Scuola Internazionale Superiore di Studi Avanzati (SISSA)

Gibbs-Markov-Young Structures and Decay of Correlations

Candidate: Marks B. Ruziboev

Supervisor: Prof. Stefano Luzzatto

A thesis presented for the degree of
Doctor of Philosophy

Abstract iii

Dedication v

Acknowledgments vii

Introduction 1
1. Basic definitions 2
2. Young Towers 4

Chapter 1. Direct Product Systems 9
1. Ergodic properties of product systems 9
2. Geometric properties of product systems 12
3. A tower for the product 14
4. Return time asymptotics 22
5. Rates of decay 24
6. Proof in the non-invertible setting 33
7. Proof in the invertible setting 35

Chapter 2. Decay of Correlations for Continuous observables 37
1. Decay of correlation for invertible systems 37
2. Applications 39
3. Proof of Theorem 2.1 41
4. Proofs of Theorems 2.5 and 2.7 44
5. Proof of Theorem 2.2 46
6. No decay for too big classes of observables 48

Bibliography 51

Contents

Abstract

In this work we study mixing properties of discrete dynamical systems and related to them geometric structure. In the first chapter we show that the direct product of maps with Young towers admits a Young tower whose return times decay at a rate which is bounded above by the slowest of the rates of decay of the return times of the component maps. An application of this result, together with other results in the literature, yields various statistical properties for the direct product of various classes of systems, including Lorenz-like maps, multimodal maps, piecewise C^2 interval maps with critical points and singularities, Hénon maps and partially hyperbolic systems.

The second chapter is dedicated to the problem of decay of correlations for continuous observables. First we show that if the underlying system admits Young tower then the rate of decay of correlations for continuous observables can be estimated in terms of modulus of continuity and the decay rate of tail of Young tower. In the rest of the second chapter we study the relations between the rates of decay of correlations for smooth observables and continuous observables. We show that if the rates of decay of correlations is known for C^r, observables ($r \geq 1$) then it is possible to obtain decay of correlations for continuous observables in terms of modulus of continuity.

Dedication

I dedicate this thesis to my mother Dilorom Ruzimova and to my wife Assem Kaliyeva for their support and love.

Acknowledgments

I am highly indebted to my supervisor Stefano Luzzatto for his encouragement and continuous attention. I owe my deepest gratitude to ICTP and SISSA opening joint PhD program.

Thanks to all my colleagues and friends with whom I enjoyed many useful discussions and conversations during my PhD, in particular Oliver Butterley, Shokhrukh Kholmatov, Benjamin Kikwai, Maksim Smirnov, Sina Türeli, Khadim M. War.

Introduction

In the present work we consider ergodic properties of discrete time dynamical systems and related geometric structure to them. In the 70's Sinai [**63**], Ruelle [**61**] and Bowen [**15**] introduced particular geometric structure - Markov partitions to study smooth uniformly hyperbolic systems. Markov partitions are used to conjugate the system to symbolic dynamics over finite alphabet and this allowed to use powerful methods of statistical mechanics in dynamical systems. Since then many ergodic and statistic properties including exponential decay of correlations, various limit theorems, invariance principles has been proven for uniformly hyperbolic systems, and the theory is more or less complete now.

Subsequently, the notions of non-uniform hyperbolicity and partial hyperbolicity were introduced [**54, 55, 56, 57**]. But generalizing the above results in this setting of weaker hyperbolicity turned out to be very challenging problem and it only had partial success. The results obtained for particular classes of systems which are not uniformly hyperbolic (for example, quadratic family on interval, billiards, Hénon maps) motivated systematic study of non-uniformly hyperbolic systems and a natural analogue of Markov partitions for non-uniformly hyperbolic systems was introduced by L.-S. Young in [**68, 69**]. Nowadays it is often referred as Gibbs-Markov-Young (GMY)-structure. The key difference from Markov partition is that it is a local structure with countable partition. More precisely, it is defined as a countable partition of a reference set and each element is assigned return time in such a way that the system maintain uniformly hyperbolic behaviour at return times and partition satisfies the Markov property for the return map. She shows that existence of such a structure with integrable return time implies existence of SRB-measures and the rate of decay of correlations is related to the tail of the return time. Moreover gives constructions for some examples of systems such as Sinai billiards, maps with indifferent fixed point and Hénon maps. Later I. Melbourne and M.Nicol showed the existence of GMY-structure with rapidly decaying tail implies the large deviations principles and almost sure invariant principles [**49, 50**]. These results show that to obtain statistical properties of a dynamical systems it is sufficient to show the existence of GMY-structures, and naturally many works devoted to the constructions of GMY-structures for various classes of maps, see for example [**17, 25, 26**] for one dimensional maps with critical and singular points, [**2, 3, 5, 30**] for non-uniformly expanding maps, partially hyperbolic maps with "weakly" expanding direction [**1, 4**]. But still there is no complete understanding of the class of systems which admit GMY-structures and is still active area of research.

One way of constructing new dynamical system from given systems is taking the direct product of given systems. It is one the few general constructions in ergodic theory, where the properties of new system can be obtained in terms of properties of given ones [**24**]. The second chapter of the current works is devoted to construction of Young towers for direct product of systems which admit Young towers. The direct product of Young towers itself is not in general itself a Young tower, and so it is not immediately obvious that the direct product of systems which admit a Young tower also admits a Young tower. The main result of chapter 1 is to show that, Young tower can be constructed for the product system if each component admits a Young tower, and that we can obtain some estimates for the decay of the return times of the tower for the product in terms of the rates of decay for the individual towers. We will also discuss various applications of this result to products of systems for which Young towers are known to exist.

It is known that the decay rates of correlation for Hölder continuous observables can be related to the decay rates of tail of the return time of GMY-tower [**6, 68, 69**]. In chapter 2 we show that these results can be extended to uniformly continuous observables defined on any compact metric spaces. Precisely we obtain decay rates of correlations for uniformly continuous observables for the systems that admit Young towers with certain tails in terms of modulus of continuity of the observables. In the second part of chapter 2 we show that for a system defined on a compact manifolds if the rates of decay of correlations is known for smooth observables then by a simple approximation argument it is possible to obtain the rates of decay of correlations for uniformly continuous observables provided that the cofactor of the bound of correlations has special form, namely bounded by the norms of observables in the corresponding Banach spaces.

In the rest of this chapter we give main definitions used in the work.

1. Basic definitions

Let (M, μ) be a measure space and $f : M \to M$ is a measurable map preserving μ. The measure μ is called mixing if for any μ- measurable sets A and B

$$\mu(f^{-n}(A) \cap B) \to \mu(A)\mu(B) \quad \text{as} \quad n \to \infty.$$

The existence of mixing measure shows the chaotic behaviour of the system. A natural way to classify the mixing systems is through the rates of mixing i.e. the rates at which the above convergence occurs. The rates of mixing defined as follows.

Definition 0.1. Let $\mathcal{B}_1$, $\mathcal{B}_2$ be Banach spaces of measurable observables defined on M. We denote the *correlation* of observables $\varphi \in \mathcal{B}_1$ and $\psi \in \mathcal{B}_2$ with respect to μ by

$$C_n(\varphi, \psi; \mu) := \left| \int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu \right|.$$

Let $\{\gamma_n\}$ be a sequence of positive numbers such that

$$\lim_{n\to\infty} \gamma_n = 0.$$

We say that (f, μ) has *decay of correlations at rate* $\{\gamma_n\}$ with respect to μ for observables in $\mathcal{B}_1$ against observables in $\mathcal{B}_2$ if there exists constant $C = C(\varphi, \psi) > 0$ such that for any $\varphi \in \mathcal{B}_1$, $\psi \in \mathcal{B}_2$ the inequality

$$C_n(\varphi, \psi; \mu) \leq C\gamma_n$$

holds for all $n \in \mathbb{N}$.

Remark 0.2. The system (f, μ) is mixing if and only if for any $\varphi, \psi \in L^2(M, \mu)$

$$C_n(\varphi, \psi; \mu) \to 0 \quad \text{as} \quad n \to \infty.$$

Remark 0.3. As standard counterexamples show that can be no specific rate of if the spaces $\mathcal{B}_1 =$ and $\mathcal{B}_2$ are too big, for example $\mathcal{B}_1$ $\mathcal{B}_2 = L^2(M, \mu)$. This can be formulated in this case by saying that there exist observables $\varphi, \psi \in L^2(\mu)$ such that $C_n(\varphi, \psi; \mu)$ decays arbitrarily slowly. Nevertheless, requiring some regularity of observables sometimes it is possible to obtain the rates od decay of correlations. It turns out that for non-invertible systems it is sufficient if one of the observables is "regular" (usually Hölder continuous or function of bounded variation), whereas for invertible systems it is necessary both of the observables to be regular (see chapter 3).

Rates of decay of correlations have been extensively studied and are well known for many classes of systems and various families of observables, indeed the literature is much too vast to give complete citations. We just mention that the first results on rates of decay of correlations go back at least to [**15, 38, 61, 63**] in the 70's and since then results have been obtained in [**7, 10, 15, 17, 18, 22, 30, 31, 36, 33, 40, 39, 44, 52, 60, 66, 67, 68, 69**] amongst others.

Notice that in the above works observables are assumed to be Hölder continuous, or functions of bounded variation. Hence a natural question to ask is to what extend can we generalize the classes of observables that still admit some decay rate? Several papers address this question in the context of one-sided subshifts of finite type on a finite alphabet, for example see [**16, 29, 37, 34, 59**]. For a comprehensive discussion of shift maps and their ergodic properties we refer to [**9**]. Moreover, there are results on non-invertible Young Towers, for example [**18, 37, 46, 48**], and results that apply directly to certain non-uniformly expanding systems [**60**]. We emphasize that all of the results we mention above are for non-invertible maps and in the invertible case the only reference we found is [**71**], which gives interesting estimates for billiards with non-Hölder observables.

