor *for the IFS if*

$$F = \bigcup_{i=1}^{m} S_i(F).$$

*The dynamics of the IFS on the attractor are given as follows: for a point* $x \in F$ *and an element* $\omega \in \Omega = \{1, \dots, m\}^{\mathbb{N}}$*, we consider the sequence of compositions* $S_{\omega_1} \circ \cdots \circ S_{\omega_n}(x)$*. Note that contrary to sequential systems, the order of the composition is reversed. This is due to the fact that we are applying different contractions to the underlying set, which correspond to the inverse branches of expanding maps in the sequential case.*

The next theorem gives a first recurrence result for dynamical systems with invariant measures.

**Theorem 2.1.12** (Poincare)**.** *Suppose that* $\mu$ *is a probability measure and that* $T$ *is* $\mu$*-invariant. Then for any measurable set* $E$ *with* $\mu(E) > 0$*, we have that*

$$\mu(x \in E : T^n(x) \in E \text{ i.o. }) = \mu(E).$$

**Proof.** This is theorem 1.14 in [Wal00]. ■

This result says that for each set of positive measure, almost all of its points return to it infinitely often. In the next section we formulate a quantitative version of this result, fundamental in ergodic theory.

## 2.2 Ergodicity and mixing

We have defined a category in which we can do dynamics, namely, the category of measure preserving transformations. Within this category, we can find transformations which are in some sense, the irreducible components of measure preserving transformations.

**Definition 2.2.1.** *We say that a measure preserving transformation $T: X \to X$ is* ergodic *if $T^{-1}A = A$ for a measurable set A, implies that $\mu(A) = 0$ or $\mu(X \setminus A) = 0$.*

Note that proving that a transformation is not ergodic requires exhibiting a decomposition of the space $X$ into two subspaces $X = X_1 \cup X_2$ of positive measure such that $T^{-1}X_i \subseteq X_i$.
On the other hand, proving that a transformation is ergodic in general is a hard problem, and it may require more sophisticated techniques. Ergodicity can be characterized in terms of the observables of the system

**Proposition 2.2.2.** *The measure $\mu$ is ergodic with respect to $T$ if and only if the only if for all $f \in L^2$ such that $f \circ T = f$ almost everywhere, then $f$ is constant.*

**Proof.** This is theorem 1.6 from [Wal00]. ■

For the rest of the section we assume that the measures we work with are probability measures (we can assume they are finite and then normalize to obtain a probability measure). Despite the notion of ergodicity being about irreducibility of the dynamics, it has strong consequences on the asymptotic behavior of the dynamics.
We formulate now one of the foundational results on ergodic theory, namely, the Birkhoff Ergodic Theorem:

**Theorem 2.2.3** (Birkhoff). *Suppose that the measure $\mu$ is $T$-invariant and that $f$ is an integrable observable. Then there exists a $T$-invariant function $\widehat{f} \in L^1$ such that*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) = \widehat{f}(x) \quad a.e.,$$

*and $\int_X f\,d\mu = \int_X \widehat{f}\,d\mu$. If $\mu$ is ergodic, then $\widehat{f}$ is constant and equal to $\int_X f\,d\mu$.*

**Proof.** This is theorem 1.4 from [Wal00]. ■

In most applications we will consider ergodic systems. If that is the case and we take $f = 1_A$ for a positive measure measurable set $A$, then the conclusion of the ergodic theorem is that asymptotically, the ratio of the time spent in $A$ by the orbit of a generic point is proportional to the measure of $A$. As this quantitative result only depends on the measure of $A$, we can think of it as an *equidistribution* result. In the words of G.D. Birkhoff (see [Bir31])

> ‘ The Ergodic Theorem then says: for any such measure-preserving transformation $T$, and for each individual point $P$ (except possibly an exceptional set of measure 0), there is a definite probability that its iterates under $T$, from $P$ on, namely
>
> $$P, T(P), T^2(P), \ldots \text{ and } P, T^{-1}(P), T^{-2}(P), \ldots$$
>
> fall in any given measurable set $M$.’

Furthermore, then he adds:

> ‘What the Ergodic Theorem means, roughly speaking, is that for a discrete measure-preserving transformation or a measure-preserving flow of a finite volume, probabilities and weighted means tend toward limits when we start from a definite state $P$ (not belonging to a possible exceptional set of measure 0), and, furthermore, the limiting value is the same in both directions.’

We present now one of the main examples of ergodic dynamics that we will use throughout this work.

**Proposition 2.2.4.** *Let $X = [0,1]$ be the unit interval and $T, \mu$ the Gauss map and the Gauss measure as defined in example 2.1.9. Then $\mu$ is ergodic with respect to $T$.*

**Proof.** This proof is from [VO16], but we include it as some of its ideas are fundamental for our investigation. By computing derivatives, it is possible to prove that

$$|T'(x)| \geq 1 \ , \ |(T^2)'(x)| \geq 2 \ , \ |T''(x)/T'(x)^2| \leq 2$$

for all $x \in (0,1]$. Let $I(k) = (\frac{1}{k+1}, \frac{1}{k})$ for all $k \geq 1$, and denote by $g_k$ the local inverse of $T$ restricted to $I(k)$. For any finite sequence $(a_1, \ldots, a_p) \in \mathbb{N}^p$ with $p \geq 2$, denote $I(a_1, \ldots, a_p) = \{x \in [0,1] : T^{k-1}(x) \in I(a_k), \ k = 1, \ldots, p\}$. Then the inverse of $T^p$ for the interval $I(a_1,\dots,a_p)$ is given by $g_{a_p}\circ\cdots\circ g_{a_1}$. The previous metric estimates imply that for for any $z\in I(k)$,

$$|(\log|T'\circ g_k(z)|)'|=\left|\frac{T''(g_k(z))g_k'(z)}{T'(g_k(z))}\right|=\left|\frac{T''(g_k(z))}{T'(g_k(z))^2}\right|\le 2.$$

Then for any two points $x,y\in I(a_1,\dots,a_p)$, by using the mean value theorem and the previous estimate, we have

$$\begin{aligned}\log\frac{|(T^p)'(x)|}{|(T^p)'(y)|}&=\sum_{j=1}^{p}\log|T'\circ g_{a_j}(T^j(x))|-\log|T'\circ g_{a_j}(T^j(y))|\\&\le 2\sum_{j=1}^{p}|T^j(x)-T^j(y)|=2\sum_{i=0}^{p-1}|T^{p-i}(x)-T^{p-i}(y)|\\&\le\sum_{j=1}^{p}2^{1-[i/2]}|T^p(x)-T^p(y)|\le 8.\end{aligned}$$

Take two measurable sets $E_1,E_2\subset I(a_1,\dots,a_p)$, and then by integrating with respect to the Lebesgue measure $m$ twice, we obtain

$$\frac{m(T^p(E_1))}{m(T^p(E_2))}=\frac{\int_{E_1}|(T^p)'|dm}{\int_{E_2}|(T^p)'|dm}\le e^8\frac{m(E_1)}{m(E_2)}.$$

Now, the density of density of the Gauss measure $\mu$ with respect to the Lebesgue measure is bounded above and below:

$$\frac{1}{2\log 2}\le\frac{1}{(1+x)\log 2}\le\frac{1}{\log 2},$$

we obtain the same inequality for the Gauss measure

$$\frac{\mu(T^k(E_1))}{\mu(T^k(E_2))}=\frac{\int_{E_1}|(T^k)'|dm}{\int_{E_2}|(T^k)'|dm}\le K\frac{\mu(E_1)}{\mu(E_2)}\tag{2.1}$$

for some constant $K>0$. Let $A\subset(0,1)$ be a $T$-invariant set of positive Gauss measure (and hence, positive Lebesgue measure). Then by the Lebesgue's density theorem (corollary 2.14 in [Mat99]), almost every point of $A$ has density 1, that is, for almost every $a\in A$,

$$\lim_{r\to 0}\frac{\mu(A\cap B(a,r))}{\mu(B(a,r))}=1.$$

Fix such point $a$, and let $\{I(a_1,\dots,a_m)\}_{m\ge 1}$ be the sequence of intervals such that $T^{k-1}(a)\in I(a_k)$ for all $k\ge 1$. We apply the estimate 2.1 to the sets $E_1=A^c\cap I(a_1,\dots,a_k)$ and $E_2=I(a_1,\dots,a_k)$ and obtain

$$\mu(A^c)=\frac{\mu(A^c)}{\mu(0,1)}\le K\frac{\mu(A^c\cap I(a_1,\dots,a_k))}{\mu(I(a_1,\dots,a_k))}\to 0$$

as $k\to\infty$. This proves that $\mu(A)=1$, and hence $\mu$ is ergodic with respect to $T$. ■

The previous proof is important not only because it shows that the Gauss measure is ergodic, but also because it introduces a series of techniques which are fundamental throughout this work, namely, using symbolic coding for the dynamics, the bounded distortion estimates density arguments.

**Proposition 2.2.5.** *Let $X$ be a symbolic space with alphabet $\mathscr{A} = \{1,\dots,m\}$ and let $T$ be the left shift on $X$. If $\mu$ is a Bernoulli measure on $X$, then $\mu$ is ergodic with respect to $T$.*

We postpone the proof of this proposition, as it is an easy consequence of a later proposition.
An important consequence of the ergodic theorem is that it characterizes ergodicity in terms of asymptotic independence of sets.

**Proposition 2.2.6.** *A measure $\mu$ is ergodic with respect to $T$ if and only if*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu\left(T^{-i}A \cap B\right) = \mu(A)\mu(B)$$

*for all measurable sets $A,B$.*

**Proof.** This is Corollary 1.14.2 in [Wal00]. ■

We will refer to this property as *asymptotic independence of the means*. In what follows we will define strongers notions of asymptotic independence, which will be central to our investigation.

**Definition 2.2.7.** *We say that the measure $\mu$ is* mixing *with respect to $T$ if*

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

*for all measurable sets $A,B$.*

This notion is weaker than probabilistic independence but stronger than ergodicity. The intuition of this definition is the following: we take two subsets of the phase space $A,B$, and we let one of them evolve according to the dynamics $T$. If we wait long enough, its evolved state $T^{-n}A$ is virtually independent of $B$. In particular, if we take $B = A$ we obtain that the system has a *memory loss* property.

**Corollary 2.2.8.** *Mixing implies ergodicity.*

If the sigma-algebra admits a generator, then we only need to check decay for the elements of the generating algebra:

**Proposition 2.2.9.** *Suppose $\mathscr{E}$ is an algebra of sets such that $\mathscr{B} = \sigma(\mathscr{E})$. Then a measure $\mu$ is mixing with respect to $T$ if and only if*

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

*for all sets $A, B \in \mathscr{E}$.*

**Proof.** This is Lemma 7.1.2 in [VO16]. ■

Mixing can also be characterized in terms of integrals:

**Proposition 2.2.10.** *A measure $\mu$ is mixing with respect to $T$ if and only if*

$$\lim_{n\to\infty} \int_X f \circ T^n \cdot g d\mu - \int_X f d\mu \int_X g d\mu = 0$$

*for all $f, g \in L^2$.*

**Proof.** Mixing implies the equality above for indicator functions, and by linearity, it holds for simple functions. For $L^1$ functions, an approximation argument yields the result. The opposite direction is obvious. ■

The quantity above is of great importance by itself, so we will give it a name:

**Definition 2.2.11.** *Let $f, g$ be functions in $L^2$. We define their* correlation function *with respect to the dynamics $T$ and the measure $\mu$ by*

$$C_n(f, g) = \int_X f \circ T^n \cdot g d\mu - \int_X f d\mu \int_X g d\mu.$$

One of the main problems of our investigation is to study the rate at which the correlation function decays with respect to $n$ for a certain class of functions. For more general dynamics we will need an adapted notion of decay of correlations or loss of memory which we will introduce in chapter 4.
We will prove now that the symbolic dynamical system is mixing.

**Proposition 2.2.12.** *Let $X$ be a symbolic space with alphabet $\mathscr{A} = \{1, \dots, m\}$ and let $T$ be the left shift on $X$. If $\mu$ is a Bernoulli measure on $X$, then $\mu$ is mixing with respect to $T$. Consequently, $\mu$ is ergodic.*

**Proof.** Let $A, B$ be two cylinders in $X$, that is, there exist two finite sequences $a_1, \ldots, a_k$ and $b_1, \ldots, b_p$ such that $A = [a_1, \ldots, a_k]$ and $B = [b_1, \ldots, b_p]$. Then for any $n \geq 1$ we obtain

$$T^{-(n+p)}A \cap B = \bigcup_{(i_1,\ldots,i_n)\in\mathscr{A}^n} [b_1, \ldots, b_p, i_1, \ldots, i_n, a_1, \ldots, a_k].$$

Since the cylinder sets of the same length are disjoint, it follows that

$$\begin{aligned}
\mu(T^{-(n+p)}A \cap B) &= \sum_{(i_1,\ldots,i_n)\in\mathscr{A}^n} \mu[b_1, \ldots, b_p, i_1, \ldots, i_n, a_1, \ldots, a_k] \\
&= \sum_{(i_1,\ldots,i_n)\in\mathscr{A}^n} \mu[b_1, \ldots, b_p]\mu[i_1, \ldots, i_n]\mu[a_1, \ldots, a_k] \\
&= \mu[A]\mu[B],
\end{aligned}$$

proving the first part of the proposition. The ergodicity of $\mu$ follows from corollary 2.2.8. ■

We formulate now a related notion of mixing from the topological point of view:

**Definition 2.2.13.** *Let $X$ be a topological space and $T: X \to X$ be a continuous transformation. We say that $T$ is* topologically mixing *if for any two non-empty open sets $U, V$, there exists $n_0 \in \mathbb{N}$ such that $T^{-n}U \cap V \neq \emptyset$ for all $n \geq n_0$.*

The structure of the set of ergodic measures on a space for a given transformation is quite complicated. We give one of its properties which will be useful when considering multiple different ergodic measures for a fixed transformation:

**Proposition 2.2.14.** *Suppose that the sigma-algebra of $X$ can be generated by a countable set. Let $\{\mu_i\}$ be a family of distinct invariant ergodic measures with respect to $T$. Then the measures are mutually singular: there exist a family of subsets of $X$ $\{P_i\}$, pairwise disjoint, invariant under $T$ such that $\mu_i(P_i) = 1$ for all $i$.*

**Proof.** This is lemma 4.3.3 from [VO16]. ■

## 2.3 Dimension theory

In this section we introduce the dimension theory elements we will study throughout this work. Recall the diameter of a set $U \subset \mathbb{R}^n$ is given by

$$|U| = \sup\{\|x - y\| : x, y \in U\},$$

where the norm is the Euclidian norm of $\mathbb{R}^n$. For a cover $\mathscr{U}$ of a set $X \subset \mathbb{R}^n$, its diameter is given by

$$\operatorname{diam}\mathscr{U} = \sup\{|U| : U \in \mathscr{U}\}.$$

**Definition 2.3.1.** *Given $X \subset \mathbb{R}^n$ and $\alpha \geq 0$, the $\alpha-$*dimensional Hausdorff measure *of $X$ is given by*

$$m(X,\alpha) = \lim_{\delta\to 0} \inf_{\mathscr{U}} \sum_{U\in\mathscr{U}} |U|^\alpha,$$

*where the infimum is taken over finite or countable covers $\mathscr{U}$ of $X$ with* $\operatorname{diam}\mathscr{U} \leq \delta$.

It is possible to prove (see section 3.2 of [Fal04]) that there exists a number $s \in [0,\infty]$ such that $m(X,\alpha) = \infty$ for $t < s$ and $m(X,\alpha) = 0$ for $t > s$, since $m(X,\alpha)$ is decreasing in $\alpha$ for a fixed set $X$.

**Definition 2.3.2.** *The unique number*

$$\dim_H X = \inf\{\alpha \in [0,\infty] \mid m(X,\alpha) = 0\}$$

*is called the* Hausdorff dimension *of $X$.*

We extend the notion of Hausdorff dimension to finite Borel measures on $\mathbb{R}^n$:

**Definition 2.3.3.** *Let $\mu$ be a finite Borel measure on $\mathbb{R}^n$. The* Hausdorff dimension *of $\mu$ is defined by*

$$\dim_H \mu = \inf\{\dim_H(Z) \mid \mu(\mathbb{R}^n \setminus Z) = 0\}.$$

We define now the analogue notion of Packing dimension:

**Definition 2.3.4.** *We say that a collection of balls $\{U_n\}_n \subset \mathbb{R}^n$ is a $\delta-$*packing *of the set $E \subset \mathbb{R}^n$ if the diameter of the balls is less than or equal to $\delta$, they are pairwise disjoint and their centres belong to $E$. For $\alpha \in \mathbb{R}^n$, the $\alpha-$dimensional pre-packing measure of $E$ is given by*

$$P(E,\alpha) = \lim_{\delta\to 0} \sup\left\{\sum_n \operatorname{diam}(U_n)^\alpha\right\}$$

*where the supremum is taken over all $\delta-$packings of $E$. The $\alpha-$*dimensional packing measure *of $E$ is defined by*

$$p(E,\alpha) = \inf\left\{\sum_i P(E_i,\alpha)\right\}$$

*where the infimum is taken over all covers* $\{E_i\}$ *of* $E$. *Finally, we define the* packing dimension *of* $E$ *by*

$$\dim_p(E) = \sup\{s \mid p(E,\alpha) = \infty\} = \inf\{s \mid p(E,\alpha) = 0\}.$$

We extend the notion of packing dimension to finite Borel measures on $\mathbb{R}^n$.

**Definition 2.3.5.** *Let* $\mu$ *be a finite Borel measure on* $\mathbb{R}^n$. *The* Packing dimension *of* $\mu$ *is defined by*

$$\dim_p \mu = \inf\{\dim_p(Z) \mid \mu(\mathbb{R}^n \setminus Z) = 0\}.$$

Bounding the Hausdorff dimension from above or the Packing dimension from below usually involves the use of a single suitable cover of the space, while for bounds from below and above respectively, we have to deal with *every* cover of the space. There are several tools to help with this problem, and we will make use of the so called (local) Mass Distribution Principles. For this, we introduce the notion of local dimension.

**Definition 2.3.6.** *The* lower *and* upper pointwise dimensions/local dimension *of the measure* $\mu$ *at a point* $x \in X$ *are given by*

$$\underline{d}_\mu(x) = \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}, \quad \overline{d}_\mu(x) = \limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}.$$

*When both limits coincide, we call the common value the pointwise dimension/local dimension of* $\mu$ *at* $x$ *and denote it by* $d_\mu(x)$ *and say that* $\mu$ *is* exact dimensional *if* $\underline{d}_\mu(\cdot) = \overline{d}_\mu(\cdot)$ $\mu$*-almost everywhere.*

If $d_\mu(x) = d$, then $\mu(B(x,r)) \sim r^d$ for small values of $r$. We state now the local version of the Mass Distribution Principle.

**Proposition 2.3.7.** *Let* $X \subset \mathbb{R}^n$ *and* $\alpha \in (0,\infty]$, *then*

1. *If* $\underline{d}_\mu(x) \geq \alpha$ *for* $\mu$*–almost every* $x \in X$, *then* $\dim_H \mu \geq \alpha$*;*
2. *If* $\underline{d}_\mu(x) \leq \alpha$ *for every* $x \in X$, *then* $\dim_H X \leq \alpha$*,*
3. *If* $\overline{d}_\mu(x) \geq \alpha$ *for* $\mu$*–almost every* $x \in X$, *then* $\dim_p \mu \geq \alpha$*;*
4. *If* $\overline{d}_\mu(x) \leq \alpha$ *for every* $x \in X$, *then* $\dim_p X \leq \alpha$*,*

5. *We have*

$$\dim_H \mu = \operatorname{ess\,sup}\{\underline{d}_\mu(x) \mid x \in X\},$$
$$\dim_p \mu = \operatorname{ess\,sup}\{\overline{d}_\mu(x) \mid x \in X\},$$

**Proof.** This follows from Proposition 2.3 of [Fal97]. ■

In particular, if $\underline{d}_\mu(\cdot)$ is constant almost everywhere, then $\dim_H \mu$ is equal to that constant value. Analogously, if $\overline{d}_\mu(\cdot)$ is constant almost everywhere, then $\dim_p \mu$ is equal to that constant value.

## 2.4 Entropy, Lyapunov exponent

In this section we introduce the notions of entropy and Lyapunov exponent, which are central for the investigation of chapter 3. In this section, $(X, \mathscr{B}, \mu)$ denotes a probability space. When it is clear from the context, we will not mention $\mathscr{B}$ and $\mu$.

**Definition 2.4.1.** *A* partition *of $X$ is a finite/countable set $\mathscr{P} \subset \mathscr{B}$ such that*

$$\mu\left(\bigcup_{P\in\mathscr{P}} P\right) = 1$$

*and $P \cap Q = \emptyset$ for $P, Q \in \mathscr{P}$, $P \neq Q$.*

If we have two partitions of $X$, we can compare them with the following relation:

**Definition 2.4.2.** *Let $\mathscr{P}, \mathscr{Q}$ be two partitions of $X$. We say that $\mathscr{Q}$ is* finer *than $\mathscr{P}$ (or $\mathscr{P}$ is* coarser *than $\mathscr{Q}$) if every set in $\mathscr{Q}$ is contained in some set of $\mathscr{P}$, up to a set of measure zero. In this case we denote $\mathscr{P} \prec \mathscr{Q}$.*

If $x$ is a point of $X$ and $\mathscr{P}$ is a partition of $X$, we denote by $\mathscr{P}(x)$ the element of $\mathscr{P}$ containing $x$. Note that the function $x \mapsto \mathscr{P}(x)$ is defined $\mu$-almost everywhere.

**Definition 2.4.3.** *If $\{\mathscr{P}_i\}_{i\in I}$ is a finite/countable family of partitions, we define their* join *by*

$$\bigvee_{i\in I} \mathscr{P}_i := \left\{ \bigcap_{i\in I} P_i : P_i \in \mathscr{P}_i \right\}.$$

*If the collection $\{\mathscr{P}_i\}_{i\in I}$ consists of finitely many partitions $\mathscr{P}_1,\dots,\mathscr{P}_n$, we denote their join by $\mathscr{P}_1 \vee \dots \vee \mathscr{P}_n$.*

**Proposition 2.4.4.** *The join of a family of partitions is a partition. It also holds that $\mathscr{P},\mathscr{Q} \prec \mathscr{P} \vee \mathscr{Q}$.*

**Proof.** For all $i \in I$, there exists a set $X_i \subset X$ of full measure such that $\bigcup_{P\in\mathscr{P}_i} P = X_i$. Define $\widehat{X} = \bigcap_{i\in I} X_i$, which is a subset of full measure. Then by construction we have that $\bigcup_{P\in\bigvee_{i\in I}} P = \bar{X}$. On the other hand, if we take two elements of $\bigvee_{i\in I}\mathscr{P}_i$, $\bigcap_{i\in I} P_i$ and $\bigcap_{i\in I} Q_i$ with $(\bigcap_{i\in I} P_i)\cap(\bigcap_{i\in I} Q_i) \neq \emptyset$, then taking $x \in (\bigcap_{i\in I} P_i)\cap(\bigcap_{i\in I} Q_i)$ we have that $x \in P_i \cap Q_i$ for all $i \in I$. So $P_i = Q_i$ for all $i$, giving a contradiction. The second assertion follows immediately from the definition. ■

We define now the entropy of a partition:

**Definition 2.4.5.** *For a partition $\mathscr{P}$ of $X$, we define its* entropy *by*

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

*Here we follow the convention $0\cdot\infty = 0$.*

In general, the function the entropy of a join is less than the sum of the individual entropies of the partitions being considered:

**Lemma 2.4.6.** *If $\mathscr{P},\mathscr{Q}$ are partitions of $X$, then we have that*

$$H_\mu(\mathscr{P}\vee\mathscr{Q}) \leq H_\mu(\mathscr{P}) + H_\mu(\mathscr{Q}).$$

**Proof.** This is the remark after lemma 9.1.5 in [VO16]. ■

We introduce now the entropy of a measure preserving transformation. For this, we need a notion of how to *dynamize* partitions.

**Lemma 2.4.7.** *If $T$ is a measure preserving transformation and $\mathscr{P}$ is a partition of $X$, then $T^{-n}\mathscr{P} := \{T^{-n}(P) : P \in \mathscr{P}\}$ is also a partition of $X$ for all $n \geq 1$. In this case, we have that $H_\mu(T^{-n}\mathscr{P}) = H_\mu(\mathscr{P})$.*

**Proof.** Note that since $T$ preserves $\mu$, we have that

$$\mu\left(\bigcup_{P\in\mathscr{P}} T^{-1}(P)\right) = \mu\left(T^{-1}\bigcup_{P\in\mathscr{P}} P\right) = \mu\left(\bigcup_{P\in\mathscr{P}} P\right) = 1.$$

On the other hand, for $P,Q \in \mathscr{P}$ we have that $T^{-1}P \cap T^{-1}Q = T^{-1}(P \cap Q)$ so if $P \neq Q$ then $T^{-1}P \cap T^{-1}Q = \emptyset$, proving that $T^{-1}\mathscr{P}$ is a partition of $X$. Iterating this, the conclusion follows for $T^{-n}\mathscr{P}$. The second part of the lemma is trivial from the definition of entropy of a partition. ■

Now we can define the dynamical iterates of a partition.

**Definition 2.4.8.** *Let $T\colon X \to X$ be a measure preserving transformation on $X$, and suppose $\mathscr{P}$ is a partition of $X$. We define the $n$-th* dynamical iterate *of $\mathscr{P}$ by*

$$\mathscr{P}^n := \bigvee_{i=0}^{n-1} T^{-n}\mathscr{P}.$$

*For $x \in X$, the element of $\mathscr{P}^n$ containing $x$ is denoted by $\mathscr{P}^n(x)$.*

We are interested in the sequence of entropies $H_\mu(\mathscr{P}^n)$ for a given partition $\mathscr{P}$. This sequence is subadditive:

**Lemma 2.4.9.** *For any $n,m \geq 1$ we have $H_\mu(\mathscr{P}^{n+m}) \leq H_\mu(\mathscr{P}^m) + H_\mu(\mathscr{P}^n)$.*

**Proof.** This follows immediately from the observation that $\mathscr{P}^{n+m} = \mathscr{P}^m \vee T^{-m}(\mathscr{P}^n)$ and lemmas 2.4.6 and 2.4.7. ■

The previous lemma implies that the sequence of real numbers $a_n = H_\mu(\mathscr{P}^n)$ is subadditive, that is, $a_{n+m} \leq a_n + a_m$ for all $n,m \geq 1$.

**Lemma 2.4.10** (Fekete)**.** *If $a_n$ is a subadditive sequence of real numbers, then the limit $\lim_n \frac{a_n}{n}$ exists in $[-\infty,\infty)$ and is equal to $\inf_n \frac{a_n}{n}$.*

**Proof.** This is lemma 3.3.4 in [VO16]. ■

With this lemma we can consider the asymptotic growth rate of the entropies $H_\mu(\mathscr{P}^n)$:

**Definition 2.4.11.** *The* entropy of $T$ with respect to $\mathscr{P}$ *is defined as $h_\mu(T,\mathscr{P}) = \lim_n \frac{1}{n} H_\mu(\mathscr{P}^n)$. The* entropy of $T$ with respect to $\mu$ *is defined as $h_\mu(T) = \sup_{\mathscr{P}} h_\mu(T,\mathscr{P})$, where the supremum is taken over all partitions with finite entropy. When there is no risk of confusion, we will denote the entropy by $h_\mu$.*

**Example 2.4.12.** *Consider again $(X,T,\mu)$ to be a symbolic space with alphabet $\mathscr{A}$, $T$ the left-shift and $\mu$ a Bernoulli measure with probability vector $p = (p_1,\ldots) \in (0,1)^{\mathscr{A}}$. There is a natural partition $\mathscr{P}$ on $X$ given by $\mathscr{P} = \{C(a) = [a] : a \in \mathscr{A}\}$, where the sets*

*$C(a)$ are defined as in example 2.1.8. The iterates under $T$ of this partition are given by $\mathscr{P}^n = \{C(\alpha) = [a_1,\dots,a_n] : \alpha = (a_1,\dots,a_n) \in \mathscr{A}^n\}$. We can compute the entropy of the partition $\mathscr{P}^n$:*

$$\begin{aligned}
H_\mu(\mathscr{P}^n) &= -\sum_{(a_1,\dots,a_n)\in\mathscr{A}^n} p_{a_1}\cdot\dots\cdot p_{a_n}\log(p_{a_1}\cdot\dots\cdot p_{a_n}) \\
&= -\sum_j \sum_{(a_1,\dots,a_n)\in\mathscr{A}^n} p_{a_1}\cdot\dots\cdot p_{a_j}\cdot\dots\cdot p_{a_n}\log p_{a_j} \\
&= -\sum_j\sum_{a_j} p_{a_j}\log p_{a_j}\sum_{a_i,i\neq j} p_{a_1}\cdots p_{a_{j-1}}p_{a_{j+1}}\cdot p_{a_n} \\
&= -n\sum_a p_a\log p_a,
\end{aligned}$$

*from where it follows that $h_\mu(T,\mathscr{P}) = -\sum_{a\in\mathscr{A}} p_a \log p_a$.*

Computing $h_\mu(T)$ is in general much harder, as we have to compute a supremum over *all* partitions. For this, we have the following theorem:

**Theorem 2.4.13** (Kolmogorov-Sinai)**.** *Suppose that $\mathscr{P}$ is a partition with finite entropy such that $\sigma(\bigcup_n \mathscr{P}^n) = \mathscr{B}$. Then $h_\mu(T) = h_\mu(T,\mathscr{P})$.*

**Proof.** This is corollary 9.2.5 from [VO16]. ■

Using the previous theorem we can see that for a Bernoulli measure on the symbolic space, we have that $h_\mu(T) = -\sum_{a\in\mathscr{A}} p_a\log p_a$. We formulate now a notion of *local entropy*.

**Definition 2.4.14.** *For a partition $\mathscr{P}$ with finite entropy, we define the function*

$$x \mapsto h_\mu(T,\mathscr{P},x) = \lim_{n\to\infty} -\frac{1}{n}\log\mu(\mathscr{P}^n(x))$$

*and call it the* local entropy *function.*

The local entropy function is well-defined almost everywhere thanks to the Shannon-McMillan-Breiman theorem:

**Theorem 2.4.15** (Shannon-McMillan-Breiman)**.** *The limit defining $h_\mu(T,\mathscr{P},x)$ exists for $\mu$-almost every $x\in X$. The function $x\mapsto h_\mu(T,\mathscr{P},x)$ is integrarble and one has*

$$\int_X h_\mu(T,\mathscr{P},x)d\mu = h_\mu(T,\mathscr{P}).$$

*If $T$ is ergodic with respect to $\mu$, then $h_\mu(T,\mathscr{P},x) = h_\mu(T,\mathscr{P})$ for almost every $x\in X$.*

**Proof.** This is Theorem 9.3.1 from [VO16] ■

The previous result, combined with Kolmogorov-Sinai's theorem give a way to compute the entropy of a transformation $T$ using local analysis of a convenient partition.
Finally, we introduce the notion of Lyapunov exponent. While we introduced entropy in a general setting, for the Lyapunov exponent we restrict ourselves to maps of the unit interval, so we can give a definition simple enough for our purposes. For the rest of this section, $X = [0,1]$, $\mu$ a probability measure on $X$ and $T\colon [0,1] \to [0,1]$ a $\mu$-invariant, piece-wise differentiable map with $-\log T'$ integrable with respect to $\mu$.

**Definition 2.4.16.** *We define the* Lyapunov exponent *of $T$ with respect to $\mu$ by*

$$\lambda_\mu = \int_X \log\left|T'\right| d\mu.$$

Note that the Lyapunov exponent represents the *average* exponent of the derivative of the map $T$. In general this is a measure of the expansion properties of the dynamics. By the chain rule and the ergodic theorem, we have that

$$\lambda = \lim_{n\to\infty} -\frac{1}{n}\log\left|(T^n)'(x)\right|$$

for Lebesgue-almost every point $x \in [0,1]$.

## 2.5 Gibbs measures

In this section we introduce the concept of Gibbs measures, which provides a generalization of Bernoulli measures in symbolic spaces. We give a brief motivation for this in a probabilistic context: suppose that $(X, \mathscr{B}, \mu, T)$ is a probability preserving dynamical system and $\varphi\colon X \to \mathbb{R}$ a measurable function. We can construct a time series $Z_n = \varphi \circ T^n$, which by invariance of $T$, is stationary. This process is in general not i.i.d., but in some concrete cases it can be. Consider for instance $(X, \mathscr{B}, \mu, T)$ be a Bernoulli shift, and $\varphi(x) = x_1$, where $x = (x_1, \dots) \in X = \mathscr{A}^{\mathbb{N}}$. We can think of this random variable as sampling the first digit of random sequences in $X$, according to the probability measure $\mu$. In this setting, the process $Z_n = \varphi \circ T^n$ corresponds to sampling the $n$-th digit of random sequences, and as such, it is an i.i.d. process. Gibbs measures allow the time series $Z_n$ to have a certain degree of dependence. We proceed to introduce their definition:

**Definition 2.5.1.** *Let $(X, \mathscr{B}, T)$ be a symbolic space with alphabet $\mathscr{A}$. A measure $\mu$ is a* Gibbs measure *for an observable $\varphi\colon X \to \mathbb{R}$ if there exist constants $C > 0$ and $P = P(\varphi) \in \mathbb{R}$ such that*

$$C^{-1} \le \frac{\mu([a_1, \dots, a_n])}{\exp\left(-nP(\varphi) + \sum_{k=0}^{n-1} \varphi\left(T^k x\right)\right)} \le C$$

*for any* $(a_1,\ldots,a_n) \in \mathscr{A}^n$ *and* $x \in [a_1,\ldots,a_n]$. *We call the number* $P(\varphi)$ *the* topological pressure *of* $\varphi$.

The regularity of the observable $\varphi$ will play an essential role in the investigation of chapter 3. In general, Gibbs measures posses nice ergodic properties:

**Lemma 2.5.2.** *Gibbs measures are mixing, and consequently, ergodic.*

**Proof.** This is lemma 1.13 from [Bow08]. ∎

In general, we will consider Gibbs measures in symbolic spaces and project them to the unit interval $[0,1]$ and study the properties of these projected measures.

## 2.6 Iterated function systems

In example 2.1.11 we introduced the notion of iterated function system in $\mathbb{R}^n$. Let $\{S_1,\ldots,S_m : X \to X\}$ be an IFS with attractor $F$. The structure of $F$ can be described as follows: let $\mathscr{S}$ be the set of all non-empty compact subsets of $X$, and define a function $S : \mathscr{S} \to \mathscr{S}$ by

$$S(E) = \bigcup_{k=1}^{m} S_k(E).$$

**Lemma 2.6.1.** *For any non-empty compact set* $E \in \mathscr{S}$ *with* $\mathscr{S}_i(E) \subset E$ *for all* $i$, *we have that*

$$F = \bigcap_{k=0}^{\infty} S^k(E).$$

**Proof.** This is Theorem 9.1 from [Fal04]. ∎

**Example 2.6.2.** *Consider* $X = \mathbb{R}$, $S_1 = \frac{1}{3}x$ *and* $S_2 = \frac{1}{3}x + \frac{2}{3}$. *The the associated attractor* $F$ *is the usual Cantor set. In figure 2.2 the process of applying the different contractions to the interval* $[0,1]$.