1.1. Modulus of continuity. In some parts of the present work the regularity of the observables will be crucial. We define the regularity in terms of the *modulus continuity* of observables. Let M be a metric space and $\mathcal{C}(M)$ be the space of continuous functions defined on it. Define modulus of continuity for $\varphi \in \mathcal{C}(M)$ as

$$\mathcal{R}_\varphi(\varepsilon) = \sup\{|\varphi(x) - \varphi(y)| : \ d(x, y) < \varepsilon\}.$$

Obviously, $\mathcal{R}_\varphi(\varepsilon) \to 0$ as $\varepsilon \to 0$ if and only if φ is uniformly continuous. If φ is Hölder continuous with Hölder exponent α then $\mathcal{R}_\varphi(\varepsilon) = \mathcal{O}(\varepsilon^\alpha)$. In general, $\mathcal{R}_\varphi(\varepsilon)$ might converge to 0 very slowly for continuous observables. For example if we take $\varphi : [0, 1] \to \mathbb{R}$ as

$$\varphi(x) = \begin{cases} |\log x|^{-\gamma} & \text{if} \quad x > 0 \\ 0 \quad \text{if} \quad x = 0 \end{cases}$$

for some $\gamma > 1$ then $\mathcal{R}_\varphi(\varepsilon) = |\log x|^{-\gamma}$. We refer to [**32**] for the examples of various slow rates of convergence.

Given observables $\varphi, \psi \in \mathcal{C}(M)$ define

$$\mathcal{R}_{\varphi,\psi}(\varepsilon) = \max\{\mathcal{R}_\varphi(\varepsilon), \mathcal{R}_\psi(\varepsilon)\}. \tag{1}$$

In the second chapter we will show that the rate of decay of correlations can be estimated in terms of $\mathcal{R}_{\varphi,\psi}$.

2. Young Towers

The definition of Young tower is slightly different for expanding systems and for the systems with contracting directions. Since, expanding systems on compact manifolds are non-invertible and most of the well studied examples of systems with contracting directions are diffeomorphisms we refer to expanding systems as non-invertible systems and for the systems with contracting directions as invertible systems.

2.1. Gibbs-Markov-Young Towers. We start with the formal definition of Young Tower for non-invertible maps. To distinguish this case from the tower for invertible maps, this structure sometimes is referred to as a Gibbs-Markov-Young (GMY) structure or GMY-tower. Let $f : M \to M$ be a $C^{1+\alpha}$, $\alpha \in (0, 1)$ local diffeomorphism (outside some critical/singular set) of a Riemannian manifold M on which we have a Riemannian volume which we will refer to as Lebesgue measure. To define a GMY-structure for f we start with a measurable set Δ_0 with finite positive Lebesgue measure, we let m_0 denote the restriction of Lebesgue measure to Δ_0, a mod 0 partition $\mathcal{P} = \{\Delta_{0,i}\}$ of Δ_0, and a return time function $R : \Delta_0 \to \mathbb{N}$ that is constant on the partition elements, i.e. $R|\Delta_{0,i} = R_i$ such that the induced map $f^R : \Delta_0 \to \Delta_0$ is well defined. Then, for any two point $x, y \in \Delta_0$ we can define the *separation time* $s(x, y)$ as the smallest $k \geq 0$ such that $(f^R)^k(x)$ and $(f^R)^k(y)$ lie in different partition elements and assume that there exists $\beta \in (0, 1)$ and $D > 0$ such that the following conditions are satisfied.

(G1) **Markov**: for any $\Delta_{0,i} \in \mathcal{P}$ the map $f^{R_i} : \Delta_{0,i} \to \Delta_0$ is a bijection.

(G2) **Uniform expansion**: $\|(Df^R)^{-1}(x)\| \leq \beta$ for m_0 a.e. x.

(G3) **Bounded distortion**: for a.e. pair $x, y \in \Delta_0$ with $s(x, y) < \infty$ we have

$$\left| \frac{\det Df^R(x)}{\det Df^R(y)} - 1 \right| \leq D\beta^{s(f^R x, f^R y)}. \tag{2}$$

(G4) **Integrability**: $\int R\, dm_0 < \infty$.

(G5) **Aperiodicity**: $\gcd\{R_i\} = 1$.

Young showed that the first three assumptions (G1)-(G3) imply that the map f admits an f-invariant measure μ which is absolutely continuous with respect to Lebesgue. Condition (G4) implies that this measure is finite, and therefore can be taken to be a probability measure, and condition (G5) implies that it is mixing. We note that the integrability and the aperiodicity assumption are not always included in the definition of a Young Tower. We include them here because we actually require the existence of a mixing probability measure for each component for our argument to work in the proof of the Theorem and also because we are interested in applications to the problem of the rate of decay of correlations for product systems and this requires the measures involved to be mixing probability measures.

We note also that the return time R is not generally a *first return* time of point to Δ_0. It is therefore often useful to work with an "extension" of f in which the returns to Δ_0 are first return times. This extension is precisely what we refer to as a *GMY-tower*. The formal construction of this extension proceeds as follows. We let

$$\Delta = \{(z, n) \in \Delta_0 \times \mathbb{Z}_0^+ |\ R(z) > n\}, \tag{3}$$

where $\mathbb{Z}_0^+$ denotes the set of all nonnegative integers. For $\ell \in \mathbb{Z}_0^+$ the subset $\Delta_\ell = \{(\cdot, \ell) \in \Delta\}$ of Δ is called its *ℓth level*. By some slight abuse of notation, we let Δ_0 denote both the subset of the Riemannian manifold M on which the induced map f^R is defined and the 0'th level of tower Δ. The collection $\Delta_{\ell,i} := \{(z, \ell) \in \Delta_\ell |\ (z, 0) \in \Delta_{0,i}\}$ forms a partition of Δ that we denote by η. The set $\Delta_{R_i-1,i}$ is called the *top level* above $\Delta_{0,i}$. We can then define a map $F : \Delta \to \Delta$

letting

$$F(z,\ell)=\begin{cases}(z,\ell+1) & \text{if } \ell+1<R(z),\\ (f^{R(z)}(z),0) & \text{if } \ell+1=R(z).\end{cases} \tag{4}$$

There exists a natural projection $\pi:\Delta\to M$ defined by $\pi(x,\ell)=f^{\ell}(x_0)$ for $x\in\Delta$ with $F^{-\ell}(x)=x_0\in\Delta_0$. Notice that π is a semi-conjugacy $f\circ\pi=\pi\circ F$.

For future reference we note that we can extend the return time and the separation time to all of Δ. Indeed, for any $x\in\Delta$ we can define a *first hitting time* by

$$\hat{R}(x):=\min\{n\geq 0: F^n(x)\in\Delta_0\}. \tag{5}$$

Notice that if $x\in\Delta_0$ then $\hat{R}(x)=0$. We also extend the separation time to Δ by setting $s(x,y)=0$ if x and y belong to different elements of η and $s(x,y)=s(\tilde{x},\tilde{y})$ if x and y are in the same element of η, where $\tilde{x}$ and $\tilde{y}$ are the corresponding projections to Δ_0.

We define a reference measure m on Δ as follows. Let $\mathcal{A}$ be the Borel σ-algebra on Δ_0 and let m_0 denote the restriction of Lebesgue measure to Δ_0 where, as mentioned above, we are identifying $\Delta_0\subset M$ with the 0'th level of the tower. For any $\ell\geq 0$ and $A\subset\Delta_\ell$ such that $F^{-\ell}(A)\in\mathcal{A}$ define $m(A)=m_0(F^{-\ell}(A))$. Notice that with this definition, the restriction of m to Δ_0 is exactly m_0, whereas the restriction of m to the upper levels of the tower is not equal to the Riemannian volume of their projections on M because the tower map $F:\Delta\to\Delta$ is by definition an isometry between one level and the next, except on the top level where it maps back to the base with an expansion which corresponds to the "accumulated" expansion of all the iterates on the manifold. Correspondingly, for every $x\in\Delta_{R_i-1,i}\subset\Delta$, the Jacobian of F at x is $JF(x)=\det Df^R(x_0)$, where $x_0=F^{-R_i+1}x\in\Delta_0$, and $JF(x)=1$ otherwise.

2.2. Young Towers for invertible systems. We start with the definition of Young towers in this setting, following [**6**] which generalizes the definition of [**68**] (see Remarks 2.3-2.5 from [**6**]). We change the notation slightly to distinguish this case from non-invertible case. Let M be a Riemannian manifold. If $\gamma\subset M$ is a submanifold, then m_γ denotes the restriction of the Riemannian volume to γ. Consider $f:M\to M$ such that $f:M\setminus\mathcal{C}\to M\setminus f(\mathcal{C})$ is a $C^{1+\varepsilon}$ diffeomorphism on each connected component of $M\setminus\mathcal{C}$ for some $\mathcal{C}\subset M$ with the following properties.

(A1) There exists $\Lambda\subset M$ with hyperbolic product structure, i.e. there are families of stable and unstable manifolds $\Gamma^s=\{\gamma^s\}$ and $\Gamma^u=\{\gamma^u\}$ such that $\Lambda=(\cup\gamma^s)\cap(\cup\gamma^u)$; $\dim\gamma^s+\dim\gamma^u=\dim M$; each γ^s meets each γ^u at a unique point; stable and unstable manifolds are transversal with angles bounded away from 0; $m_{\gamma^u}(\gamma^u\cap\Lambda)>0$ for any γ^u.