**Example 2.6.3.** *Consider now* $X = \mathbb{R}^2$ *and the contractions*

$$\begin{aligned}
S_1 &= \left(\frac{1}{2}x + \frac{1}{4}, \frac{1}{2}y + \frac{\sqrt{3}}{4}\right) \\
S_2 &= \left(\frac{1}{2}x + \frac{1}{2}, \frac{1}{2}y\right) \\
S_3 &= \left(\frac{1}{2}x, \frac{1}{2}y\right).
\end{aligned}$$

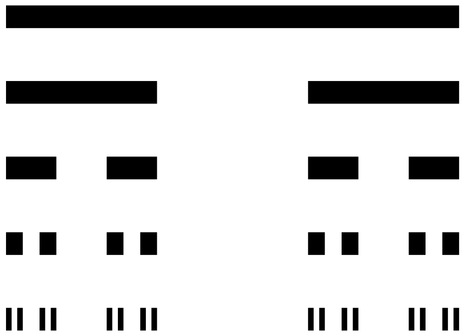

Figure 2.2: First four iterations of the construction of the Cantor set

*Then the associated attractor $F$ is the Sierpinski triangle. In figure 2.3 we show the fourth iteration of the process to construct $F$, that is, we show the set $\bigcup_{i\in\mathscr{I}_4} S^i(E)$ with $E$ being an equilateral triangle.*

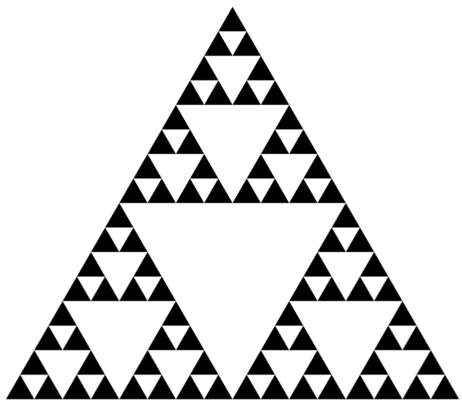

Figure 2.3: Fourth iteration of the construction of the Sierpinski triangle.

We can describe the attractor $F$ by making use of the idea of coding. Denote the set of finite words length $n$ by $\mathscr{I}_n := \{1,\dots,m\}^n$. Then

$$S^k(E) = \bigcup_{\mathscr{I}_k} S_{i_1} \circ \cdots \circ S_{i_k}(E).$$

From this and lemma 2.6.1 it follows that if $S_i(E) \subset E$ for all $i$, then for each $x \in F$ and there exists a sequence $(i_1,\dots) \in \{1,\dots,m\}^{\mathbb{N}}$ such that $x \in S_{i_1} \circ \cdots \circ S_{i_k}(E)$ for all $k \geq 1$. Thus, we have that

$$x = \bigcap_{k=1}^{\infty} S_{i_1} \circ \cdots \circ S_{i_k}(E).$$

We are interested in studying the geometric properties of the attractor of IFS. This problem is in general hard, but for a certain class of IFS much progress has been made.

**Definition 2.6.4.** *Let $\{S_1,\dots,S_m\}$ be an IFS in $\mathbb{R}^n$. We say the maps $S_i$ are* similarities *if there exist numbers $r_i \in (0,1)$ such that*

$$|S_i(x) - S_i(y)| = r_i |x - y|$$

*for all $x, y \in \mathbb{R}^n$. In this case we call the attractor $F$ of the IFS a* self-similar *set and the numbers $r_i$ the* similarity ratios.

In this case, given that the system satisfies a separation condition (see definition 5.2.1), then the Hausdorff dimension satisfies an equation in terms of the ratios $r_i$ (see theorem 5.2.2). A more complicated case is when the maps $S_i$ are not similarities but linear maps.

**Definition 2.6.5.** *We say that a transformation $S : \mathbb{R}^n \to \mathbb{R}^n$ is* affine *if it can be written as*

$$S(x) = Ax + b,$$

*where $A$ is an $n \times n$ matrix and $b \in \mathbb{R}^n$. If an IFS consists of contracting affine transformations, we say that the attractor is a* self-affine *set.*

In order to study the dimension theoretical properties of self-affine sets, Falconer introduced in [Fal97] the singular value function. This function keeps track of the rate of contraction in the strongest directions for products of matrices, and enables us to obtain precise information about the scale and number of elements of optimal coverings used to obtain effective bounds for the Hausdorff dimension of attractors.

## 2.7 Sequential and random dynamical systems

In this section we introduce the general theory of sequential and random dynamical systems. While for the sake of clarity we do most of it in great generality, we will work in a context of one dimensional piecewise continuous transformations of the unit interval. We will limit to the case of dynamics indexed by a discrete set, but a more general definition can be given. In this section $(X, \mathscr{B}, m)$ represents a probability space. We will refer to $m$ as the *reference measure*.

**Definition 2.7.1.** *A* (random) dynamical system *consists of a probability space* $(\Sigma, \mathcal{M}, \mathbb{P})$*, a measurable transformation* $\sigma\colon \Sigma \to \Sigma$ *such that* $\sigma \circ \mathbb{P} = \mathbb{P}$*, a measurable space* $(X, \mathscr{B})$ *and for each* $\omega \in \Sigma$*, a measurable transformation* $T_\omega\colon X \to X$ *of the space* $(X, \mathscr{B})$*. With this, we define a random dynamics by*

$$\begin{aligned} &F\colon \Sigma \times X \to \Sigma \times X \\ &F(\omega, x) = (\sigma(\omega), T_\omega(x)). \end{aligned}$$

We will often refer to $X$ as the *phase space* of the dynamics, and to the maps given by the individual actions of $F(s, \cdot)$ as the *dynamics*. For random dynamical systems, we will call $\Sigma$ the *noise space*.
The way we think of random dynamical systems is as follows: we sample an element $\omega$ of the noise space $\Sigma$ according to the probability distribution $\mathbb{P}$, and this element defines a sequence $\{\sigma^n(\omega)\}$. Associated to each element of the sequence, there is a map $T_{\sigma^n(\omega)}$, and we compose these maps sequentially: $T_\omega,\ T_{\sigma(\omega)} \circ T_\omega,\ T_{\sigma^2(\omega)} \circ T_{\sigma(\omega)} \circ T_\omega$. The second component of the iterates of $F$ gives the composition of this sequences of maps. We define the $n$-fold composition of $F$ with itself by

$$F^n(\omega, x) = (\sigma^n(\omega), T_{\sigma^n(\omega)} \circ \cdots \circ T_\omega).$$

The next example is the main example of random dynamical systems that we are going to work with:

**Example 2.7.2.** *Suppose* $\{T_1, \ldots, T_m\}$ *is a finite family of measurable transformations on* $X$*, and* $p = (p_1, \ldots, p_m) \in (0,1)^m$ *is a probability vector which defines a probability measure* $P$ *on* $\{1, \ldots, m\}$*. Define the noise space* $\Sigma = \{1, \ldots, m\}^{\mathbb{N}}$ *with the probability measure* $\mathbb{P} = P^{\otimes \mathbb{N}}$ *and* $\sigma(\omega)_n = \omega_{n+1}$*, that is, the left shift on* $\Sigma$*. For each element* $\omega \in \Sigma$*, define* $T_\omega = T_{\omega_1}$*, so then* $F^n(\omega, x) = (\sigma^n(\omega), T_{\omega_n} \circ \cdots \circ T_{\omega_1})$*. We will call this type of random dynamical system an i.i.d. random dynamical system.*

For random dynamical systems, it is often not the case that there exists an invariant measure for all maps. We need a notion which takes into account the random nature of random dynamical systems:

**Definition 2.7.3.** *A measure* $\mu$ *on* $X$ *is called a* stationary measure *for the random dynamical system* $F\colon \Sigma \times X \to \Sigma \times X$ *if*

$$\mu(A) = \int_\Sigma \mu\left(T_\omega^{-1}(A)\right) d\mathbb{P}(\omega)$$

*for all measurable subsets* $A$ *of* $X$*, or equivalently, the measure* $\mathbb{P} \otimes \mu$ *on* $\Sigma \times X$ *is invariant under* $F$*.*

The previous definition means that the measure $\mu$ is invariant *in average*. Now we show a basic model of sequential and random dynamical systems:

**Example 2.7.4.** *Let* $X = [0,1]$ *and the transformations* $T_1(x) = 2x \pmod 1$ *and* $T_2(x) = 3x \pmod 1$*. For a given sequence* $\omega \in \{1,2\}^{\mathbb{N}}$*, define the* sequential composition *of the maps* $T_1, T_2$ *according to the sequence* $\omega$ *as the sequence of maps* $\mathcal{T}^k = T_{\omega_k} \circ \cdots \circ T_{\omega_1}$*. Note that the Lebesgue measure* $m$ *is invariant for both maps* $T_1, T_2$*, and so is for the sequential compositions* $\mathcal{T}^k$*. This will not be the case in general for the maps we will study in chapter 4.*
*We can also construct an i.i.d. random dynamical system* $F : \Sigma \times X \to \Sigma \times X$ *with this set of transformations. A stationary measure for this system is given by* $\mathbb{P} \otimes m$*.*

In section 2.1 we introduced the notion of non-singular transformations: a measurable function $T : X \to X$ is non-singular with respect to $m$ if $m(T^{-1}A) = 0$ if and only if $m(A) = 0$ for $A \in \mathscr{B}$. This means that the transformation $T$ does not turn sets of measure 0 into sets of positive measure. One of the consequences of this is that the pushforward measure $T_* m$ is absolutely continuous with respect to $m$, where $T_* m(A) = m(T^{-1}A)$ . Moreover:

**Lemma 2.7.5.** *If* $\mu_f$ *is an absolutely continuous measure with respect to* $m$*, with density* $f \in L^1(m)$*, then* $T_* \mu_f$ *is also absolutely continuous with respect to* $m$*.*

**Proof.** This follows immediately from the non-singularity of $T$ with respect to $m$. ∎

By Radon-Nikodym's theorem (Theorem 8.9 in [Bar14]), the density of the measure $T_* \mu_f$ is a function in $L^1(m)$.

**Definition 2.7.6.** *We define the* transfer operator *of* $T$ *with respect to* $m$ *as the function*

$$\begin{aligned} P : L^1(m) &\to L^1(m) \\ f &\mapsto \frac{dT_* \mu_f}{dm}. \end{aligned}$$

*We also define the* Koopman operator *of* $T$ *with respect to* $m$ *as the function*

$$\begin{aligned} U : L^\infty(m) &\to L^\infty(m) \\ g &\mapsto g \circ T. \end{aligned}$$

The transfer operator and the Koopman operator satisfy a duality relation:

**Proposition 2.7.7.** *For every $f \in L^1(m)$ and $g \in L^\infty(m)$ we have that*

$$\int_X Pf \cdot g dm = \int_X f \cdot U g dm.$$

*Moreover, $Pf$ is the unique element of $L^1(m)$ such that this equality holds for all $g \in L^\infty(m)$.*

**Proof.** First we check that the equality holds: for such choices of $f, g$ we have that

$$\begin{aligned}\int_X Pf \cdot g dm &= \int_X \frac{dT_* \mu_f}{dm} g dm \\ &= \int_X g d(T_* \mu_f) \\ &= \int_X g \circ T d\mu_f \\ &= \int_X f \cdot g \circ T dm,\end{aligned}$$

proving the assertion. To check that $Pf$ is the unique function satisfying the equality, suppose that there are two functions $h_1, h_2 \in L^1(m)$ such that the equality holds. Then, taking $g = \operatorname{sign}(h_1 - h_2)$ we have that

$$\begin{aligned}\int_X |h_1 - h_2| dm &= \int_X (h_1 - h_2) g dm \\ &= \int_X h_1 g dm - \int_X h_2 g dm \\ &= \int_X f g \circ T dm - \int_X f g \circ T dm \\ &= 0,\end{aligned}$$

from where it follows that $h_1 = h_2$ in $L^1(m)$ as claimed. ■

From this proposition it is easy to see that the transfer operator associcated to $T^n$ is $P^n$, the $n$-fold composition of $P$ with itself. In this section we will consider compositions of different maps, and consequently, we will have different transfer operators. More precisely, consider a collection of non-singular transformations $\{T_\alpha\}_\alpha$ on $X$, with associated transfer operators $\{P_\alpha\}_\alpha$. For any sequence $(\alpha_1, \dots, \alpha_n)$, we consider the composition $T_{\alpha_n} \circ \dots \circ T_{\alpha_1}$. By applying 2.7.7 successively to $T_{\alpha_n} \circ \dots \circ T_{\alpha_1}$ we obtain that its transfer operator is given by $P_{\alpha_n} \circ \dots \circ P_{\alpha_1}$.
Suppose now that the family of maps is finite, so we list it as $\{T_{\alpha_1}, \dots, T_{\alpha_m}\}$, and define $\Omega = \{\alpha_1, \dots, \alpha_m\} \subset (0, \alpha)$. Given a probability distribution on $\mathbb{P} = (p_1, \dots, p_m)$ on $\Omega$, define a Bernoulli measure $\mathbb{P}^{\otimes \mathbb{N}}$ on $\Sigma = \Omega^{\mathbb{N}}$ by $\mathbb{P}^{\otimes \mathbb{N}}\{\omega : \omega_{j_1} = \alpha_{j_1}, \dots, \alpha_{j_k} = a_k\} = \prod_{i=1}^k p_{\alpha_{j_i}}$ for every

finite cylinder and extending to the sigma-algebra generated by the cylinders of $\Sigma$ by Kolmogorov's extension theorem. This measure is invariant and ergodic with respect to the shift operator $\tau$ on $\Sigma$ , $\tau : \Sigma \to \Sigma$ acting on sequences by $(\tau(\omega))_k = \omega_{k+1}$ (see section 2.1). We will denote $\mathbb{P}^{\otimes \mathbb{N}}$ by $\nu$ from now on.
For $\omega \in \Sigma$ define $\mathcal{T}_\omega^n := T_{(\tau^n\omega)_1} \circ \cdots \circ T_{\omega_1} = T_{\omega_n} \circ \cdots \circ T_{\omega_1}$. We define the random dynamical system $F : \Sigma \times X \to \Sigma \times X$ by $F(\omega, x) = (\tau\omega, T_{\omega_1}x)$. We will also use $\Omega$-indexed subscripts for random transfer operators associated to the maps $T_{\omega_i}$ so that $P_{\omega_i} := P_{T_{\omega_i}}$. We will also abuse notation and write $P_\omega$ for $P_{\omega_1}$ if $\omega = (\omega_1, \omega_2, \ldots, \omega_n, \ldots)$. The analogue concept of invariance for a random dynamical system is stationarity, as defined in section 2.1. In this setting, it is equivalent to

$$\mu(A) = \int_\Sigma \mu\left(T_{\omega_1}^{-1}(A)\right) d\nu(\omega)$$

for every measurable set $A$. We now define a transfer operator for the random dynamical system, which captures the average behavior of the transfer operators of the individual transformations:

**Definition 2.7.8.** *The* annealed transfer operator $P : L^1(m) \to L^1(m)$ *is defined by averaging over all the transformations:*

$$P = \sum_{\omega \in \Omega} p_\omega P_\omega = \int_\Sigma P_\omega \, d\nu(\omega).$$

*The* annealed Koopman operator $U : L^\infty(m) \to L^\infty(m)$ *defined by*

$$U\varphi(x) = \sum_{\omega \in \Omega} p_\omega \varphi(T_\omega x) = \int_\Sigma \varphi(T_\omega x) d\nu(\omega).$$

The annealed operators satisfy the same duality relationship as the usual operators:

**Lemma 2.7.9.**

$$\int_X (U\varphi) \cdot \psi \, dm = \int_X \varphi \cdot P\psi \, dm$$

*for all observables $\varphi \in L^\infty(m)$ and $\psi \in L^1(m)$.*

Our main interest is in establishing statistical laws for the sequence of compositions $T_{\alpha_n} \circ \cdots \circ T_{\alpha_1}$, either for a fixed sequence $\{\alpha_i\}_{i \geq 1}$ or random choices of it according to a probability distribution on $\Omega$. In general we distinguish three different regimes, which we informally describe as:

1. We say that a result is *sequential* if it holds for all choices of sequences in $\Sigma$;

2. We say that a result is *annealed* if it holds for the average sequence in $\Sigma$;

3. We say that a result is *quenched* if it holds for almost every choice of sequence in $\Sigma$.

It may seem that quenched results are weaker than sequential ones, but in general for sequential results, rates and constants depend on the choice of the sequence of maps while quenched results hold *uniformly* for all sequences in a set of total probability.

CHAPTER 3

## DIMENSION OF MEASURES WITH INFINITE ENTROPY

### 3.1 Introduction

In this chapter we study the dimension of measures invariant under a certain class of maps of the unit interval $[0,1]$: Expanding Markov Renyi (EMR) maps. These maps $T\colon [0,1] \to [0,1]$ admit representations by means of symbolic dynamics, and satisfy smoothness properties that allow us to use ergodic theoretic methods to study their geometric properties. Given an ergodic T-invariant probability measure $\mu$, we are interested in the pointwise behavior of the *local dimension*

$$d(x) = \lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r},$$

where $B(x,r)$ denotes the open ball of center $x$ and radius $r$. This limit in general may not exist, in which case we study the corresponding limit superior and limit inferior (see definition 2.3.6). When the limit exists almost everywhere, we say that the measure is *exact dimensional*. If this is the case, by ergodicity of $\mu$ the value of the local dimension is constant almost everywhere (see proposition 2.2.2). Knowledge of the almost sure behavior of the local dimension yields information about the *Hausdorff* and the *packing* dimension of the measure (see proposition 2.3.7).

There are two dynamical quantities which are particularly relevant when studying the local dimension of such measures: the *metric entropy* $h_\mu$ (or simply the entropy) and the *Lyapunov exponent* $\lambda_\mu$ of $(T,\mu)$ (see section 3.4 for the ad-hoc definitions of $h_\mu$, $\lambda_\mu$). Formulae relating the dynamical invariants $h_\mu, \lambda_\mu$ and the local dimension have been extensively studied for the last few decades in the case $h_\mu < \infty$. For Bernoulli measures invariant under the Gauss map, Kinney and Pitcher proved in [KP66] that if the measure $\mu$ is defined by a probability vector $p = \{p_i\}$, the Hausdorff dimension of $\mu$ can be computed with the formula

$$\dim_H \mu = \frac{-\sum_{n=1}^{\infty} p_n \log p_n}{2\int_0^1 |\log x| \mathrm{d}\mu(x)}$$

provided that $\sum_{n=1}^{\infty} p_n \log n < \infty$.

For more general maps of the interval, in [LM85] the authors proved that for a $\mathscr{C}^1$ map $T\colon [0,1] \to [0,1]$ where $T$ and $T'$ are piecewise monotonic and the Lyapunov exponent $\lambda_\mu$ is positive, if $\mu$ is an invariant ergodic probability measure, then [LM85, corollary in the appendix] we have that the measure is exact dimensional and

$$\lim_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} = \frac{h_\mu}{\lambda_\mu}$$

$\mu$-almost everywhere. In particular, $\dim_H \mu = h_\mu/\lambda_\mu$ by proposition 2.3.7. Other versions of the formula were proved by Young and Hofbauer, Raith in [You82] and [HR92], among others. In all of these examples, it is assumed $0 < \lambda_\mu < \infty$. In the context of countable Markov systems, Mauldin and Urbanski proved ([MU03, theorem 4.4.2]) the following theorem:

**Theorem 3.1.1** (Volume Lemma). *Let $(X,T)$ be a countable Markov shift coded by the shift in countably many symbols $(\Sigma,\sigma)$. Suppose that $\mu$ is a Borel shift-invariant ergodic probability measure on $\Sigma$ such that at least one of the numbers $H_\mu(\alpha)$ or $\lambda_\mu$ is finite, where $H_\mu(\alpha)$ is the entropy of $\mu$ with respect to the natural partition $\alpha$ in cylinders of $\Sigma$ (see definition 2.4.5). Then $\mu$ is exact dimensional and*

$$\dim_H(\mu \circ \pi^{-1}) = \frac{h_\mu}{\lambda_\mu},$$

*where $\pi\colon \Sigma \to X$ is the coding map.*

The case when $\lambda_\mu = 0$ was studied by Ledrappier and Misiurewicz in [LM85], wherein they constructed a $\mathscr{C}^r$ map of the interval and a non-atomic ergodic invariant measure which has zero Lyapunov exponent and is such that the local dimension does not exist almost everywhere. More precisely, they show that the lower local dimension and upper local dimension are not equal ([LM85, theorem 1]):

$$\underline{d}_\mu(x) = \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} < \limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} = \overline{d}_\mu(x)$$

almost everywhere. For this construction, the authors consider a class of unimodal maps (Feigenbaum's maps).
The dimension of Bernoulli measures for the Gauss map $G$ was studied by Kifer, Peres and Weiss in [KPW01], where they show that there is a universal constant $\varepsilon_0 > 10^{-7}$ so that

$$\dim_H(\mu_p \circ \pi^{-1}) \leq 1 - \varepsilon_0$$

for every Bernoulli measure on the symbolic space coding the Gauss map, where $\pi$ is the coding map. This inequality holds even for the case where the entropy of the measure is infinite. They also show that for an infinite entropy Bernoulli measure $\mu$, the Hausdorff dimension satisfies $\dim_H \mu \leq 1/2$. Their method relies on estimating the dimension of the sets of points for which the frequency of a sequence of digits in their continued fraction expansion differs from the expected value by a certain threshold is uniformly (with respect to the sequence of digits) bounded from 1, and a bound on the dimension of points that lie in unusually short cylinders. This situation has been recently studied by Jurga and Baker (see [Jur18] and [BJ18]) using different methods. Concretely, in [Jur18] the author uses ideas of the Hilbert-Birkhoff cone theory and extract information about the dynamics through the transfer operator. On the other hand, in [BJ18]) the authors construct a Bernoulli measure $\mu_q$ such that $\sup_p \dim_H \mu_p = \dim_H \mu_q$, where the supremum is taken over all Bernoulli measures. This in conjunction with the Variational Principle (see [Wal00]) yield their result.
The focus of our investigation is twofold: in the first place, we investigate the Hausdorff dimensions of invariant ergodic measures for piecewise expanding maps of the interval with countably many branches. In particular, we focus on maps exhibiting similar properties to the Gauss map and measures with infinite entropy and infinite Lyapunov exponent. In the second place, we show that the measures considered are not exact dimensional, by showing that the upper dimension is positive while the lower dimension is zero almost everywhere.
The main result of this chapter is:

**Theorem 3.1.2.** *Let $T\colon [0,1] \to [0,1]$ be a Gauss-like map and $\mu$ be an infinite entropy Gibbs measure of controlled decay, and such that the decay ratio $s$ exists . Then $\underline{d}_\mu(x) = 0, \overline{d}_\mu(x) = s$ $\mu$-almost everywhere.*

This shows that there is a dimension gap for this class of maps and measures. For the Gauss map, $s = 1/2$. The Gibbs assumption on the measure implies that a certain sequence of observables can be seen as a non-integrable stationary ergodic process and allows us to use some tools of infinite ergodic theory developed by Aaronson and Nakada (see [Aar77], [AN03]). In particular, the pointwise behavior of the Birkhoff sums *excluding the biggest term* of such sums (trimmed sums) plays a fundamental role in our arguments. We remark that the methods used in the context of finite entropy fail, as they rely on the fact that the measure and diameter of the iterates of the natural Markov partition decrease at an exponential rate given by $h_\mu$ and $\lambda_\mu$ respectively, enabling the use of coverings by balls of different scales. To tackle this problem, we make use of more refined coverings of balls, which are capable of detecting the asymptotic interaction between the Gibbs measure and the Lebesgue measure.

## 3.2 Expanding Markov-Renyi maps

Recall the proof that the Gauss measure $\mu$ is ergodic with respect to the Gauss map $T$ (see proposition 2.2.4). The essential ingredients of the proof are the metric estimates

$$\left|T'(x)\right| \geq 1 \ , \ \left|(T^2)'(x)\right| \geq 2 \ , \ \left|T''(x)/T'(x)^2\right| \leq 2,$$

the fact that the invariant measure $\mu$ is absolutely continuous with respect to the Lebesgue measure $m$, and that $T$ is locally a bijection. We define a general class of maps which satisfy these properties, and hence, the same results apply to them.

**Definition 3.2.1.** *We say that a map* $T\colon I \to I$ *of the interval* $I = [0,1]$ *is an* EMR (expanding Markov-Renyi) map *if there is a countable collection of closed intervals* $\{I(n)\}$ *(with disjoint interiors* $\operatorname{int} I(n)$*) such that:*

1. *The map is* $\mathscr{C}^2$ *on* $\bigcup_n \operatorname{int} I(n)$,
2. *Some power of* $T$ *is uniformly expanding, i.e., there is a positive integer* $r$ *and a constant* $\alpha > 1$ *such that* $|(T^r)'(x)| \geq \alpha$ *for all* $x \in \bigcup_n \operatorname{int} I(n)$,
3. *The map is Markov and can be coded by a full shift (see next subsection),*
4. *The map satisfies Renyi's condition: there is a constant* $E > 0$ *such that*

$$\sup_{n \in \mathbb{N}} \sup_{x,y,z \in I(n)} \frac{|T''(x)|}{|T'(y)||T'(z)|} \leq E,$$

Under these condition, the following results hold (see theorem in 1 of section 4, chapter 7 of [CFS12]):

**Theorem 3.2.2.** *Suppose that $T$ is an EMR map. Then there exists a $T$-invariant Borel probability measure $\mu$, which is equivalent to the Lebesgue measure $m$, and $K^{-1} \leq \frac{d\mu}{dm} \leq K$ for a positive constant $K$. The measure $\mu$ is mixing, and in particular, ergodic.*

The construction of the measure follows a standard argument of taking a weak limit of the sequence pushforwards of the reference measure through the dynamics. The metric condition 4 of the definition of EMR map is stronger than the metric properties of the proof of the ergodicity of the Gauss measure. Condition 2 is a uniform hyperbolicity condition for some iterate of the map.
This class of maps was first introduced in [PW99] in the context the multifractal analysis of the Lyapunov exponent for the Gauss map. Renyi's condition provides good estimates for the Lebesgue measure of the cylinders associated to the Markov structure of the map (see next subsection). While the absolutely continuous invariant measure is an object of interest in dynamical systems, we will study measures which are mutually singular with respect to this measure. The above properties of the map are useful to give metric estimates in terms of symbolic dynamics (see section 3.3).
For simplicity, we will assume that the maps are orientation preserving (the orientation reversing case only differs in the relative position of the cylinders). The set of branches must accumulate at least at one point, and we assume that it accumulates at exactly one point: we also assume that the branches accumulate on the left endpoint of $I$ (the case when the branches accumulate in the right endpoint of $I$ is analogous). Re-indexing if necessary, we can assume that $I(n+1) < I(n)$ for all $n$, in the sense that $x < y$ for all $x \in \operatorname{int} I(n+1)$ and $y \in \operatorname{int} I(n)$. Let $r_n = |I(n)|$.

**Definition 3.2.3.** *We say that an EMR map $T$ is a* Gauss-like map *if it satisfies the following conditions:*

1. $r_n > 0$ *for every* $n \in \mathbb{N}$,
2. $r_{n+1} \leq r_n$,
3. $\sum_n r_n = 1$,
4. $0 < K \leq r_{n+1}/r_n \leq K' < \infty$ *for some constants* $K, K'$,
5. $\{r_n\}$ *decays at most polynomially as n goes to infinity (see definition* (3.6.3)*).*

In figure 3.1 we show a generic example of a Gauss-like map.

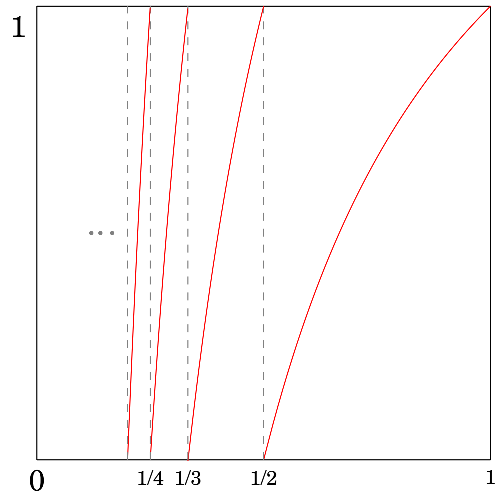


Figure 3.1: Example of an orientation-preserving Gauss-like map with the choice $I(n) = \left[\frac{1}{n+1}, \frac{1}{n}\right]$.

We want to keep in mind piecewise linear functions as the main example, as for this class of maps, calculations are simplified. We will also keep in mind the example of the Gauss map.

## 3.3 Markov structure and symbolic coding

We describe now the Markov structure of the maps considered. Given a finite sequence of natural numbers $(a_1,\ldots,a_n) \in \mathbb{N}^n$, the *n-th level cylinder* associated to $(a_1,\ldots,a_n)$ is the set $I(a_1,\ldots,a_n) = I_{a_1} \cap T^{-1}(I(a_2)) \cap \ldots \cap T^{-(n-1)}(I(a_n))$. Let $\mathscr{O} = \bigcup_n \bigcup_k T^{-n}(\partial I(k))$, then given $x \in [0,1] \setminus \mathscr{O}$ and $n \in \mathbb{N}$, there exists a unique sequence $(a_1(x), a_2(x), \ldots) \in \mathbb{N}^{\mathbb{N}}$ such that $x \in I(a_1(x),\ldots,a_n(x))$ for every $n$ (here $\partial I(k)$ represents the boundary of the set $I(k)$, in this case, consisting of the two endpoints of the interval). We denote this sequence by $(a_1, a_2, \ldots)$ when $x$ is clear from the context. We also denote $I_n(x) = I(a_1,\ldots,a_n)$ and we say $x$ is coded by the sequence $(a_n)$. From now on, whenever we say $x \in I$, we mean $x \in I \setminus \mathscr{O}$.

Let $\Sigma = \mathbb{N}^{\mathbb{N}}$ and $\sigma: \Sigma \to \Sigma$ be the left shift over $\mathbb{N}$: $(\sigma(x_n))_n = x_{n+1}$. Then the cylinders in the symbolic space are defined by

$$C(a_1, a_2, \ldots, a_n) = \left\{(x_n) \in \Sigma \mid x_j = a_j \text{ for } j = 1, \ldots, \right\}.$$

We endow the space $\Sigma$ with the topology generated by the cylinders defined above. Then the map $\pi: \Sigma \to I \setminus \mathscr{O}$ given by $\pi((x_n)) = \bigcap_n I(x_1,\ldots,x_n)$ is a continuous bijection.

Given $x \in I$ with coding sequence $(a_n)$ and $n \geq 1$, denote by $I_n^l(x) = I(a_1, \dots, a_{n-1}, a_n + 1)$ (resp $I_n^r(x) = I(a_1, \dots, a_{n-1}, a_n - 1)$ if $a_n \geq 2$) the level $n$ cylinder on the left (resp right) of $I_n(x)$. Also, denote by $\widehat{I}_n(x) = I_n(x) \cup I_n^r(x) \cup I_n^l(x)$. If there is no risk of confusion, we omit the dependence on $x$.
Renyi's condition introduced in the previous subsection implies that the length of each cylinder is comparable to the derivative of the iterates of the map at any point of the cylinder. More precisely,

**Lemma 3.3.1.** *There exists a constant $D > 0$ such that*

$$0 < D^{-1} \leq \left|(T^n)'(x)\right| \cdot |I(a_1, \dots, a_n)| \leq D$$

*for every finite sequence $(a_1, \dots, a_n) \in \mathbb{N}^n$ and $x \in I(a_1, \dots, a_n)$.*

**Proof.** The proof of the ergodicity of the Gauss map 2.2.4 applies *mutatis mutandis* in this setting and shows that

$$\frac{\left|(T^n)'(x)\right|}{\left|(T^n)'(y)\right|} \leq C_1$$

for some constant $C_1$, for all $x, y \in I(a_1, \dots, a_n)$, $(a_1, \dots, a_n) \in \mathbb{N}^n$. Integrating with respect to the Lebesgue measure over $y$ yields the result. ∎

A notion of dimension which is more adapted to the underlying structure of our dynamical system is the symbolic dimension, which we proceed to define.

**Definition 3.3.2.** *Given $x \in I$, we define the* lower symbolic dimension *of $\mu$ at $x$ by*

$$\underline{\delta}(x) = \liminf_{n \to \infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|},$$

*and the* upper symbolic dimension *of $\mu$ at $x$ by*

$$\overline{\delta}(x) = \limsup_{n \to \infty} \frac{\log \mu(I_n(x))}{\log |I_n(x)|},$$

*If $\overline{\delta}(x) = \underline{\delta}(x)$, then we define the* symbolic dimension *of $\mu$ at $x$ as the common value, denote it by $\delta(x)$, and we say that $\mu$ is* symbolic exact dimensional *if $\underline{\delta}(x) = \overline{\delta}(x)$ almost everywhere with respect to $\mu$.*

We recall now the definition of Gibbs measures:

**Definition 3.3.3.** *Let $\mu$ be an invariant probability measure with respect to $T$. Then we say that $\mu$ is a* Gibbs measure *associated to the potential* $\log\varphi\colon \Sigma \to \mathbb{R}$, *that is, there exists a constant $C > 0$ so that*

$$C^{-1} \le \frac{\mu(C(a_1,\dots,a_n))}{\exp\left(-nP(\log\varphi) + S_n(\log\varphi)(x)\right)} \le C,$$

*where $x$ is any point in $C(a_1,\dots,a_n)$, $(a_1,\dots,a_n,\dots)$ is any sequence in $\Sigma$, $S_n f(x)$ is the Birkhoff sum of $f$ at the point $x$, and $P(\log\varphi)$ is a constant (depending on the potential).*

The constant $P(\log\varphi)$ is usually called the *topological pressure* of $\log\varphi$. In this chapter we will not call such constant pressure, as it does not carry the same meaning it does in the finite entropy case. Throughout this work we will assume that $P(\log\varphi) = 0$, otherwise we can take the potential $\log\varphi - P(\log\varphi)$ for which $P\left(\log\varphi - P(\log\varphi)\right) = 0$. It is important to note that it is not trivial that this will not affect our computations, and we will show later how we can overcome that difficulty. The sequence $p_n = \mu(I(n))$ will be of particular relevance for our computations.
We can project this measure to $I$ by setting $\widehat{\mu} = \pi^{-1} \circ \mu$. We assume these measures are invariant and ergodic with respect to $T$. We will denote by $\mu$ both the measure in the symbolic space and the projected measure.
We define the *n-th variation* of the potential $\log\varphi$ by

$$\mathrm{var}_n(\log\varphi) = \sup\{|\log\varphi(x) - \log\varphi(y)| \mid x, y \in I(a_1,\dots,a_n), (a_1,\dots,a_n) \in \mathbb{N}^n\}.$$

**Definition 3.3.4.** *Let $x_n$ be the unique fixed point of $T$ in $I(n)$. We define then the* decay ratio *by*

$$s = \lim_{n\to\infty} \frac{\log\varphi(x_n)}{\log r_n} = \lim_{n\to\infty} \frac{\log p_n}{\log r_n},$$

*whenever any of these limits exists. Similary, the* tail decay ratio *is defined by*

$$\widehat{s} = \lim_{n\to\infty} \frac{\log\sum_{m\ge n}\varphi(x_m)}{\log\sum_{m\ge n} r_m} = \lim_{n\to\infty} \frac{\log\sum_{m\ge n} p_m}{\log\sum_{m\ge n} r_m}.$$

Whenever these two quantities exists, both definitions given for $s$ and $\widehat{s}$ agree since $\mu$ is a Gibbs measure. Throughout this chapter we will assume that the decay ratio exists for our measure. Note also that the definitions above are independent of the choice of the point $x_n$ representing each cylinder if $\mathrm{var}_1(\log\varphi) < \infty$. By the Cersàro-Stolz theorem (see [Fur13, Appendix B]) we can write the decay ratio as

$$s = \lim_{n\to\infty} \frac{\sum_{k=1}^n p_n \log p_n}{\sum_{k=1}^n p_n \log r_n}.$$

**Definition 3.3.5.** *Assume that* $\mathrm{var}_1(\log\varphi) < \infty$. *Suppose that for the sequence* $q = \{q_n\}_n = \{\varphi(x_n)\}_n$ *we have*

$$0 < K \le p_{n+1}/p_n \le 1$$

*for every* $n \in \mathbb{N}$, *for a constant* $K$. *Then we say that* $\mu$ *has* controlled decay.

This condition prevents the existence of large jumps for the potential along sufficiently sparse subsequences of $\{x_n\}$. By the Gibbs property, the properties hold if we replace $p_n$ by $q_n$.

## 3.4 Entropy and Lyapunov exponent

Our main results are for a class of measures with infinite entropy. This condition can be expressed by saying that the potential $-\log\varphi$ is not integrable with respect to $\mu$.
Our definition of entropy differs from the conventional (see section 2.4), as we deal with partitions with infinite entropy. For this, recall the Shannon-McMillan-Breiman theorem adapted to our system, which in the case of Gibbs measures, is equivalent to the Ergodic theorem:

**Theorem 3.4.1** (Shannon-McMillan-Breiman, infinite entropy)**.** *For any Gibbs measure* $\mu$ *associated to a potential with finite first variation, the limit*

$$\lim_{n\to\infty} -\frac{1}{n}\log\big(\mu(I_n(x))\big) \tag{3.1}$$

*exists* $\mu$*-almost everywhere and is constant. If* $\sum_n -p_n \log p_n < \infty$, *then such constant is finite; otherwise, it is equal to infinity.*

The proof for the case when the series is finite is the usual for Shannon-McMillan-Breiman theorem, see section 9.3 in [VO16]. The proof for the infinite case follows from lemma 3.4.2 and lemma 3.4.3, using that the measures have the Gibbs property. We define then the entropy $h_\mu$ as the almost sure value of the limit in theorem 3.4.1.

**Lemma 3.4.2.** *For a Gibbs measure with finite first variation, the entropy* $h_\mu$ *is finite if and only if any of the series*

$$-\sum_{n=1}^{\infty} q_n \log q_n \ , \ -\sum_{n=1}^{\infty} p_n \log p_n$$

*converges.*

**Proof.** The partition of $[0,1]$ by cylinders $\{I(n)\}$ is a generating partition, and hence theorem 2.4.13 allows us to compute the entropy of $\mu$ using the entropy of this partition. The entropy of $\mu$ with respect to this partition is given by

$$H(\mu,\alpha) = -\sum_{n=1}^{\infty} p_n \log p_n.$$

The convergence of this series is equivalent to the convergence of $-\sum_{n=1}^{\infty} q_n \log q_n$ since we have

$$\exp(-\operatorname{var}_1 \log\varphi) < \frac{q_n}{\varphi(x)} < \exp(\operatorname{var}_1 \log\varphi),$$
$$C^{-1} \leq \frac{p_n}{\varphi(x)} \leq C$$

for any $x \in I(n)$. ■

We prove a well known fact about non-integrable observables.

**Lemma 3.4.3.** *Let $f\colon [0,1] \to \mathbb{R}$ be a bounded below measurable function such that $\int_0^1 f \,\mathrm{d}\mu = \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) = \infty$$

*for $\mu$ almost every point.*

**Proof.** The proof is an standard application of the Monotone Convergence Theorem. Assume $f$ is positive (otherwise, decompose $f$ into its positive and negative part) and let $M > 0$. Then

$$\begin{aligned}\liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) &\geq \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \min\{f \circ T^k, M\}(x) \\ &= \int_0^1 \min\{f, M\}(x) \mathrm{d}\mu(x)\end{aligned}$$

by Birkhoff's Ergodic Theorem applied to $\min\{f,M\}$. By the Monotone Convergence Theorem,

$$\lim_{M\to\infty} \int_0^1 \min\{f,M\}(x)\mathrm{d}\mu(x) = \int_0^1 f \,\mathrm{d}\mu(x) = \infty$$

from where we conclude the result. ■

This result implies in particular that we can assume that the pressure of our potential is zero, as $S_n(\log\varphi)$ dominates $-nP(\log\varphi)$ when $\log\varphi$ is not integrable.
Now we can finish the proof of theorem 3.4.1 by noting that the Gibbs property of $\mu$ implies that

$$-\frac{1}{n}\log C - \frac{1}{n}S_n(\log\varphi)(x) \le -\frac{1}{n}\log(\mu(I_n(x))) \le \frac{1}{n}\log C - \frac{1}{n}S_n(\log\varphi)(x),$$

and using that the first variance of $(\log\varphi)$ is finite, we can also bound the integral of $\log\varphi$ by

$$\int_X \log\varphi d\mu \ge \sum_{k=1}^{\infty}\mu(I(k))\min_{x\in I(k)}\log\varphi(x) \ge C\sum_{k=1}^{\infty}p_k\log p_k + \sum_{k=1}^{\infty}p_k\log C,$$

and by applying lemmas 3.4.2 and lemma 3.4.3 the result follows.

## 3.5 Symbolic dimension

We formulate a lemma regarding the metric and measure theoretic properties of the cylinders associated to the map. This will allow us to write geometric quantities in ergodic theoretic terms. Its proof is a standard application of the bounded distortion and Gibbs properties.

**Lemma 3.5.1.** *For every finite sequence $(a_1,\dots,a_n)\in\mathbb{N}^n$ and $j\in\mathbb{N}$, we have that*

*(a)* $\left|\log|I(a_1,\dots,a_n)| - \sum_{k=1}^n \log r_{a_k}\right| \le nD_1 + D_2$

*(b)* $\left|\log\left|\bigcup_{m=0}^j I(a_1,\dots,a_{n-1},a_n+m)\right| - \sum_{k=1}^{n-1}\log r_{a_k} - \log\left(\sum_{k=0}^j r_{a_n+k}\right)\right| \le nD_1 + D_2,$

*(c)* $\left|\log\left|\bigcup_{m=0}^\infty I(a_1,\dots,a_{n-1},a_n+m)\right| - \sum_{k=1}^{n-1}\log r_k - \log\left(\sum_{k=0}^\infty r_{a_n+k}\right)\right| \le nD_1 + D_2,$

*(d)* $\left|\log\mu(I(a_1,\dots,a_n)) - \sum_{k=1}^n \log p_{a_k}\right| \le nG_1 + G_2,$

*(e)* $\left|\log\mu\left(\bigcup_{m=0}^j I(a_1,\dots,a_{n-1},a_n+m)\right) - \sum_{k=1}^{n-1}\log p_{a_k} - \log\left(\sum_{k=0}^j p_{a_n+k}\right)\right| \le nG_1 + G_2,$

*(f)* $\left|\log\mu\left(\bigcup_{m=0}^\infty I(a_1,\dots,a_{n-1},a_n+m)\right) - \sum_{k=1}^{n-1}\log p_{a_k} - \log\left(\sum_{k=0}^\infty p_{a_n+k}\right)\right| \le nG_1 + G_2,$

*where $D_1,D_2$ are distortion constants and $G_1,G_2$ are constants arising from the Gibbs property.*

**Proof.** We prove only the first part as the proof of the rest is similar. Fix a sequence $(a_1,\ldots,a_n)\in\mathbb{N}^n$, then by 3.3.1 we have that

$$\begin{aligned}0<D^{-1}\leq&\left|(T^n)'(x)\right|\cdot|I(a_1,\ldots,a_n)|\leq D\\ D^{-1}\leq& r_{a_k}|T'(x)|\leq D\end{aligned}$$

for any $x\in I(a_k)$. Using the chain rule, we have that

$$|(T^{n-1})'(x)|=|T'(T^{n-1}(x))|\cdot|T'(T^{n-2}(x))|\cdot\ldots\cdot|T'(T^1(x))|\cdot|T'(x)|,$$

from where it follows that

$$D^{-(n+1)}\leq(r_{a_1}\cdot\ldots\cdot r_{a_n})^{-1}\cdot|I(a_1,\ldots,a_n)|\leq D^{n+1}$$

which yields the first part. ■

We proceed to compute the symbolic dimension of our system.

**Theorem 3.5.2.** *Let $T$ be an EMR map and $\mu$ a Gibbs measure with controlled decay and infinte entropy. Then if the decay ratio $s$ exists, we have that $\mu$ is symbolic-exact dimensional and for $\mu$-almost every $x\in I$,*

$$\delta(x)=s.$$

**Proof.** By Lemma 3.4.3 applied to the observables $\log\varphi$ and $\log r_{a_1}$ and Lemma 3.5.1, we have

$$\begin{aligned}\underline{\delta}(x)\leq\liminf_{n\to\infty}\frac{S_n(\log\varphi)(x)}{-nD_1-D_2+S_n(\log r_{a_1})(x)}&=\liminf_{n\to\infty}\frac{\log(q_{a_1}\ldots q_{a_n})}{\log(r_{a_1}\ldots r_{a_n})},\\ \underline{\delta}(x)\geq\liminf_{n\to\infty}\frac{S_n(\log\varphi)(x)}{nD_1+D_2+S_n(\log r_{a_1})(x)}&=\liminf_{n\to\infty}\frac{\log(q_{a_1}\ldots q_{a_n})}{\log(r_{a_1}\ldots r_{a_n})},\end{aligned}$$

for almost every $x\in I$, and analogously for the upper symbolic dimension

$$\overline{\delta}(x)=\limsup_{n\to\infty}\frac{\log(q_{a_1}\ldots q_{a_n})}{\log(r_{a_1}\ldots r_{a_n})}$$

where $(a_1,a_2,\ldots)$ is the sequence coding $x$. With a similar argument, we can also show that the same holds true if we switch $q_n$ for $p_n$:

$$\underline{\delta}(x)=\liminf_{n\to\infty}\frac{\log(p_{a_1}\ldots p_{a_n})}{\log(r_{a_1}\ldots r_{a_n})},$$

and analogously for the upper symbolic dimension.
For $x \in I$ and $n,k \geq 1$, define

$$f_{n,k}(x) = \#\{i \in \{1,\ldots,n\} \mid a_i(x) = k\},$$

that is, the number of times the orbit of $x$ visits the interval $I_k$ in the first $n$ steps. Recall that from the Birkhoff Theorem, we have that for every $k$,

$$\lim_{n\to\infty} \frac{f_{n,k}}{n} = p_k$$

for $\mu-$almost every $x \in I$. In particular, the orbit of almost every $x \in I$ visits every cylinder $I(n)$ infinitely many times. Fix $x$ in the set where the convergence holds, and then define $m : \mathbb{N} \to \mathbb{N}$ by $m(n) = \max\{k_i(x) \mid i \in \{1,\ldots,n\}\}$, where $k_i(x)$ is the $i$-th digit of the expansion of $x$. The previous remark shows that $m$ is unbounded, and it is clearly non-decreasing. Thus, we can write

$$-\log(r_{k_1}\ldots r_{k_n}) = -\sum_{j=1}^{n} \log r_{k_j} = -\sum_{j=1}^{m(n)} f_{n,j} \log r_j.$$

Given $\epsilon > 0$, there exists $n_1$ such that

$$\left|\frac{\log p_k}{\log r_k} - s\right| < \epsilon$$

for every $k \geq n_1$, and consequently, $(-\log p_k) < (\epsilon + s)(-\log r_k)$ for $k \geq n_1$. For $n$ large enough so that $m(n) > n_1$, we write

$$\frac{\log(p_{k_1}\ldots p_{k_n})}{\log(r_{k_1}\ldots r_{k_n})} = \frac{\sum_{k=1}^{n_1} f_{n,k}(-\log p_k) + \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log p_k)}{\sum_{k=1}^{n_1} f_{n,k}(-\log r_k) + \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)}.$$

We split the sum in two different parts:

$$A(n) = \frac{\sum_{k=1}^{n_1} f_{n,k}(-\log p_k)}{\sum_{k=1}^{n_1} f_{n,k}(-\log r_k) + \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)},$$
$$B(n) = \frac{\sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log p_k)}{\sum_{k=1}^{n_1} f_{n,k}(-\log r_k) + \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)}.$$

For $k = 1,\ldots,n_1$ taking $\epsilon_k = p_k/2$ in the definition of $p_k$ as a limit, there exists $n_2 \geq n_1$ such that

$$\frac{np_k}{2} \leq f_{n,k} \leq \frac{3np_k}{2}$$

for every $n \geq n_2$. Thus, the terms $\sum_{k=1}^{n_1} f_{n,k}(-\log p_k)$ and $\sum_{k=1}^{n_1} f_{n,k}(-\log r_k)$ are asymptotically linear in $n$ for $n$, that is, when divided by $n$ they have a finite limit. We will show that $\sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)$ grows faster than linear as a function of $n$.
Given $M > 0$, since the Lyapunov exponent is infinite, there exists $n_3$ such that

$$\sum_{k=n_1+1}^{m(n)} p_k(-\log r_k) > 2M$$

for every $n \geq n_3$. Now, for $k = n_1+1, \ldots, m(n_3)$, take $\epsilon_k = p_k/2$ and so there exists $n_4 \geq n_3$ such that

$$f_{n,k} \geq \frac{np_k}{2}$$

for every $n \geq n_4$ and $k = n_1+1, \ldots, m(n_3)$. Thus

$$\begin{aligned}
\frac{1}{n} \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k) &= \frac{1}{n} \sum_{k=n_1+1}^{m(n_4)} f_{n,k}(-\log r_k) + \frac{1}{n} \sum_{k=m(n_4)+1}^{m(n)} f_{n,k}(-\log r_k) \\
&\geq \frac{1}{n} \sum_{k=n_1+1}^{m(n_4)} \frac{np_k}{2}(-\log r_k) \\
&= \frac{1}{2} \sum_{k=n_1+1}^{m(n_4)} p_k(-\log r_k) > M
\end{aligned}$$

for every $n \geq n_4$. This shows that $A(n) \to 0$ as $n \to \infty$. To estimate $B(n)$, we note that

$$B(n) \leq (s+\epsilon) \cdot \frac{\sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)}{\sum_{k=1}^{n_1} f_{n,k}(-\log r_k) + \sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)}$$

Using the same argument as above, we obtain that $\sum_{k=n_1+1}^{m(n)} f_{n,k}(-\log r_k)$ grows faster than linear, so $\lim B(n) \leq s+\epsilon$. This shows that

$$\delta(x) \leq s.$$

The proof of the opposite inequality is analogous. ∎

## 3.6 The decay ratio

Now we proceed to study the properties of the decay ratio. In fact, we show that for infinite entropy measures, it is completely determined by the properties of the partition $\{I(n) \mid n \in \mathbb{N}\}$:

**Definition 3.6.1.** *The* convergence exponent *of the partition* $\{r_n\}$ *of* $I$ *is defined by*

$$s_\infty = \inf\left\{ s \geq 0 \mid \sum_{n=1}^{\infty} r_n^s < \infty \right\}.$$

**Proposition 3.6.2.** *In general, we have that* $s_\infty \leq s$. *Under the assumption that* $h_\mu = \infty$, *we also have* $s = s_\infty$.

**Proof.** Given $\epsilon > 0$, there exists $n_1$ such that

$$(\epsilon + s)\log r_n < \log p_n < (s - \epsilon)\log r_n$$

for every $n \geq n_1$, and thus $r_n^{s+\epsilon} < p_n$ for every $n \geq n_1$. Summing over $n$ we get

$$\sum_{n=1}^{\infty} r_n^{s+\epsilon} = \sum_{n=1}^{n_1-1} r_n^{s+\epsilon} + \sum_{n=n_1}^{\infty} r_n^{s+\epsilon} \leq \sum_{n=1}^{n_1-1} r_n^{s+\epsilon} + \sum_{n=n_1}^{\infty} p_n < \infty.$$

Hence, $s_\infty \leq s + \epsilon$ for every $\epsilon > 0$ and so $s_\infty \leq s$.
Now, assuming that $h_\mu = \infty$, suppose that $s_\infty < s$, and hence, there is $\alpha > 0$ such that $s_\infty \leq s_\infty + \alpha < s$ and

$$\sum_{n=1}^{\infty} r_n^{s_\infty + \alpha} < \infty.$$

Let $\epsilon = (s - s_\infty - \alpha)/2 > 0$, then there is an integer $n_0$ such that

$$r_n^{s+\epsilon} \leq p_n \leq r_n^{s-\epsilon}$$

for all $n \geq n_0$. This implies that

$$\sum_{n=n_0}^{\infty} p_n(-\log p_n) \leq (s+\epsilon) \sum_{n=n_0}^{\infty} r_n^{s-\epsilon}(-\log r_n).$$

Recall the one sided limit criterion for convergence of series: let $a_b, b_n > 0$ sequences such that

$$\limsup_{n\to\infty} \frac{a_n}{b_n} = c \in [0, \infty)$$

and $\sum b_n < \infty$. Then $\sum a_n < \infty$.
Let $f : [0, \infty) \to \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} 0, & \text{for } x = 0, \\ x^\epsilon(-\log x), & \text{for } x > 0. \end{cases}$$

It is easy to see that $f$ is continuous. Taking $a_n = r_n^{s-\epsilon}(-\log r_n)$ and $b_n = r_n^{s_\infty+\alpha}$ and using the continuity of $f$, we get that

$$\limsup_{n\to\infty} \frac{a_n}{b_n} = \lim_{n\to\infty} r_n^{\epsilon}(-\log r_n) = 0.$$

We conclude that

$$\sum_{n=n_0}^{\infty} p_n(-\log p_n) \le (s+\epsilon) \sum_{n=n_0}^{\infty} r_n^{s-\epsilon}(-\log r_n) < \infty,$$

contradicting the fact that the entropy is infinite. ■

We give now a definition for the asymptotic decay of the sequence $\{r_n\}$.

**Definition 3.6.3.** *The* asymptotic rate *of the sequence* $\{r_n\}$ *is defined as*

$$\alpha = \sup\{t \ge 0 \mid \lim_{n\to\infty} n^t r_n < \infty\}.$$

*We say that* $\{r_n\}$ *decays* polynomially *if* $\alpha > 1$*, and we say that* $\{r_n\}$ *decays* superpolynomially *if* $\alpha = \infty$.

We will assume that the supremum is achieved for our sequences $\{r_n\}$. Note that if $r_n$ has polynomial decay with asymptotic $\alpha$, then $s_\infty = 1/\alpha$. For simplicity, we will assume that the supremum of the definition is achieved for our partitions. If we know the asymptotic of $\{r_n\}$, we can compute the asymptotic of the tail of the series of $\{r_n\}$:

**Lemma 3.6.4.** *If the asymptotic of* $\{r_n\}$ *is* $\alpha > 1$*, then the asymptotic of* $\{R_n = \sum_{m\ge n} r_n\}$ *is* $\alpha - 1$.

**Proof.** It suffices to show that the sets $A = \{t \ge 1 \mid \lim_{n\to\infty} n^t r_n < \infty\}$ and $A' = \{t \ge 0 \mid \lim_{n\to\infty} n^{t-1} R_n < \infty\}$ are the same. Let $t \in A$, then $\lim_{n\to\infty} n^t r_n = d$, and so given $\epsilon$, there is $n_0 \in \mathbb{N}$ such that

$$\frac{(d-\epsilon)}{n^t} < r_n < \frac{(d+\epsilon)}{n^t}.$$

for $n \ge n_0$. Hence, for $n \ge n_0$,

$$\frac{(d-\epsilon)}{(t-1)} \le \sum_{m=n}^{\infty} \frac{n^{t-1}(d-\epsilon)}{m^t} \le n^{t-1} R_n \le \sum_{m=n}^{\infty} \frac{n^{t-1}(d+\epsilon)}{m^t} \le \frac{n^{t-1}(d+\epsilon)}{(n+1)^{t-1}(t-1)}$$

from which follows that $t-1 \in A'$. Now, if $t \in A'$, we have that $\lim_{n\to\infty} n^{t-1} R_n = d' < \infty$, and thus, given $\epsilon > 0$, there is $n_1 \in \mathbb{N}$ such that

$$\frac{-\epsilon + d'}{n^t} \le \sum_{m\ge n} r_n \le \frac{\epsilon + d'}{n^t}.$$

This implies that

$$\frac{(-\epsilon+d')}{n^t} - \frac{(\epsilon+d')}{(n+1)^t} \leq r_n \leq \frac{(\epsilon+d')}{n^t} - \frac{(-\epsilon+d')}{(n+1)^t}.$$

from which follows that $t+1 \in A$, proving the assertion. ∎

## 3.7 Infinite ergodic theory

In this section we explore the consequences of the non-integrability of the functions $-\log r_{a_1}$ and $-\log p_{a_1}$ (or equivalently, $h_\mu = \lambda_\mu = \infty$). Using tools of infinite ergodic theory we can prove that the diameter of the cylinders decreases faster than exponentially from a given level to the next.
We start by showing one of the usual arguments used to compute Hausdorff dimensions and remark how it fails in our case.

**Lemma 3.7.1.** *Let $T$ be an EMR map and $\mu$ a Gibbs measure. Then for almost every $x \in I$ and every $r > 0$ there exists $n$ such that*

$$\frac{\log C_1 \mu(I_{n-1}(x))}{\log C_2 |I_n(x)|} \leq \frac{\log \mu(B(x,r))}{\log r} \leq \frac{\log \mu(I_n(x))}{\log |I_{n-1}(x)|}. \tag{3.2}$$

*for constants $C_1, C_2$.*

This of this Lemma uses a well known argument and can be found for instance in [Pes08]. Note that if $\lambda_\mu < \infty$, then inequality (3.2) and the Ergodic Theorem would immediately imply that $s = \dim_H \mu = \dim_p \mu$. However, since in our case $\lambda_\mu = \infty$, the previous argument does not work. In fact, here lies the main difficulty of the infinite entropy and Lyapunov exponent case. The following lemma shows that the situation is as bad as it can get: for almost every point, the diameter of the cylinders decreases arbitrarily from one level to the next.

**Theorem 3.7.2.** *Let $T$ be a Gauss-like map and $\mu$ an infinite entropy Gibbs with controlled decay. Then for almost every $x \in I$, we have that*

$$\liminf_{n\to\infty} \frac{\log |I_n(x)|}{\log |I_{n-1(x)}|} = 1,$$

*and*

$$\limsup_{n\to\infty} \frac{\log |I_n(x)|}{\log |I_{n-1}(x)|} = \infty.$$

The proof of the first equality is an immediate consequence of Poincare's recurrence theorem: since most orbits will visit any cylinder infinitely often, in particular they will visit the cylinder $[1,1]$ infinitely often. Whenever this happens, let us say at time $n+1$, the ratio of the sizes (Lebesgue) of the two cylinders $I_n(x)$ and $I_{n+1}(x)$ is a constant ($r_1$). In particular the limit of the logarithms is equal to one. We postpone the proof of the second equality. We will return to this issue once we set up the appropriate tools to prove this result. A corollary to the previous theorem is the following:

**Corollary 3.7.3.** *For almost every $x \in I$, we have that $\underline{d}(x) \leq s$ and hence $\dim_H \mu \leq s$.*

The main tool that we will use to prove Theorem 3.7.2 are results about the pointwise behavior of trimmed sums.

In this section we introduce some infinite ergodic theory notions and results. Define $\{g_n = -\log r_1 \circ T^{n-1}\}$. The tail of the cumulative distribution function of $g_1$ is $\mathscr{F}(t) = \mu(g_1 \geq t)$, and it can be seen that $\mu(g_1) := \int_0^1 g_1 d\mu = \lambda_\mu$. By invariance of the measure, the cumulative distribution of $g_n$ is the same as $\mathscr{F}$. As we saw in Lemma 3.4.3, the Ergodic Theorem fails to provide non-trivial information. This result was vastly generalized by Robbins and Chow for i.i.d. random variables in [CR61] and in the ergodic stationary case by Aaronson in [Aar77] who proved the following theorem:

**Theorem 3.7.4.** *[Aar77, theorem 1] Let $f : [0,1] \to \mathbb{R}$ be a non-negative measurable function. If $\mu(f) = \infty$ then for any sequence $\{b_n\}$ of positive numbers, either*

$$\limsup_{n\to\infty} \frac{1}{b_n} \sum_{k=0}^{n-1} f \circ T^k = \infty \quad a.e.$$

*or*

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=0}^{n-1} f \circ T^k = 0 \quad a.e..$$

It is possible to prove that the lack of convergence in the previous theorem is due to a finite number of terms which are not comparable in size to the rest of the terms of the sum. This was proved in the i.i.d. case by Mori in [Mor76],[Mor77] and in the stationary ergodic case by Aaronson and Nakada in [AN03]. We formulate the result by Aaronson and Nakada in a setting appropriate for our purposes.

We denote the ergodic sum of a function $f$ by $S_n(f)(x)$ and define $\tilde{S}_n(f)(x) = S_n(f)(x) - \max\{f,\ldots,f\circ T^{n-1}\}(x)$. When the dependence of $S_n(f)/\tilde{S}_n(f)$ on $f$ is clear, we drop it from the notation and write $S_n$. We refer to $\tilde{S}_n$ as the *trimmed ergodic sum* of $f$.

**Definition 3.7.5.** *We say that the sequence* $\{f \circ T^k\}$ *has trimmed convergence if there exists a sequence* $\{b_n\}$ *such that*

$$\lim_{n\to\infty} \frac{\tilde{S}_n(x)}{b_n} = 1$$

*almost surely.*

In order to ensure trimmed convergence, it is necessary that the process satisfies certain mixing condition.

**Definition 3.7.6.** *For a stationary process* $(X_1, X_2, \ldots)$*, for* $k \geq 1$ *and* $k < N+1 \leq \infty$*, denote by* $\sigma_k^N = \sigma(X_k, \ldots, X_{N+1})$*, that is, the sigma-algebra generated by the random variables* $(X_k, \ldots, X_{N+1})$*. Define also*

$$\vartheta(n) = \sup\left\{\left|\frac{\mu(A \cap B)}{\mu(A)\mu(B)} - 1\right| : A \in \sigma_1^k, B \in \sigma_{k+n}^{\infty}, \mu(A)\mu(B) > 0, k \geq 1\right\}.$$

*We say that the process is* continued fraction mixing *(cf-mixing) if* $\vartheta(1) < \infty$ *and* $\vartheta(n) \to 0$ *as* $n \to \infty$*.*

**Theorem 3.7.7.** *[AN03, theorem 1.1] Let* $(X_1, X_2, \ldots)$ *be a non-negative, ergodic stationary process with* $L(t) = \mu(\min\{X, t\})$*, and set* $\varepsilon(t) := t(\log L)'(t)$*. Suppose that the process is* continued fraction mixing *with exponential rate and that*

$$\sum_{n=1}^{\infty} \frac{\varepsilon^2(n)}{n} < \infty.$$

*Then* $\{X_n\}$ *has trimmed convergence.*

As remarked in [AN03], any Gibbs-Markov map is CF-mixing with exponential rate (corollary 4.7.8 from [Aar97]). Here by exponential rate we mean that there exist constants $c_1 \geq 0$ and $\theta \in (0,1)$ such that $\vartheta(n) \leq c\theta^n$ for all $n \geq 1$. For our particular sequence, the series in the previous theorem can be explicitly expressed in terms of the sequences $\{p_n\}$ and $\{r_n\}$:

**Lemma 3.7.8.** *Suppose that*

$$\sum_{n=1}^{\infty} (\log r_n)^2 (p_n^2 + 2p_n p_{n+1}) < \infty.$$

*Then the sequence* $\{g_n = -\log r_1 \circ T^{n-1}\}$ *has trimmed convergence.*

**Proof.** We show that if

$$\sum_{n=1}^{\infty}(\log r_n)^2(p_n^2+2p_np_{n+1})<\infty,$$

then

$$\sum_{n=1}^{\infty}\frac{\varepsilon^2(n)}{n}<\infty.$$

Let $\mathscr{F}(t)=\mu(X\geq t)$ and note that

$$(\log L)'(t)=\frac{\mathscr{F}(t)}{L^2(t)},$$

and hence

$$\sum_{n=1}^{\infty}\frac{\varepsilon^2(n)}{n}=\sum_{n=1}^{\infty}\frac{n\mathscr{F}^2(n)}{L^4(n)}\leq c\sum_{n=1}^{\infty}n\mathscr{F}^2(n),$$

since $L(n)$ is bounded away from zero. We compare the above sum to the corresponding integral. We can then see that if $x\in[0,-\log r_1)$ then $\mathscr{F}(x)=1$, while if $x\in[-\log r_n,-\log r_{n+1})$ for $n\geq 1$ then

$$\mathscr{F}(x)=\sum_{k=n+1}^{\infty}p_k,$$

so then the integral is

$$\begin{aligned}\int_0^{\infty}x\,(\mathscr{F}(x))^2\mathrm{d}x&=\int_0^{-\log r_1}x\left(\sum_{k=1}^{\infty}p_k\right)^2\mathrm{d}x+\sum_{n=1}^{\infty}\left(\int_{-\log r_n}^{-\log r_{n+1}}x\left(\sum_{k=n}^{\infty}p_k\right)^2\mathrm{d}x\right)\\&=\frac{(\log r_1)^2}{2}+\frac{1}{2}\sum_{n=1}^{\infty}\left(\int_{-\log r_n}^{-\log r_{n+1}}x\left(\sum_{i,j=n}^{\infty}p_ip_j\right)\mathrm{d}x\right)\\&=\frac{(\log r_1)^2}{2}+\frac{1}{2}\sum_{n=1}^{\infty}\left((\log r_{n-1})^2-(\log r_n)^2\right)\left(\sum_{i,j=n}^{\infty}p_ip_j\right).\end{aligned}$$

Call now

$$a_n=(\log r_n)^2\quad,\quad b_n=\sum_{i,j=n}^{\infty}p_ip_j.$$

Then, the above expression has the form

$$\sum_{n=1}^{\infty}(a_{n+1}-a_n)b_n$$

which can be written as

$$-a_1 b_1 + \sum_{n=1}^{\infty} a_{n+1}(b_n - b_{n+1}).$$

Note that

$$\begin{aligned} b_{n+1} - b_n &= 2p_n p_{n+1} + p_n^2 \\ b_1 &= 1. \end{aligned}$$

With this, the integral becomes

$$\begin{aligned} \int_0^{\infty} x\,(\mathscr{F}(x))^2 \mathrm{d}x &= \frac{(\log r_1)^2}{2} + \frac{1}{2}\left(-(\log r_1)^2 + \sum_{n=1}^{\infty}(\log r_n)^2(p_n^2 + 2p_n p_{n+1})\right) \\ &= \sum_{n=1}^{\infty}(\log r_n)^2(p_n^2 + 2p_n p_{n+1}) \end{aligned}$$

as we wanted to prove. ■

We show now that the trimmed convergence condition is satisfied by systems for which $\{r_n\}$ decays polynomially or slower.

**Lemma 3.7.9.** *Suppose that*

$$\lim_{n\to\infty} \frac{1}{n}(\log r_n)^2 = c \in [0,\infty).$$

*Then the sequence* $\{g_n = -\log r_1 \circ T^{n-1}\}$ *has trimmed convergence.*

**Proof.** Since $p_n$ and $p_{n+1}$ are comparable (as the measure is of controlled decay), it suffices to prove that

$$\sum_{n=1}^{\infty}(\log r_n)^2 p_n^2 < \infty.$$

Note that $\{p_n\} \subset \ell^2$ and we have that

$$1 = \left(\sum_{n=1}^{\infty} p_n\right)^2 = \sum_{i,j=1}^{\infty} p_i p_j.$$

Since the sequence $\{p_n\}$ is decreasing, we have that

$$\sum_{j=2}^{\infty} p_j^2(j-1) = \sum_{j=2}^{\infty} p_j \sum_{i=1}^{j-1} p_j \le \sum_{j=2}^{\infty} p_j \sum_{i=1}^{j-1} p_i \le \sum_{j=2}^{\infty} p_j \sum_{i=1}^{\infty} p_i < \infty.$$

Comparing in the limit the series of the left hand side to the series $\sum_n p_n^2(\log r_n)^2$, we get that this series converges. ■

**Corollary 3.7.10.** *If $T$ is a Gauss-like map and $\mu$ is a measure with infinite entropy and controlled decay, then the sequence $g_n = -\log r_1 \circ T^{n-1}$ has trimmed convergence.*

Now we are in position to prove theorem 3.7.2:

**Proof of theorem 3.7.2** By Theorem 3.7.7 there exists a sequence $\{b_n\}$ such that

$$\lim_{n\to\infty} \frac{\tilde{S}_n(x)}{b_n} = 1 \quad \text{a.e..}$$

Now, by Theorem 3.7.4 we also have that

$$\limsup_{n\to\infty} \frac{S_n(x)}{b_n} = \infty \quad \text{a.e.}$$

or

$$\liminf_{n\to\infty} \frac{S_n(x)}{b_n} = 0 \quad \text{a.e..}$$

Since the trimmed sum is $\Omega(b_n)$, the first condition must hold in a set of full measure. Let $(a_n)$ be the coding sequence of $x$. With an argument analogous to the one used in the proof of Theorem 3.5.2, the limit in question is equivalent to

$$\limsup_{n\to\infty} \frac{\sum_{k=1}^{n} \log r_{a_k}}{\sum_{k=1}^{n-1} \log r_{a_k}} = 1 + \limsup_{n\to\infty} \frac{\log r_{a_n}}{\log(r_{a_1} \dots r_{a_{k_{n-1}}})} = 1 + \limsup_{n\to\infty} \frac{g_n(x)}{S_{n-1}(x)}$$

Given $1 > \varepsilon > 0$, there exists $n_0$ such that

$$\left| \frac{\tilde{S}_n(x)}{b_n} - 1 \right| < \varepsilon$$

for every $n \geq n_0$ at $x$. Since $\limsup \frac{S_n}{b_n} = \infty$, given an integer $M > 0$ there exists $n_1 \geq n_0$ such that

$$\frac{S_{n_1}(x)}{b_{n_1}} > 2M + 1$$

at $x$. Combining these two inequalities, we obtain

$$\left| \frac{\max\{g_1 \dots, g_{n_1}\}(x)}{b_{n_1}} \right| = \left| \frac{S_{n_1}(x)}{b_{n_1}} - \frac{\tilde{S}_{n_1}(x)}{b_{n_1}} \right| > 2M.$$

Now, there exists an index $j \in \{1, \dots, n_1\}$ such that $g_j = \max\{g_1 \dots, g_{n_1}\}$ at $x$, and so $\tilde{S}_j(x) = S_{j-1}(x)$. Since the $g_i$ are positive, we have that

$$S_{j-1}(x) = \tilde{S}_j(x) \leq \tilde{S}_{n_1}(x) < b_{n_1}(1+\varepsilon) < 2b_{n_1} < \frac{\max\{g_1 \dots, g_{n_1}\}(x)}{M} = \frac{g_j(x)}{M},$$

and hence

$$M < \frac{g_j(x)}{S_{j-1}(x)}.$$

This implies that

$$\limsup_{n\to\infty} \frac{g_n(x)}{S_{n-1}(x)} = \infty$$

and so

$$\limsup_{n\to\infty} \frac{\log|I_n|}{\log|I_{n-1}|} = \infty$$

as we wanted to prove. ■

## 3.8 Upper bound for $\dim_H \mu$

With the tools developed in the previous sections, we proceed with the dimension computations.