Let Γ^s and Γ^u be the defining families of Λ. A subset $\Lambda_0 \subset \Lambda$ is called s-subset if Λ_0 also has a hyperbolic structure and its defining families can be chosen as Γ^u and $\Gamma_0^s \subset \Gamma^s$. Similarly, we define u-subsets. For $x \in \Lambda$ let $\gamma^\theta(x)$ denote the element of Γ^θ containing x, where $\theta = u, s$.

(A2) There are pairwise disjoint s-subsets $\Lambda_1, \Lambda_2, ..., \subset \Lambda$ such that $m_{\gamma^u}((\Lambda \setminus \cup\Lambda_i) \cap \gamma^u) = 0$ on each γ^u and for each Λ_i, $i \in \mathbb{N}$ there is R_i such that $f^{R_i}(\Lambda_i)$ is u-subset; $f^{R_i}(\gamma^s(x)) \subset \gamma^s(f^{R_i}(x))$ and $f^{R_i}(\gamma^u(x)) \supset \gamma^u(f^{R_i}(x))$ for any $x \in \Lambda_i$.

(A3) There exist constants $C \geq 1$ and $\beta \in (0,1)$ such that $\text{dist}(f^n(x), f^n(y)) \leq C\beta^n$, for all $y \in \gamma^s(x)$ and $n \geq 0$;

(A4) Regularity of the stable foliation: given $\gamma, \gamma' \in \Gamma^u$ define $\Theta : \gamma' \cap \Lambda \to \gamma \cap \Lambda$ by $\Theta(x) = \gamma^s(x) \cap \gamma$. Then

(a) Θ is absolutely continuous and

$$u(x) := \frac{d(\Theta_* m_{\gamma'})}{dm_\gamma}(x) = \prod_{i=0}^{\infty} \frac{\det Df^u(f^i(x))}{\det Df^u(f^i(\Theta^{-1}(x)))};$$

(b) There exists $C > 0$ and $\beta < 1$ such that, letting the separation time $s(x,y)$ be the smallest k where $(f^R)^k(x)$ and $(f^R)^k(y)$ lie in different partition elements, we have

$$\log \frac{u(x)}{u(y)} \leq C\beta^{s(x,y)} \quad \text{for} \quad x, y \in \gamma' \cap \Lambda.$$

(A5) Bounded distortion: for $\gamma \in \Gamma^u$ and $x, y \in \Lambda \cap \gamma$

$$\log \frac{\det D(f^R)^u(x)}{\det D(f^R)^u(y)} \leq C\beta^{s(f^R(x), f^R(y))}.$$

(A6) $\int R \, dm_0 < \infty$, where m_0 is the restriction of Lebesgue measure to Λ.

(A7) $\gcd\{R_i\} = 1$.

Given such a structure we can define Young-tower as we define Δ in the non-invertible case. This time we denote the tower by $\mathcal{T}$.

CHAPTER 1

Direct Product Systems

1. Ergodic properties of product systems

Let $f_i : M_i \to M_i$, $i = 1, ..., \ell$, be a family of maps defined on a family of Riemannian manifolds. Define the product map $f = f_1 \times ... \times f_\ell$ on $M = M_1 \times ... \times M_\ell$ by

$$f(x_1, ..., x_\ell) = (f_1(x_1), .., f_\ell(x_\ell)). \tag{6}$$

Dynamical properties of the product system f can be partially but not completely deduced from the dynamical properties of its components, as the product system may exhibit significantly richer dynamics. Even in the simplest setting of the product of two identical maps $g \times g : M \to M$, if p, q are periodic points for g, then the set-theoretic product of the periodic *orbits* $\mathcal{O}^+(p) = \{p_1, ..., p_m\}$ and $\mathcal{O}^+(q) = \{q_1, ..., q_n\}$ may consist of several periodic orbits[1], which are in some sense *new* periodic orbits which do not exist in either of the original systems. Similarly, from an ergodic-theoretic point of view, the space $\mathcal{M}_{g\times g}$ of invariant probability measures for the product map contains many invariant measures which are not products of invariant measures for g such as measures supported on the "new" periodic orbits mentioned above [2] and measures supported on the "diagonal" $\{(x, x) : x \in M\} \subset M \times M$ which is invariant for the product.[3]

However, if we consider product system with product measure we can say much more about the behaviour of the system. More precisely we have the following

[1]In fact, if $\gcd\{m, n\} = 1$, then (p, q) is periodic point of period mn. On the other hand, if $\gcd\{m, n\} = k > 1$, then (p, q) is periodic point with period mn/k, but $\mathcal{O}^+(p) \times \mathcal{O}^+(q)$ is a union of k periodic orbits. As a simple example, consider the case $m = 2$ and $n = 4$. In this case, the product of the orbits splits into two orbits for the product map: $\mathcal{O}^+(p) \times \mathcal{O}^+(q) = \mathcal{O}^+(p_1 \times q_1) \cup \mathcal{O}^+(p_2 \times q_1)$.

[2]As above let p and q be periodic points of period m and n respectively for f and g. Then the Dirac measures $\delta_{\mathcal{O}^+(p)} = \frac{1}{m} \sum \delta_{p_i}$ and $\delta_{\mathcal{O}^+(q)} = \frac{1}{n} \sum \delta_{q_j}$ preserved by f and g respectively. In the case $\gcd\{m, n\} = k > 1$, the Dirac measure defined as $\delta_{\mathcal{O}^+(p\times q)} = \frac{k}{mn} \sum \delta_{p_i \times q_j}$, $p_i \times q_j \in \mathcal{O}^+(p \times q)$ is preserved by $f \times g$, but it is not a product of measures $\delta_{\mathcal{O}^+(p)}$ and $\delta_{\mathcal{O}^+(q)}$.

[3]Thanks to M. Blank for this interesting observation.

PROPOSITION 1.1. *Let $f_i : M_i \to M_i$, $i = 1, ..., \ell$ be a map of probability space $(M_i, \mathcal{B}_i, \mu_i)$ for $i = 1, ..., \ell$. Let $f : M \to M$ denote the product map and $\mu = \mu_1 \times \cdots \times \mu_\ell$ product measure on product space M.*

(i) *If μ_i is invariant for f_i for all f_i, $i = 1, ..., \ell$ then it is invariant for f.*

(ii) *If μ_i is ergodic f_i for all $i = 1, ..., \ell$ then μ is ergodic for f.*

(iii) *If (f, μ_i) is mixing for all $i = 1, ..., \ell$ then (f, μ) is mixing.*

PROOF. Note that it is sufficient to prove the proposition only for measurable rectangles i.e. the sets of form $A_1 \times \cdots \times A_2$ where $A_i \in \mathcal{B}_i$, $i = 1, ..., \ell$ and then the results can be extended to all measurable sets in the product space by standard extension theorem [**14**].

(i) Assume that μ_i is invariant for f_i, $i = 1, ..., \ell$. Take any $A_i \in \mathcal{B}_i$, $i = 1, ..., \ell$ then $f^{-1}(A_1 \times \cdots \times A_\ell) = f_1^{-1}(A_1) \times \cdots \times f_\ell^{-1}(A_\ell)$. Therefore,

$$\mu(f^{-1}(A_1 \times \cdots \times A_\ell)) = \prod_{i=1}^{\ell} \mu_i(f_i^{-1}(A_i)) = \prod_{i=1}^{\ell} \mu_i(A_i) = \mu(A_1 \times \cdots \times A_\ell)$$

by the invariance of μ_i for f_i and the definition of product measure.

(iii) Suppose that the systems (f_i, μ_i), $i = 1, ..., \ell$ are mixing and let μ be the product measure. Let $A_1 \times \cdots \times A_\ell$ and $B_1 \times \cdots \times B_\ell$ be two measurable rectangles in the product space. Then for any n we have

$$\mu\left(f^{-n}(A_1 \times \cdots \times A_\ell) \cap (B_1 \times \times \cdots \times B_\ell)\right) = \prod_{i=1}^{\ell} \mu_i(f_i^{-n}(A_i) \cap B_i).$$

Since the individual systems are mixing we obtain

$$\mu\left(f^{-n}(A_1 \times \cdots \times A_\ell) \cap (B_1 \times \times \cdots \times B_\ell)\right) \to \mu(A_1 \times \cdots \times A_\ell))\mu(B_1 \times \cdots \times B_\ell)$$

as $n \to \infty$. This proves mixing for measurable rectangles.

The proof of item (ii) relies on spectral properties of direct product and its components and we refer to [**24**] for the proof.

$\square$

With little more work it is possible to obtain more deeper properties of product systems with product measure such as the rate of decay of correlations, if the rates of decay of correlations known for component systems.

Let $f_i : (M_i, \mu_i) \circlearrowleft$, $i = 1, ..., p$, be a family of maps defined on compact metric spaces M_i, preserving Borel probability measures μ_i as above let $f : M \to M$ denote the direct product map on product space M and μ the product measure. We let $\mathcal{A}_i$, $\mathcal{B}_i$ denote Banach spaces of functions on M_i and $\mathcal{A}$, $\mathcal{B}$ denote Banach spaces of functions on M. For completeness we state the minimal requirements on the spaces $\mathcal{A}_i, \mathcal{B}_i, \mathcal{A}, \mathcal{B}$ needed for the calculations to work. We assume these spaces satisfy the following properties:

(1) For all $1 \leq i \leq p$ $\mathcal{A}_i, \mathcal{B}_i \subset L^2(M_i, \mu_i)$ and $\mathcal{A}, \mathcal{B} \subset L^2(M,\mu)$.