Now we prove an upper bound for $\dim_H \mu$. This bound is related to the tail decay ratio $\widehat{s}$. We prove two necessary lemmas to give the desired bound. The first lemma shows that $\{p_n\}$ decays slower than any polynomial, while the second lemma, shows the existence of $\widehat{s}$ and that $\widehat{s} = 0$ for Gauss-like maps.

**Lemma 3.8.1.** *Suppose that the decay ratio exists and it is equal to s, the sequence $\{r_n\}$ decays polynomially and the measure μ has infinite entropy. Then for all $\delta > 0$, there exist constants $C, n_0$ such that*

$$p_n \geq \frac{C}{n^{1+\delta}}$$

*for all $n \geq n_0$.*

**Proof.** Let $\alpha > 0$ be the polynomial decay of $r_n$. Then by proposition 3.6.2, $s = s_\infty = 1/\alpha$, we can take $\epsilon > 0$ small enough so that $\epsilon\alpha + \epsilon s + \epsilon^2 < \delta$. Then there exists $C > 0$ and $n_0 \in \mathbb{N}$ such that

$$\frac{C}{n^{\alpha+\epsilon}} \leq r_n$$
$$\log r_n^{s+\epsilon} \leq \log p_n$$

for all $n \geq n_0$. This implies that

$$\frac{C^{s+\epsilon}}{n^{1+\delta}} \leq \frac{C^{s+\epsilon}}{n^{(\alpha+\epsilon)(s+\epsilon)}} \leq p_n$$

for all $n \geq n_0$ as we wanted. ■

**Lemma 3.8.2.** *Under the same assumptions of the previous lemma, the tails decay ratio $\widehat{s}$ (definition 3.3.4) exists and is equal to zero.*

**Proof.** By the lemma above, for $\delta > 0$, there are constants $C, n_0$ such that

$$p_n \geq \frac{C}{n^{1+\delta}}$$

for all $n \geq n_0$. This implies that

$$\sum_{m=n}^{\infty} p_m \geq \frac{C}{\delta n^{\delta}}$$

for $n \geq n_0$. On the other hand, if we take $\epsilon < \alpha - 1$, there exists $n_1$ such that

$$r_n \leq \frac{C}{n^{\alpha-\epsilon}}$$

for $n \geq n_1$ and consequently,

$$\sum_{m=n}^{\infty} r_m \leq \frac{C}{(\alpha-\epsilon-1)n^{\alpha-\epsilon-1}}$$

for $n \geq n_1$. Hence

$$\frac{\log \sum_{m=n}^{\infty} p_m}{\log \sum_{m=n}^{\infty} r_m} \leq \frac{\log C - \log\delta - \delta\log n}{\log C - \log(\alpha-\epsilon-1) - (\alpha-\epsilon-1)\log n}$$

for $n \geq \max\{n_0, n_1\}$. This implies that

$$\limsup_{n\to\infty} \frac{\log \sum_{m=n}^{\infty} p_m}{\log \sum_{m=n}^{\infty} r_m} \leq \frac{\delta}{(\alpha-\epsilon-1)}.$$

Letting $\delta \to 0$ we conclude the result. ■

Now we can compute the lower local dimension, and consequently, obtain the Hausdorff dimension of the measure.

**Theorem 3.8.3.** *Suppose $T$ is a Gauss-like map and $\mu$ is an infinite entropy Gibbs measure with controlled decay. Then*

$$\liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} = 0$$

*for $\mu$ almost every $x \in I$.*

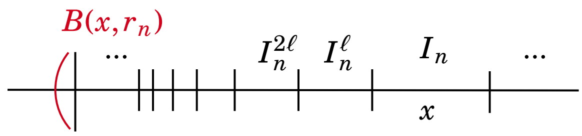


Figure 3.2: Relative position of intervals and ball.

**Proof.** Let $x$ be a point where Theorems 3.5.2, 3.7.2 and 3.7.4 hold (such a set is of full measure). Given such $x$ and $n \in \mathbb{N}$, take

$$t_n = \left| \bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x) \right|,$$

where $I_n^{m\cdot\ell}(x) = I(a_1(x),\ldots,a_{n-1}(x),a_n(x)+m)$. In figure 3.2 we show the intervals contained in the ball $B(x,t_n)$.

Then

$$\bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x) \subseteq B(x,t_n)$$

and so

$$\frac{\log \mu(B(x,t_n))}{\log t_n} \leq \frac{\log \mu\left(\bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x)\right)}{\log \left|\bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x)\right|}.$$

Note now that the above inequality can be expressed in terms of the sequences $\{p_n\},\{r_n\}$ using Lemma 3.5.1

$$\log \mu\left(\bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x)\right) \geq \sum_{k=1}^{n-1} \log p_{a_k} + \log\left(\sum_{m=0}^{\infty} p_{a_n+m}\right) - nG_1 - G_2$$
$$\log \left|\bigcup_{m=0}^{\infty} I_n^{m\cdot\ell}(x)\right| \leq \sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{m=0}^{\infty} r_{a_n+m}\right) + nD_1 + D_2$$

where $G_1, G_2$ are constants arising from the Gibbs property and the finite first variation of the potential, and $D_1, D_2$ are constants arising from the bounded distortion property. Thus, we have

$$\frac{\log \mu(B(x,t_n))}{\log t_n} \leq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \log\left(\sum_{m=0}^{\infty} p_{a_n+m}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{m=0}^{\infty} r_{a_n+m}\right) + nD_1 + D_2}.$$

For $\epsilon > 0$ and $n$ large enough, we have that

$$\frac{-\log \sum_{m=0}^{\infty} p_{n+m}}{-\log \sum_{m=0}^{\infty} r_{n+m}} < \epsilon$$

and

$$-\epsilon + s < \frac{-\sum_{k=1}^{n-1} \log p_{a_k}}{-\sum_{k=1}^{n-1} \log r_{a_k}} < s + \epsilon.$$

Thus, if $a_n$ is large enough, we have

$$\frac{\log \mu(B(x,t_n))}{\log t_n} \leq \frac{(s+\epsilon)\left(\sum_{k=1}^{n-1} \log r_{a_k}\right) + \epsilon \log\left(\sum_{m=0}^{\infty} r_{a_n+m}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{m=0}^{\infty} r_{a_n+m}\right) + nD_1 + D_2}.$$

If $\alpha > 1$ is the polynomial decaying ratio of $\{r_n\}$, then by Lemma 3.6.4 we get that the tail decay asymptotic of $\sum_{m=0}^{\infty} r_{n+m}$ is $\alpha - 1$. We can then rewrite the above inequality as

$$\frac{\log \mu(B(x,t_n))}{\log t_n} \leq \frac{(s+\epsilon)\left(\sum_{k=1}^{n-1} \log r_{a_k}\right) + \epsilon K(\alpha-1)\log\left(r_{a_n}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + K(\alpha-1)\log\left(r_{a_n}\right) + nD_1 + D_2}.$$

where $K$ is the constant implied in the tail asymptotic for $\{r_n\}$. By Theorem 3.7.4 and Theorem 3.7.2, we can take an increasing subsequence $a_{n_k}$ so that

$$\lim_{k\to\infty} \frac{-\log r_{a_{n_k}}}{-\sum_{k=1}^{n_k-1} \log r_{a_k}} = \infty,$$
$$\lim_{k\to\infty} -\frac{1}{n_k} \log r_{a_{n_k}} = \infty.$$

We get then

$$\lim_{k\to\infty} \frac{\log \mu(B(x,t_{n_k}))}{\log t_{n_k}} \leq \epsilon.$$

Letting $\epsilon \to 0$ we conclude that $\underline{d}(x) \leq \widehat{s}$ as we wanted. ■

From the above result and proposition 2.3.7, we can conclude that for such measures, $\dim_H \mu = 0$.

## 3.9 Computation of $\dim_p \mu$

In the previous section we completely determined the Hausdorff dimension of the measures of our interest. Now we proceed to compute the packing dimension. First we give a lower bound for the upper local dimension. The proof uses similar ideas to the proof of Theorem 3.8.3: we choose a particular cover of the ball and use that the Birkhoff sums for the potentials $-\log p_{a_1}, -\log r_{a_1}$ grow faster than linear.

**Proposition 3.9.1.** *Suppose $T$ is a Gauss-like map and $\mu$ is an infinite entropy Gibbs measure with controlled decay. Then*

$$\limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \geq s$$

*for $\mu$ almost every $x \in I$.*

**Proof.** By Birkhoff's Ergodic Theorem, we have that

$$\lim_{n\to\infty} \frac{f_{n,1}}{n} = p_1$$

almost everywhere, where $f_{n,k}$ is as defined in the proof of 3.5.2. Lemma 3.4.3 and Theorem 3.5.2 hold in a set of full measure as well. We pick a point $x$ where the three results hold. Since $p_1 < 1$, we can pick a subsequence $k_n \nearrow \infty$ such that $a_{k_n} \neq 1$ for every $n$. Then, for all $n$, take $t_n = \min\{|I_{k_n}|, |I^r_{k_n}|, |I^\ell_{k_n}|\} = |I^\ell_{k_n}|$.

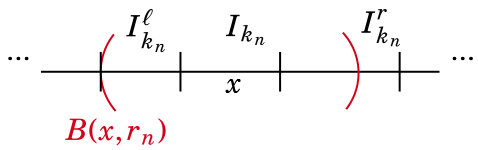


Figure 3.3: Relative position of the intervals and ball.

Here we denote $I^\ell_n = I(a_1,\dots,a_{n-1},a_n+1)$ and $I^r_n = I(a_1,\dots,a_{n-1},a_n-1)$ whenever $a_n > 1$. This choice of $r_n$ implies that $B(x,t_n) \subseteq I^\ell_{k_n} \cup I_{k_n} \cup I^r_{k_n}$. In figure 3.3 we show this inclusion. From the Gibbs property and the fact that $\varphi(x_n), \varphi(x_{n+1})$, and $r_n, r_{n+1}$ are comparable, it follows that there are constants $C_1, C_2 > 0$ such that $\mu(I^\ell_n \cup I_n \cup I^r_n) \leq C_1 \mu(I_n)$ and $|I^\ell_n| \geq C_2 |I_n|$ for every $n$. Using this and Lemma 3.5.1 we have that

$$\begin{aligned}
\frac{\log \mu(B(x,t_n))}{\log r_n} &\geq \frac{\log(C_1 \mu(I_{k_n}))}{\log(C_2 |I_{k_n}|)} \\
&\geq \frac{\log C_1 + k_n G_1 + G_2 + \sum_{i=1}^{k_n} \log p_{a_i}}{\log C_2 - D_2 - k_n \log D_1 + \sum_{i=1}^{k_n} \log r_{a_i}}.
\end{aligned}$$

By Lemma 3.4.3 and Theorem 3.5.2, the last expression converges to $s$, as desired. ■

Giving an upper bound for the upper local dimension requires a more involved analysis of the geometric structure of the partition and its relation to the geometry of the balls. We will need the following lemma:

**Lemma 3.9.2.** *Suppose that $\{r_n\}$ decays polynomially with degree $\alpha > 1$. Then, for every $0 < \delta < \min\{1/3, (\alpha-1)/(\alpha+1)\}$, $0 < \eta < 1/2$ there exists $k_0 \in \mathbb{N}$ such that*

$$\frac{\log \sum_{m=k}^{n+k} p_m}{\log \sum_{m=k-1}^{n+k+1} r_m} \leq \frac{1+\delta}{\alpha-\delta} + \eta$$

*for all $k \geq k_0$ and $n \in \mathbb{N}$.*

**Proof.** Recall that for such sequence $\{r_n\}$, we have that $s = 1/\alpha$. Fix $0 < \delta < \min\{1/3, s(\alpha - 1)/(\alpha+1)\}$, $0 < \eta < 1/2$. Note that this implies that

$$\frac{\delta}{\alpha - 1 - \delta} < s = \frac{1}{\alpha} < \frac{1+\delta}{\alpha - \delta}.$$

Now, since

$$\lim_{k\to\infty} \frac{(1+\delta)\log 2 + \delta \log k}{\log(\alpha-1-\delta) + (\alpha-1-\delta)\log(k-2)} = \frac{\delta}{\alpha-1-\delta} < \frac{1+\delta}{\alpha-\delta}$$

and

$$\lim_{k\to\infty} \frac{(1+\delta)\log(2k)}{(\alpha-\delta)\log(k-1) - \log 3} = \frac{1+\delta}{\alpha-\delta},$$

we can find $k_0 \in \mathbb{N}$ such that

$$\begin{aligned} \frac{(1+\delta)\log 2 + \delta \log k}{\log(\alpha-1-\delta) + (\alpha-1-\delta)\log(k-2)} &< \frac{1+\delta}{\alpha-\delta} + \eta \\ \frac{(1+\delta)\log(2k)}{(\alpha-\delta)\log(k-1) - \log 3} &< \frac{1+\delta}{\alpha-\delta} + \eta \end{aligned}$$

for all $k \geq k_0$. It can be proved using calculus that for $\delta < (\alpha-1)/2$, the inequality

$$(1+\delta)\log(2k) \leq (\alpha-\delta)\log(k-1) - \log 3$$

holds for sufficiently large $k$, so we can take $k_0$ large enough so that this holds. Finally, we can take $k_0$ large enough so that we also have

$$\begin{aligned} r_k &\leq \frac{1}{k^{\alpha-\delta}} \\ \frac{1}{k^{1+\delta}} &\leq p_k \end{aligned}$$

for all $k \geq k_0$. Let $n \in \mathbb{N}$. We divide in two cases:

**Case 1:** $n \geq k$. Then

$$\sum_{m=k}^{n+k} p_m \geq \frac{n}{(2k)^{1+\delta}} \geq \frac{1}{2^{1+\delta} k^{\delta}}$$

and

$$\sum_{k-1}^{n+k+1} r_m \leq \sum_{m=k-1}^{n+k+1} \frac{1}{m^{\alpha-\delta}} \leq \sum_{m=k-1}^{\infty} \frac{1}{m^{\alpha-\delta}} \leq \frac{1}{\alpha-1-\delta}\left(\frac{1}{(k-2)^{\alpha-1-\delta}}\right)$$

for all $k \geq k_0$. Then

$$\frac{\log \sum_{m=k}^{n+k} p_m}{\log \sum_{m=k-1}^{n+k+1} r_m} \leq \frac{(1+\delta)\log 2 + \delta \log k}{\log(\alpha-1-\delta)+(\alpha-1-\delta)\log(k-2)} \leq \frac{1+\delta}{\alpha-\delta}+\eta$$

for all $k \geq k_0$.

**Case 2:** $n < k$. Then

$$\sum_{m=k}^{n+k} p_m \geq \frac{n+1}{(2k)^{1+\delta}}$$

and

$$\sum_{k-1}^{n+k+1} r_m \leq \sum_{m=k-1}^{n+k+1} \frac{1}{m^{\alpha-\delta}} \leq \frac{n+3}{(k-1)^{\alpha-\delta}} \leq 3\frac{(n+1)}{(k-1)^{\alpha-\delta}}.$$

Hence

$$\frac{\log \sum_{m=k}^{n+k} p_m}{\log \sum_{m=k-1}^{n+k+1} r_m} \leq \frac{(1+\delta)\log(2k)-\log(n+1)}{(\alpha-\delta)\log(k-1)-\log 3-\log(n+1)}.$$

We use the following Lemma:

**Lemma 3.9.3.** *For $a,b,c>0$ such that $a-c, b-c>0$, we have that*

$$\frac{a-c}{b-c} \leq \frac{a}{b}$$

*if and only if $b \geq a$.*

We can use this with $a=(1+\delta)\log(2k)$, $b=(\alpha-\delta)\log(k-1)-\log 3$ and $c=\log(n+1)$. This implies that

$$\frac{\log \sum_{m=k}^{n+k} p_m}{\log \sum_{m=k-1}^{n+k+1} r_m} \leq \frac{(1+\delta)\log(2k)}{(\alpha-\delta)\log(k-1)-\log 3} \leq \frac{1+\delta}{\alpha-\delta}+\eta.$$

for all $k \geq k_0$, as we wanted to prove. ■

With the previous lemma, we can now prove the upper bound for the upper local dimension. The proof is based on carefully choosing the covers of the balls; such covers must be fine enough so they are not affected by Theorem 3.7.2. This means that we want to cover the ball with cylinders of the same scale, otherwise, the cover would yield trivial bounds.

**Proposition 3.9.4.** *Suppose $T$ is a Gauss-like map and $\mu$ is an infinite entropy Gibbs measure with controlled decay. Then*

$$\limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r} \le s$$

*for $\mu$ almost every $x \in I$.*

**Proof.** Let $x$ be a point where Theorem 3.5.2 and Lemma 3.4.3 applied to $f = -\log r_{a_1}$ hold. Given $r > 0$, there exists a unique natural number $n = n(r)$ such that

$$|I_n(x)| < r \le |I_{n-1}(x)|.$$

Note that $n \to \infty$ as $r \to 0$. Let $\delta > 0$ and $\eta$ as in Lemma 3.9.2. Then there exists $k_0 \in \mathbb{N}$ such that

$$\frac{\log \sum_{m=k}^{n+k} p_m}{\log \sum_{m=k-1}^{n+k+1} r_m} \le \frac{1+\delta}{\alpha - \delta} + \eta$$

for all $k \ge k_0$. Recall that by $I_n^{m\cdot r}(x)$ we denote the cylinder $I(a_1,\ldots,a_{n-1},a_n - m)$, where $(a_n)$ is the sequence coding $x$ and $m < a_n$. We separate the proof in two cases:
**Case 1:**

$$I_n = I(a_1,\ldots,a_{n-1},k_0) \subset B(x,r).$$

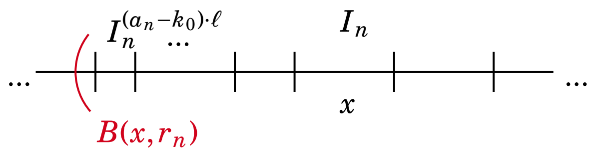


Figure 3.4: Case 1

This case is shown in figure 3.4. In this case, using Lemma 3.5.1 we have that

$$\begin{aligned}\log(\mu(B(x,r))) &\ge \log(\mu(I(a_1,\ldots,a_{n-1},k_0)))\\ &\ge \sum_{k=1}^{n-1} \log p_{a_k} + \log p_{k_0} - nG_1 - G_2.\end{aligned}$$

We get then

$$\begin{aligned}\frac{\log \mu(B(x,r))}{\log r} &\le \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \log p_{k_0} - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + nD_1 + D_2}\\ &\le \frac{(s+\delta)\sum_{k=1}^{n-1} \log r_{a_k} + \log p_{k_0} - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + nD_1 + D_2}. \qquad (3.3)\end{aligned}$$

**Case 2:**

$$I(a_1,\ldots,a_{n-1},k_0) \not\subset B(x,r).$$

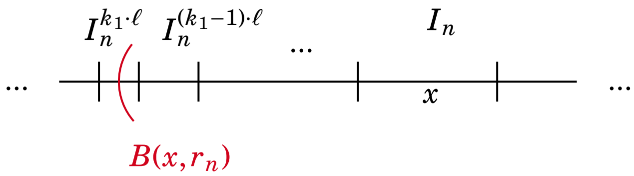


Figure 3.5: Case 2

This implies that there exists $k_1 \in \mathbb{N}$ such that

$$\bigcup_{m=1}^{k_1-1} I_n^{k\cdot\ell}(x) \subset B(x,r),$$
$$\left|\bigcup_{m=0}^{k_1} I_n^{k\cdot\ell}(x)\right| > r$$

as shown in figure 3.5, and consequently

$$\log(\mu(B(x,r))) \geq \sum_{k=1}^{n-1} \log p_{a_k} + \log\left(\sum_{k=1}^{k_1-1} p_{a_n-k}\right) - nG_1 - G_2$$
$$\log r \leq \sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) + nD_1 + D_2.$$

We obtain then

$$\frac{\log(\mu(B(x,r)))}{\log r} \leq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \log\left(\sum_{k=1}^{k_1-1} p_{a_n-k}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) + nD_1 + D_2}.$$

Using inequality (3.9.2)

$$\frac{\log(\mu(B(x,r)))}{\log r} \leq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \left(\frac{1+\delta}{\alpha-\delta} + \eta\right)\log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) + nD_1 + D_2}.$$

For $\delta > 0$, there exist $n_0 \in \mathbb{N}$ such that

$$\frac{-\sum_{k=1}^{n-1} \log q_{a_k}}{-\sum_{k=1}^{n-1} \log r_{a_k}} < s + \delta$$

for all $n \geq n_0$. We obtain

$$\frac{\log(\mu(B(x,r)))}{\log r} \leq \frac{(s+\delta)\sum_{k=1}^{n-1}\log r_{a_k} + \left(\frac{1+\delta}{\alpha-\delta}+\eta\right)\log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1}\log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) + nD_1 + D_2}$$

$$\leq \max\left\{(s+\delta), \left(\frac{1+\delta}{\alpha-\delta}+\eta\right)\right\} \cdot \frac{\sum_{k=1}^{n-1}\log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) - nG_1 - G_2}{\sum_{k=1}^{n-1}\log r_{a_k} + \log\left(\sum_{k=0}^{k_1} r_{a_n-k}\right) + nD_1 + D_2}. \tag{3.4}$$

By Lemma 3.4.3 we have that the right hand side of (3.3) and (3.4) converge to

$$(s+\delta)\ ,\ \max\left\{(s+\delta), \left(\frac{1+\delta}{\alpha-\delta}+\eta\right)\right\}$$

respectively. We conclude that

$$\limsup_{r\to 0} \frac{\log(\mu(B(x,r)))}{\log r} \leq \max\left\{(s+\delta), \left(\frac{1+\delta}{\alpha-\delta}+\eta\right)\right\}.$$

Letting $\delta \to 0$ and $\eta \to 0$, we obtain the desired result. ■

**Corollary 3.9.5.** *For an infinite entropy Gibbs measure $\mu$ with infinite entropy and controlled decay, associated to a Gauss-like map, we have that $0 = \underline{d}(x) < s = \overline{d}(x)$ for almost every point, and hence $\mu$ is not exact dimensional.*

**Proof of theorem 3.1.2** The previous corollary gives us the almost sure behavior of the local dimensions, and hence, we have obtained values for both the packing and the Hausdorff dimension. ■

## 3.10 A measure with positive dimension

In this section we construct an EMR map $T$ and a measure $\mu$ with infinite entropy and positive Hausdorff dimension. For a sequence $\{r_n\} \subset \mathbb{R}_+$, such that $\sum_n r_n = 1$, consider the partition of the unit interval $[0,1]$ given by $I(n) = [b_{n+1}, b_n]$, where

$$b_1 = 1,$$
$$b_n = 1 - \sum_{k=1}^{n-1} r_k \text{ for } n \geq 2.$$

By construction, the diameter of each interval is $|I(n)| = r_n$. Define the map $T$ by

$$T(x) = (b_n - b_{n+1})^{-1}x - b_{n+1}\cdot(b_n - b_{n+1})^{-1} \text{ for } x \in I(n).$$

Thus, $T$ restricted to each interval $I(n)$ is a linear bijection to $[0,1]$. Since the map $T$ is piecewise linear, it is easy to prove that

$$|I(a_1,\dots,a_n)| = r_{a_1}\cdot\ldots\cdot r_{a_n}.$$

for every finite sequence $(a_1,\dots,a_n)\in\mathbb{N}^n$. For the measure, we take the Bernoulli measure $\mu$ constructed by assigning measure$\{p_n\}$ to the cylinders $I(n)$ in the following way:

1. $s_1 = \dfrac{\log p_1}{\log r_1} = 1$,
2. $p_n = \dfrac{K_1}{(n+1)(\log(n+1))^2}$, for all $n \geq 3$,
3. $r_n = \dfrac{K_2}{(n+1)(\log(n+1))^{3/2}}$ , for all $n \geq 3$.

Here $K_1, K_2$ are constants that ensure that the first condition holds as well as $\sum_n p_n = \sum_n r_n = 1$, and their values are $K_1 \approx 0.52$ and $K_2 \approx 0.48$. We can modify the value of $p_2$ and $r_2$ if it is necessary in order for the conditions above to hold. It is easy to see that there is a constant $C \in (0,1)$ such that

$$0 < C \leq \frac{p_{n+1}}{p_n}, \frac{r_{n+1}}{r_n} \leq 1$$

for all $n \geq 1$, and that $p_n \leq r_n$ for all $n \geq 3$.
Observe that for our map and measure $(T,\mu)$, the value of the decay ratio is $s = 1$. The following lemma proves that almost every point does not have long sequences of digits equal to 1. This implies that the orbits of the typical points do not spend much time near the right end of the unit interval. The proof is an easy application of the Borel-Cantelli lemma:

**Lemma 3.10.1.** *For $\mu$-almost every $x \in I$, $T^n(x) \in A_n := I(1,\dots,1)$ (where the cylinder consists of $n-1$ consecutive 1's) for finitely many indices $n \geq 2$. In particular, there exists $n_0 = n_0(x)$ such that for every $n \geq n_0$, there is $k \in \{\lfloor n/2\rfloor,\dots,n-1\}$ such that $a_k \neq 1$.*

**Proof.** Let $B_n = T^{-n}A_n$, and note that $\mu(B_n) = \mu(A_n) = p_1^{n-1}$, where $p_1 = \mu(I(1))$. Since $p_1 \in (0,1)$, we have that $\sum_n \mu(B_n) < \infty$. By the Borel-Cantelli lemma, $\mu(\{x \in I \mid x \in B_n \text{ i.o.}\}) = 0$, proving the result. ■

The next lemma shows that the decay ratio not only gives the asymptotic logarithmic comparison between the measure and the diameter of the cylinder $I(n)$, but also of blocks of neighboring cylinders of the form $I(n)\cup I(n+1)\cup\ldots\cup I(n+m)$:

**Lemma 3.10.2.** *For every $\epsilon > 0$, there exists $n_0 = n_0(\epsilon)$ such that*

$$\frac{\log \sum_{k=n}^{n+m+2} p_k}{\log \sum_{k=n}^{n+m} r_k} \geq 1 - \epsilon$$

*for all $n \geq n_0$ and $m \geq 0$ (including when $m = \infty$).*

**Proof.** Fix $\epsilon > 0$ and define $n_0 = \max\{3, n_1\}$, where $n_1$ is given by

$$n_1 = \min\left\{ n : \frac{\log 3}{-\log\left(\frac{K_2/2}{(\log(n+1))^{1/2}}\right)} \leq \epsilon \right\}.$$

Then, for $n \geq n_0$ and $m \geq 0$ we have

$$\log \sum_{k=n}^{n+m} r_k \leq \log \sum_{k=n}^{\infty} r_k \leq \log\left(\frac{K_2/2}{(\log(n+1))^{1/2}}\right) \leq \frac{\log 3}{\epsilon}$$

since $n \geq n_1$ and

$$\frac{\log \sum_{k=n}^{n+m} p_k}{\log \sum_{k=n}^{n+m} r_k} \geq 1$$

since $n \geq 3$. This implies that

$$\frac{\log \sum_{k=n}^{n+m+2} p_k}{\log \sum_{k=n}^{n+m} r_k} \geq \frac{\log\left(3 \sum_{k=n}^{n+m} p_k\right)}{\log \sum_{k=n}^{n+m} r_k} = \frac{\log 3}{\log \sum_{k=n}^{n+m} r_k} + \frac{\log \sum_{k=n}^{n+m} p_k}{\log \sum_{k=n}^{n+m} r_k} \geq 1 - \epsilon$$

for all $n \geq n_0$ and $m \geq 0$ as we wanted to prove. ■

Note that as a consequence of the previous lemma, we have that $\widehat{s} = 1$ for our system. We proceed to prove the main result:

**Proposition 3.10.3.** *Let $T$ be the unique orientation preserving piecewise linear map defined by the sequence $\{r_n\}$, and $\mu$ be the unique Bernoulli measure defined by the sequence $\{p_n\}$. Then $\underline{d}(x) = 1$ for $\mu$-almost every point. In particular, $\mu$ is exact dimensional and $\dim_H \mu = \dim_p \mu = 1$.*

**Proof.** Let $x$ be a point where Theorem 3.5.2 and lemma 3.10.1 hold, $1 > r > 0$ and $\epsilon > 0$. We write $(a_1, a_2, \ldots)$ for the sequence coding $x$. Then, there exists $n = n(x, r)$ such that

$$|I_n(x)| < Cr \leq |I_{n-1}(x)|.$$

This implies that $r \leq C^{-1}|I_{n-1}(x)| = C^{-1}\prod_{k=1}^{n-1} p_{a_k} \leq p_{a_{n-1}+1}\prod_{k=1}^{n-2} p_{a_k} = I(a_1,\dots,a_{n-2},a_{n-1}+1)$.

We need to find a suitable cover of $B(x,r)$. We start by proving a few reductions; these rule out the cases when the radius of the ball is of a similar scale as $|I_n|$.

**Reduction 1:** we can assume $a_n > 1$. Suppose $a_n = 1$. By lemma 3.10.1, if $r$ is small enough, we have that there exists $k = k(x,n) \in \{\lfloor n/2 \rfloor,\dots,n-1\}$ such that $a_k \neq 1$, and define $\widehat{k} = \sup\{k \in \{1,\dots,n-1\} \mid a_k \neq 1\}$. Then we have

$$B(x,r) \subset I(a_1,\dots,a_{n-2},a_{n-1}) \cup I(a_1,\dots,a_{n-2},a_{n-1}+1) \cup I(a_1,\dots,a_{\widehat{k}}-1).$$

Since the first two cylinders are strictly to the left of the third one and correspond to strictly smaller scales, we can bound the measure of the ball by

$$\mu(B(x,r)) \leq 3\mu(I(a_1,\dots,a_{\widehat{k}}-1)).$$

Noting that for all $k \in \{\widehat{k}+1,\dots,n\}$, $a_k = 1$ and that $\widehat{k} \geq \lfloor n/2 \rfloor$, we have that

$$\frac{\log \mu(B(x,r))}{\log r} \geq \frac{\log 3 + \log C^{-1} + \sum_{k=1}^{\widehat{k}} \log p_{a_k}}{\sum_{k=1}^{\widehat{k}} \log r_{a_k} + \widehat{k}\log r_1}.$$

**Reduction 2:** we can assume

$$I(a_1,\dots,a_{n-1},1) \not\subset B(x,r).$$

Otherwise, we have that $r \geq |I(a_1,\dots,a_{n-1},1)|$, and giving the same argument as in the previous reduction, we obtain the bound

$$\frac{\log \mu(B(x,r))}{\log r} \geq \frac{\log 3 + \log C^{-1} + \sum_{k=1}^{\widehat{k}} \log p_{a_k}}{\sum_{k=1}^{\widehat{k}} \log r_{a_k} + (\widehat{k}+1)\log r_1}$$

where $\widehat{k}$ is defined in the same way as in the previous reduction.

**Reduction 3**: we can assume

$$I(a_1,\dots,a_{n-1}+1,1) \not\subset B(x,r).$$

Otherwise, we can bound the radius of the ball by $r > |I(a_1,\dots,a_{n-1}+1,1)|$ and consequently obtain the bound

$$\frac{\log \mu(B(x,r))}{\log r} \geq \frac{\log 3 + \log C^{-1} + \sum_{k=1}^{\widehat{k}} \log p_{a_k}}{\log C + \sum_{k=1}^{\widehat{k}} \log r_{a_k} + (\widehat{k}+1)\log r_1}$$

where $\widehat{k}$ is chosen in the same way as in reduction 1.

**Reduction 4**: we can assume

$$I(a_1,\dots,a_{n-1}+1,k_0)\not\subset B(x,r),$$

where $k_0 = k_0(\epsilon)$ is the constant $n_0(\epsilon)$ as in lemma 3.10.2. Similarly as in the previous reduction, if that was not the case we have the bound

$$\frac{\log\mu(B(x,r))}{\log r}\geq\frac{\log 3+\log C^{-1}+\sum_{k=1}^{\widehat{k}}\log p_{a_k}}{\log C+\sum_{k=1}^{\widehat{k}}\log r_{a_k}+\widehat{k}\log r_1+\log r_{k_0}}.$$

The three first reductions imply that we can assume

$$B(x,r)\subset I(a_1,\dots,a_{n-1})\cup I(a_1,\dots,a_{n-1}+1).$$

We divide now in cases, according on where does the left end of the ball falls. Denote by $\partial^{\ell}(I(a_1,\dots,a_n)),\partial^{r}(I(a_1,\dots,a_n))$ the left and right endpoints of the cylinder $I(a_1,\dots,a_n)$ respectively.

**Case 1:** $x-r>\partial^{\ell}(I(a_1,\dots,a_n))$.

This implies there exist $a_n>m,m'\geq 0$ such that

$$\partial^{\ell}(I(a_1,\dots,a_n-m))<x+r\leq\partial^{r}(I(a_1,\dots,a_n-m))$$
$$\partial^{\ell}(I(a_1,\dots,a_n+m'))\leq x-r<\partial^{r}(I(a_1,\dots,a_n+m')).$$

By the reduction 4 we can assume $a_n-m>k_0$. We can now cover then the ball by

$$B(x,r)\subset\bigcup_{k=a_n-m}^{a_n+m'}I(a_1,\dots,a_{n-2},a_{n-1},k)$$

and bound the radius by

$$r>\frac{1}{2}\left|\bigcup_{k=a_n-m+1}^{a_n+m'-1}I(a_1,\dots,a_{n-2},a_{n-1},k)\right|.$$

We obtain the bound

$$\begin{aligned}\frac{\log\mu(B(x,r))}{\log r}&\geq\frac{\sum_{k=1}^{n-1}\log p_{a_k}+\log\sum_{k=a_n-m}^{a_n+m'}p_k}{\sum_{k=1}^{n-1}\log r_{a_k}+\log\sum_{k=a_n-m+1}^{a_n+m-1'}r_k-\log 2}\\&\geq\frac{\sum_{k=1}^{n-1}\log p_{a_k}+\log\sum_{k=a_n-m}^{a_n+m'}p_k}{\sum_{k=1}^{n-1}\log r_{a_k}+\log\sum_{k=a_n-m}^{a_n+m-2'}r_k+\log C-\log 2}\\&\geq(1-\epsilon)\frac{\sum_{k=1}^{n-1}\log p_{a_k}+\log\sum_{k=a_n-m}^{a_n+m'-2}r_k}{\sum_{k=1}^{n-1}\log r_{a_k}+\log\sum_{k=a_n-m}^{a_n+m'-2}r_k+\log C-\log 2}\end{aligned}$$

**Case 2:** $x-r=\partial^{\ell}(I(a_1,\dots,a_n))$

This is the limit of the previous case. In this case there exists $m<a_n$ such that

$$\partial^{\ell}(I(a_1,\dots,a_n-m))<x+r\le\partial^{r}(I(a_1,\dots,a_n-m)).$$

Again, we can assume $a_n-m>k_0$. Thus, we can cover the ball by

$$B(x,r)\subset\bigcup_{k=a_n-m}^{\infty}I(a_1,\dots,a_{n-2},a_{n-1},k)$$

and bound the radius of the ball by

$$r>\frac{1}{2}\left|\bigcup_{k=a_n-m+1}^{\infty}I(a_1,\dots,a_{n-2},a_{n-1},k)\right|.$$

By using lemma 3.10.2 we obtain the bound

$$\begin{aligned}\frac{\log\mu(B(x,r))}{\log r}&\ge\frac{\sum_{k=1}^{n-1}\log p_{a_k}+\log\sum_{k=a_n-m}^{\infty}p_k}{\sum_{k=1}^{n-1}\log r_{a_k}+\log\sum_{k=a_n-m}^{\infty}r_k+\log C-\log 2}\\&\ge(1-\epsilon)\frac{\sum_{k=1}^{n-1}\log p_{a_k}+\log\sum_{k=a_n-m}^{\infty}r_k}{\sum_{k=1}^{n-1}\log r_{a_k}+\log\sum_{k=a_n-m}^{\infty}r_k+\log C-\log 2}\end{aligned}$$

**Case 3:** $x-r<\partial^{\ell}(I(a_1,\dots,a_n))$. In this case, there exist $m<a_n-k_0$ and $m'>0$ such that

$$\partial^{r}(I(a_1,\dots,a_n-m))<x+r\le\partial^{r}(I(a_1,\dots,a_n-m))$$
$$\partial^{\ell}(I(a_1,\dots,a_{n-1}+1,\underbrace{1,\dots,1}_{m'+1}))>x-r\ge\ \partial^{\ell}(I(a_1,\dots,a_{n-1}+1,\underbrace{1,\dots,1}_{m'})).$$

This can be interpreted as finding the shortest sequence of ones such that the ball contains the cylinder $I(a_1,\dots,a_{n-1}+1,1,\dots,1)$, where the sequence of ones if of length $m'+1$. This implies we can cover the ball by

$$B(x,r)\subset\bigcup_{k=a_n-m}^{\infty}I(a_1,\dots,a_{n-2},a_{n-1},k)\cup I(a_1,\dots,a_{n-1}+1,\underbrace{1,\dots,1}_{m'})$$

and we can estimate the radius by

$$r>\frac{1}{2}\left|\bigcup_{k=a_n-m+1}^{\infty}I(a_1,\dots,a_{n-2},a_{n-1},k)\right|+\frac{1}{2}|I(a_1,\dots,a_{n-1}+1,\underbrace{1,\dots,1}_{m+1})|.$$

With this, we obtain the bound

$$
\begin{aligned}
\frac{\log \mu(B(x,r))}{\log r} &\geq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \log\left(C p_1^{m'} + \sum_{k=a_n-m}^{\infty} p_k\right)}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(C r_1^{m'+1} + \sum_{k=a_n-m}^{\infty} r_k\right) + \log C - \log 2} \\
&\geq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \log\left(C r_1^{s_1 m'} + \left(\sum_{k=a_n-m}^{\infty} r_k\right)^{\widehat{s}-\epsilon}\right)}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(C r_1^{m'+1} + \sum_{k=a_n-m}^{\infty} r_k\right) + \log C - \log 2} \\
&\geq \frac{\sum_{k=1}^{n-1} \log p_{a_k} + \min\{s_1, \widehat{s}-\epsilon\} \log\left(C r_1^{m'} + \sum_{k=a_n-m}^{\infty} r_k\right)}{\sum_{k=1}^{n-1} \log r_{a_k} + \log\left(C r_1^{m'+1} + \sum_{k=a_n-m}^{\infty} r_k\right) + \log C - \log 2}
\end{aligned}
$$

We finish the proof by observing that in all reductions and cases we have a lower bound for $\log(\mu(B(x,r)))/\log r$ in terms of a fraction having numerators and denominators of comparable growth, up to multiplicative coefficients. Since such bound must hold for all choices of $r$, we obtain that

$$\liminf_{r\to\infty} \frac{\log \mu(B(x,r))}{\log r} \geq 1 - \epsilon$$

for almost every $x$, from where we obtain the result. ■

## 3.11 Final remarks of the chapter

Theorem 3.8.3 implies that for maps such that $\{r_n\}$ decays polynomially, the Hausdorff dimension of ergodic invariant measures with infinite entropy is equal to zero under mild independence and regularity assumptions on the measure.

**Question 1.** *Is there an ergodic invariant measure $\mu$ for a Gauss-like map with $h_\mu = \lambda_\mu = \infty$, and $\dim_H \mu > 0$?*

We believe that the infinite entropy condition and the polynomial decay of the size of the partition forces the Hausdorff dimension to drop to zero. The example from the previous section shows that this might not be the case when $r_n$ does not decay polynomially fast. We also formulate two questions for a more general case:

**Question 2.** *What can be said about the almost sure value of the symbolic dimension when $\mu$ is only assumed to be ergodic?*

**Question 3.** *What can be said about $\dim_H \mu$ when $\mu$ is only assumed to be ergodic?*

The main difficulty with questions 2 and 3 is that our methods rely on the asymptotic independence of the digits in the symbolic space. This implies that we can write the measure and diameter of cylinders in the form of Birkhoff sums, allowing us to use ergodic theoretic methods to study the almost sure behavior of such sums. For measures which do not satisfy any kind of independence assumption, we are not able to use such techniques. Based on these remarks, we finish the chapter with the following conjecture:

**Conjecture 1.** *For the Gauss map, there are no invariant ergodic probability measures with infinite entropy and positive Hausdorff dimension.*

CHAPTER 4

# LIMIT LAWS FOR SEQUENTIAL AND RANDOM DYNAMICAL SYSTEMS

## 4.1 Introduction

In the previous chapter we studied asymptotic properties of iterations of a single map and their associated invariant probability measures. In this chapter, we focus on a problem where we consider compositions of possibly different maps, belonging to a common family. We study a particular family of non-uniformly hyperbolic maps and their limit laws with respect to a reference measure (Lebesgue). In particular, we are interested in studying large deviations and central limit theorems for both deterministic and random compositions of such maps. The complications are twofold: on one hand, the lack of uniform hyperbolicity of the maps makes the traditional methods used to obtain limit laws (quasi-compactness of the transfer operators) obsolete, as they rely on the quasi-compactness of the transfer operator on certain spaces, a property which no longer holds when the maps are not uniformly-hyperbolic. On the other hand, the fact that there is no common invariant measure for the family makes the time series associated to a given potential non-stationary. This implies that in order to study the fluctuations of the Birkhoff sums around their expected value, we must randomly center the sums.

The theory of limit laws and rates of decay of correlations for uniformly hyperbolic and some non-uniformly hyperbolic sequential and random dynamical systems has recently seen major progress. For expanding and uniformly hyperbolic maps, the works of Kifer, [Kif91] and [Kif92] are foundational. Other results in this area include: in [CR07] strong laws of large numbers and centered central limit theorems for sequential expanding maps; in [AHN$^+$15], polynomial decay of correlations for sequential intermittent systems; in [NTV18], sequential and quenched (self-centering) central limit theorems for intermittent systems; in [ANV15], annealed versions of a central limit theorem, large deviations principle, local limit theorem and almost sure invariance principle are proven for random expanding dynamical systems, as well as quenched versions of a central limit theorem, dynamical Borel-Cantelli lemmas, Erdős-Rényi laws and concentration inequalities; in [AA16], necessary and sufficient conditions are given for a central limit theorem without random centering for uniformly expanding maps; and in [BB16b] mixing rates and central limit theorems are given for random intermittent maps using a Tower construction. Recently the preprint [BBR19] considered quenched decay of correlation for slowly mixing systems and the preprint [AF18] used martingale techniques to obtain large deviations for systems with stretched exponential decay rates.

More precisely, we consider in the first instance a fixed deterministically chosen sequence of maps $\dots T_{\alpha_n}, \dots, T_{\alpha_1}$ in the sequential case, or a randomly drawn sequence $\dots T_{\omega_n}, \dots, T_{\omega_1}$ with respect to a Bernoulli measure $\nu$ on $\Sigma := \{T_1, \dots, T_k\}^{\mathbb{N}}$, where each of the maps $T_j$ is a Liverani-Saussol-Vaienti [LSV99] intermittent map of form

$$T_{\alpha_j}(x) = \begin{cases} x + 2^{\alpha_j} x^{1+\alpha_j}, & 0 \le x \le 1/2, \\ 2x - 1, & 1/2 \le x \le 1 \end{cases},$$

for numbers $0 < \alpha_j \le \alpha < 1$. In figure 4.1 the graph of $T_\alpha$ is shown for a particular choice of $\alpha$.

We consider the asymptotic behavior of the centered (that is, after subtracting their expectation) sums

$$S_n := \sum_{k=1}^{n} \varphi \circ (T_{\alpha_k} \circ \dots \circ T_{\alpha_1})$$
$$\widehat{S}_n := S_n - \mathbb{E}(S_n) = \sum_{k=1}^{n} \varphi \circ (T_{\alpha_k} \circ \dots \circ T_{\alpha_1}) - \sum_{k=1}^{n} m(\varphi \circ T_{\alpha_k} \circ \dots \circ T_{\alpha_1})$$

for sufficiently regular observables $\varphi$. Denote by $m$ Lebesgue measure on $X := [0,1]$, and by $m(\varphi)$ the integral of $\varphi$ with respect to $m$. We will also consider the measure $\widetilde{m}$ given by $d\widetilde{m}(x) = x^{-\alpha} dm$, where $0 < \alpha_j \le \alpha < 1$. The motivation for introduction of this measure is that in the case of a stationary system, if $\alpha_k = \alpha$ for each $k$, then a natural and convenient measure to use is the invariant measure $\mu_\alpha$ for $T_\alpha$, which behaves near

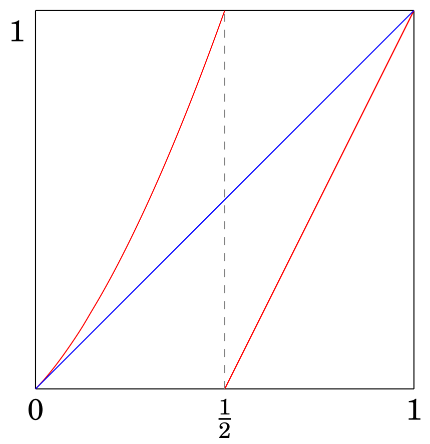


Figure 4.1: In red, $T_\alpha$ for $\alpha = 0.8$. In blue, the identity map.

0 as $x^{-\alpha}$. In the stationary case large deviation estimates are given with respect to $\mu_\alpha$ and $m$ in [MN08].

In the sequential case of a fixed realization we are interested in the large deviations of the self-centered sums $\widehat{S}_n$. In particular, we obtain a bound of the form

$$m\left\{x : \left|\widehat{S}_n\right| > n\epsilon\right\} \le C_{\alpha,\varphi,\epsilon}\, p(n)^{-1}.$$

for $\epsilon > 0$ and $n \ge 1$, where $C_{\alpha,\varphi,\epsilon}$ is a constant and $p$ is a function of polynomial growth. We also obtain large deviations with respect to $\widetilde{m}$, which are in a sense sharper. In the sequential case centering is clearly necessary.

In the annealed case we consider the random dynamical system (RDS) $F : \Sigma \times [0,1] \to \Sigma \times [0,1]$ given by $F(\omega, x) = (\tau\omega, T_{\alpha_1}x)$ for $\omega = (\alpha_1, \alpha_2, \ldots) \in \Sigma$, where $\tau$ is the left-shift operator on $\Sigma$. For $\nu$ a Bernoulli measure on $\Sigma$, we suppose $\mu$ is a stationary measure for the stochastic process on $[0,1]$, that is, a measure such that $\nu \otimes \mu$ is $F$ invariant. This assumption is valid in the setting we consider. If $\varphi$ is an observable such that $\mu(\varphi) = 0$, we obtain a bound of the form

$$\nu \otimes \mu\{(\omega, x) : |S_n| > n\epsilon\} \le C_{\alpha,\varphi,\epsilon}\, p(n)^{-1}$$

And similarly, in the quenched case, once again assuming $\mu(\varphi) = 0$, we give bounds for

$$m\{x : |S_n| > n\epsilon\} \le C_{\alpha,\varphi,\epsilon}\, p(n)^{-1}$$

for $\nu$-almost every realization $\omega \in \Sigma$. Note that in both the annealed and quenched case the centering is only on average; there is no random (sampling) centering.

Since the maps we are considering are not uniformly hyperbolic, spectral methods used to obtain limits laws are not immediately available. Our techniques to establish large deviations estimates are based on those developed for stationary systems, in particular the martingale methods of [MN08, Mel09].
Using recent work of [AA16] and [HS20] we extend the results of [NTV18] on quenched central limit theorems (CLT) for centered observables over random compositions of intermittent maps in two ways, first by enlarging the parameter range over which the quenched CLT holds and second by showing as a consequence of results in [HS20] that the variance in the quenched CLT is almost surely constant and equal to the variance of the annealed CLT.
We also study the necessity of centering to achieve a quenched CLT using ideas of [AA16] and [ANV15]. The work of [ANV15] together with our observations show that centering is necessary 'generically' (in a sense made precise later) to obtain the quenched CLT in fairly general hyperbolic situations.
More concretely, we improve some earlier results of [NTV18]. We describe these results in what follows:

**Theorem 4.1.1** ([NTV18]). *Let $\varphi$ be a $C^1([0,1])$ function, and assume that*

$$\sigma_n^2 := \mathrm{var}\left(\widehat{S}_n\right) = \mathbb{E}\left[(S_n - \mathbb{E}[S_n])^2\right] \gtrsim n^{\beta}.$$

*If*

$$0 < \alpha < \frac{1}{9} \text{ and } \beta > \frac{1}{2(1-2\alpha)},$$

*then*

$$\frac{\widehat{S}_n}{\sigma_n} \Rightarrow N(0,1).$$

The previous result is with respect to the Lebesgue measure. Our improvements to this theorem are the following:

- we show that the sequential CLT in [NTV18, Theorem 3.1], [HL19], holds for the sharp $\alpha < 1/2$ (from $\alpha < 1/9$) if the variance grows at the rate specified.
- we show that the CLT holds not only with respect to Lebesgue measure $m$ but also for $d\widetilde{m} = x^{-\alpha} dm$, which scales at the origin as the invariant measure of $T_\alpha$.
- in the case of quenched CLT's of [NTV18, Theorem 3.1], using results of Hella and Stenlund [HS20] we show that the variance $\sigma_\omega^2$ is almost-surely the same for any sequence of maps and equal to the annealed variance $\sigma^2$.

## 4.2 Notation and assumptions

Throughout this chapter, $m$ denotes the Lebesgue measure on $X := [0,1]$ and $\mathscr{B}$ the Borel $\sigma$-algebra on $[0,1]$; by $\lceil x \rceil$ will denote the smallest integer greater or equal to $x$. We consider the family of intermittent maps given by

$$T_\alpha(x) = \begin{cases} x + 2^\alpha x^{1+\alpha}, & 0 \le x \le 1/2, \\ 2x - 1, & 1/2 \le x \le 1 \end{cases}, \tag{4.1}$$

for $\alpha \in (0,1)$.
For $\beta_k \in (0,1)$ denote by $P_{\beta_k} = P_k : L^1(m) \to L^1(m)$ the transfer operator (or Ruelle-Perron-Frobenius operator) with respect to $m$ associated to the map $T_{\beta_k} = T_k$, defined as the "pre-dual" of the Koopman operator $f \mapsto f \circ T_k$, acting on $L^\infty(m)$. The duality relation is given by

$$\int_X P_k f \; g \; dm = \int_X f \; g \circ T_k \; dm$$

for all $f \in L^1(m)$ and $g \in L^\infty(m)$ [BG97, Proposition 4.2.6]. For a fixed sequence $\{\beta_k\}$ such that $0 < \beta_k \le \alpha$ for all $k$, define

$$\begin{aligned} \mathscr{T}^\infty &:= \dots, T_{\beta_n}, \dots, T_{\beta_1} \\ \mathscr{T}_m^n &:= T_{\beta_n} \circ \dots \circ T_{\beta_m}, \qquad \mathscr{T}^n := \mathscr{T}_1^n \\ \mathscr{P}_m^n &:= P_{\beta_n} \circ \dots \circ P_{\beta_m}, \qquad \mathscr{P}^n := \mathscr{P}_1^n \end{aligned}$$

We will often write, for ease of exposition when there is no ambiguity, $T_{\beta_n} \circ \dots \circ T_{\beta_m}$ as $T_n \circ \dots \circ T_m$ and $P_{\beta_n} \circ \dots \circ P_{\beta_m}$ as $P_n \circ \dots \circ P_m$.
Since $L^1(m)$ is invariant under the action of the transfer operators, the duality relation extends to compositions

$$\int_X \mathscr{P}_k^n f \; g \; dm = \int_X f \; g \circ \mathscr{T}_k^n \; dm.$$

We will write $\mathbb{E}_m[\varphi|\mathscr{F}]$ for the conditional expectation of $\varphi$ on a sub-$\sigma$-algebra $\mathscr{F}$ with respect to the measure $m$. To simplify notation we might write $\mathbb{E}$ for $\mathbb{E}_m$.

**Remark 4.2.1.** *In [CR07, NTV18] it is shown that*

$$\mathbb{E}_m[\varphi \circ \mathscr{T}^\ell | \mathscr{T}^{-k}\mathscr{B}] = \frac{P_k \circ \dots \circ P_{\ell+1}(\varphi \cdot \mathscr{P}^\ell(\mathbf{1}))}{\mathscr{P}^k(\mathbf{1})} \circ \mathscr{T}^k \tag{4.2}$$

*for* $0 \le \ell \le k$.

One of the main tools to study sequential and random systems of intermittent maps is the use of cones (see [LSV99], [AHN+15], [NTV18] ). Define the cone $\mathscr{C}_2$ by

$$\mathscr{C}_2 := \{f \in C^0((0,1]) \cap L^1(m) \mid f \geq 0, f \text{ non-increasing }, X^{\alpha+1} f \text{ increasing }, f(x) \leq a x^{-\alpha} m(f)\},$$

where $X(x) = x$ is the identity function and $m(f)$ is the integral of $f$ with respect to $m$. In [AHN+15] it is proven that for a fixed value of $\alpha \in (0,1)$, provided that the constant $a$ is big enough, the cone $\mathscr{C}_2$ is invariant under the action of all transfer operators $P_\beta$ with $0 < \beta \leq \alpha$.

**Notation.** *In general we will denote the transfer operator with respect to a non-singular*[1] *measure $\mu$ (not necessarily Lebesgue measure) by $P_\mu$. Similarly, the (conditional) expectation will be denoted by $\mathbb{E}_\mu$.*
*Denote the centering with respect to $\mu$ of a function $\varphi \in L^1(X,\mu)$ by*

$$[\varphi]^\mu := \varphi - \frac{1}{\mu(X)} \int_X \varphi \, d\mu \tag{4.3}$$

*In particular, for $g(x) := x^{-\alpha}$, denote the measure $gm$ by $\widetilde{m}$, the corresponding transfer operator by $\widetilde{P} := P_{gm}$, and the (conditional) expectation by $\mathbb{E}_{\widetilde{m}} := \mathbb{E}_{gm}$.*

## Random dynamical systems.

Now we introduce a randomized choice of maps: consider a finite family of intermittent maps of the form (4.1), indexed by a set $\Omega = \{\beta_1, \ldots, \beta_m\} \subset (0,\alpha)$. Given a probability distribution $\mathbb{P} = (p_1, \ldots, p_m)$ on $\Omega$, define a Bernoulli measure $\mathbb{P}^{\otimes \mathbb{N}}$ on $\Sigma := \Omega^{\mathbb{N}}$ by $\mathbb{P}^{\otimes \mathbb{N}}\{\omega : \omega_{j_1} = \beta_{j_1}, \ldots, \omega_{j_k} = \beta_{j_k}\} = \prod_{i=1}^k p_{j_i}$ for every finite cylinder and extend to the sigma-algebra generated by the cylinders of $\Sigma$ by Kolmogorov's extension theorem. This measure is invariant and ergodic with respect to the shift operator $\tau$ on $\Sigma$, $\tau : \Sigma \to \Sigma$ acting on sequences by $(\tau(\omega))_k = \omega_{k+1}$. We will denote $\mathbb{P}^{\otimes \mathbb{N}}$ by $\nu$ from now on.
For $\omega = (\omega_1, \omega_2, \ldots) \in \Sigma$ define $\mathscr{T}_\omega^n := T_{(\tau^n \omega)_1} \circ \cdots \circ T_{\omega_1} = T_{\omega_n} \circ \cdots \circ T_{\omega_1}$. The random dynamical system is defined as

$$\begin{aligned} F : \Sigma \times X &\to \Sigma \times X \\ (\omega, x) &\mapsto \left(\tau\omega, T_{\omega_1} x\right). \end{aligned}$$

The iterates of $F$ are given by $F^n(\omega, x) = (\tau^n(\omega), \mathscr{T}_\omega^n(x))$.

[1]The measure $\mu$ is non-singular for the transformation $T$ if $\mu(A) > 0 \implies \mu(T(A)) > 0$.

We will also use $\Omega$-indexed subscripts for random transfer operators associated to the maps $T_{\omega_i}$, so that $P_{\omega_i} := P_{T_{\omega_i}}$. We will also abuse notation and write $P_\omega$ for $P_{\omega_1}$ if $\omega = (\omega_1, \omega_2, \ldots, \omega_n, \ldots)$.
A probability measure $\mu$ on $X$ is said to be stationary with respect to the RDS $F$ if

$$\mu(A) = \int_\Sigma \mu\left(T_{\omega_1}^{-1}(A)\right) d\nu(\omega) = \sum_{\beta \in \Omega} p_\beta \mu\left(T_\beta^{-1}(A)\right)$$

for every measurable set $A$, where $p_\beta$ is the $\mathbb{P}$-probability of the symbol $\beta$. This is equivalent to the measure $\nu \otimes \mu$ being invariant under the transformation $F\colon \Sigma \times X \to \Sigma \times X$.
See Remark 4.4.3 about the existence and ergodicity of such a stationary measure in our setting.
The annealed transfer operator $P\colon L^1(m) \to L^1(m)$ is defined by averaging over all the transformations:

$$P = \sum_{\beta \in \Omega} p_\beta P_\beta = \int_\Sigma P_\omega \, d\nu(\omega).$$

This operator is "pre-dual" to the annealed Koopman operator $U\colon L^\infty(m) \to L^\infty(m)$ defined by

$$(U\varphi)(x) := \sum_{\beta \in \Omega} p_\beta \varphi(T_\beta x) = \int_\Sigma \varphi(T_\omega x) d\nu(\omega) = \int_\Sigma F(\tilde{\varphi})(\omega, x) d\nu(\omega)$$

where $\tilde{\varphi}(\omega, x) := \varphi(x)$. The annealed operators satisfy the duality relationship

$$\int_X (U\varphi) \cdot \psi \, dm = \int_X \varphi \cdot P\psi \, dm$$

for all observables $\varphi \in L^\infty(m)$ and $\psi \in L^1(m)$.

## 4.3 Background results and the Martingale approximation

In this section we describe the main technique used to prove some of the limit law results: the martingale approximation, introduced by Gordin [Gor69]. Since there is no common invariant measure for the set of maps $\{T_k\}$, for a given $C^1$ observable $\varphi$ we center along the orbit by

$$[\varphi]_k(\omega, x) := \varphi(x) - \int_X \varphi \circ \mathcal{T}_\omega^k \, dm,$$

with $\mathcal{T}_\omega^k = Id$ for $k = 0$.
This implies that $\mathbb{E}_m([\varphi]_k \circ \mathcal{T}^k) = 0$ and consequently the centered Birkhoff sums

$$\widehat{S}_n := \sum_{k=1}^{n} [\varphi]_k \circ \mathcal{T}^k,$$

have zero mean with respect to $m$. Following [NTV18], define

$$H_1 := 0 \text{ and } H_n \circ \mathcal{T}^n := \mathbb{E}_m\left(\widehat{S}_{n-1}|\mathcal{B}_n\right) \text{ for } n \geq 2 \tag{4.4}$$

and the (reverse) martingale sequence $\{M_n\}$ by

$$M_0 := 0 \text{ and } \widehat{S}_n = M_n + H_{n+1} \circ \mathcal{T}^{n+1},$$

where the filtration here is $\mathcal{B}_n = \mathcal{T}^{-n}\mathcal{B}$. Define $\psi_n \in L^1(m)$ by setting

$$\psi_n = [\varphi]_n + H_n - H_{n+1} \circ T_{n+1},$$

then $M_n - M_{n-1} = \psi_n \circ \mathcal{T}^n$ and we have that $\mathbb{E}(M_n|\mathcal{B}_{n+1}) = 0$. Thus $\{\psi_n \circ \mathcal{T}^n\}$ is a reverse martingale difference scheme. An explicit expression for $H_n$ is given by

(4.5)
$$H_n = \frac{1}{\mathcal{P}^n \mathbf{1}} \left[P_n\left([\varphi]_{n-1}\mathcal{P}_{n-1}\mathbf{1}\right) + P_n P_{n-1}\left([\varphi]_{n-2}\mathcal{P}_{n-2}\mathbf{1}\right) + \cdots + P_n P_{n-1} \cdots P_1\left([\varphi]_0 \mathcal{P}_0 \mathbf{1}\right)\right].$$

**Remark 4.3.1.** *The formulas derived so far with $m$ being the Lebesgue measure actually hold for any measure $\mu$ that is non-singular for the transformations $T_\beta$ considered. The conditional expectations $\mathbb{E}_\mu$ will be with respect to $\mu$ and the transfer operator $P_\mu$ will be with respect to the measure space $(X, \mu)$. In particular the centering will have the form*

$$[\varphi]_k(\omega, x) := \varphi(x) - \frac{1}{\mu(X)} \int_X \varphi \circ \mathcal{T}_\omega^k \, d\mu,$$

*but all other equations are the same, with the notational changes just described.*

We collect and extend some results from [NTV18] concerning the properties of $H_n$, as well as the non-stationary decay of correlations for the sequential system.
We state first a few formulas for changing from a measure $m$ to the measure $g(x)\,dm(x)$ with $g \in L^1(m)$; for simplicity, we denote this new measure as $gm$ when there is no possibility of confusion.

**Lemma 4.3.2** (Change of measure)**.** *We state this result only for the situation we need, but it holds also for any measure $\mu$ non-singular with respect to $T$ in place of $m$ the Lebesgue measure, and instead of $g(x) = x^{-\alpha}$ for any $g \in L^1(\mu)$, $g > 0$.*

*Note that $L^1(gm) = g^{-1} \cdot L^1(m)$, so all formulas below make sense for $\varphi$ in the appropriate $L^1$-space.*
*We have:*

$$\begin{aligned} m(\varphi) &= m(P_m\varphi) \\ P_{gm}(\varphi) &= g^{-1}P_m(g\varphi) \\ g\,[\varphi]^{gm} &= [g\varphi]^m - \frac{m(g\varphi)}{m(g)}[g]^m \\ \mathbb{E}_{gm}(\varphi|\mathscr{B}) &= \mathbb{E}_m(g\varphi|\mathscr{B})/\mathbb{E}_m(g|\mathscr{B}) \end{aligned} \tag{4.6}$$

*Therefore*

$$(\mathscr{P}_{gm})_\ell^k\left([\varphi]^{gm}\right) = g^{-1}(\mathscr{P}_m)_\ell^k\left([g\varphi]^m - \frac{m(g\varphi)}{m(g)}[g]^m\right) \tag{4.7}$$

**Proof.** The first two properties are standard and follow from the definition of the transfer operator. The third is a direct computation using the notation (4.3).
For the fourth, $\mathbb{E}_{gm}(\varphi|\mathscr{B})$ is the function $\Phi$ that is $\mathscr{B}$-measurable and $\int \Phi\psi\, d(gm) = \int \varphi\psi\, d(gm)$ for each $\psi \in L^\infty(\mathscr{B})$. Expanding the LHS,

$$\int \Phi\psi\, d(gm) = \int \Phi\psi g\, dm = \int \Phi\psi\mathbb{E}_m(g|\mathscr{B})\, dm$$

whereas the RHS becomes

$$\int \varphi\psi\, d(gm) = \int \varphi\psi g\, dm = \int \mathbb{E}_m(g\varphi|\mathscr{B})\psi\, dm$$

Thus $\Phi\mathbb{E}_m(g|\mathscr{B}) = \mathbb{E}_m(g\varphi|\mathscr{B})$, as claimed. ∎

**Proposition 4.3.3** ([NTV18])**.** *If $\varphi, \psi$ are both in the cone $\mathscr{C}_2$ and have the same mean, $\int_X \varphi dm = \int_X \psi dm$, then by [NTV18, Theorem 1.2]*

$$\left\|\mathscr{P}^n(\varphi) - \mathscr{P}^n(\psi)\right\|_{L^1(m)} \le C_\alpha\left(\|\varphi\|_{L^1(m)} + \|\psi\|_{L^1(m)}\right) n^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}}$$

*Moreover [NTV18, Remark 2.5 and Corollary 2.6], for $\varphi \in C^1$, $h \in \mathscr{C}_2$ and any sequence of maps $\mathscr{T}^\infty$:*

$$\left\|\mathscr{P}^n\left([h\varphi]^m\right)\right\|_{L^1(m)} \le C_\alpha \mathscr{F}\left(\|\varphi\|_{C^1} + m(h)\right) n^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}}$$

*where $C_\alpha$ depends only on the map $T_\alpha$, and $\mathscr{F}: \mathbb{R} \to \mathbb{R}$ is an affine function.*

The decay result of Proposition 4.3.3 for products of elements in the cone with $C^1$ observables (see also [LSV99, Theorem 4.1]), follows from Lemma 4.3.4, which was stated in [LSV99, proof of Theorem 4.1]. The proof of Lemma 4.3.4 is given next; a different – less transparent – proof is given in [NTV18, Lemma 2.4].

**Lemma 4.3.4.** *Suppose $\varphi \in C^1$ and $h \in \mathscr{C}_2$. Then there exist constants $\lambda, A, B \in \mathbb{R}$ such that $(\varphi + A + \lambda x)h + B$ and $(A + \lambda x)h + B$ both are in $\mathscr{C}_2$ and hence if $\int \varphi h dm = 0$ then $\left\|\mathscr{P}^j(\varphi h)\right\|_{L^1(m)} \leq C\rho(j)\|\varphi h\|_{L^1(m)}$ where $\rho(j)$ is the $L^1(m)$-decay for centered functions from the cone $\mathscr{C}_2$.*

Note that in our setting $\rho(j) = j^{-\frac{1}{\alpha}+1}(\log j)^{\frac{1}{\alpha}}$.

**Proof of Lemma 4.3.4** Let $f_1 = (\varphi + \lambda x + A)h + B$ and $f_2 = (A + \lambda x)h + B$.
First we show that $f_1 \in \mathscr{C}_2$. It is clear that $f_1 \in C^0(0,1] \cap L^1(m)$. Choose $\lambda < 0$ such that $|\lambda| > \|\varphi'\|_{L^\infty}$ and $A > 0$ large enough so that

$$\varphi + \lambda x + A > 0.$$

This ensures that $f_1 \geq 0$ for any value of $B \geq 0$. Note now that

$$(\varphi + \lambda x + A)' = \varphi' + \lambda \leq 0$$

so $\varphi + \lambda x + A$ is decreasing. Since both $\varphi + \lambda x + A$ and $h$ are positive and decreasing, we obtain that $f_1$ is decreasing as well. We show now that $x^{\alpha+1} f_2$ is increasing. Since $h \in \mathscr{C}_2$, $h$ is non-increasing so $h'$ exists $m$-a.e. and $h' \leq 0$ $m$-a.e. Then $(x^{\alpha+1}h)'$ exists $m$-a.e. as well, and we can compute this derivative as

$$(x^{\alpha+1}h)' = (\alpha+1)x^\alpha h + x^{\alpha+1}h' \geq 0$$

because it is increasing.
We compute now the derivative of $x^{\alpha+1} f_2$:

$$\begin{aligned}(x^{\alpha+1}[(\varphi + \lambda x + A)h + B])' = (\alpha+1)x^\alpha \varphi h + x^{\alpha+1}\varphi' h + x^{\alpha+1}\varphi h' + (\alpha+2)x^{\alpha+1}h\lambda + \\ \lambda x^{\alpha+2}h' + (\alpha+1)Ax^\alpha h + Ax^{\alpha+1}h' + (\alpha+1)x^\alpha B.\end{aligned}$$

We group terms conveniently: note that

$$(\alpha+1)x^\alpha \varphi h + (\alpha+1)Ax^\alpha h + x^{\alpha+1}\varphi h' + Ax^{\alpha+1}h' = (\varphi + A)[(\alpha+1)x^\alpha h + h' x^{\alpha+1}] \geq 0$$

$m$-a.e., since the term in the square brackets corresponds to $(x^{\alpha+1}h)' \geq 0$. The term $\lambda x^{\alpha+2}h'$ is non-negative $m$-a.e. since $\lambda, h' \leq 0$. Since $0 \leq h(x)x^\alpha \leq am(h)$, we have $0 \leq$

$-x^{\alpha+1}h' \le (\alpha+1)x^\alpha h \le (\alpha+1)am(h)$ and then the terms $(\alpha+2)\lambda x^{\alpha+1}h + x^{\alpha+1}h\varphi'$ are bounded. Thus, we can take $B > 0$ big enough so that

$$(\alpha+1)x^\alpha B \ge (\alpha+2)\lambda x^{\alpha+1}h + x^{\alpha+1}h\varphi'.$$

With this, we have that $(x^{\alpha+1}h)' \ge 0$ and so $x^{\alpha+1}h$ is increasing.
Finally, we check that $f_1(x)x^\alpha \le am(f_1)$. Using that $h(x)x^\alpha \le am(h)$,

$$[(\varphi+\lambda x+A)h+B]x^\alpha \le (\varphi+\lambda x+A)hx^\alpha + B \le \sup(\varphi+\lambda x+A)am(h)+B.$$

On the other hand, $am((\varphi+\lambda x+A)h+B) \ge a\inf(\varphi+\lambda x+A)m(h)+aB$, so it suffices to have

$$\begin{aligned} &\sup(\varphi+\lambda x+A)am(h)+B \le a\inf(\varphi+\lambda x+A)m(h)+aB \\ \iff &B \ge \frac{a}{a-1}\left[\sup(\varphi+\lambda x+A)-\inf(\varphi+\lambda x+A)\right]m(h). \end{aligned}$$

Thus, we see that $f_1 \in \mathscr{C}_2$. The proof that $f_2 \in \mathscr{C}_2$ is the same, take $\varphi(x) \equiv 0$. ∎

A consequence of Proposition 4.3.3 is the non-stationary decay of correlations ([NTV18, Page 1130])

$$\begin{aligned} &\left|\int_X \varphi\cdot\psi\circ T_{\omega_n}\circ\ldots\circ T_{\omega_1}dm - m(\varphi)\cdot m(\psi\circ T_{\omega_n}\circ\ldots\circ T_{\omega_1})\right| \\ &\qquad\qquad \le \|\psi\|_\infty \left\|\mathscr{P}^n_\omega(\varphi)-\mathscr{P}^n_\omega\left(\mathbf{1}\int_X\varphi dm\right)\right\|_{L^1(m)} \end{aligned}$$

We derive next decay estimates with respect to the measure $\widetilde{m}$, which are better in $L^p$, $p>1$, than those for $m$.