(2) For any $\varphi \in \mathcal{A}$, $\psi \in \mathcal{B}$ for all $1 \leq i \leq p$ and $\mu_1 \times \cdots \times \mu_{i-1} \times \mu_{i+1} \times \cdots \mu_p$ - almost every $(x_1, ..., x_{i-1}, x_{i+1}, ..., x_p)$ we have $\hat{\varphi}_i := \varphi(x_1, ..., x_{i-1}, \cdot, x_{i+1}, ..., x_p) \in \mathcal{A}_i$, and $\hat{\psi}_i := \psi(x_1, ..., x_{i-1}, \cdot, x_{i+1}, ..., x_p) \in \mathcal{B}_i$.

(3) There exists $C > 0$ such that for any $\varphi \in \mathcal{A}, \psi \in \mathcal{B}$ and any $1 \leq i \leq p$, we have
$$\|\hat{\varphi}_i\|_{\mathcal{A}_i} \leq C\|\varphi\|_{\mathcal{A}} \quad \text{and} \quad \|\hat{\psi}_i\|_{\mathcal{B}_i} \leq C\|\psi\|_{\mathcal{B}}.$$

It is easy to check that all these conditions are satisfied for some of the commonly considered classes of observables such as Hölder continuous or essentially bounded functions.

THEOREM 1.2. *Suppose $Cor_\mu(\varphi_i, \psi_i \circ f_i^n) \leq C\gamma_n$ for all non-zero $\varphi_i \in \mathcal{A}_i$ and $\psi_i \in \mathcal{B}_i$ for all $i = 1, ..., p$. Then there exists a constant $\bar{C} > 0$ such that for all non-zero $\varphi \in \mathcal{A}$,$\psi \in \mathcal{B}$, and for all $n \geq 0$ we have*
$$Cor_\mu(\varphi, \psi \circ f^n) \leq \bar{C}\gamma_n.$$

We now suppose that
$$\mathrm{Cor}_\mu(\varphi_i, \psi_i \circ f_i^n) \leq C\gamma_n$$
for all non-zero $\varphi_i \in \mathcal{A}_i$ and $\psi_i \in \mathcal{B}_i$ for all $i = 1, ..., p$ and obtain a bound for the correlation function of the product. For simplicity we give the calculation for $p = 2$, the general case follows by successive applications of the argument. Let $\varphi \in \mathcal{A}$ be such that $\int \varphi d\mu = 0$ and $\psi \in \mathcal{B}$. Moreover, let $\bar{\varphi}(x_1) = \int \varphi(x_1, y)d\mu_2(y)$. If we fix the first coordinate, by Fubini's theorem we have
$$\begin{aligned}
&\int \varphi(x_1, x_2)\psi(f_1^n(x_1), f_2^n(x_2))d\mu_1 d\mu_2 \\
&= \int \left(\int \psi(f_1^n(x_1), f_2^n(x_2))[\varphi(x_1, x_2) - \bar{\varphi}(x_1)]d\mu_2 \right) d\mu_1 \\
&+ \int \bar{\varphi}(x_1)\psi(f_1^n(x_1), f_2^n(x_2))d\mu_1 d\mu_2.
\end{aligned} \tag{7}$$

Since μ_2 is f_2-invariant, we can write the first term of the right hand side as
$$\begin{aligned}
I := &\int \psi(f_1^n(x_1), f_2^n(x_2))\varphi(x_1, x_2)d\mu_2 - \bar{\varphi}(x_1) \int \psi(f_1^n(x_1), f_1^n(x_2))d\mu_2 \\
&\leq C\|\varphi(x_1, \cdot)\|_{\mathcal{A}_2}\|\psi(x_1, \cdot)\|_{\mathcal{B}_2}\gamma_n.
\end{aligned}$$

The inequality above follows since the left hand side gives the correlations of the second component with respect to μ_2. From the third assumption we obtain $I \leq C\|\varphi\|_{\mathcal{A}}\|\psi\|_{\mathcal{B}}$. Again by the invariance of μ_2 we can write the second summand of the equation (7) as

$$\int \psi(f_1^n(x_1), f_2^n(x_2))\bar{\varphi}(x_1)d\mu_1 d\mu_2 = \int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1 d\mu_2.$$

Note that $\int \bar{\varphi}(x_1)d\mu_1 = 0$ by the choice of φ. Then this expression can be written as

$$\begin{aligned}
&\int \left(\int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1 \right) d\mu_2 \\
&= \int \left(\int \psi(f_1^n(x_1), x_2)[\bar{\varphi}(x_1)d\mu_1 - \int \bar{\varphi}(x_1)d\mu_1] \right) d\mu_2 \\
&= \int \left(\int \psi(f_1^n(x_1), x_2)\bar{\varphi}(x_1)d\mu_1 - \int \psi(x_1, x_2)d\mu_1 \int \bar{\varphi}(x_1)d\mu_1 \right) d\mu_2.
\end{aligned}$$

The expression under the integral with respect to μ_2 is exactly the correlation with respect to μ_1. Again using the third property of the Banach spaces $\mathcal{A}$ and $\mathcal{B}$ we have

$$\mathrm{Cor}_\mu(\varphi, \psi \circ f_1 \times f_2) \leq 2C\gamma_n.$$

This completes the proof of the required statement. [4]

It seems it is impossible to obtain more finer properties such as large deviations principle, local limit theorem to obtain directly by this method. Hence we need to use some geometric properties of product system, which will be the topic of next section.

2. Geometric properties of product systems

Now we turn to study of geometric properties of direct product systems. Suppose that each of our maps $f_i : M_i \to M_i$, $i = 1, ..., \ell$ admit GMY- towers with reference measures $m_0^{(i)}$ defined on bases $\Delta_0^{(i)}$ with return time functions $R^{(i)} : \Delta_0^{(i)} \to \mathbb{N}$. Let $f : M \to M$ denote the corresponding product system (6), and $\bar{m}_0 = m_0^{(1)} \times \cdots \times m_0^{(\ell)}$ denote the product measure on $\bar{\Delta}_0 = \Delta_0^{(1)} \times \cdots \times \Delta_0^{(\ell)}$. Finally, let

$$\mathcal{M}_n = \max_{i=1,..,\ell} m_0^{(i)}\{R^{(i)} > n\}. \tag{8}$$

We always assume that either all of the maps are invertible or all are noninvertible so that the product system is of the same type.

[4] Thanks to Professor Liverani for pointing out this proof.

THEOREM 1.3. *The product system f admits a tower with reference measure $\bar{m}_0$ on $\bar{\Delta}_0$ and return time $T : \Delta_0 \to \mathbb{N}$ satisfying the following bounds.*

Exponential Decay: *if $\mathcal{M}_n = \mathcal{O}(e^{-\tau n})$ for some $\tau > 0$, then*

$$\bar{m}_0\{T > n\} = \mathcal{O}(e^{-\tau' n}).$$

for some $\tau' > 0$.

Stretched Exponential Decay *if $\mathcal{M}_n = \mathcal{O}(e^{-\tau n^{\theta}})$ for some $\tau, \theta > 0$, then*

$$\bar{m}_0\{T > n\} = \mathcal{O}(e^{-\tau' n^{\theta'}});$$

for all $0 < \theta' < \theta$ and some $\tau' = \tau'(\theta') > 0$.

Polynomial Decay *if $\mathcal{M}_n = \mathcal{O}(n^{-\alpha})$ for some $\alpha > \ell$, then*

$$\bar{m}_0\{T > n\} = \mathcal{O}(n^{\ell-\alpha}).$$

The idea of constructing a Young tower for a product of Young towers is sketched, in the non-invertible setting with exponential or polynomial return times, without a fully developed proof, in the PhD thesis of Vincent Lynch [**45**]. Our construction and estimates lead to a "loss" of one exponent for each component of the product system in the case of polynomial rates of decay, and also do not allow us to obtain results if the decay is slower than polynomial, such as in the interesting examples of Holland [**32**] which exhibit decay at rates of the form $(\log \circ ... \circ \log n)^{-1}$. It is not clear to us if this is just a technical issue or if there might be some deeper reasons.

2.1. Applications. As we pointed out previously, Young towers have been shown to imply a variety of statistical properties such as decay of correlations, invariance principles, limit theorems which in some cases can also be quantified in terms of the rate of decay of the tail of the return times associated to the tower [**6, 20, 21, 30, 46, 49, 50, 68, 69**]. An immediate consequence is that the dynamical systems which are direct products of systems which admit Young towers satisfy the statistical properties corresponding to the tail estimates of the product as given in our Theorem. Some examples of these systems include the following. *Lorenz-like interval maps* which are uniformly expanding and have a single singularity with dense preimages satisfying $|f'(x)| \approx |x|^{-\beta}$ for some $\beta \in (1/2, 1)$ admit a Young Tower with exponential tail, see [**25**] for the precise technical conditions; *Multimodal maps* for which the decay rate of the return times was obtained in terms of the growth rate of the derivative along the critical orbits [**17**]; *Maps with critical points and singularities* in one and also higher dimensions [**5, 26, 7**]; *Planar periodic Lorentz gas* was introduced by Sinai [**64**] and admits Young Towers with exponential tails [**69, 21**]; *Hénon maps*, for certain choices of parameters (a, b), the maps $H_{a,b} : \mathbb{R}^2 \to \mathbb{R}^2$ given by $H_{a,b}(x, y) = (1 - ax^2 + y, bx)$ admit a Young Tower with exponential tail [**11, 12**]; *Partially hyperbolic systems* under certain additional conditions [**4, 6**].