**Proposition 4.3.5.** *For $\varphi:[0,1]\to\mathbb{R}$ bounded, $h\in\mathscr{C}_2$ and $1\le p\le\infty$:*

$$\left\|\widetilde{\mathscr{P}}^n(\varphi)\right\|_{L^\infty(\widetilde{m})} \le m(g)\|\varphi\|_{L^\infty(\widetilde{m})} \tag{4.8}$$

*For $\varphi\in C^1([0,1])$, $h\in\mathscr{C}_2$ and $1\le p\le\infty$:*

$$\left\|\widetilde{\mathscr{P}}^n\left(\left[(g^{-1}h)\varphi\right]^{\widetilde{m}}\right)\right\|_{L^1(\widetilde{m})} \le C_\alpha\mathscr{F}_1\left(\|\varphi\|_{C^1}+m(h)\right)n^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}} \tag{4.9}$$

$$\left\|\widetilde{\mathscr{P}}^n\left(\left[(g^{-1}h)\varphi\right]^{\widetilde{m}}\right)\right\|_{L^p(\widetilde{m})} \le C_\alpha\mathscr{F}_p\left(\|\varphi\|_{C^1}+m(h)\right)n^{\frac{1}{p}(-\frac{1}{\alpha}+1)}(\log n)^{\frac{1}{p\alpha}} \tag{4.10}$$

*where $C_\alpha$ depends only on $T_\alpha$ and $\mathscr{F}_p$ are affine functions.*
*Note that the $L^1$ and $L^p$ bounds are relevant only for $\varphi\in C^1$.*

**Proof.** To prove the $L^\infty$ estimate (4.8) note that by the invariance of the cone $\mathscr{C}_2$, $\mathscr{P}^n(g) \in \mathscr{C}_2$, so $\mathscr{P}^n(g) \le x^{-\alpha} m(\mathscr{P}^n(g)) = x^{-\alpha} m(g)$. That is, using (4.6),

$$\widetilde{\mathscr{P}}^n(\mathbf{1}) = g^{-1}\mathscr{P}^n(g) \le m(g)$$

Since $-\|\varphi\|_{L^\infty}\mathbf{1} \le \varphi \le \|\varphi\|_{L^\infty}\mathbf{1}$ and $\widetilde{\mathscr{P}}^n$ are positive operators, we obtain (4.8).
For (4.9) assume that $\varphi \in C^1$ (otherwise it is clearly satisfied). In view of (4.7):

$$\begin{aligned} \left\|\widetilde{\mathscr{P}}^n\left([(g^{-1}h)\varphi]^{\widetilde{m}}\right)\right\|_{L^1(\widetilde{m})} &= \left\|g^{-1}\mathscr{P}^n\left([h\varphi]^m\right) - \frac{m(g\varphi)}{m(g)} g^{-1}\mathscr{P}^n([g]^m)\right\|_{L^1(\widetilde{m})} \\ &= \left\|\mathscr{P}^n\left([h\varphi]^m\right) - \frac{m(g\varphi)}{m(g)}\mathscr{P}^n([g]^m)\right\|_{L^1(m)} \\ &\le \left\|\mathscr{P}^n\left([h\varphi]^m\right)\right\|_{L^1(m)} + \left|\frac{m(g\varphi)}{m(g)}\right| \left\|\mathscr{P}^n([g]^m)\right\|_{L^1(m)} \end{aligned} \tag{4.11}$$

By [NTV18, Theorem 1.2], there is an affine function $\mathscr{F}: \mathbb{R} \to \mathbb{R}$ such that for $\varphi \in C^1([0,1])$ and $h \in \mathscr{C}_2$ can write $[\varphi h]^m = \Psi_1 - \Psi_2$ with $\Psi_1, \Psi_2 \in \mathscr{C}_2$ and $\|\Psi_{1,2}\|_{L^1(m)} \le \mathscr{F}(\|\varphi\|_{C^1} + m(h))$. By [NTV18, Theorem 1.2], for an observable $\psi$ in the cone $\mathscr{C}_2$ and for any sequence of maps $\mathscr{T}^\infty$, we have

$$\int_X \left|\mathscr{P}^n\left([\psi]^m\right)\right| dm \le C_\alpha \|\psi\|_{L^1(m)} n^{-\frac{1}{\alpha}+1} (\log n)^{\frac{1}{\alpha}}$$

where $C_\alpha$ depends only on $T_\alpha$. Applying these to (4.11), we obtain (4.9).
Finally, note that the $L^1$ and $L^\infty$ bounds give (4.10), since

$$\|f\|_{L^p} \le \|f\|_{L^\infty}^{1-\frac{1}{p}} \|f\|_{L^1}^{\frac{1}{p}} \tag{4.12}$$

because

$$\int |f|^p \le \int \|f\|_{L^\infty}^{p-1} |f| = \|f\|_{L^\infty}^{p-1} \|f\|_{L^1}.$$

■

**Lemma 4.3.6.** *Let $\varphi \in C^1$ and $\alpha < 1$. Then*

$$\|H_n \circ \mathscr{T}^n\|_{L^p(m)} \le \begin{cases} C_{p,\alpha,\|\varphi\|_{C^1}+m(g)} & \text{if } 1 \le p < \frac{1}{\alpha} - 1 \\ C_{p,\alpha,\|\varphi\|_{C^1}+m(g)} n^{1+\frac{1}{p}(1-\frac{1}{\alpha})} (\log n)^{\frac{1}{p\alpha}} & \text{if } p > \max\{1, \frac{1}{\alpha} - 1\} \end{cases}$$

*and the same bounds hold for $\|\widetilde{H}_n \circ \mathscr{T}^n\|_{L^p(\widetilde{m})}$, where*

$$H_n \circ \mathscr{T}^n := \mathbb{E}_m\left([S_{n-1}]^m \,|\mathscr{B}_n\right),\ \widetilde{H}_n \circ \mathscr{T}^n := \mathbb{E}_{\widetilde{m}}\left([S_{n-1}]^{\widetilde{m}} \,|\mathscr{B}_n\right),\ \mathscr{B}_n := \mathscr{T}^{-n}\mathscr{B}.$$

**Proof.** We prove the statement for $\widetilde{H}_n$. The one for $H_n$ is obtained the same way, using Proposition 4.3.3 instead of (4.9).
Using the definition of $\widetilde{H}_n$:
(4.13)
$$\|\widetilde{H}_n \circ \mathcal{T}^n\|_{L^p(\widetilde{m})} = \left\| \sum_{k=1}^{n-1} \mathbb{E}_{\widetilde{m}} \left( \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} | \mathcal{B}_n \right) \right\|_{L^p(\widetilde{m})} \le \sum_{k=1}^{n-1} \left\| \mathbb{E}_{\widetilde{m}} \left( \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} | \mathcal{B}_n \right) \right\|_{L^p(\widetilde{m})}$$

We will bound each term of the above sum in both $L^1$ and $L^\infty$, and then use (4.12) to obtain an $L^p$-bound.
In $L^\infty$ we have
$$\left\| \mathbb{E}_{\widetilde{m}} \left( \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} | \mathcal{B}_n \right) \right\|_{L^\infty(\widetilde{m})} \le \left\| \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} \right\|_{L^\infty(\widetilde{m})} \le 2\|\varphi\|_{L^\infty(\widetilde{m})}.$$

In $L^1$ we use (4.2) to compute the conditional expectation. Since the conditional expectation preserves the expected value, one can check that the centering holds as written below[2]. We can then use (4.9) for the decay, with $h = \mathcal{P}^k(g)$, because $\widetilde{\mathcal{P}}^k(\mathbf{1}) = g^{-1}\mathcal{P}^k(g)$.

$$\left\| \mathbb{E}_{\widetilde{m}} \left( \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} | \mathcal{B}_n \right) \right\|_{L^1(\widetilde{m})} = \left\| \frac{\widetilde{P}_n \circ \cdots \circ \widetilde{P}_{k+1} \left( \left[ \varphi \cdot \widetilde{\mathcal{P}}^k(\mathbf{1}) \right]^{\widetilde{m}} \right)}{\widetilde{\mathcal{P}}^n(\mathbf{1})} \circ \mathcal{T}^n \right\|_{L^1(\widetilde{m})}$$
$$= \left\| \widetilde{P}_n \circ \cdots \circ \widetilde{P}_{k+1} \left( \left[ \varphi \cdot \widetilde{\mathcal{P}}^k(\mathbf{1}) \right]^{\widetilde{m}} \right) \right\|_{L^1(\widetilde{m})} = \left\| \widetilde{P}_n \circ \cdots \circ \widetilde{P}_{k+1} \left( \left[ \varphi \cdot g^{-1}\mathcal{P}^k(g) \right]^{\widetilde{m}} \right) \right\|_{L^1(\widetilde{m})}$$
$$\le C_\alpha \mathcal{F}_1 \left( \|\varphi\|_{C^1} + m(\mathcal{P}^k(g)) \right) (n-k)^{-\frac{1}{\alpha}+1} (\log(n-k))^{\frac{1}{\alpha}}.$$

Note that $m(\mathcal{P}^k(g)) = m(g)$, so the coefficient above does not depend on $k$.
Apply now (4.12) to obtain for $1 \le p \le \infty$ that
$$\left\| \mathbb{E}_{\widetilde{m}} \left( \left[ \varphi \circ \mathcal{T}^k \right]^{\widetilde{m}} | \mathcal{B}_n \right) \right\|_{L^p(\widetilde{m})} \le C_{p,\alpha,\|\varphi\|_{C^1}+m(g)} \left[ (n-k)^{-\frac{1}{\alpha}+1} (\log(n-k))^{\frac{1}{\alpha}} \right]^{\frac{1}{p}}$$
which gives the desired bound upon summing over $k = 1, \dots, n-1$. ■

A useful remark is the following lower bound for functions in the cone $\mathcal{C}_2$:

**Proposition 4.3.7** ([LSV99, Lemma 2.4]). *For every function $f \in \mathcal{C}_2$ one has*
$$\inf_{x \in [0,1]} f(x) = f(1) \ge \min \left\{ \alpha, \left[ \frac{\alpha(1+\alpha)}{a^\alpha} \right]^{\frac{1}{1-\alpha}} \right\} m(f).$$
*Denote the constant in the above expression by $D_\alpha$. Then $\mathcal{P}^n \mathbf{1} \ge D_\alpha > 0$ for all $n \ge 1$.*

[2] $\widetilde{m}(\varphi \cdot \widetilde{\mathcal{P}}^k(\mathbf{1})) = \widetilde{m}(\varphi \circ \mathcal{T}^k)$ because, by the definition of the transfer operator, $\int \varphi \cdot \widetilde{\mathcal{P}}^k(\mathbf{1}) d\widetilde{m} = \int \varphi \circ \mathcal{T}^k \cdot \mathbf{1} d\widetilde{m}$

We will also use Rio's inequality, taken from [MPU+06]. This is a concentration inequality that allows us to bound the moments of Birkhoff sums.

**Proposition 4.3.8** ([MPU+06, Rio17])**.** *Let* $\{X_i\}$ *be a sequence of* $L^2$ *centered random variables with filtration* $\mathscr{F}_i = \sigma(X_1, \dots, X_i)$. *Let* $p \geq 1$ *and define*

$$b_{i,n} = \max_{i \leq u \leq n} \|X_i \sum_{k=i}^{u} \mathbb{E}(X_k | \mathscr{F}_i)\|_{L^p},$$

*then*

$$\mathbb{E}|X_1 + \cdots + X_n|^{2p} \leq \left(4p \sum_{i=1}^{n} b_{i,n}\right)^p.$$

Finally, we recall a theorem of Liverani which allows us to establish distributional convergence of stationary systems.

**Theorem 4.3.9** (special case of [Liv96, Theorem 1.1])**.** *Assume* $T : Y \to Y$ *preserves the probability measure* $\eta$ *on the* $\sigma$*-algebra* $\mathscr{B}$. *Denote by* $P$ *its transfer operator.*
*If* $\varphi \in L^\infty(\eta)$ *with* $\eta(\varphi) = 0$ *and* $\sum_k \|P^k \varphi\|_{L^1(\eta)} < \infty$ *then a central limit theorem holds for* $S_n\varphi := \sum_{k=1}^{n} \varphi \circ T^k$ *with respect to the measure* $\eta$, *that is,* $\frac{1}{\sqrt{n}} S_n \varphi$ *converges in distribution to* $N(0, \sigma^2)$. *The variance is given by*

$$\sigma^2 = -\eta(\varphi^2) + 2 \sum_{k=0}^{\infty} \eta(\varphi \cdot \varphi \circ T^k).$$

*In addition,* $\sigma^2 = 0$ *iff* $\varphi \circ T$ *is a measurable coboundary, that is* $\varphi \circ T = g - g \circ T$ *for a measurable* $g$.

Note that this theorem uses essentially bounded observables; we will apply this theorem to continuous observables in compact domains.

## 4.4 Polynomial large deviations estimates

In this section we prove the large deviation estimates for sequential, annealed and quenched cases.

### 4.4.1 Sequential dynamical systems

Recall we fixed a sequence $\mathscr{T}^\infty = \dots T_{\alpha_n}, \dots, T_{\alpha_1}$ where each of the maps is of the form

$$T_{\alpha_j}(x) = \begin{cases} x + 2^{\alpha_j} x^{1+\alpha_j}, & 0 \leq x \leq 1/2, \\ 2x - 1, & 1/2 \leq x \leq 1 \end{cases},$$

for $0 < \alpha_j \le \alpha < 1$. In the first part of this section we prove that for such a fixed sequence of maps $\mathcal{T}^\infty$, a polynomial large deviations bound holds for the centered sums.

**Theorem 4.4.1** (Sequential LD). *Let $0 < \alpha < 1$ and $\varphi \in C^1([0,1])$. Then the centered sums satisfy the following large deviations upper bound: for any $\epsilon > 0$ and $p > \max\{1, \frac{1}{\alpha} - 1\}$,*

$$m\left\{x : \sum_{j=1}^{n}[\varphi(\mathcal{T}^j)(x) - m(\varphi(\mathcal{T}^j))] > n\epsilon\right\} \le C_{\alpha,p,\|\varphi\|_{C^1}} n^{1-\frac{1}{\alpha}}(\log n)^{\frac{1}{\alpha}}\epsilon^{-2p}$$

*where $C = C_{\alpha,p,\|\varphi\|_{C^1}}$ is a constant depending on $\alpha$, $p$ and the $C^1$ norm of $\varphi$, but* not *on the sequence $\mathcal{T}^\infty$.*
*The same estimate (by the same proof) holds for the measure $\tilde{m}$.*

**Remark 4.4.2.** *In [MN08] these bounds are shown to be basically optimal since in the case of a single map $T_\alpha$ being iterated there exists an open and dense set of $C^1$ observables $\varphi$ such that for any $\delta > 0$, $\mu\left\{x : \sum_{j=1}^{n}[\varphi(\mathcal{T}^j)(x) - m(\varphi(\mathcal{T}^j))] > n\epsilon\right\} \ge C_\epsilon n^{1-\frac{1}{\alpha}-\delta}$ infinitely often for the absolutely continuous invariant measure $\mu$.*

**Proof of Theorem 4.4.1** We prove the estimate for $m$, the one for $\tilde{m}$ is obtained the same way.
Fix $n$ and for $i \in \{1,\dots,n\}$, define the sequence of $\sigma-$algebras $\mathcal{F}_{i,n} = \mathcal{F}_i = \mathcal{T}^{-(n-i)}(\mathcal{B})$. Note that $\mathcal{F}_i \subset \mathcal{F}_{i+1}$ hence $\{\mathcal{F}_i\}_{i=1}^n$ is an increasing sequence of $\sigma-$algebras. Take $X_i = [\varphi]_{n-i} \circ \mathcal{T}^{n-i}$, so that $X_i$ is $\mathcal{F}_i$ measurable. Recall that $\psi_j = [\varphi]_j + H_j - H_{j+1} \circ T_{j+1}$ for all $j \ge 0$. We define $Y_i = \psi_{n-i} \circ \mathcal{T}^{n-i}$, $h_i = H_{n-i} \circ \mathcal{T}^{n-i}$ for $i \in \{1,\dots,n\}$. Hence $Y_i = X_i + h_i - h_{i-1}$. Note also that $\mathcal{G}_i := \sigma(X_1,\dots,X_i) \subset \sigma(\mathcal{F}_1,\dots,\mathcal{F}_i) = \mathcal{F}_i$, as $\sigma(X_i) \subset \mathcal{F}_i$ for all $i$. Since $\mathbb{E}(\psi_i \circ \mathcal{T}^i | \mathcal{T}^{-i-1}\mathcal{B}) = 0$, $\mathbb{E}(Y_i|\mathcal{F}_j) = 0$ for all $j \ge i$. Hence $\mathbb{E}(Y_i|\mathcal{G}_j)) = \mathbb{E}(\mathbb{E}(Y_i|\mathcal{F}_j)|\mathcal{G}_j) = 0$ for $j \ge i$.
For $p \ge 1$ define $b_{i,n}$ as in Rio's inequality, with $\mathcal{G}_i$, $X_i$ as described above so that

$$b_{i,n} = \max_{i \le u \le n}\left\| X_i \sum_{k=i}^{u}\mathbb{E}(X_k|\mathcal{G}_i)\right\|_{L^p(m)}.$$

Here all the expectations are taken with respect to $m$.
Recalling the expression we have for the martingale difference, we can write the sum inside the $p$-norm as

$$\begin{aligned}\sum_{k=i}^{u}\mathbb{E}(X_k|\mathcal{G}_i) &= \sum_{k=i}^{u}[\mathbb{E}(Y_k|\mathcal{G}_i) - \mathbb{E}(h_k|\mathcal{G}_i) + \mathbb{E}(h_{k-1}|\mathcal{G}_i)] \\ &= \left[\sum_{k=i}^{u}\mathbb{E}(Y_k|\mathcal{G}_i)\right] + \mathbb{E}(h_{i-1}|\mathcal{G}_i) - \mathbb{E}(h_u|\mathcal{G}_i).\end{aligned}$$

If $k > i$, then $\mathbb{E}(Y_k|\mathscr{G}_i) = 0$. This reduces the expression above to

$$\mathbb{E}(Y_i|\mathscr{G}_i) + \mathbb{E}(h_{i-1}|\mathscr{G}_i) - \mathbb{E}(h_u|\mathscr{G}_i).$$

We note that $\|E[f|\mathscr{G}]\|_p \le \|f\|_p$ for any $f \in L^p(m)$, $p \ge 1$. Therefore, we may bound $b_{i,n}$ by $\max_{i\le u\le n} \|X_i\|_\infty(\|Y_i\|_p + \|h_{i-1}\|_p + \|h_u\|_p)$.
We now pick $p > \max\{1, \frac{1}{\alpha} - 1\}$. Since $\|X_i\|_\infty$ is uniformly bounded by $2\|\varphi\|_\infty$ and $Y_i = X_i + h_i - h_{i-1}$, we may bound $\max_{i\le u\le n} \|X_i\|_\infty(\|Y_i\|_p + \|h_{i-1}\|_p + \|h_u\|_p)$ by $C_{\alpha,p,\|\varphi\|_{C^1}} n^{1+\frac{1}{p}(1-\frac{1}{\alpha})}(\log n)^{\frac{1}{p\alpha}}$ where $C_{\alpha,p,\|\varphi\|_{C^1}}$ is independent of $n$. This is a consequence of Proposition 4.3.6.
Therefore $(4p\sum_{i=1}^n b_{i,n})^p \le C_{\alpha,\varphi,p} n^{2p+(1-\frac{1}{\alpha})}(\log n)^{\frac{1}{\alpha}}$. By Rio's inequality $E|X_1 + X_2 + \cdots + X_n|^{2p} \le C_{\alpha,\varphi,p} n^{2p+(1-\frac{1}{\alpha})}(\log n)^{\frac{1}{\alpha}}$. Thus, by Markov's inequality,

$$m(|X_1 + \cdots + X_n|^{2p} > n^{2p}\epsilon^{2p}) \le C_{\alpha,\varphi,p}(n^{-2p}\epsilon^{-2p})n^{2p+(1-\frac{1}{\alpha})}(\log n)^{\frac{1}{\alpha}} = C_{\alpha,\varphi,p} n^{1-\frac{1}{\alpha}}(\log n)^{\frac{1}{\alpha}}\epsilon^{-2p}$$

■

### 4.4.2 Random dynamical systems

Now we prove large deviations estimates for the randomized systems. First we recall some notation. The annealed transfer operator $P: L^1(m) \to L^1(m)$ is defined by averaging over all the transformations:

$$P = \sum_{\beta\in\Omega} p_\beta P_\beta = \int_\Sigma P_\omega \, d\nu(\omega).$$

This operator is dual to the annealed Koopman operator $U: L^\infty(m) \to L^\infty(m)$ defined by

$$(U\varphi)(x) = \sum_{\beta\in\Omega} p_\beta \varphi(T_\beta x) = \int_\Sigma \varphi(T_\omega x) d\nu(\omega) = \int_\Sigma \tilde{\varphi}(F(\omega,x)) d\nu(\omega)$$

where $\tilde{\varphi}(\omega,x) := \varphi(x)$. The annealed operators satisfy the duality relationship

$$\int_X (U\varphi)\cdot\psi \, dm = \int_X \varphi\cdot P\psi \, dm$$

for all observables $\varphi \in L^\infty(m)$ and $\psi \in L^1(m)$.

**Remark 4.4.3.** *It is easy to see that the averaged transfer operator $P$ has no worse rate of decay in $L^1$ then the slowest of the maps (so better than $n^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}}$, by Proposition 4.3.3). By taking a limit point of $\frac{1}{n}\sum_{k=1}^n P^k(\mathbf{1})$, there is an invariant vector $h$ for $P$ in the cone $\mathscr{C}_2$, see [LSV99]. The measure $\mu = hm$ is stationary for the RDS; by Proposition 4.3.7, $h \ge D_\alpha > 0$.*
*Moreover, Bahsoun and Bose [BB16b, BB16a] have shown that there exists a unique absolutely continuous (with respect to the Lebesgue measure) stationary measure $\mu$, and $\nu\otimes\mu$ is mixing — so also ergodic.*

Using the same idea as in the proof of Theorem 4.4.1, we can obtain an annealed result for the random dynamical system. Note that $P_\mu$, the transfer operator with respect to the stationary measure $\mu$, satisfies $P_\mu\mathbf{1} = \mathbf{1}$ and so $\|P_\mu\varphi\|_\infty \le P_\mu(\|\varphi\|_\infty) = \|\varphi\|_\infty\|P_\mu\mathbf{1}\|_\infty = \|\varphi\|_\infty$. An easy calculation shows that $P_\mu(\varphi) = \frac{1}{h}P(h\varphi)$ where $h \in \mathscr{C}_2$ is the density of the invariant measure $\mu$ and hence $h \ge D_\alpha m(h)$ is bounded below. As before this observation allows us to bootstrap in some sense the $L^1(\mu)$ decay rate to $L^p(\mu)$ for $p \ge 1$, a technique used in [MN08, Mel09].

**Theorem 4.4.4** (Annealed LD). *Let $\varphi \in C^1([0,1])$ with $\mu(\varphi) = 0$ and let $0 < \alpha < 1$. Then the Birkhoff averages have annealed large deviations with respect to the measure $\nu \otimes \mu$ with rate*

$$(\nu \otimes \mu)\{(\omega, x) : \left|\sum_{j=1}^{n} \varphi \circ \mathcal{T}_\omega^j(x)\right| \ge n\epsilon\} \le C_{\alpha,p,\|\varphi\|_{C^1}} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

*for any $p > \max\{1, \frac{1}{\alpha} - 1\}$.*
*Note that the Birkhoff sums above* are not centered for a given realization *$\omega$, only on average over $\Sigma$.*

**Proof.** To prove this result we will use the construction used to prove the annealed CLT in [ANV15]: let $\Sigma_X := X^{\mathbb{N}_0}$, endowed with the $\sigma$-algebra $\mathscr{G}$ generated by the cylinders, and the left shift operator $\tau\colon \Sigma_X \to \Sigma_X$.
Denote by $\pi$ the projection from $\Sigma_X$ onto the 0-th coordinate, that is, $\pi(x) = x_0$ for $x = (x_0, x_1, \ldots)$. We can lift any observable $\varphi\colon X \to \mathbb{R}$ to an observable on $\Sigma_X$ by setting $\varphi_\pi := \varphi \circ \pi\colon \Sigma_X \to \mathbb{R}$.
Following [ANV15, §4], one can introduce a $\tau$-invariant probability measure $\mu_c$ on $\Sigma_X$ such that $\mathbb{E}_\mu(\varphi) = \mathbb{E}_{\mu_c}(\varphi_\pi)$, and the law of $S_n(\varphi)$ on $\Sigma \times X$ under $\nu \otimes \mu$ is the same as the law of the $n$-th Birkhoff sum of $\varphi_\pi$ on $\Sigma_X$ under $\mu_c$ and $\tau$; thus it suffices to establish large deviations for the latter.
Define now

$$H_n := \sum_{k=1}^{n} P_\mu^k(\varphi) : X \to \mathbb{R}$$

From the relation $P_\mu(.) = \frac{1}{h}P(.h)$, we have that $\|P_\mu^n(\varphi)\|_{L^1(\mu)} \le C_{\alpha,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{1/\alpha}$ because $\mu(\varphi) = 0$. We calculate $\mathbb{E}_\mu|P_\mu^i(\varphi)|^p = \mathbb{E}_\mu[|P_\mu^i(\varphi)|^{p-1}|P_\mu^i(\varphi)|] \le \|P_\mu^i(\varphi)\|_\infty^{p-1}\|P_\mu^i(\varphi)\|_{L^1(\mu)}$. Hence $\|P_\mu^k(\varphi)\|_{L^p(\mu)} \le Ck^{(1-1/\alpha)/p}(\log k)^{1/(p\alpha)}$ and thus $\|H_n\|_{L^p(\mu)}$ satisfies the bounds of Lemma 4.3.6.

We lift $\varphi$ and $H_n$ to $\Sigma_X$ and denote them by $\varphi_\pi$ and $H_{n,\pi}$ respectively, and define

$$\chi_n := \varphi_\pi + H_{n,\pi} - H_{n,\pi} \circ \tau : \Sigma_X \to \mathbb{R}.$$

We now continue as in the proof of Theorem 4.4.1, applying Rio's inequality. For $i = 1, \dots, n$ take the sequences $\{X_i = \varphi_\pi \circ \tau^{n-i}\}$, $\{Y_i = \chi_{n-i} \circ \tau^{n-i}\}$ and $\mathcal{G}_i = \tau^{-(n-i)}\mathcal{G}$. We have $\mathbb{E}_{\mu_c}[Y_i|\mathcal{G}_k] = 0$ for $k > i$ and so, for $p > \max\{1, \frac{1}{\alpha} - 1\}$,

$$b_{i,n} = \max_{i \le u \le n} \left\| X_i \sum_{k=i}^{u} \mathbb{E}_{\mu_c}(X_k|\mathcal{G}_i) \right\|_{L^p(\mu_c)} \le C n^{1+\frac{1}{p}(1-\frac{1}{\alpha})} (\log n)^{\frac{1}{p\alpha}}$$

which gives, as in Theorem 4.4.1,

$$\mu_c(|X_1 + \cdots + X_n|^{2p} > n^{2p}\epsilon^{2p}) \le C_{\alpha,\varphi,p} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

■

Using similar ideas, it is possible to obtain an annealed central limit theorem. This has been established already by Young Tower techniques in [BB16a, Theorem 3.2]. We include the statement of the annealed central limit and an alternative proof for completeness and to give an expression for the annealed variance.

**Proposition 4.4.5** (Annealed CLT). *If $\alpha < \frac{1}{2}$ and $\varphi \in C^1$ with $\mu(\varphi) = 0$ then a central limit theorem holds for $S_n\varphi$ on $\Sigma \times X$ with respect to the measure $\nu \otimes \mu$, that is, $\frac{1}{\sqrt{n}} S_n\varphi$ converges in distribution to $N(0, \sigma^2)$, with variance $\sigma^2$ given by*

$$\sigma^2 = -\mu(\varphi^2) + 2\sum_{k=0}^{\infty} \mu(\varphi U^k \varphi)$$

**Proof.** We will use the results of [ANV15, Section 4] and [Liv96, Theorem 1.1] (see Theorem 4.3.9). We proceed as in Theorem 4.4.4, using the averaged operators $U$ and $P$. As in [ANV15, Section 4], to $U$ corresponds a transition probability on $X$ given by $U(x, A) = \sum_\beta \{p_\beta : T_\beta x \in A\}$. The stationary measure $\mu$ is invariant under $U$. Extend $\mu$ to the unique probability measure $\mu_c$ on $\Sigma_X := X^{N_0} = \{\underline{x} = (x_0, x_1, x_2, \dots, x_n, \dots)\}$, endowed with the $\sigma$-algebra $\mathcal{G}$ given by cylinder sets, such corresponding to $\mu$ such that $\{x_n\}_{n\ge 0}$ is a Markov chain on $(\Sigma_X, \mathcal{G}, \mu_c)$ (where $x_n$ is the $n$-th coordinate of $\underline{x}$) induced by the random dynamical system. The left shift $\tau$ on $\Sigma_X$ preserves $\mu_c$. Given $\varphi : X \to \mathbb{R}$, $\mu(\varphi) = 0$, we define $\varphi_\pi$ on $\Sigma_X$ by $\varphi_\pi(x_0, x_1, x_2, \dots, x_n, \dots) := \varphi(x_0)$. As in [ANV15, Section 4], to prove the CLT for $S_n(\varphi)$ with respect to $\nu \otimes \mu$ on $\Sigma \times X$ it suffices to prove the CLT for the Birkhoff sum $\sum_{j=0}^{n} \varphi_\pi \circ \tau^k$ with respect to $\mu_c$ on $\Sigma_X$.

We introduce the Koopman operator $\widetilde{U}$ and transfer operator $\widetilde{P}$ for the map $\tau$ on the probability space $(\Sigma_X, \mathscr{G}, \mu_c)$. We define the decreasing sequence of $\sigma$-algebras $\mathscr{G}_k = \tau^{-k}\mathscr{G}$, and note that $\widetilde{P}$, $\widetilde{U}$ satisfy $\widetilde{P}^k\widetilde{U}^k f = f$ and $\widetilde{U}^k\widetilde{P}^k f = \mathbb{E}_{\mu_c}(f|\mathscr{G}_k)$ for every $\mu_c$-integrable $f$. We note that $\varphi_\pi \in L^\infty(\mu_c)$. As in [ANV15, Lemma 4.2] we have $\widetilde{P}^n(\varphi_\pi) = (P^n\varphi)_\pi$. Thus $\sum_{k=0}^\infty \widetilde{P}^k\varphi_\pi$ converges in $L^1(\mu_c)$ if $\alpha < \frac{1}{2}$ and therefore $\sum_{k=0}^\infty |\int \varphi_\pi \widetilde{U}^k\varphi_\pi d\mu_c| < \infty$. Thus the result for $\sum_{j=0}^n \varphi_\pi \circ \tau^k$ follows from [Liv96, Theorem 1.1]. The stated formula for $\sigma^2$ is also given in [Liv96, Theorem 1.1]. ∎

We will use the annealed and sequential results to obtain quenched large deviations for random systems of intermittent maps. We denote the Birkhoff sums by $S_{n,\omega}(x)$ to stress the dependence on the realization $\omega$.

**Theorem 4.4.6** (Quenched LD)**.** *Suppose $\varphi \in C^1$ and $\mu(\varphi) = 0$. Fix $0 < \alpha < 1$. Then, given $p > \max\{1, \frac{1}{\alpha} - 1\}$ and $\kappa := \lceil \frac{4p}{1-\alpha} \rceil$ for $\nu$-almost every realization $\omega \in \Sigma$ the Birkhoff averages have large deviations with polynomial rate,* even without centering*: there is an $N(\omega)$ such that for each $\epsilon > 0$*

$$m\{x : S_{n,\omega}\varphi > 4n\epsilon\} \le C_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-\kappa} \text{ for } n \ge N(\omega).$$

*Note that the Birkhoff sums $S_{n,\omega}\varphi$ above* are not centered with respect to the realization *$\omega$, only on average over $\Sigma$.*

**Remark 4.4.7.** *The point of the above Theorem, compared to the sequential Theorem 4.4.1, is that for almost each realization the large deviation estimates hold* even without centering. *That is, the contribution of the means (with respect to the measure $m$ on $X$) can be ignored for almost each realization $\omega$.*

**Proof of Theorem 4.4.6** Choose $p > \max\{1, \frac{1}{\alpha} - 1\}$ and $\epsilon > 0$. By Theorem 4.4.1, for all $\omega \in \Sigma$,

$$m\left\{x : \left|\frac{1}{n}S_{n,\omega}\varphi(x) - \frac{1}{n}\sum_{j=1}^n m(\varphi \circ T_\omega^j)\right| \ge \epsilon\right\} \le C_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

with $C_{\alpha,\varphi,\delta}$ independent of $\omega$. Integrating over $\Sigma$ with respect to $\nu$ we obtain

$$\nu \otimes m\left\{(\omega, x) : \left|\frac{1}{n}S_{n,\omega}\varphi(x) - \frac{1}{n}\sum_{j=1}^n m(\varphi \circ T_\omega^j)\right| \ge \epsilon\right\} \le C_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

By Theorem 4.4.4, we also have the annealed estimate for the non-centered sums:

$$\nu \otimes m\left\{(\omega, x) : \left|\frac{1}{n}S_{n,\omega}\varphi(x)\right| \ge \epsilon\right\} \le C_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

Theorem 4.4.4 refers to the measure $\nu \otimes \mu$ but since $\frac{dm}{d\mu} = \frac{1}{h} \le \frac{1}{D_\alpha}$, the large deviations estimate applies also to $\nu \otimes m$. Observe now that

$$\left\{(\omega,x) : \left|\frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| > 2\epsilon\right\}$$

$$\subset \left\{(\omega,x) : \left|\frac{1}{n}S_{n,\omega}\varphi(x)\right| < \epsilon, \left|\frac{1}{n}S_{n,\omega}\varphi(x) - \frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| \ge \epsilon\right\}$$

$$\bigcup \left\{(\omega,x) : \left|\frac{1}{n}S_{n,\omega}\varphi(x)\right| > \epsilon\right\}.$$

Thus

$$\nu \otimes m \left\{(\omega,x) : \left|\frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| > 2\epsilon\right\} \le K_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p}$$

and, as there is no dependence on $x \in X$, this means

$$\nu \left\{\omega : \left|\frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| > 2\epsilon\right\} \le K_{\alpha,p,\varphi} n^{1-\frac{1}{\alpha}} (\log n)^{\frac{1}{\alpha}} \epsilon^{-2p} \tag{4.14}$$

Denote $\beta := \frac{1}{\alpha} - 1 > 0$.

The proof we give does not give an optimal value of $\kappa$. In the case $\beta > 1$ a simpler proof may be given but the resulting exponent $\kappa$ is also not optimal and no better than the estimate we give.

Let $\tau = \frac{2}{\beta}$ and $\delta > 0$ small. Choose $\gamma = \frac{1}{2p}(\beta - \frac{1}{\tau}) - \delta = \frac{\beta}{4p} - \delta$ and $\kappa = \lceil (1+\beta^{-1})(4p) \rceil = \lceil \frac{4p}{1-\alpha} \rceil$. Then $(2p\gamma - \beta)\tau < -1$ and $\gamma\kappa > \beta$ for $\delta > 0$ small enough.

For $\epsilon = n^{-\gamma}$ the bound (4.14) becomes

$$\nu \left\{\omega : \left|\frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| > 2n^{-\gamma}\right\} \le K_{\alpha,p,\varphi} n^{2p\gamma} n^{-\beta} (\log n)^{\frac{1}{\alpha}}$$

Consider the subsequence $n_k := k^\tau$. As $(2p\gamma - \beta)\tau < -1$, for $\nu$ almost every $\omega$ there exists an $N(\omega)$ such that for all $n_k > N(\omega)$,

$$\left|\frac{1}{n_k}\sum_{j=1}^{n_k} m(\varphi \circ T_\omega^j)\right| \le 2n_k^{-\gamma}$$

If $n_k \le n < n_{k+1}$ then

$$\left|\frac{1}{n}\sum_{j=1}^{n} m(\varphi \circ T_\omega^j)\right| \le \frac{1}{n_k}\left|\sum_{j=1}^{n_k} m(\varphi \circ T_\omega^j) + \sum_{j=n_k+1}^{n} m(\varphi \circ T_\omega^j)\right|$$

$$\le 2n_k^{-\gamma} + \frac{\|\varphi\|_\infty}{n_k}|n_{k+1} - n_k|$$

There is $K>0$, independent of $\omega$, depending only on $\tau$, $\gamma$ and $\|\varphi\|_\infty$, such that

$$2n_k^{-\gamma}+\frac{\|\varphi\|_\infty}{n_k}|n_{k+1}-n_k|<3n^{-\gamma}\qquad \text{if } k\geq K.$$

Indeed, $\lim_{k\to\infty}\frac{n_{k+1}}{n_k}=1$, $\frac{1}{n_k}|n_{k+1}-n_k|=O(\frac{1}{k})$, $\frac{1}{k}=O(\frac{1}{n^{1/\tau}})$ and $n^{-1/\tau}<n^{-\gamma}$ because $1/\tau>\gamma$. Increase $N(\omega)$ such that $n>N(\omega)$ implies $n\geq K^\tau$ and $C_{\alpha,p,\varphi}n^{\gamma\kappa-\beta}(\log n)^{1/\alpha}>1$.
We will show that for $n>N(\omega)$

$$m\left(x:\left|\frac{1}{n}S_{n,\omega}\varphi(x)\right|\geq 4\epsilon\right)\leq C_{\alpha,p,\varphi}\epsilon^{-\kappa}n^{-\beta}(\log n)^{1/\alpha}.$$

Suppose $\epsilon<n^{-\gamma}$. Then $C_{\alpha,p,\varphi}\epsilon^{-\kappa}n^{-\beta}(\log n)^{1/\alpha}\geq C_{\alpha,p,\varphi}n^{\gamma\kappa-\beta}(\log n)^{1/\alpha}>1$ and there is nothing to prove.
If $\epsilon\geq n^{-\gamma}$ and $n>N(\omega)$ then, as $|\frac{1}{n}\sum_{j=1}^n m(\varphi\circ T_\omega^j)|<3\epsilon$,

$$\left\{x:\left|\frac{1}{n}S_{n,\omega}\varphi(x)\right|\geq 4\epsilon\right\}\subset\left\{x:\left|\frac{1}{n}S_{n,\omega}\varphi(x)-\frac{1}{n}\sum_{j=1}^n m(\varphi\circ T_\omega^j)\right|\geq\epsilon\right\}$$

Hence the result holds by Theorem 4.4.1, as

$$m\left(x:\left|\frac{1}{n}S_{n,\omega}\varphi(x)-\frac{1}{n}\sum_{j=1}^n m\left(\varphi\circ T_\omega^j\right)\right|\geq\epsilon\right)\leq C_{\alpha,p,\varphi}\epsilon^{-2p}n^{-\beta}(\log n)^{1/\alpha}$$

and $2p<\kappa$. ■

We remark that the methods used to prove these results in the uniformly expanding case are not applicable here, as they rely on the quasi-compactness of the transfer operator. In the uniformly expanding case, which has exponential large deviations for Hölder observables, it is possible to obtain a rate function.

## 4.5 The Role of Centering in the Quenched CLT for RDS

In this section we discuss two results: Proposition 4.5.1, that the quenched variance is the same for almost all realizations $\omega\in\Sigma$, and Theorem 4.5.5, that generically one must center the observations in order to obtain a CLT (as opposed to LD Theorem 4.4.6, where centering did not affect the quenched LD). Note that these hinge on the rate of growth of the mean of the Birkhoff sums; we see that it is $o(n)$ but not $o(\sqrt{n})$. We use the recent paper by Hella and Stenlund [HS20] to extend and clarify results of [NTV18].

In [NTV18, Theorem 3.1] a self-norming quenched CLT is obtained for $\nu$-a.e. realization $\omega$ of the random dynamical system of Theorem 4.4.4. More precisely, recalling the definition of the centered observables $[\varphi]_k(\omega,x) = \varphi(x) - m(\varphi \circ \mathcal{T}_\omega^k)$ and $\sigma_n^2(\omega) := \int \left[\sum_{k=1}^n [\varphi]_k(\omega, \mathcal{T}_\omega^k x)\right]^2 dx$ it is shown that $\frac{1}{\sigma_n(\omega)} \sum_{k=1}^n [\varphi]_k(\omega,\cdot) \circ \mathcal{T}_\omega^k \to N(0,1)$ provided $\sigma_n^2 \approx n^\beta$, with $\alpha < \frac{1}{9}$ and $\beta > \frac{1}{2(1-2\alpha)}$. Various scenarios under which $\sigma_n^2(\omega) > n^\beta$ are given in [NTV18]. See also [HL19].

If the maps $T_{\omega_i}$ preserved the same invariant measure then it suffices to consider observables with mean zero, since the mean would be the same along each realization. In the setting of [ALS09] this is the case, namely all realizations preserve Haar measure, and the authors address the issue of whether the variance $\sigma_n^2(\omega)$ can be taken to be the "same" for almost every quenched realization in the setting of random toral automorphisms. They show that for almost every quenched realization the variance in the quenched CLT may be taken as a uniform constant. The technique they use is adapted from random walks in random environments and consists in analyzing a random dynamical system on a product space.

A natural question is whether in our setup of random intermittent maps, after centering, $\sigma_n(\omega)$ can be taken to be "uniform" over $\nu$-a.e. realization. Recent results of Hella and Stenlund [HS20] give conditions under which $\frac{1}{n}\sigma_n^2(\omega) \to \sigma^2$ for $\nu$-a.e. $\omega$, as well as information about rates of convergence. Note that this is also true in the context of uniformly expanding maps considered by [AA16] using the same method used in [HS20].

A related question is whether we need to center at all. For example, if $\mu(\varphi) = 0$, where $\mu$ is the stationary measure on $X$, then for $\nu$-a.e $\omega$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n \left[\varphi(\mathcal{T}_\omega^j x) - m(\varphi(\mathcal{T}_\omega^j))\right] \to 0 \qquad \text{for } \mu\text{-a.e. } x$$

by the ergodicity of $\nu \otimes \mu$, but also

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^n m\left(\varphi(\mathcal{T}_\omega^j)\right) \to 0 \qquad \text{for } \nu\text{-a.e. } \omega,$$

by the proof of Theorem 4.4.6. So for the strong law of large numbers centering is not necessary. Using ideas of [AA16] we consider the related question of whether centering is necessary to obtain a quenched CLT with almost surely constant variance. We show the answer to this is positive: to obtain an almost surely constant variance in the quenched CLT we need to center.

### 4.5.1 Non-random quenched variance

For Proposition 4.5.1, we verify that our system satisfies the conditions SA1, SA2, SA3 and SA4 of [HS20]; then, by [HS20, Theorem 4.1], the quenched variance is almost surely the same, equal to the annealed variance.

**Proposition 4.5.1.** *Let $\alpha < \frac{1}{2}$, $\varphi \in C^1$ and define the annealed variance*

$$\sigma^2 := \lim_{n\to\infty} \frac{1}{n} \left\| [S_n]^{\nu\otimes m} \right\|^2_{L^2(\nu\otimes m)} = \lim_{n\to\infty} \frac{1}{n} \left\| S_n - \int_{\Sigma\times X} S_n d\, \nu \otimes m \right\|^2_{L^2(\nu\otimes m)}$$
$$= \sum_{k=0}^{\infty} (2-\delta_{0k}) \lim_{i\to\infty} \int_{\Sigma} \left[ m(\varphi_i \varphi_{i+k}) - m(\varphi_i) m(\varphi_{i+k}) \right] d\nu$$

*If $\sigma^2 > 0$ then for $\nu$-a.e. $\omega$*

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left[ \varphi(\mathcal{T}^j_\omega \cdot) \right]^m \to^d N(0, \sigma^2)$$

*in distribution with respect to $m$.*

**Remark 4.5.2.** *Proposition 4.4.5 shows that the annealed CLT holds for $\alpha < \frac{1}{2}$ and under the usual genericity conditions the annealed variance satisfies $\sigma^2 > 0$. Thus Proposition 4.5.1 extends [NTV18, Theorem 5.3] from the parameter range $\alpha < \frac{1}{9}$ to $\alpha < \frac{1}{2}$. Note that [HL19], proved the CLT for $\alpha < \frac{1}{3}$.*

**Proof of Proposition 4.5.1** We will verify conditions SA1, SA2, SA3 and SA4 of [HS20, Theorem 4.1] in our setting, with $\eta(k) = Ck^{-\frac{1}{\alpha}+1}(\log k)^{\frac{1}{\alpha}}$ in the notation of [HS20].
**SA1**: If $j > i$ then

$$\left| \int \varphi \circ \mathcal{T}^i_\omega(x) \varphi \circ \mathcal{T}^j_\omega(x) dm - \int \varphi \circ \mathcal{T}^i_\omega(x) dm \int \varphi \circ \mathcal{T}^j_\omega(x) dm \right|$$

$$= \left| \int \varphi \circ \mathcal{T}^{j-i+1}_\omega(\mathcal{T}^i_\omega x) \varphi(x) P^i_\omega \mathbf{1} dm - \int \varphi \mathcal{P}^i_\omega \mathbf{1} dm \int \varphi(x) \mathcal{P}^j_\omega \mathbf{1} dm \right| \le C(j-i)^{-\frac{1}{\alpha}+1}(\log(j-i))^{\frac{1}{\alpha}}$$

by the same argument as in the proof of [NTV18, Proposition 1.3].
**SA2**: Our underlying shift $\sigma : \Sigma \to \Sigma$ is Bernoulli hence $\alpha$-mixing.
**SA3**: We need to check [HS20, equations (4), (5)] that

$$\left| \int \varphi(T_{\omega_k} T_{\omega_{k-1}} \cdots T_{\omega_1} x) dm - \int \varphi(T_{\omega_k} T_{\omega_{k-1}} \cdots T_{\omega_{r+1}} x) dm \right| \le C\eta(k-r).$$

This follows since

$$\left| \int \varphi(T_{\omega_k} T_{\omega_{k-1}} \cdots T_{\omega_1} x) dm - \int \varphi(T_{\omega_k} T_{\omega_{k-1}} \cdots T_{\omega_{r+1}} x) dm \right|$$

$$= \left| \int \varphi(x) P_{\omega_k} P_{\omega_{k-1}} \cdots P_{\omega_1} \mathbf{1} dm - \int \varphi(x) P_{\omega_k} P_{\omega_{k-1}} \cdots P_{\omega_{r+1}} \mathbf{1} dm \right|$$

$$\leq \|\varphi\|_\infty \left\| P_{\omega_k} P_{\omega_{k-1}} \cdots P_{\omega_{r+1}} [\mathbf{1} - P_{\omega_r} \cdots P_{\omega_1} \mathbf{1}] \right\|_{L^1}$$

Since $\mathbf{1}$ and $P_{\omega_r} \cdots P_{\omega_1} \mathbf{1}$ both lie in the cone and have the same $m$-mean, we have

$$\left\| P_{\omega_k} P_{\omega_{k-1}} \cdots P_{\omega_{r+1}} [\mathbf{1} - P_{\omega_r} \cdots P_{\omega_1} \mathbf{1}] \right\|_{L^1} \leq C(k-r)^{-\frac{1}{\alpha}+1} (\log(k-r))^{\frac{1}{\alpha}}$$

by [NTV18, Theorem 1.2].

**SA4**: $(\sigma, \Sigma, \nu)$ is stationary so SA4 is automatic. ■

### 4.5.2 Centering is generically needed in the CLT

Now we address the question of the necessity of centering in the quenched central limit theorem. We show that if $\int \varphi d\mu_{\beta_i} \neq \int \varphi d\mu_{\beta_j}$ for two maps $T_{\beta_i}$, $T_{\beta_j}$, where $\mu_{\beta_i}$ is the invariant measure of $T_{\beta_i}$, then centering is needed: although

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \left[ \varphi(\mathcal{T}_\omega^j) - m(\varphi(\mathcal{T}_\omega^j)) \right] \to^d N(0, \sigma^2)$$

for $\nu$-a.e. $\omega$, it is not the case that

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \varphi(\mathcal{T}_\omega^j) \to^d N(0, \sigma^2)$$

for $\nu$-a.e. $\omega$.

Our proof has the same outline as that of [AA16], adapted to our setting of polynomial decay of correlations. First we suppose that the maps $T_{\beta_i}$ do not preserve the same measure. After reindexing we can suppose that $T_{\beta_1}$ and $T_{\beta_2}$ have different invariant measures and that $\int \varphi d\mu_{\beta_1} \neq \int \varphi d\mu_{\beta_2}$, a condition satisfied by an open and dense set of observables. Recall that the RDS has the stationary measure $d\mu = h dm$, $h \geq D_\alpha > 0$ and we have assumed $\mu(\varphi) = 0$, $\varphi \in \mathscr{C}^1$.

Our proof can be summarized in the following steps:

- First, construct a product random dynamical system on $X \times X$ and prove that it satisfies an annealed CLT for $\tilde{\varphi}(x, y) = \varphi(x) - \varphi(y)$ with distribution $N(0, \tilde{\sigma}^2)$;
- then, observe that almost every uncentered quenched CLT has the same variance only if $2\sigma^2 = \tilde{\sigma}^2$, where the original RDS with stationary measure $d\mu = h dm$ satisfies an annealed CLT for $\varphi$ with distribution $N(0, \sigma^2)$;

- next, observe that the conclusions of [AA16, Theorem 9] hold in our setting and $\tilde{\sigma}^2 = 2\sigma^2$ if and only if $\lim_{n\to\infty}\frac{1}{n}\int_\Sigma \left(\sum_{k=1}^{n-1}\int_X \varphi\circ\mathcal{T}_\omega^k h dm\right)^2 d\nu = 0$;
- finally, using ideas of [AA16], we show the limit above is zero only if a certain function $G$ on $\Sigma$ is a Hölder coboundary, which in turn implies $\int\varphi d\mu_{\beta_1} = \int\varphi d\mu_{\beta_2}$, a contradiction.

Let $\varphi\colon X\to\mathbb{R}$ be $\mathscr{C}^1$, with $\int_X \varphi d\mu = 0$, and define $S_n(\varphi) = \sum_{k=0}^{n-1}\varphi(\mathcal{T}_\omega^k x)$ on $\Sigma\times X$. Recall the standard expression (e.g. see [AA16]) for the annealed variance,

$$\sigma^2 = \lim_{n\to\infty}\frac{1}{n}\int_\Sigma\int_X [S_n(\varphi)]^2\, d\mu d\nu.$$

We also consider the product random dynamical system $(\widetilde{\Sigma} := \Sigma\times X\times X, \tilde{\nu} := \nu\otimes\mu\otimes\mu, \tilde{T})$ defined on $X^2$ by $\tilde{T}_\omega(x,y) = (T_\omega x, T_\omega y)$. For an observable $\varphi$, define $\tilde{\varphi}\colon X^2\to\mathbb{R}$ by $\tilde{\varphi}(x,y) = \varphi(x)-\varphi(y)$, and its Birkhoff sums $S_n(\tilde{\varphi})$. In Theorem 4.5.3 and Corollary 4.5.4 we show $\frac{1}{\sqrt{n}}\sum_{j=1}^n \tilde{\varphi}\circ\tilde{T}^j \to^d N(0,\tilde{\sigma}^2)$ with respect to $\nu\otimes\mu\otimes\mu$ for some $\tilde{\sigma}^2\geq 0$.
The following lemma from [ANV15] is general and does not depend upon the underlying dynamics. It is a consequence of Levy's continuity theorem (Theorem 6.5 in [Kar93]).

**Lemma** ([ANV15, Lemma 7.2])**.** *Assume that $\sigma^2>0$ and $\tilde{\sigma}^2>0$ are such that*

1. $\frac{S_n(\varphi)}{\sqrt{n}}$ *converges in distribution to $N(0,\sigma^2)$ under the probability $\nu\otimes\mu$,*
2. $\frac{S_n(\tilde{\varphi})}{\sqrt{n}}$ *converges in distribution to $N(0,\tilde{\sigma}^2)$ under the probability $\nu\otimes\mu\otimes\mu$,*
3. $\frac{S_{n,\omega}(\varphi)}{\sqrt{n}}$ *converges in distribution to $N(0,\sigma^2)$ under the probability $\mu$, for $\nu$ almost every $\omega$.*

*Then $2\sigma^2 = \tilde{\sigma}^2$.*

We will show that the system $\widetilde{F}(\omega,x,y) = (\tau\omega, T_{\omega_1}x, T_{\omega_1}y)$ with respect to the measure $\nu\otimes\mu^2$ on $\Sigma\times[0,1]^2$ (recall that $\nu := \mathbb{P}^{\otimes\mathbb{N}}$ and $\mu$ is a stationary measure of the RDS) has summable decay of correlations in $L^2$ for $\alpha<\frac{1}{2}$, and as a corollary it satisfies the CLT.

**Theorem 4.5.3.** *Suppose that for $\omega\in\Sigma$, $h = \frac{d\mu}{dm}\in\mathscr{C}_2$ and each $\varphi\in\mathscr{C}^1$ with $m(\varphi h) = 0$*

$$\|P_{\omega_n}\dots P_{\omega_1}(\varphi h)\|_{L^1(m)} \leq C\rho(n)(\|\varphi\|_{\mathscr{C}^1}+m(h)).$$

*Then there is a constant $\widetilde{C}$, independent of $\omega$, such for each $\psi\in\mathscr{C}^1(X\times X)$ and $\varphi\in L^\infty(X\times X)$ with $(\mu\otimes\mu)(\psi) = 0$, one has*

$$\left|\int\varphi(\mathcal{T}_\omega^n x, \mathcal{T}_\omega^n x)\psi(x,y)d\mu(x)d\mu(y)\right| \leq \widetilde{C}\rho(n)\|\varphi\|_{L^\infty}(\|\psi\|_{\mathscr{C}^1}+1)$$

**Proof.** Since $X \times X$ is compact, $\psi$ is uniformly $\mathscr{C}^1$ in both variables in the sense that $\psi(x_0, y)$ is uniformly $\mathscr{C}^1$ for each $x_0$ and similarly for $\psi(x, y_0)$. We want to estimate

$$I := \int \varphi(\mathcal{T}_\omega^n x, \mathcal{T}_\omega^n y)\psi(x,y)d\mu(x)d\mu(y).$$

Define

$$\overline{\psi}(x) := \int \psi(x,y)d\mu(y), \qquad h_x(y) := \psi(x,y) - \overline{\psi}(x).$$

Then $\overline{\psi}, h_x \in \mathscr{C}^1(X)$, with $\mathscr{C}^1$-norms bounded by $2\|\psi\|_{\mathscr{C}^1}$, uniformly with respect to $x$. We can write $I$ as

$$I = \underbrace{\int \varphi(\mathcal{T}_\omega^n x, \mathcal{T}_\omega^n y)\left[\psi(x,y) - \overline{\psi}(x,y)\right]d\mu(x)d\mu(y)}_{:=I_1}$$
$$+ \underbrace{\int \varphi(\mathcal{T}_\omega^n x, \mathcal{T}_\omega^n y)\overline{\psi}(x,y)d\mu(x)d\mu(y)}_{:=I_2}.$$

Define now $g_{\omega,x}(y) := \varphi(\mathcal{T}_\omega^n x, y)$. Then (note that $\int h_x(y)h(y)dm(y) = 0$)

$$\begin{aligned}
|I_1| = \left|\int\left(\int g_{\omega,x}(\mathcal{T}_\omega^n y)h_x(y)dm(y)\right)d\mu(x)\right| &= \left|\int\left(\int g_{\omega,x}(y)\mathscr{P}_\omega^n(h_x(y)h(y))dm(y)\right)d\mu(x)\right| \\
&\le \|\varphi\|_{L^\infty}\sup_x \left\|\mathscr{P}_\omega^n(h_x(y)h(y))\right\|_{L^1(m(y))} \\
&\le C'\|\varphi\|_{L^\infty}(\|\psi\|_{\mathscr{C}^1} + m(h))\rho(n).
\end{aligned}$$

by the hypothesis.
Similarly, define $k_{\omega,y}(x) := \varphi(x, \mathcal{T}_\omega^n y)$ so then (again, $\int \overline{\psi}(x)h(x)dm(x) = 0$)

$$\begin{aligned}
|I_2| &= \left|\int\left(\int k_{\omega,y}(\mathcal{T}_\omega^n x)\overline{\psi}(x)d\mu(x)\right)d\mu(y)\right| \\
&= \left|\int\left(\int k_{\omega,y}(x)\mathscr{P}_\omega^n(\overline{\psi}(x)h(x))dm(x)\right)d\mu(y)\right| \\
&\le \|\varphi\|_{L^\infty}\left\|\mathscr{P}_\omega^n(\overline{\psi}(x)h(x))\right\|_{L^1(m(x))} \\
&\le C'\|\varphi\|_{L^\infty}(\|\psi\|_{\mathscr{C}^1} + m(h))\rho(n).
\end{aligned}$$

These imply that $|I| \le 2C'\|\varphi\|_{L^\infty}(\|\psi\|_{\mathscr{C}^1} + m(h))\rho(n)$. ∎

**Corollary 4.5.4.** *Under the assumptions of Theorem 4.5.3, for $\psi \in \mathscr{C}^1(X \times X)$ with $(\mu \otimes \mu)(\psi) = 0$, $\frac{1}{\sqrt{n}}\sum_{k=1}^n \psi \circ \widetilde{F}^k(\omega, x, y)$ satisfies a CLT with respect to $\nu \otimes \mu \otimes \mu$, that is*

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n \psi \circ \widetilde{F}^k(\omega, x, y) \to^d N(0, \widetilde{\sigma}^2)$$

*in distribution for some $\widetilde{\sigma}^2 \ge 0$.*

**Proof.** Let $Q$ be the adjoint of $\tilde{F}(\omega,x,y) = (\sigma\omega, T_{\omega_1}x, T_{\omega_1}y)$ with respect to the invariant measure $\nu\otimes\mu\otimes\mu$ on $\Sigma\times X^2$ so that

$$\int \varphi\circ\tilde{F}(\omega,x,y)\psi(\omega,x,y)d\mu(x)d\mu(y)d\nu(\omega) = \int \varphi(\omega,x,y)(Q\psi)(\omega,x,y)d\mu(x)d\mu(y)d\nu(\omega).$$

for $\varphi\in L^\infty(\Sigma\times X\times X)$. Iterating we have

$$\int \varphi\circ\tilde{F}^n(\omega,x,y)\psi(\omega,x,y)d\mu(x)d\mu(y)d\nu(\omega) = \int \varphi(\omega,x,y)(Q^n\psi)(\omega,x,y)d\mu(x)d\mu(y)d\nu(\omega).$$

Taking $\varphi = \mathrm{sign}\,(Q^n\psi)$, we see from Theorem 4.5.3 that $\|Q^n\psi\|_{L^1}\le C'\rho(n)$.
The proof now follows, as in Proposition 4.4.5, from [Liv96, Theorem 1.1] (see Theorem 4.3.9). ■

Suppose two of the maps $T_{\beta_1}$ and $T_{\beta_2}$ have different invariant measures. It is possible to find a $\mathscr{C}^1$ $\varphi$ such that $\int\varphi d\mu_{\beta_1}\neq\int\varphi d\mu_{\beta_2}$. In fact, $\int\varphi d\mu_{\beta_1}\neq\int\varphi d\mu_{\beta_2}$ for a $\mathscr{C}^2$ open and dense set of $\varphi$.

**Theorem 4.5.5.** *Let $\varphi\in C^1$ with $\mu(\varphi)=0$ and suppose that $\int\varphi\, d\mu_{\beta_1}\neq\int\varphi\, d\mu_{\beta_2}$. Then it is not the case that*

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\varphi(\mathscr{T}_\omega^j.)\to N(0,\sigma^2)$$

*for almost every $\omega\in\Sigma$. Hence, the Birkhoff sums need to be centered along each realization.*

**Proof.** We follow the counterexample method of [AA16, Section 4.3]. We show that in the uncentered case $2\sigma^2\neq\tilde{\sigma}^2$. To do this we use [AA16, Theorem 9] which holds in our setting, namely $\tilde{\sigma}^2 = 2\sigma^2$ if and only if

$$\lim_{n\to\infty}\int_\Sigma\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n-1}\int_X\varphi P_{\omega_k}\dots P_{\omega_n}(h)dm\right)^2 d\nu = 0 \tag{4.15}$$

(as in [AA16, Section 4.3] we change the time direction and replace $(\omega_1,\omega_2,\dots,\omega_n)$ by $(\omega_n,\omega_2,\dots,\omega_1)$; this does not affect integrals with respect to $\nu$ over finitely many symbols). Note that the sequence $P_{\omega_1}P_{\omega_2}\dots P_{\omega_n}h$ is Cauchy in $L^1$, as $\alpha<\frac{1}{2}$ and

$$\|P_{\omega_1}P_{\omega_2}\dots P_{\omega_n}(h) - P_{\omega_1}P_{\omega_2}\dots P_{\omega_n}\dots P_{\omega_{n+k}}(h)\|_1\le Cn^{-\frac{1}{\alpha}+1}(\log n)^{\frac{1}{\alpha}}$$

by Proposition 4.3.3. Thus $P_{\omega_1}P_{\omega_2}\dots P_{\omega_n}h\to h_\omega$ in $L^1$ for some $h_\omega\in\mathscr{C}_2$. This limit defines $h_\omega$, in terms of $\bar{\omega} := (\dots,\omega_n,\omega_2,\dots,\omega_1)$, i.e. $\omega$ reversed in time. We define $G(\omega) :=$

$\int_X \varphi h_\omega dm$. Note also that $\|P_{\omega_1}P_{\omega_2}\dots P_{\omega_n}h - h_\omega\|_1 \le Cn^{-1-\delta}$ for some $\delta > 0$, uniformly for $\omega \in \Sigma$. Hence

$$\int_\Sigma \left(\sum_{k=1}^{n-1} \frac{1}{\sqrt{n}} \int_X \varphi P_{\omega_k} \dots P_{\omega_n} h dm\right)^2 d\nu$$
$$= \int_\Sigma \left(\sum_{k=1}^{n-1} \frac{1}{\sqrt{n}} \left(\int_X \varphi h_{\tau^k \omega}\, dm + O\left(\sum_{k=1}^{n-1} \frac{1}{(n-k)^{1+\delta}}\right)\right)\right)^2 d\nu$$

which gives, using (4.15), that

$$\lim_{n\to\infty} \int_\Sigma \left(\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n-1} G(\tau^k\omega)\right)\right)^2 d\nu = 0. \tag{4.16}$$

We put a metric on $\Sigma$ by defining $d(\omega,\omega^{'}) = s(\omega,\omega^{'})^{-1-\frac{\epsilon}{2}}$ where $s(\omega,\omega^{'}) = \inf\{n : \omega_n \neq \omega^{'}_n\}$. With this metric $\Sigma$ is a compact and complete metric space. Note that $\|h_\omega - h_{\omega'}\|_{L^1} \le Cs(\omega,\omega^{'})^{-\frac{\epsilon}{2}}$ hence $G(\omega)$ is Hölder with respect to our metric.
As in the Abdulkader-Aimino counterexample, (4.16) implies that $G = H - H\circ\tau$ for a Hölder function $H$ on the Bernoulli shift $(\tau,\Sigma,\nu)$: by [Liv96, Theorem 1.1] (see Theorem 4.3.9) $G$ is a measurable coboundary, and therefore a Hölder coboundary, by the standard Livšic regularity theorem (see for instance [VO16, Section 12.2]). Now consider the points $\beta_1^* := (\beta_1,\beta_1,\cdots)$ and $\beta_2^* := (\beta_2,\beta_2,\cdots)$ in $\Sigma$; they are fixed points for $\tau$, and correspond to choosing only the map $T_{\beta_1}$, respectively only the map $T_{\beta_2}$. This implies $G(\beta_1^*) = G(\beta_2^*) = 0$ which in turn implies $\int \varphi d\mu_{\beta_1} = \int \varphi d\mu_{\beta_2}$, a contradiction. ■

CHAPTER 5

# STATISTICAL PROPERTIES OF KÄENMÄKI MEASURES

## 5.1 Introduction

In this chapter we prove some results concerning the statistical properties of Käenmäki measures. This is part of an ongoing project at the time of submission of this thesis. These measures provide the right generalization of Gibbs measures (see definition 2.5.1) for self-affine systems when studying the dimension of the corresponding attractors. While we introduce the theory in a general setting, we will work in a particular case of IFS consisting of matrices which are diagonal or anti-diagonal. In this case, the Käenmäki measures measures admit an explicit description in terms of Gibbs measures on an associated graph directed system. We prove that such measures are not mixing (theorem 5.3.4), and that they satisfy a 0-1 law for the measure of shrinking targets associated to cylinders (theorem 5.3.7).

## 5.2 General theory

Recall the definition of self-affine system from chapter 2: a set of contractions $\{S_i : X \to X, i = 1, \dots, m\}$ on a closed set $X \subset \mathbb{R}^n$ such that $S_i(x) = A_i x + b_i$ for linear transformations $A_i$. In this setting, there exists a set $E$ called the attractor of the self-affine system, such that

$$E = \bigcup_{i=1}^{m} S_i(E).$$

A historical problem is the investigation of the fractal properties of the set $E$ for different classes of self-affine systems. When the IFS consists of similarities, then the problem becomes much easier: if the maps $S_i$ are such that

$$|S_i(x) - S_i(y)| = r_i |x - y|$$

for all $x, y \in X$. It is possible to compute the dimension of the attractor of an IFS consisting of similarities if the attractor $F$ can be written as

$$F = \bigcup_{i=1}^{m} S_i(F),$$

where the union is *disjoint enough*. We make this notion more precise:

**Definition 5.2.1.** *We say that the IFS $\{S_1, \dots, S_m\}$ satisfies the* open set condition *(OSC) if there exists a non-empty open set $V$ such that*

$$V \subset \bigcup_{i=1}^{m} S_i(V),$$

*where the union is disjoint.*

Under this condition it is possible to compute the Hausdorff dimension of the attractor (Theorem 3(i) in [Hut81]).

**Theorem 5.2.2.** *Suppose that $\{S_1, \dots, S_m\}$ is an IFS consisting of similarities satisfying the OSC. Then the Hausdorff dimension of $F$ is equal to $s$, where $s$ is given by the only solution of*

$$\sum_{i=1}^{m} r_i^s = 1.$$

**Example 5.2.3.** *For the Cantor set, the similarity ratios are $r_1 = r_2 = \frac{1}{3}$. The OSC holds with $V = (0,1)$. Thus the Hausdorff dimension of the Cantor set is $s = \dim_H F = \frac{\log 2}{\log 3}$.*

The techniques used to prove this have become standard in the conformal setting and resemble some of the ideas used in chapter 3, in particular, the use of symbolic coding, bounded distortion properties and Birkhoff and Shannon-McMillan-Breiman theorems. These techniques rely heavily on the fact that for similarities (or even conformal maps), we see the same rate of contraction in all directions. When we study self-affine sets, this is not the case, as we may have directions where the contraction is much stronger than others.

Given $n \times n$ non-singular matrix $A$ we denote by $1 > \alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n > 0$ the singular values of $A$, defined as the lengths of the semi-axes of the ellipsoid $T(B)$ where $B = B_0(1) = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$.

**Definition 5.2.4.** *Let $0 \le s \le n$ and $A$ be a non-singular $n \times n$ matrix. The* singular value *function is defined by*

$$\varphi^s(A) = \alpha_1 \alpha_2 \cdots \alpha_{r-1} \alpha_r^{s-r+1},$$

*where $r$ is the integer satisfying $r-1 < s \le r$.*

One of the most important properties of $\varphi^s$ is that it is *submultiplicative*:

**Lemma 5.2.5.** *If $A, B$ are non-singular $n \times n$ matrices, then we have that $\varphi^s(AB) \le \varphi^s(A)\varphi^s(B)$.*

**Proof.** This is lemma 2.1 from [Fal88]. ∎

Consider now an IFS $\{S_1, \dots, S_m\}$ consisting of affine transformations $S_i(x) = A_i x + b_i$, where the $A_i$ are $n \times n$ non-singular matrices. If we define the sums

$$\Sigma_k^s = \sum_{(i_1,\dots,i_k)\in\mathscr{I}_k} \varphi^s(A_{i_1} \dots A_{i_k}),$$

submultiplicativity of the singular value function implies that the sequence $\frac{1}{n}\log \Sigma_k^s$ is subadditive, and consequently, there exists a number $\Sigma_\infty^s$ such that $(\Sigma_k^s)^{1/k} \to \Sigma_\infty^s$ for each $s$.

**Proposition 5.2.6.** *The following numbers exist and are equal:*

1. *the unique $s$ such that $\Sigma_\infty^s = 1$;*

2. $$\inf\left\{ s : \sum_{k=1}^{\infty} \sum_{\mathscr{I}_k} \varphi^s(A_{i_1} \dots A_{i_k}) < \infty \right\} = \sup\left\{ s : \sum_{k=1}^{\infty} \sum_{\mathscr{I}_k} \varphi^s(A_{i_1} \dots A_{i_k}) = \infty \right\}.$$

**Proof.** This is proposition 4.1 from [Fal88]. ∎

**Definition 5.2.7.** *The number $s$ from proposition 5.2.6 is called the* affinity dimension *of the IFS, and is denoted by $d(A_1, \dots, A_m)$.*

One of the foundational results in the theory of dimension of self-affine sets is the following theorem:

**Theorem 5.2.8** (Falconer)**.** *For all $(b_1, \dots, b_m) \in \mathbb{R}^{mn}$ we have that $\dim_H F \le d(A_1, \dots, A_m)$. If the contraction ratios $r_1, \dots, r_m$ of the matrices $A_1, \dots, A_m$ are such $r_i \le \frac{1}{3}$ for all $i$, then we have that $\dim_H F = \min\{n, d(A_1, \dots, A_m)\}$ for almost every $(b_1, \dots, b_m) \in \mathbb{R}^{mn}$.*

**Proof.** This is the main theorem of [Fal88]. ■

Similarly to the case of one dimensional maps, we can define a topological pressure for higher dimensional systems.

**Definition 5.2.9.** *For $s \in \mathbb{R}$, define the* Pressure *of the IFS by*

$$P(s) = \lim_{n\to\infty} \frac{1}{n} \log\left(\sum_{\mathscr{I}_n} \varphi^s(A_{i_1}\cdots A_{i_n})\right).$$

Let $\{S_i(x) = A_i x + b_i, i = 1,\dots,m\}$ be a self-affine system, and consider the symbolic space $\Sigma^{\mathbb{N}} = \{1,\dots,m\}^{\mathbb{N}}$ equipped with the product topology (or equivalently the topology generated by the cylinder sets), and the dynamics of the left shift $\sigma: \Sigma \to \Sigma$. Let $\mathscr{M}_\sigma$ be the set of all $\sigma$-invriant Borel probability measures on $\Sigma$.

**Definition 5.2.10.** *A measure $\nu \in \mathscr{M}_\sigma$ is called a $\varphi^s$-*equilibrium *of the self-affine system $\{S_i\}$ if*

$$\sup_{\mu\in\mathscr{M}_\sigma}\left[h_\mu + \lim_{n\to\infty}\frac{1}{n}\int_\Sigma \log\varphi^s(A_{i_n}\cdots A_{i_1})d\mu\right]$$

*is achieved at $\nu$. Here $h_\mu$ represents the entropy of the measure $\mu$ (see definition 2.4.11). When $s$ is equal to the affinity dimension of the system, we call the corresponding measure a* Käenmäki measure.