We emphasize that the construction of Young towers, and especially the estimation of the decay of the return times, is in general highly non-trivial and relies on the specific geometric and dynamical properties of the system under consideration. The geometry of the direct product of any of the systems mentioned above is extremely complicated and it is doubtful that a tower construction could be achieved without taking advantage of the information that each component admits a tower with certain decay rates and applying our theorem.

In certain cases, combining our result with existing literature, it is possible to deduce statistical properties of the product system directly from the statistical properties of the component systems without assuming a priori that these admit Young towers. Indeed, in certain settings, such as that of non-uniformly expanding systems, statistical properties such as decay of correlations or large deviations imply the existence of Young tower with corresponding tail estimates see e.g. [**3**, Theorem 4.2]. Thus, taking the direct product of any finite number of such systems we can apply our result and those of [**49**] to conclude that Large Deviations for the component systems implies Large Deviations for the product system.

3. A tower for the product

In this section we begin the proof of our main result. To simplify the notation we will assume that we have a product of only two systems, the general case follows immediately by iterating the argument.

3.1. Basic ideas and notation. We begin by introducing some basic notions. Suppose $F : (\Delta, m) \circlearrowleft$ is a GMY-tower as defined in (3) and (4) above and let η be the partition of Δ into $\Delta_{\ell,i}$'s. Then, for $n \geq 1$, let

$$\eta_n := \bigvee_{j=0}^{n-1} F^{-j}\eta := \{A_1 \cap F^{-1}(A_2) \cap ... \cap F^{1-n}(A_n) |\ A_1, ..., A_n \in \eta\} \tag{9}$$

be the refinements of the partition η defined by the map F. For $x \in \Delta$ let $\eta_n(x)$ be the element containing x. From (9) it is easily seen that $\eta_n(x)$ has the form

$$\eta_n(x) = \left(\bigvee_{j=0}^{n-1} F^{-j}\eta\right)(x) = \eta(x) \cap F^{-1}\eta(F(x)) \cap ... \cap F^{1-n}\eta(F^{n-1}(x)).$$

REMARK 1.4. To get a better feeling for the partitions η_n notice that from the definition of tower for $x \notin \Delta_0$ we have $F^{-1}(\eta(x)) = \eta(F^{-1}(x))$, which shows that the element $\eta(x)$ gets refined only when $F^{-j}(x) \in \Delta_0$ for some j, $j = 0, ..., n-2$.

It may be instructive to consider more in detail the cases $n = 2$ and $n = 3$ (notice that $\eta_1 = \eta$). For $n = 2$, from (9) we have $F^{-1}\eta \vee \eta$. In this case only the elements on the top levels (recall definition just before equation (4)) get refined so that the new elements are mapped by F bijectively onto $\Delta_{0,i} \subset \Delta_0$, for some i. All the

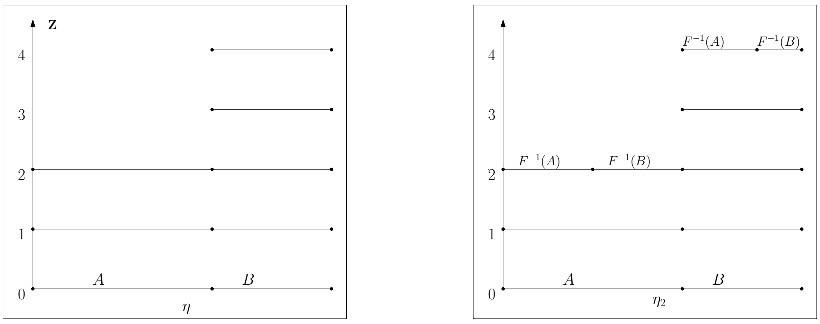

FIGURE 1. η and η_2

other elements remain unchanged, see Figure 1 (for simplicity, the pictures are drawn when the partition of base contains only two elements and has return times 3 and 5, in particular, for $x \in \Delta$ with $F^2(x) \in \Delta_0$ we have $F^2(\eta_2(x)) = \Delta_0$. This is because $\eta_2(x) = \eta(x)$ and $F(\eta(x))$ is top level element of η). For $n = 3$ all the

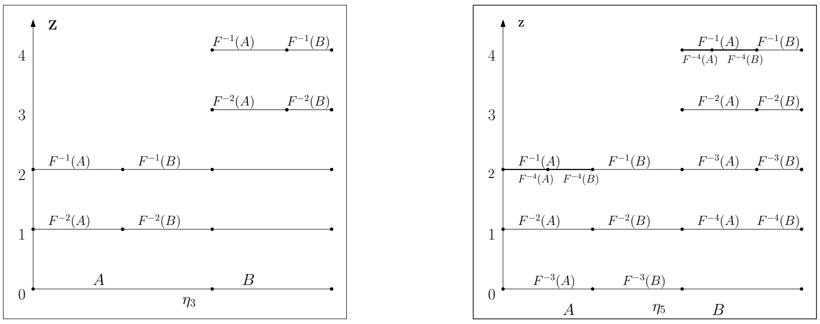

FIGURE 2. η_3 and η_5

elements on the top levels and on the levels "immediately below" the top levels get refined so that the top level elements of the new partition are mapped onto some $\Delta_{0,i}$ by F; the elements belonging to the levels immediately below the top levels are mapped onto some $\Delta_{0,i}$ by F^2 and other elements remain unchanged, see the left hand side of Figure 2. On the right hand side of Figure 2 is illustrated the situation for $n = 5$ in this simple example, where $n = \min_i\{R_i\} + 2$, and therefore the refinement procedure "reaches" Δ_0. After this time, the top levels undergo a "second round" of refinements. In Figure 2 the bold elements of η_5 are the elements of η which have been refined twice. In general, for each n the refinement procedure affect $n-1$ levels below the top levels. If n is sufficiently large, some of the partition elements might get refined several times.

The following statement follows almost immediately from the observation above. Recall also the definition of the first hitting time in (5).

LEMMA 1.5. *For any $x \in F^{-n}(\Delta_0)$ the map $F^n : \eta_n(x) \to \Delta_0$ is a bijection and $F^n(\eta(x)) = \Delta_0$. Moreover if $\hat{R}(x) = n$ then $\hat{R}|\eta_n(x) \equiv n$.*

PROOF. The proof is by induction on n. For $n = 1$ we have no refinement, $\eta_1 = \eta$, and therefore the conclusion follows from the definition of tower. Assume that the assertion is true for $n = k$. From equality (9) we obtain $\eta_{k+1} = F^{-1}(\eta_k) \vee \eta$. Let $x \in \Delta$ be a point, such that $F^{k+1}(x) \in \Delta_0$ then using the relation $\eta_{k+1}(x) = F^{-1}\eta_k(F(x)) \cap \eta(x)$ we obtain $F^{k+1}(\eta_{k+1}(x)) = F^k(\eta_k(F(x))) \cap F^{k+1}(\eta(x))$. Since $F^k(\eta_k(F(x))) = \Delta_0$ from the inductive assumption and $\Delta_0 \subset F^{k+1}(\eta(x))$ from the definition of tower we get $F^{k+1}(\eta_{k+1}(x)) = \Delta_0$. Since the first return time is constant on the elements of η, the second assertion follows. □

3.2. Distortion estimates. We collect here a few simple estimates which hold in general for any Young tower, and which are used mainly in Section 4. Recall the definition of the partitions η_n in (9) and for any $n \geq 1$, let

$$\eta_n^0 = \{A \in \eta_n |\ F^n(A) = \Delta_0\}.$$

LEMMA 1.6. *For any $A \in \eta_n^0$ and $x, y \in A$ the following inequality holds*

$$\left| \frac{JF^n(x)}{JF^n(y)} - 1 \right| \leq D,$$

where D as in (2).

PROOF. The collection η_n^0 is a partition of $F^{-n}\Delta_0$ and for any $x \in \Delta_0$ each $A \in \eta_n^0$ contains a single element of $\{F^{-n}x\}$. For $x \in A$ let $j(x)$ be the number of visits of its orbit to Δ_0 up to time n. Since the images of A before time n will remain in an element of η, all the points in A have the same combinatorics up to time n and so $j(x)$ is constant on A. Therefore $JF^n(x) = (JF^R)^j(\tilde{x})$, for the projection $\tilde{x}$ of x into Δ_0 (i.e. if $x = (z, \ell)$ then $\tilde{x} = (z, 0)$). Thus for any $x, y \in \Delta_0$ from (2) we obtain

$$\left| \frac{JF^n(x)}{JF^n(y)} - 1 \right| = \left| \frac{(JF^R)^j(\tilde{x})}{(JF^R)^j(\tilde{y})} - 1 \right| \leq D. \tag{10}$$

□

Corollary 1.7. *For any $A \in \eta_n^0$ and $y \in A$ we have*

$$JF^n(y) \geq \frac{m(\Delta_0)}{m(A)(1+D)}. \tag{11}$$

Proof. Lemma 1.6 implies $JF^n(x) \leq (1+D)JF^n(y)$. Integrating both sides of this inequality with respect to x over A gives

$$m(\Delta_0) = \int_A JF^n(x)dm \leq JF^n(y)(1+D)m(A).$$

Hence, for any $y \in A$ we have the statement. □

Lemma 1.8. *There exists $M_0 \geq 1$ such that for any $n \in \mathbb{N}$*

$$\frac{dF_*^n m}{dm} \leq M_0.$$

Proof. Let $\nu_n = F_*^n m$. We will estimate the density $dF_*^n m/dm$ at different point $x \in \Delta$ and consider three different cases according to the position of x. First of all, for any $x \in \Delta_0$, from Corollary 1.7 we have

$$\begin{aligned}\frac{d\nu_n}{dm}(x) = \sum_{y \in F^{-n}x} \frac{1}{JF^n y} &\leq (D+1) \sum_{A \in \eta_0^{(n)}} \frac{m(A)}{m(\Delta_0)} \\ &\leq (D+1)\frac{m(\Delta)}{m(\Delta_0)} := M_0.\end{aligned}$$