Under certain *irreducibility* conditions on the set of matrices defining the IFS, the corresponding $\varphi^s$-equilibrium measures have a *Gibbs*-like property.

**Definition 5.2.11.** *We say that the family of matrices $\{A_1,\dots,A_m\}$ is* irreducible *if there are no proper subspaces $V \subset \mathbb{R}^n$ such that $A_i V \subset V$ for all $i$.*

**Theorem 5.2.12** ([FL02], [FK10])**.** *For a self-affine system $\{S_i(x) = A_i x + b_i, i = 1,\dots,m\}$ with $2\times 2$ irreducible matrices $\{A_i\}$. Then $\{S_i\}$ has a unique $\varphi^s$-equilibrium $m^s$, which is ergodic and satisfies the following Gibbs property:*

$$C^{-1}e^{-P(s)k}\varphi^s(\mathbf{i}) \leqslant m^s([\mathbf{i}]) \leqslant Ce^{-P(s)k}\varphi^s(\mathbf{i}),$$

*where $P(s)$ is the pressure of the IFS.*

We will assume that each map $S_i$ maps the unit square into itself. Under our assumptions, there exists a unique non-empty compact set $F \subset [0,1]^2$ such that

$$F = \bigcup_{i\in\mathscr{I}} S_i(F)$$

which we call the self-affine set associated to the system $\{S_i\}_{\mathscr{I}}$. We can describe the structure of the system by using a symbolic coding. Let $\Sigma = \mathscr{I}^{\mathbb{N}}$ and $\sigma : \Sigma \to \Sigma$ the left shift operator, that is, for $\mathbf{i} = (i_1, i_2, \dots) \in \Sigma$, $\sigma(\mathbf{i}) = (i_2, i_3, \dots)$. The topology of $\Sigma$ is generated by the cylinder sets $[\mathbf{i}] = \{x \in \Sigma : x_1 = \mathbf{i}_1, \dots, x_k = \mathbf{i}_k\}$ for $\mathbf{i} \in \mathscr{I}^k$. For any integer $k$ and sequence $\mathbf{i} = (i_1, i_2, \dots)$, we denote $\mathbf{i}|_k = (i_1, \dots, i_k)$. For $\mathbf{i} \in \mathscr{I}^* = \bigcup_k \mathscr{I}^k$, we write

$$S_{\mathbf{i}} = S_{i_1} \circ \dots \circ S_{i_k}.$$

With this, we can write a projection map from the symbolic space to the self-affine set by

$$\Pi(\mathbf{i}) = \bigcap_{k=1}^{\infty} S_{\mathbf{i}_k}\left([0,1]^2\right)$$

and then $F = \Pi(\Sigma)$. We use the same coding for matrices of our system: give a word $\mathbf{i} \in \Sigma$, denote $A(\mathbf{i}_n) = A_{i_1} \cdot \dots \cdot A_{i_n}$, or $A(\mathbf{i})$ when $\mathbf{i} \in \mathscr{I}^*$. In this case, we denote the length of $\mathbf{i}$ by $|\mathbf{i}|$.

## 5.3 Diagonal-antidiagonal systems

In this section we will consider self-affine systems of the form $S_i(x) = A_i x + t_i$ indexed by a finite alphabet $\mathscr{I} = \{1, \dots, d\}$. Here, the linear parts of the maps are contracting non-singular $2 \times 2$ matrices with non-negative entries of the form

$$A_i = \begin{bmatrix} a_i & 0 \\ 0 & b_i \end{bmatrix}$$

for $i < l$ and

$$A_i = \begin{bmatrix} 0 & a_i \\ b_i & 0 \end{bmatrix}$$

for $i \geq l$.

In this section we describe the construction as well as the structure of Kaenmaki measures. For a linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$, recall that its singular values can be defined as the length of the semiaxes of the ellipsoid $T(B(0,1))$. If we write them in descending order, denote them by

$$1 > \alpha_1(\mathbf{i}) \geqslant \alpha_2(\mathbf{i}) > 0.$$

For $s \in (0,2]$, let $\varphi^s : \mathscr{I}^* \to \mathbb{R}^+$ be the Falconer potential:

$$\varphi^s(\mathbf{i}) = \begin{cases} \alpha_1(\mathbf{i})^s & \text{if } s \in (0,1) \\ \alpha_1(\mathbf{i})\alpha_2(\mathbf{i})^{s-1} & \text{if } s \in [1,2] \end{cases}$$

It is easy to see that $\varphi^s$ is subadditive: $\varphi^s(\mathbf{ij}) \le \varphi^s(\mathbf{i})\varphi^s(\mathbf{j})$. With this we can define the subadditive pressure by

$$P(s) = \lim_{n\to\infty} \frac{1}{n} \log\left( \sum_{\mathbf{i}\in\mathscr{I}^n} \varphi^s(\mathbf{i}) \right).$$

Assuming that $P(2) \le 0$, the affinity dimension is defined as the value $s^*$ such that $P(s^*) = 0$. From 5.2.12 it follows that for each $s$, there is a unique ergodic Borel probability measure $m^s$ and a constant $C \ge 1$ such that

$$C^{-1}e^{-P(s)k}\varphi^s(\mathbf{i}) \leqslant m^s([\mathbf{i}]) \leqslant Ce^{-P(s)k}\varphi^s(\mathbf{i})$$

for $\mathbf{i} \in \mathscr{I}^k$. The measure $m^s$ can be projected to a measure $\mu^s$ on $F$ by setting $\mu^s(A) = m^s(\Pi^{-1}(A))$ for every measurable set $A$.

The structure of the Kaenmaki measures for these particular systems was described in [FJJ18]: the authors construct a graph directed system (GDS) on the alphabet $\{1,\dots,2d\}$ and with matrix

$$A(i,j) = \begin{cases} 1 & \text{if } i \in \{1,\dots,l-1\} \cup \{d+l,\dots,2d\} \text{ and } j \leqslant d \\ 1 & \text{if } i \in \{l,\dots,d+l-1\} \text{ and } j > d \\ 0 & \text{otherwise.} \end{cases}$$

We denote the symbolic space of this GDS by $\Sigma_A$. With this GDS, it is possible to keep track of the orientation of the linear part of $S_{\mathbf{i}}$ for each $\mathbf{i} \in \mathscr{I}^*$. We send each sequence $\mathbf{i} \in \Sigma$ to an element in $\Sigma_A$ given by $\tau(\mathbf{i}) = \tau(i_1 i_2 \dots) = (\tau_1(\mathbf{i})\tau_2(\mathbf{i})\dots)$ where $\tau_1(\mathbf{i}) = i_1$ and

$$\tau_m(\mathbf{i}) = \begin{cases} i_m & \text{if card}\left\{1 \leqslant j \leqslant m-1 : i_j \geqslant l\right\} \text{ is even} \\ i_m + d & \text{if card}\left\{1 \leqslant j \leqslant m-1 : i_j \geqslant l\right\} \text{ is odd} \end{cases}$$

If the linear part of $S_{\mathbf{i}|_k}$ is diagonal, then $\tau_k(\mathbf{i}) = i_k$, otherwise $\tau_k(\mathbf{i}) = i_k + d$. In this way, if the linear part of $S_{\mathbf{i}|_k}$ is diagonal, the last digit of $\tau(\mathbf{i})|_k$ is less or equal to $d$, while if it is anti-diagonal, it is greater than $d$. It is easy to see that $\tau$ is injective but not surjective. The image of $\tau$ is the set of sequences with first digit at most $d$. The complementary function to $\tau$, defined by $\omega(\mathbf{i}) = \omega(i_1 i_2 \cdots) = (\omega_1(\mathbf{i})\omega_2(\mathbf{i})\dots)$, $\omega_1(\mathbf{i}) = i_1 + d$ and

$$\omega_m(\mathbf{i}) = \begin{cases} i_m + d & \text{if card}\left\{1 \leqslant j \leqslant m-1 : i_j \geqslant l\right\} \text{ is even} \\ i_m & \text{if card}\left\{1 \leqslant j \leqslant m-1 : i_j \geqslant l\right\} \text{ is odd} \end{cases}$$

Similarly to $\tau$, $\omega$ is injective but not surjective; in fact, its image consists of all the sequences in $\Sigma_A$ starting with digits greater than $d$. It follows that $\Sigma_A = \tau(\Sigma) \cup \omega(\Sigma)$.

As observed in [FJJ18], the system $(\Sigma_A,\sigma)$ is mixing. In particular, this implies the existence of Gibbs measures for regular enough potentials (see [Bow08]). Define the locally constant potentials $f_{1,s}, f_{2,s} : \Sigma_A \to \mathbb{R}$ by

$$f_{1,s}(\mathbf{i}) = \begin{cases} s\log a_{i_1} & \text{if } i_1 \leq d \\ s\log b_{i_1-d} & \text{if } i_1 \geq d+1 \end{cases}$$

and

$$f_{2,s}(\mathbf{i}) = \begin{cases} s\log b_{i_1} & \text{if } i_1 \leq d \\ s\log a_{i_1-d} & \text{if } i_1 \geq d+1 \end{cases}$$

for $s \in (0,1)$, and

$$f_{1,s}(\mathbf{i}) = \begin{cases} \log a_{i_1} + (s-1)\log b_{i_1} & \text{if } i_1 \leq d \\ \log b_{i_1-d} + (s-1)\log a_{i_1-d} & \text{if } i_1 \geq d+1 \end{cases}$$

and

$$f_{2,s}(\mathbf{i}) = \begin{cases} \log b_{i_1} + (s-1)\log a_{i_1} & \text{if } i_1 \leq d \\ \log a_{i_1-d} + (s-1)\log b_{i_1-d} & \text{if } i_1 \geq d+1 \end{cases}$$

for $s \in [1,2]$. The Gibbs measures associated to these potentials are denoted $m_1$ and $m_2$ respectively. Since $\tau$ is injective, we can define a measure on $\Sigma$ by $\nu(E) = m_1(\tau(E)) + m_2(\tau(E)) = m_1(\tau(E)) + m_1(\omega(E))$ for every measurable set $E$. We summarize the properties of $\nu$ obtained in [FJJ18]:

**Theorem 5.3.1.** *The measure $\nu$ is the unique Kaenmaki measure for $\varphi^s$, that is, $\nu = m^s$. The topological pressure of the potentials $f_{1,s}$ and $f_{2,s}$ coincides with the subadditive pressure $P(s)$.*

With this description of the Kaenmaki measure in mind, we examine some statistical properties of it. We first prove that $m^s$ is not mixing. Using a different description, it was proven in [Mor17] that a similar class of systems is not mixing with respect to the equilibrium measure. To prove it in our context, we will need the following lemma:

**Lemma 5.3.2.** *For any finite length word $\mathbf{i} \in \mathscr{I}$, the image of the cylinder $[\mathbf{i}]$ under $\tau$ or $\omega$ is a cylinder in $\Sigma_A$.*

**Proof.** Fix such word $\mathbf{i}$ and consider the cylinder $[\mathbf{i}] = \{x \in \Sigma : x_1 = i_1, \ldots, x_k = i_k\}$. The image of $[\mathbf{i}]$ under $\tau$ is contained in the cylinder $[i_1, \tau_2(\mathbf{i}), \ldots, \tau_k(\mathbf{i})]$. On the other hand, if

$z \in [i_1, \tau_2(\mathbf{i}), \ldots, \tau_k(\mathbf{i})]$, then $z_1 = i_1, z_2 = \tau_2(\mathbf{i}), \ldots, z_k = \tau_k(\mathbf{i})$. Define $\mathbf{j} \in \Sigma$ by $j_1 = z_1, \ldots, j_k = z_k$ and

$$j_n = \begin{cases} z_n, & \text{if } z_n \le d \\ z_n - d, & \text{if } z_n > d. \end{cases}$$

Then $\tau(\mathbf{j}) = z$ and so $\tau[\mathbf{i}] = [i_1, \tau_2(\mathbf{i}), \ldots, \tau_k(\mathbf{i})]$. ■

From this it follows that

**Lemma 5.3.3.** *For any finite length word* $\mathbf{i} \in \mathscr{I}$,

$$\sigma^{-n}[\mathbf{i}] = \sigma_A^{-n}(\tau[\mathbf{i}]),$$

*where* $\sigma_A$ *denotes the shift operator on* $\Sigma_A$.

**Theorem 5.3.4.** *The measure* $m^s$ *is not mixing.*

**Proof.** Fix two cylinders $[\mathbf{a}] = [a_1, \ldots, a_k], [\mathbf{b}] = [b_1, \ldots, b_\ell] \subset \Sigma$ and let $\varepsilon > 0$ small enough so that

$$m_1[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})] + m_2[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})] - 2\varepsilon > 0.$$

Since $m_1, m_2$ are Gibbs measures, they are mixing (see [Bow08]), and hence, there is $n_0 \in \mathbb{N}$ such that

$$\left| m_i\left([\tau(\mathbf{a})] \cap \sigma_A^{-n}[\tau(\mathbf{b})]\right) - m_i[\tau(\mathbf{a})]m_i[\tau(\mathbf{b})] \right| < \varepsilon$$

for $i \in \{1, 2\}$ and $n \ge n_0$. Since $\tau$ is injective

$$\begin{aligned} \nu\left([\mathbf{a}] \cap \sigma^{-n}[\mathbf{b}]\right) &= m_1\left(\tau([\mathbf{a}] \cap \sigma^{-n}[\mathbf{b}])\right) + m_2\left(\tau([\mathbf{a}] \cap \sigma^{-n}[\mathbf{b}])\right) \\ &= m_1\left(\tau[\mathbf{a}] \cap \tau\sigma^{-n}[\mathbf{b}]\right) + m_2\left(\tau[\mathbf{a}] \cap \tau\sigma^{-n}[\mathbf{b}]\right) \\ &= m_1\left([\tau(\mathbf{a})] \cap \sigma_A^{-n}[\tau(\mathbf{b})]\right) + m_2\left([\tau(\mathbf{a})] \cap \sigma_A^{-n}[\tau(\mathbf{b})]\right). \end{aligned}$$

On the other hand,

$$\begin{aligned} \nu[\mathbf{a}]\nu[\mathbf{b}] &= (m_1[\tau(\mathbf{a})] + m_2[\tau(\mathbf{a}))(m_1[\tau(\mathbf{b})] + m_2[\tau(\mathbf{b})]) \\ &= m_1[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})] + m_2[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})] + m_1[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})] + m_2[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})]. \end{aligned}$$

We obtain then

$$\begin{aligned}\nu([\mathbf{a}]\cap\sigma^{-n}[\mathbf{b}])-\nu[\mathbf{a}]\nu[\mathbf{b}] &= \big(m_1([\tau(\mathbf{a})]\cap\sigma_A^{-n}[\tau(\mathbf{b})])-m_1[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})]\big)\\ &\quad+\big(m_2([\tau(\mathbf{a})]\cap\sigma_A^{-n}[\tau(\mathbf{b})])-m_2[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})]\big)\\ &\quad-(m_1[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})]+m_2[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})]).\end{aligned}$$

Thus,

$$\Big|\nu([\mathbf{a}]\cap\sigma^{-n}[\mathbf{b}])-\nu[\mathbf{a}]\nu[\mathbf{b}]\Big|>(m_1[\tau(\mathbf{a})]m_2[\tau(\mathbf{b})]+m_2[\tau(\mathbf{a})]m_1[\tau(\mathbf{b})])-2\varepsilon>0$$

and consequently $m=m^s$ cannot be mixing. ■

We consider now the problem of shrinking targets for self-affine systems. We focus on targets given by shrinking cylinders. Fix an infinite word $\mathbf{j}\in\Sigma$ and a non-decreasing sequence $\ell_k$, and define the targets by $\mathscr{B}_k=[\mathbf{j}|_{\ell_k}]$. The recurrent set associated to this family of targets is defined as

$$\begin{aligned}R(\mathbf{j}) &= \{x\in\Sigma:\sigma^k(x)\in\mathscr{B}_k \text{ for infinitely many } k\in\mathbb{N}\}\\ &= \{x\in\Sigma:\sigma^k(x)|_{\ell_k}=\mathbf{j}_{\ell_k} \text{ for infinitely many } k\in\mathbb{N}\}\\ &= \limsup_{k\to\infty} R(\mathbf{j},k),\end{aligned}$$

where $R(\mathbf{j},k)=\{x\in\Sigma:\sigma^k(x)|_{\ell_k}=\mathbf{j}_{\ell_k}\}$. The recurrent set can be projected onto the self-affine set:

$$\tilde{R}(\Pi(\mathbf{j}))=\Pi(R(\mathbf{j})).$$

In [KR18], the authors study the dimension of recurrent sets for shrinking targets associated to a self-affine system. In their setting, the maps satisfy a domination condition: there exists a constant $D$ such that

$$\varphi^s(\mathbf{ij})\geq D\varphi^s(\mathbf{i})\varphi(\mathbf{j}).$$

This condition implies that the Käenmäki measure behaves essentially like a Gibbs measure, and hence the same techniques can be used. In particular, the authors follow the ideas of [CK01] to prove the following $0-1$ law:

**Theorem 5.3.5.** *Let $\mathbf{j}\in\Sigma$ and $\ell_k$ an increasing sequence. If $m=m^t$ and $\mu=\mu^t$ are the Käenmäki measures (on $\Sigma$ and $F$ respec) corresponding to the value of $t$ such that $P(t)=0$. Then the measure of the set $\tilde{R}(\Pi(\mathbf{j}))$ is either $0$ or $1$ according to wether the sum*

$$\sum_{k=1}^{\infty}\mu([\mathbf{j}|_{\ell_k}])$$

*converges or not.*

We show that this result also holds in our setting. For this, we recall the analogue result proved for Gibbs measures by Chernov and Kleinbock. We formulate their result in our context:

**Theorem 5.3.6.** *Let $(\Sigma', \sigma)$ be a topologically transitive topological Markov shift and $\eta$ is a Gibbs measure on $\Sigma'$. Assume $\{C_k\}$ is a sequence of nested cylinders in $\Sigma'$ and*

$$\sum_{k=1}^{\infty} \mu(C_k) = \infty.$$

*Then $\{C_k\}$ is a strong Borel-Cantelli sequence: if $\chi_n$ is the indicator function of $\sigma^{-n} C_n$, then*

$$\frac{\sum_{n=1}^{N} \chi_n(x)}{\sum_{n=1}^{N} \mu(A_n)} \to 1$$

*$\mu$-almost everywhere. Moreover, if we denote $S_N = \sum_{n=1}^{N} \chi_n(x)$ and $E_N = \sum_{n=1}^{N} \mu(A_n)$,*

$$S_N = E_N + O\left(E_N^{1/2} \log^{3/2+\varepsilon} E_N\right).$$

In particular, the result above shows that the recurrent set for the shrinking targets defined by the nested cylinders has full measure if the sum converges.

**Theorem 5.3.7.** *Under the assumptions above on the self-affine system, then $\nu(\tilde{R}(\Pi(\mathbf{j})))$ and $m(\Pi(\mathbf{j}))$ are either $0$ or $1$, according to the convergence or diverngence of the series*

$$\sum_{k=1}^{\infty} m([\mathbf{j}|_{\ell_k}]).$$

**Proof.** Assume that the sum $\sum_{k=1}^{\infty} m([\mathbf{j}|_{\ell_k}])$ converges. Since $m$ is shift invariant, we have that

$$m(R(\mathbf{j}, k)) = m\left[\mathbf{j}|_{\ell_k}\right].$$

Then the Borel-Cantelli lemma implies that both $R(\mathbf{j})$ and $\tilde{R}(\mathbf{j})$ have full measure.

Suppose now that the sum $\sum_{k=1}^{\infty} m\left([\mathbf{j}|_{\ell_k}]\right)$ diverges. By theorem 5.3.1 we can write the sum as

$$\sum_{k=1}^{\infty} m\left([\mathbf{j}|_{\ell_k}]\right) = \sum_{k=1}^{\infty} m_1\left(\tau\left[\mathbf{j}_{\ell_k}\right]\right) + \sum_{k=1}^{\infty} m_2\left(\tau\left[\mathbf{j}_{\ell_k}\right]\right),$$

which implies that one of the two sums on the right hand side diverges. Without loss of generality, assume the sum with respect to the measure $m_1$ diverges, that is,

$$\sum_{k=1}^{\infty} m_1\left(\tau\left[\mathbf{j}_{\ell_k}\right]\right) = \infty.$$

Since $m_1$ is shift invariant, we have that

$$m_1\left(\sigma_A^{-n}\left[\tau(\mathbf{j})|_{\ell_k}\right]\right) = m_1\left[\tau(\mathbf{j})|_{\ell_k}\right]$$

for all $k$.
The above remark implies that the nested sequence of cylinders $\{[\tau(\mathbf{j}_{\ell_k})]\}_k$ and the Gibbs measure $m_1$ on $\Sigma_A$ satisfy the conditions of theorem 5.3.6, and hence,

$$m_1\left(\limsup_{k\to\infty} \sigma_A^{-k}\left[\tau\mathbf{j}|_{\ell_k}\right]\right) = 1.$$

Since $\tau$ is injective,

$$\limsup_{k\to\infty} \sigma_A^{-k}\left[\tau\mathbf{j}|_{\ell_k}\right] = \tau\left(\limsup_{k\to\infty} \sigma^{-k}\left[\mathbf{j}|_{\ell_k}\right]\right).$$

We obtain then

$$\begin{aligned} m\left(\limsup_{k\to\infty} \sigma^{-k}\left[\mathbf{j}|_{\ell_k}\right]\right) &= m_1\left(\tau\left(\limsup_{k\to\infty} \sigma^{-k}\left[\mathbf{j}|_{\ell_k}\right]\right)\right) + m_2\left(\tau\left(\limsup_{k\to\infty} \sigma^{-k}\left[\mathbf{j}|_{\ell_k}\right]\right)\right) \\ &\geq m_1\left(\limsup_{k\to\infty} \sigma_A^{-k}\left[\tau\mathbf{j}|_{\ell_k}\right]\right) = 1 \end{aligned}$$

from which we conclude the result. ■

# Bibliography


[AA16] Mohamed Abdelkader and Romain Aimino. On the quenched central limit theorem for random dynamical systems. *Journal of Physics A: Mathematical and Theoretical*, 49(24):244002, 2016.

[Aar77] Jon Aaronson. On the ergodic theory of non-integrable functions and infinite measure spaces. *Israel Journal of Mathematics*, 27(2):163–173, Jun 1977.

[Aar97] Jon Aaronson. *An introduction to infinite ergodic theory*. Number 50. American Mathematical Soc., 1997.

[AF18] Romain Aimino and Jorge Milhazes Freitas. Large deviations for dynamical systems with stretched exponential decay of correlations. *arXiv preprint arXiv:1812.09742*, 2018.

[AFLV11] José F Alves, Jorge M Freitas, Stefano Luzzatto, and Sandro Vaienti. From rates of mixing to recurrence times via large deviations. *Advances in Mathematics*, 228(2):1203–1236, 2011.

[AHN$^+$15] Romain Aimino, Huyi Hu, Matthew Nicol, Andrei Török, and Sandro Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete & Continuous Dynamical Systems-A*, 35(3):793, 2015.

[ALS09] Arvind Ayyer, Carlangelo Liverani, and Mikko Stenlund. Quenched clt for random toral automorphism. *Discrete and continuous dynamical systems*, 24(2):331–348, 2009.

[AN03] Jon Aaronson and Hitoshi Nakada. Trimmed sums for non-negative, mixing stationary processes. *Stochastic processes and their applications*, 104(2):173–192, 2003.

[ANV15] Romain Aimino, Matthew Nicol, and Sandro Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probability Theory and Related Fields*, 162(1-2):233–274, 2015.

[AR16] Romain Aimino and Jérôme Rousseau. Concentration inequalities for sequential dynamical systems of the unit interval. *Ergodic Theory and Dynamical Systems*, 36(8):2384–2407, 2016.

[Bar14] Robert G Bartle. *The elements of integration and Lebesgue measure*. John Wiley & Sons, 2014.

[BB16a] Wael Bahsoun and Christopher Bose. Corrigendum: Mixing rates and limit theorems for random intermittent maps (2016 *nonlinearity* 29 1417) [ MR3476513]. *Nonlinearity*, 29(12):C4, 2016.

[BB16b] Wael Bahsoun and Christopher Bose. Mixing rates and limit theorems for random intermittent maps. *Nonlinearity*, 29(4):1417, 2016.

[BBD14] Wael Bahsoun, Christopher Bose, and Yuejiao Duan. Decay of correlation for random intermittent maps. *Nonlinearity*, 27(7):1543, 2014.

[BBR19] Wael Bahsoun, Christopher Bose, and Marks Ruziboev. Quenched decay of correlations for slowly mixing systems. *Transactions of the American Mathematical Society*, 372(9):6547–6587, 2019.

[Bes35] Abram Samoilovitch Besicovitch. On the sum of digits of real numbers represented in the dyadic system. *Mathematische Annalen*, 110(1):321–330, 1935.

[BG97] Abraham Boyarsky and PawełGóra. *Laws of chaos*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, 1997. Invariant measures and dynamical systems in one dimension.

[Bil61] Patrick Billingsley. The lindeberg-levy theorem for martingales. *Proceedings of the American Mathematical Society*, 12(5):788–792, 1961.

[Bir31] George D Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660, 1931.

[BJ18] Simon Baker and Natalia Jurga. Maximising bernoulli measures and dimension gaps for countable branched systems. *arXiv preprint arXiv:1802.07585*, 2018.

[Bow08] Robert Edward Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470. Springer Science & Business Media, 2008.

[CFS12] Isaac P Cornfeld, Sergej V Fomin, and Yakov Grigorévic Sinai. *Ergodic theory*, volume 245. Springer Science & Business Media, 2012.

[CK01] Nikolai Chernov and Dmitry Kleinbock. Dynamical borel-cantelli lemmas for gibbs measures. *Israel Journal of Mathematics*, 122(1):1–27, 2001.

[CR61] Yuan-Shih Chow and Herbert Robbins. On sums of independent random variables with infinite moments and „fair" games. *Proceedings of the National Academy of Sciences of the United States of America*, 47(3):330, 1961.

[CR07] Jean-Pierre Conze and Albert Raugi. Limit theorems for sequential expanding dynamical systems. In *Ergodic Theory and Related Fields*, volume 430 of *Contemp. Math.*, pages 89–121. American Mathematical Soc., Amer. Math. Soc., Providence, RI, 2007.

[DK02] Karma Dajani and Cor Kraaikamp. *Ergodic theory of numbers*. Number 29. Cambridge University Press, 2002.

[Egg49] Harold Gordon Eggleston. The fractional dimension of a set defined by decimal properties. *The Quarterly Journal of Mathematics*, (1):31–36, 1949.

[Fal88] Kenneth J Falconer. The hausdorff dimension of self-affine fractals. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 103, pages 339–350. Cambridge University Press, 1988.

[Fal97] Kenneth Falconer. *Techniques in fractal geometry*. John Wiley, 1997.

[Fal04] Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.

[FJJ18] Jonathan M Fraser, Thomas Jordan, and Natalia Jurga. Dimensions of equilibrium measures on a class of planar self-affine sets. *Journal of Fractal Geometry*, 2018.

[FK10] De-Jun Feng and Antti Käenmäki. Equilibrium states of the pressure function for products of matrices. *Discrete & Continuous Dynamical Systems-A*, 30(3):699, 2010.

[FL02] De-Jun Feng and Ka-Sing Lau. The pressure function for products of nonnegative matrices. *Mathematical Research Letters*, 9:363–378, 2002.

[FLM10] Ai-hua Fan, Lingmin Liao, and Ji-hua Ma. On the frequency of partial quotients of regular continued fractions. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 148, pages 179–192. Cambridge University Press, 2010.

[FSU14] Lior Fishman, David Simmons, and Mariusz Urbański. Diophantine properties of measures invariant with respect to the gauss map. *Journal d'Analyse Mathématique*, 122(1):289–315, 2014.

[Fur13] Ovidiu Furdui. *Limits, Series, and Fractional Part Integrals*. Springer, 2013.

[Gor69] Mikhail Iosifovich Gordin. The central limit theorem for stationary processes. In *Doklady Akademii Nauk*, volume 188, pages 739–741. Russian Academy of Sciences, 1969.

[HL19] Olli Hella and Juho Leppänen. Central limit theorems with a rate of convergence for time-dependent intermittent maps. *Stochastics and Dynamics*, page 2050025, 2019.

[Hoe94] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. In *The Collected Works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

[HR92] Franz Hofbauer and Peter Raith. The hausdorff dimension of an ergodic invariant measure for a piecewise monotonic map of the interval. *Canadian Mathematical Bulletin*, 35(1):84–98, 1992.

[HS20] Olli Hella and Mikko Stenlund. Quenched normal approximation for random sequences of transformations. *Journal of Statistical Physics*, 178(1):1–37, 2020.

[Hut81] John E Hutchinson. Fractals and self similarity. *Indiana University Mathematics Journal*, 30(5):713–747, 1981.

[Ibr62] Ildar A Ibragimov. Some limit theorems for stationary processes. *Theory of Probability & Its Applications*, 7(4):349–382, 1962.

[IL65] Ildar A Ibragimov and Ju V Linnik. Nezavisimye stalionarno svyazannye velichiny, izdat.“, 1965.

[Jur18] Natalia Jurga. A new proof of the dimension gap for the gauss map. *arXiv preprint arXiv:1806.00841*, 2018.

[Käe04] Antti Käenmäki. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math*, 29(2):419–458, 2004.

[Kar93] Alan Karr. *Probability*. Springer Texts in Statistics. Springer-Verlag, New York, 1993.

[Kif91] Yuri Kifer. Large deviations for random expanding maps. In *Lyapunov Exponents*, pages 178–186. Springer, 1991.

[Kif92] Yuri Kifer. Averaging in dynamical systems and large deviations. *Inventiones mathematicae*, 110(1):337–370, 1992.

[Koh58] Andrei Nikolaevitch Kohnogorov. A new invariant for transitive dynamical systems. In *Dokl Akad Nauk SSSR*, volume 119, 1958.

[Kol58] Andrei Nikolaevich Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in lebesgue spaces. In *Doklady Akademii Nauk*, volume 119, pages 861–864. Russian Academy of Sciences, 1958.

[Kol59] Andrei Nikolaevitch Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. In *Dokl. Akad. Nauk SSSR*, volume 124, pages 754–755, 1959.

[KP66] John R Kinney and Tom S Pitcher. The dimension of some sets defined in terms of f-expansions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 4(4):293–315, 1966.

[KPW01] Yuri Kifer, Yuval Peres, and Benjamin Weiss. A dimension gap for continued fractions with independent digits. *Israel Journal of Mathematics*, 124(1):61–76, 2001.

[KR18] Henna Koivusalo and Felipe A. Ramírez. Recurrence to shrinking targets on typical self-affine fractals. *Proceedings of the Edinburgh Mathematical Society*, 61(2):387–400, May 2018. © 2018, The Edinburgh Mathematical Society. This is an author-produced version of the published paper. Uploaded in accordance with the publisher's self-archiving policy. Further copying may not be permitted; contact the publisher for details.

[Led81] François Ledrappier. Some properties of absolutely continuous invariant measures on an interval. *Ergodic Theory and Dynamical Systems*, 1(1):77–93, 1981.

[Liv96] Carlangelo Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems (Montevideo, 1995)*, volume 362, pages 56–75, 1996.

[LM85] François Ledrappier and Michał Misiurewicz. Dimension of invariant measures for maps with exponent zero. *Ergodic Theory and Dynamical Systems*, 5(4):595–610, 1985.

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti. A probabilistic approach to intermittency. *Ergodic theory and dynamical systems*, 19(3):671–685, 1999.

[Man81] Anthony Manning. A relation between lyapunov exponents, hausdorff dimension and entropy. *Ergodic Theory and Dynamical Systems*, 1(4):451–459, 1981.

[Mat99] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces: fractals and rectifiability*. Number 44. Cambridge university press, 1999.

[Mel09] Ian Melbourne. Large and moderate deviations for slowly mixing dynamical systems. *Proceedings of the American Mathematical Society*, 137(5):1735–1741, 2009.

[MN08] Ian Melbourne and Matthew Nicol. Large deviations for nonuniformly hyperbolic systems. *Transactions of the American Mathematical Society*, 360(12):6661–6676, 2008.

[Mor76] Toshio Mori. The strong law of large numbers when extreme terms are excluded from sums. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 36(3):189–194, 1976.

[Mor77] Toshio Mori. Stability for sums of iid random variables when extreme terms are excluded. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 40(2):159–167, 1977.

[Mor17] Ian D Morris. Ergodic properties of matrix equilibrium states. *Ergodic Theory and Dynamical Systems*, 38(6):2295–2320, 2017.

[MPU+06] Florence Merlevède, Magda Peligrad, Sergey Utev, et al. Recent advances in invariance principles for stationary sequences. *Probability surveys*, 3:1–36, 2006.

[MU03] R Daniel Mauldin and Mariusz Urbański. *Graph directed Markov systems: geometry and dynamics of limit sets*, volume 148. Cambridge University Press, 2003.

[NPPT19] Matthew Nicol, Felipe Perez Pereira, and Andrew Torok. Large deviations and central limit theorems for sequential and random systems of intermittent maps. *arXiv*, pages arXiv–1909, 2019.

[NTV18] Matthew Nicol, Andrew Török, and Sandro Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory and Dynamical Systems*, 38(3):1127–1153, 2018.

[Oks13] Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

[Ose68] Valery Iustinovich Oseledets. A multiplicative ergodic theorem. characteristic ljapunov, exponents of dynamical systems. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 19:179–210, 1968.

[Pes77] Yakov Borisovich Pesin. Characteristic lyapunov exponents and smooth ergodic theory. *Uspekhi Matematicheskikh Nauk*, 32(4):55–112, 1977.

[Pes08] Yakov B Pesin. *Dimension theory in dynamical systems: contemporary views and applications*. University of Chicago Press, 2008.

[Poi90] Henri Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270, 1890.

[PP18] Felipe Pérez Pereira. Dimension of gibbs measures with infinite entropy. *arXiv*, pages arXiv–1812, 2018.

[PW99] Mark Pollicott and Howard Weiss. Multifractal analysis of lyapunov exponent for continued fraction and manneville–pomeau transformations and applications to diophantine approximation. *Communications in mathematical physics*, 207(1):145–171, 1999.

[Rio17] Emmanuel Rio. *Asymptotic theory of weakly dependent random processes*, volume 80. Springer, 2017.

[Rok67] Vladimir Abramovich Rokhlin. Lectures on the entropy theory of transformations with invariant measure. *Uspehi Mat. Nauk*, 22(5 (137)):3–56, 1967.

[Rue78] David Ruelle. An inequality for the entropy of differentiable maps. *Boletim da Sociedade Brasileira de Matemática-Bulletin/Brazilian Mathematical Society*, 9(1):83–87, 1978.

[Sha48] Claude E Shannon. A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423, 1948.

[Sin59] Yakov Sinai. On the concept of entropy for a dynamic system. *Doklady Akademii Nauk SSSR*, 124(4):768–771, 1959.

[VO16] Marcelo Viana and Krerley Oliveira. *Foundations of ergodic theory*, volume 151. Cambridge University Press, 2016.

[Wal00] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.

[You82] Lai-Sang Young. Dimension, entropy and lyapunov exponents. *Ergodic theory and dynamical systems*, 2(1):109–124, 1982.