This proves the case $x \in \Delta_0$. For $x \in \Delta_\ell$ with $\ell \geq n$ we have $F^{-n}(x) = y \in \Delta_{\ell-n}$. Since $JF(y) = 1$ for any $y \in \Delta \setminus \Delta_0$,

$$\frac{d\nu_n}{dm}(x) = \frac{1}{JF^n(y)} = 1.$$

Finally, let $x \in \Delta_\ell$, $\ell < n$. Then for any $y \in F^{-n}x$ the equality $F^{n-\ell}y = F^{-\ell}x \in \Delta_0$ holds. Hence, $JF(F^j y) = 1$ for all $j = n-\ell..., n-1$. Therefore by the chain rule we obtain $JF^n(y) = JF^{n-\ell}(y)$. We can reduce the problem to the first case observing

$$\frac{d\nu_n}{dm}(x) = \sum_{y \in F^{-n}x} \frac{1}{JF^n y} = \sum_{y \in F^{\ell-n}(F^{-\ell}x)} \frac{1}{JF^{n-\ell}y} = \frac{d\nu_{n-\ell}}{dm}(F^{-\ell}(x)).$$

This finishes the proof. □

Lemma 1.9. *For $n > 0$, let $A \in \eta_n^0$ and $\nu = F_*^n(m|A)$. Then*

$$\left| \frac{\frac{d\nu(x)}{dm}}{\frac{d\nu(y)}{dm}} - 1 \right| \leq D,$$

for any $x, y \in \Delta_0$.

PROOF. By the assumption, $F^n : A \to \Delta_0$ is invertible. So for any $x \in \Delta_0$ there is a unique $x_0 \in A$ such that $F^n(x_0) = x$ and $\frac{d\nu}{dm}(x) = \frac{1}{m(A)}\frac{1}{JF^n x_0}$. Let $\varphi = \frac{d\nu}{dm}$ then for $x, y \in \Delta_0$, using Lemma 1.6 we obtain

$$\left|\frac{\varphi(x)}{\varphi(y)} - 1\right| = \left|\frac{JF^n(y_0)}{JF^n(x_0)} - 1\right| \le D.$$

□

The proof of the following corollary is analogous to the proof of Corollary 1.7.

COROLLARY 1.10. *For $n > 0$, let $A \in \eta_n^0$ and $\nu = F_*^n(m|A)$. Then*

$$\frac{1}{(1+D)m(\Delta_0)} \le \frac{d\nu}{dm}(x) \le \frac{1+D}{m(\Delta_0)}.$$

3.3. Return times to $\bar{\Delta}_0$. We are now ready to begin the construction of the tower for the product system. Since we are considering the product of just two systems, we will omit superfluous indexing and let f and f' be two maps that admit GMY-structure with the bases Δ_0, Δ_0' and return time functions R, R' respectively. Then we have associated GMY-towers

$$F : (\Delta, m) \circlearrowleft \quad \text{and} \quad F' : (\Delta', m') \circlearrowleft$$

with bases Δ_0 and Δ_0' and return time functions $R(x)$ and $R'(x')$ respectively. Let

$$\bar{\Delta} = \Delta \times \Delta' \quad \text{and} \quad \bar{\Delta}_0 = \Delta_0 \times \Delta_0'$$

denote the product of the two towers, and the product of their bases respectively. Letting m_0, m_0' denote the restrictions of m, m' to Δ_0, Δ_0' respectively, we let $\bar{m} = m \times m'$ and $\bar{m}_0 = m_0 \times m_0'$ denote the product measures on the corresponding products. The direct product map $\bar{F} = F \times F'$ is defined on $\bar{\Delta}$ and we will construct a tower for $\bar{F}$ with base $\bar{\Delta}_0$. We start by defining the return time function T on $\bar{\Delta}_0$. From Theorem 1 in [**69**] there exist mixing invariant probability measures μ, μ' for F and F', equivalent to m, m' respectively, and with densities which are uniformly bounded above and below. Therefore, there exist constants $c > 0$ and $n_0 > 0$ such that

$$m(F^{-n}(\Delta_0) \cap \Delta_0) > c > 0 \quad \text{and} \quad m'(F'^{-n}(\Delta_0') \cap \Delta_0') > c > 0 \tag{12}$$

for all $n \ge n_0$. We choose such n_0 and introduce a sequence $\{\tau_i\}$ of positive integers as follows. For $\bar{x} = (x, x') \in \bar{\Delta}$ let

$$\tau_0(\bar{x}) = 0 \quad \text{and} \quad \tau_1(\bar{x}) := n_0 + \hat{R}(F^{n_0}x).$$

The other elements of the sequence $\{\tau_i\}$ are defined inductively by iterating F or F' alternately depending on whether i is odd or even. More formally, for every

$j \geq 1$ let

$$\tau_{2j}(\bar{x}) := \tau_{2j-1}(\bar{x}) + n_0 + \hat{R}'(F'^{n_0+\tau_{2j-1}}x'),$$
$$\tau_{2j+1}(\bar{x}) := \tau_{2j}(\bar{x}) + n_0 + \hat{R}(F^{n_0+\tau_{2j}}x). \tag{13}$$

Remark 1.11. Notice that at every step we "wait" for n_0 iterates before defining the next term of the sequence. This implies that for any i we have $\tau_i - \tau_{i-1} \geq n_0$ and therefore, from (12), we get

$$m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0) \geq c > 0 \quad \text{and} \quad m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0) \geq c > 0$$

We use this fact in the proof of the first item in Proposition 1.13.

3.4. Initial step of construction of the partition of $\bar{\Delta}_0$. Now we can begin to define a partition $\bar{\eta}$ of $\bar{\Delta}_0$. Recall that the towers Δ, Δ' of the component systems admit by definition partitions into sets of the form $\Delta_{\ell,j}$, $\Delta'_{\ell,j}$. We will denote these given partitions by η, η' respectively and their restrictions to Δ_0, Δ'_0 by η_0, η'_0. We let

$$\xi_0 = \eta_0 \times \eta'_0$$

denote the corresponding partition of the product $\bar{\Delta}_0$. Our goal is to define a partition $\bar{\eta}$ of $\bar{\Delta}_0$ with the property that for each $\bar{x} \in \bar{\Delta}_0$ the corresponding partition element $\bar{\eta}(\bar{x}) \in \bar{\eta}$ maps bijectively to $\bar{\Delta}_0$ with some return time $T(\bar{\eta}(\bar{x}))$. Its construction requires the definition of an increasing sequence of partitions of $\bar{\Delta}_0$ denoted by

$$\xi_0 \prec \xi_1 \prec \xi_2 \prec \xi_3 \prec ....$$

A key property of these partitions will be that the sequences $\tau_1, ..., \tau_i$ are constant on elements of ξ_i. Moreover, the construction of $\bar{\eta}$ implies that all return times $T(\bar{x}) = T(\bar{\eta}((\bar{x}))$ are of the form $T(\bar{x}) = \tau_i(\bar{x})$ for some i where τ_i belongs to the sequence defined above. This property will be used below for the estimates of the tail of the return times. In particular, notice also that there may be some elements of $\eta_0 \times \eta'_0$ which map bijectively to $\bar{\Delta}_0$, and thus are candidates for elements of $\bar{\eta}$, but are not guaranteed to satisfy the requirement just stated above. The partition $\bar{\eta}$ will be defined as the union

$$\bar{\eta} = \bigcup_{i=1}^{\infty} \bar{\eta}_i \tag{14}$$

of disjoint sets $\bar{\eta}_i$ which consists of a collection of subsets which are defined in the first i steps of the construction.

The systematic construction proceeds as follows. For each $\bar{x} = (x, x') \in \bar{\Delta}_0$, let

$$\eta_{\tau_1(x)}(x) := (\bigvee_{j=0}^{\tau_1(\bar{x})-1} F^{-j}\eta)(x). \tag{15}$$

It follows immediately that $\eta_{\tau_1}(x) \subseteq \eta_0(x)$. Indeed, as the following simple Lemma proves, collection $\eta_{\tau_1} := \{\eta_{\tau_1(x)}(x) | x \in \Delta_0\}$ is in fact a partition of Δ_0.

LEMMA 1.12. *η_{τ_1} is a partition of Δ_0, $\eta_0 \prec \eta_{\tau_1}$ and τ_1 is constant on elements of η_{τ_1}.*

PROOF. First of all, note that since τ_1 depends only on the first coordinate $\eta_{\tau_1(x)}(x) \in \eta_n$, with $n = \tau_1(x)$. Then Lemma 1.5 implies that $F^{\tau_1(x)}(\eta_{\tau_1(x)}) = \Delta_0$ bijectively. In particular, all the points in $\eta_{\tau_1(x)}(x)$ have the same combinatorics up to time $\tau_1(x)$, and hence τ_1 is constant on $\eta_{\tau_1(x)}(x)$.

Now, to prove the Lemma, since η_{τ_1} clearly covers Δ_0, we just need to show every pair of sets in η_{τ_1} are either disjoint or coincide. Let $\eta_{\tau_1(x)}(x)$ and $\eta_{\tau_1(y)}(y)$ be two arbitrary elements of η_{τ_1}. If $\tau_1(x) = \tau_1(y) = n$ then $\eta_{\tau_1(x)}(x)$ and $\eta_{\tau_1(y)}(y)$ are the elements of η_n, hence they are disjoint or coincide. If $\tau_1(x) \neq \tau_1(y)$ then $\eta_{\tau_1(x)}(x) \cap \eta_{\tau_1(y)}(y) = \emptyset$ because τ_1 is constant on elements of η_{τ_1}. □

We can now define the partition ξ_1 of $\bar{\Delta}_0$ by letting, for every $\bar{x} = (x, x') \in \bar{\Delta}_0$,

$$\xi_1(\bar{x}) := \eta_{\tau_1(x)}(x) \times \eta'_0(x').$$

Notice that each element $\Gamma \in \xi_1$ has an associated value of τ_1 such that $F^{\tau_1}(x) \in \Delta_0$ for every $x \in \pi\Gamma$. On the other hand, we do not a priori have any information about the location of $F'^{\tau_1}(x')$ for $x' \in \pi'\Gamma$ (since τ_1 is defined in terms of properties of F). To study the distribution of such images, for a given $\Gamma \in \xi_1$, we consider sets of the form

$$\eta'_{\tau_1}(x') = (\bigvee_{j=0}^{\tau_1 - 1} F'^{-j}\eta')(x'). \tag{16}$$

The collection of such sets, for all $x' \in \pi'\Gamma$, form a refinement of $\pi'\Gamma$. For those points $x' \in \pi'\Gamma$ such that $F^{\tau_1}(x') \in \Delta'_0$ we then have that

$$F'^{\tau_1} : \eta'_{\tau_1}(x') \to \Delta'_0$$

bijectively. In this case we consider the set

$$\bar{\eta}_1(x, x') = \eta_{\tau_1}(x) \times \eta'_{\tau_1}(x').$$

which maps bijectively to $\bar{\Delta}_0$ by $\bar{F}^{\tau_1}$, and let $\bar{\eta}_1$ denote the collection of all sets of this form constructed at this step. This is the first collection of sets which will be included in the union (14) defined above.

3.5. General step. We now describe the general inductive step in the construction of the sequence of partitions ξ_i and the sets $\bar{\eta}_i$. The main inductive assumption is that partitions ξ_i of $\bar{\Delta}_0$ have been constructed for all $i < k$ in such a way that on each element of ξ_i the functions $\tau_1, ..., \tau_i$ are constant and such that each component of $\bar{\eta}_{i-1}$ is contained inside an element of ξ_{i-1} and contains one or more elements of ξ_i (in particular elements of the partition ξ_i either have empty intersection with $\bar{\eta}_{i-1}$ or are fully contained in some component of $\bar{\eta}_{i-1}$).

The notation is slightly different depending on whether k is even or odd, according to the different definitions of τ_k in these two cases, recall (13). For definiteness we assume that k is odd, the construction for k even is the same apart from the change in the role of the first and second components. We fix some $\Gamma \in \xi_{k-1}$ and define the partition $\xi_k|\Gamma$ as follows. For $\bar{x} \in \Gamma$, let

$$\eta_{\tau_k(x)}(\bar{x}) = (\bigvee_{j=0}^{\tau_k(\bar{x})-1} F^{-j}\eta)(x).$$

A direct generalization of Lemma 1.12 gives that the collection of sets $\eta_{\tau_k} := \{\eta_{\tau_k(x)}(x)|x \in \pi\Gamma\}$ is a partition of $\pi\Gamma$ on whose elements τ_k is constant. For every $\bar{x} \in \Gamma$ we let

$$\xi_k(\bar{x}) := \eta_{\tau_k(x)}(\bar{x}) \times \pi'\Gamma.$$

This completes the definition of the partition ξ_k and allows us define $\bar{\eta}_k$. As mentioned in the inductive assumptions above, each component of $\bar{\eta}_k$ will be contained in an element of ξ_k. Thus, generalizing the construction of such elements in the first step given above, we fix one element $\Gamma \in \xi_k$ and proceed as follows. By construction we have $F^{\tau_k}(\pi\Gamma) = \Delta_0$ and $F'^{\tau_k}(\pi'\Gamma)$ is spread around Δ'. Therefore, for every $x' \in \pi'\Gamma$ we consider sets of form

$$\eta'_{\tau_k}(x') = (\bigvee_{j=0}^{\tau_k-1} F'^{-j}\eta')(x'). \tag{17}$$

The collection of such sets, for all $x' \in \pi'\Gamma$, form a refinement of $\pi'\Gamma$. For those points $x' \in \pi'\Gamma$ such that $F^{\tau_k}(x') \in \Delta'_0$ we then have that

$$F'^{\tau_k} : \eta'_{\tau_k}(x') \to \Delta'_0$$

bijectively. In this case we consider the set

$$\bar{\eta}_k(x, x') = \eta_{\tau_k}(x) \times \eta'_{\tau_k}(x')$$

which maps bijectively to $\bar{\Delta}_0$ by $\bar{F}^{\tau_k}$, and let $\bar{\eta}_k$ denote the collection of all sets of this form constructed at this step. Moreover, for each such set we let $T = \tau_k$. Finally, notice that the construction of $\bar{\eta}_k$ through the formula (17) implies that the elements of the partition ξ_{k+1}, to be constructed in the next step, are either disjoint from $\bar{\eta}_k$ or contained components of $\bar{\eta}_k$. Indeed, the construction of ξ_{k+1}

involves a formula analogous to (17) with τ_k replaced by τ_{k+1}, clearly yielding a finer partition of $\pi'\Gamma$.

This completes the general step of the construction and in particular allows us to define the set $\bar{\eta}$ as in (14). We remark that by construction all components of $\bar{\eta}$ are pairwise disjoint but we have not yet proved that $\bar{\eta}$ is a partition of $\bar{\Delta}_0$. This will be an immediate consequence of Proposition 1.13 in the next section. For formal consistency of notation we let $T = \infty$ on all points in the complements of the elements of $\bar{\eta}$. Notice that it follows from the construction that for any $\Gamma \in \bar{\eta}$:

(1) $T|\Gamma = \tau_i|\Gamma$ for some i.
(2) $\bar{F}^T(\Gamma) = \bar{\Delta}_0$.

In particular all elements of $\bar{\eta}$ satisfy properties (G1) and (G2) in the definition of Young Tower. In Section 4 we obtain some preliminary estimates concerning the general asymptotics of the return times of the elements of $\bar{\eta}$ defined above, and in Section 5 we consider the specific cases of polynomial, stretched exponential and exponential decay rates. In Section 6 we use all these estimates to prove that $\bar{\eta}$ is indeed a partition of $\bar{\Delta}_0$ and that the return map $\bar{F}^T$ satisfies the required properties (G3)-(G5).

4. Return time asymptotics

In this section we begin the study of the asymptotics of the return time T. We will use the following general notation for conditional measures: if μ is a measure we write $\mu(B|A) := \mu(A \cap B)/\mu(B)$. Also, for every $n \geq 1$, we write

$$\bar{\mathcal{M}}_n := \sum_{j \geq n} \mathcal{M}_j \tag{18}$$

where $\mathcal{M}_j$ is the bound on the tails of the component systems, as in (8). The main result of this section is the following

PROPOSITION 1.13. *There exist constants ε_0, $K_0 > 0$ such that for any $i \geq 2$*

(1) $\bar{m}_0\{T = \tau_i | T > \tau_{i-1}\} \geq \varepsilon_0$
(2) $\bar{m}_0\{\tau_{i+1} - \tau_i \geq n | \Gamma\} \leq K_0 \bar{\mathcal{M}}_{n-n_0}$, *for any $n > n_0$ and $\Gamma \in \xi_i$.*

Recall that we have set $T = \infty$ on the complement of points belonging to some element of $\bar{\eta}$, and notice that $\{T > \tau_{i-1}\}$ is the set of points in $\bar{\Delta}_0$ which do not belong to any elements of η_j for any $j = 1, ..., i-1$ (by some slight abuse of notation we could write this as $\{T > \tau_{i-1}\} = \bar{\Delta}_0 \setminus \bigcup_{j=1}^{i-1} \bar{\eta}_j$). As an immediate consequence of item (1) we get the statement that T is finite for almost every point in $\bar{\Delta}_0$ and therefore the collection of sets $\bar{\eta}$ as in (14) is indeed a partition of $\bar{\Delta}_0$ mod 0.

The proof of Proposition 1.13 relies on some standard combinatorial estimates which is given in 3.2.

Proof of (1). By construction, $\{T > \tau_{i-1}\}$ is a union of elements of the partition ξ_i. Thus

$$\bar{m}_0\{T = \tau_i | T > \tau_{i-1}\} = \frac{1}{\bar{m}_0\{T > \tau_{i-1}\}} \sum_{\Gamma \in \xi_i, T|_\Gamma > \tau_{i-1}} \bar{m}_0\{\{T = \tau_i\} \cap \Gamma\}.$$

Thus it is sufficient to prove $\bar{m}_0\{T = \tau_i | \Gamma\} \geq \varepsilon_0$ for any $\Gamma \in \xi_i$ on which $T|_\Gamma > \tau_{i-1}$. Assume for a moment i is even and let $\Omega = \pi(\Gamma)$, $\Omega' = \pi'(\Gamma)$. Then by construction $\pi'(\{T = \tau_i\} \cap \Gamma) = \Omega'$ and

$$\bar{m}_0\{T = \tau_i | \Gamma\} = \frac{\bar{m}_0(\{T = \tau_i\} \cap \Gamma)}{\bar{m}_0(\Gamma)} = \frac{m_0(\Omega \cap F^{-\tau_i}\Delta_0) m_0'(\Omega')}{m_0(\Omega) m_0'(\Omega')} = \frac{m_0(\Omega \cap F^{-\tau_i}\Delta_0)}{m_0(\Omega)}.$$

Now recall that $F^{\tau_{i-1}}(\Omega) = \Delta_0$, which implies $\Omega \cap F^{-\tau_{i-1}}\Delta_0 = \Omega$ and therefore

$$\frac{m_0(\Omega \cap F^{-\tau_i}\Delta_0)}{m_0(\Omega)} = \frac{m_0(\Omega \cap F^{-\tau_{i-1}}(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0))}{m_0(\Omega)} = F_*^{\tau_{i-1}}(m_0|\Omega)(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0).$$

Notice that $m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0) \geq c > 0$ since $\tau_i - \tau_{i-1} \geq n_0$. Letting $\nu = F_*^{\tau_{i-1}}(m_0|\Omega)$, applying Corollary 1.10 with $n = \tau_{i-1}$ for $x, y \in \Delta_0$ we get

$$\bar{m}_0\{T = \tau_i | \Gamma\} = \nu(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0) \geq \frac{m_0(\Delta_0 \cap F^{\tau_{i-1}-\tau_i}\Delta_0)}{(1+D)m_0(\Delta_0)} \geq \frac{c}{(1+D)m_0(\Delta_0)}.$$

For odd i's we can just change F to F' and do all the calculations, that gives the estimate

$$\bar{m}_0\{T = \tau_i | \Gamma\} > \frac{c}{(1+D')m_0'(\Delta_0')}.$$

Taking $\varepsilon_0 = c \min\{\frac{1}{(1+D)m_0(\Delta_0)}, \frac{1}{(1+D')m_0'(\Delta_0)'}\}$ we get the assertion. □

Proof of (2). Assume for a moment i is even and let, as above, $\Omega = \pi(\Gamma)$, $\Omega' = \pi'(\Gamma)$. Since τ_i is constant on the elements of ξ_i we have

$$\pi\left(\{\bar{x} = (x, x') | \hat{R} \circ F^{\tau_i + n_0}(x) > n\} \cap \Gamma\right) = \{x | \hat{R} \circ F^{\tau_i + n_0}(x) > n\} \cap \Omega.$$

For convenience we begin by estimating $\bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n|\Gamma\}$. From the definition of τ_i, letting $\nu = F_*^{\tau_{i-1}}(m_0|\Omega)$, we have

$$\begin{aligned}
\bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n|\Gamma\} = \bar{m}_0\{\hat{R}\circ F^{\tau_i+n_0} > n|\Gamma\} &= \frac{\bar{m}(\{\hat{R}\circ F^{\tau_i+n_0} > n\}\cap\Gamma)}{\bar{m}(\Gamma)} \\
&= \frac{m_0'(\Omega')m_0(\pi\{\hat{R}\circ F^{\tau_i+n_0} > n\}\cap\Omega)}{m_0(\Omega)m_0'(\Omega')} \\
&= \frac{m_0(\{\hat{R}\circ F^{\tau_i+n_0} > n\}\cap\Omega)}{m_0(\Omega)} = m_0\{\hat{R}\circ F^{\tau_i+n_0} > n|\Omega\} \\
&= F_*^{\tau_i+n_0}m_0\{\hat{R} > n|\Omega\} = F_*^{\tau_i-\tau_{i-1}+n_0}\nu\{\hat{R} > n\}.
\end{aligned}$$

To bound the final term in terms of $m_0\{\hat{R} > n\}$ it is sufficient to show that the density of $F_*^{\tau_i-\tau_{i-1}+n_0}\nu$ with respect to m_0 is uniformly bounded in i. We write first

$$\frac{dF_*^k\nu}{dm}(x) = \frac{dF_*^k\nu}{dF_*^k m_0}\frac{dF_*^k m_0}{dm}(x) = \sum_{x_0\in F^{-k}(x)}\frac{\frac{d\nu}{dm_0}(x_0)}{JF^n(x_0)} \le \left\|\frac{d\nu}{dm_0}\right\|_\infty \left\|\frac{dF_*^k m_0}{dm_0}\right\|_\infty.$$

The second factor is bounded by M_0 from Lemma 1.8. Let us estimate the first one. Note that, $\frac{d\nu}{dm_0}(x) = \frac{1}{JF^{\tau_{i-1}}x_0}$, where $x_0 = (F^{\tau_{i-1}}|\Omega)^{-1}(x)$. Corollary 1.10 implies that $\|d\nu/dm\|_\infty \le (1+D)/m_0(\Delta_0)$ and so we get

$$\bar{m}_0\{\tau_{i+1} - \tau_i > n_0 + n|\Gamma\} = \bar{m}_0\{\tau_{i+1} - \tau_i - n_0 > n|\Gamma\} \le M_0\frac{1+D}{m_0(\Delta_0)}m_0\{\hat{R} > n\}.$$

For $n > n_0$, and using the definition of $\hat{R}$, we can write this is

$$\bar{m}_0\{\tau_{i+1}-\tau_i > n|\Gamma\} \le M_0\frac{1+D}{m_0(\Delta_0)}m_0\{\hat{R} > n-n_0\} = M_0\frac{1+D}{m_0(\Delta_0)}\sum_{i\ge n-n_0} m_0\{R > i\}.$$

For i odd the calculation is exactly the same and we get

$$\bar{m}_0\{\tau_{i+1} - \tau_i > n|\Gamma\} \le M_0'\frac{1+D'}{m_0'(\Delta_0')}m_0'\sum_{i\ge n-n_0} m_0'\{R' > i\}.$$

Letting $K_0 = \max\left\{\frac{M_0(D+1)}{m_0(\Delta_0)}, \frac{M_0'(D'+1)}{m_0'(\Delta_0')}\right\}$ and using (8) we get the assertion. □

5. Rates of decay

We now fix some arbitrary $n \ge 1$ and estimate $\bar{m}_0\{T > n\}$. Letting $\tau_0 = 0$ we write

$$\bar{m}_0\{T > n\} = \sum_{i\ge 1}\bar{m}_0\{T > n; \tau_{i-1} \le n < \tau_i\}. \tag{19}$$

We will estimate the right hand side of (19) using different arguments depending on whether the decay of $\mathcal{M}_n$ is exponential, stretched exponential or polynomial.

5.1. Polynomial case. We suppose that $\mathcal{M}_n = \mathcal{O}(n^{-\alpha})$ for some $\alpha > 2$ and prove that

$$\bar{m}_0\{T > n\} = \mathcal{O}(n^{1-\alpha}) \tag{20}$$

Let $K = 2\max\{\bar{m}_0(\bar{\Delta}_0), K_0, m_0(\Delta_0), m'(\Delta'_0)\}$ and $\varepsilon_0 > 0$ given by Proposition 1.13. We start with the following somewhat unwieldy estimate.

Proposition 1.14. *For any $n \in \mathbb{N}$*

$$\bar{m}_0\{T > n\} \le K \sum_{i \le \frac{1}{2}\left[\frac{n}{n_0}\right]} i(1-\varepsilon_0)^{i-3}\bar{\mathcal{M}}_{\left[\frac{n}{i}\right]-n_0} + \bar{m}_0(\bar{\Delta}_0)(1-\varepsilon_0)^{\frac{1}{2}\left[\frac{n}{n_0}\right]-1}. \tag{21}$$

Before proving Proposition 1.14 we show how it implies (20).

Proof of Theorem in the polynomial case assuming Proposition 1.14. By the definition of $\bar{\mathcal{M}}_{\left[\frac{n}{i}\right]-n_0}$ in (18) and the assumption on the asymptotics of $\mathcal{M}_n$, there exists a constant $C > 0$ such that for every $i \le \frac{1}{2}\left[\frac{n}{n_0}\right]$ we have

$$\begin{aligned}\bar{\mathcal{M}}_{\left[\frac{n}{i}\right]-n_0} &= \sum_{j \ge \left[\frac{n}{i}\right]-n_0} \mathcal{M}_j \le C \sum_{j \ge \left[\frac{n}{i}\right]-n_0} j^{\alpha} \le C\int_{\left[\frac{n}{i}\right]-n_0}^{\infty} x^{-\alpha}dx \\ &\le C\frac{i^{\alpha-1}}{n^{\alpha-1}}\left(\frac{n}{n-i(n_0+1)}\right)^{\alpha-1} \le C\frac{i^{\alpha-1}}{n^{\alpha-1}}.\end{aligned}$$

Substituting this into the statement of Proposition 1.14 we get

$$\bar{m}_0\{T > n\} \le \frac{KC}{n^{\alpha-1}} \sum_{i \le \frac{1}{2}\left[\frac{n}{n_0}\right]} i^{\alpha}(1-\varepsilon_0)^{i-3} + \bar{m}_0(\bar{\Delta}_0)(1-\varepsilon_0)^{\frac{1}{2}\left[\frac{n}{n_0}\right]-1}. \tag{22}$$

Since the series $\sum_{i=1}^{\infty}(1-\varepsilon_0)^{i-3}i^{\alpha}$ is convergent and the second term in (22) is exponentially small in n, we get (20) and thus the statement of the Theorem in the polynomial case. □

The proof of Proposition 1.14 will be broken into several lemmas. Note that

$$\sum_{i > \frac{1}{2}\left[\frac{n}{n_0}\right]} \bar{m}_0\{T > n;\ \tau_{i-1} \le n < \tau_i\} \le \bar{m}_0\{T > n;\ \tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]} \le n\},$$

which together with (19) implies

$$(23)\ \bar{m}_0\{T>n\} \le \sum_{i\le\frac{1}{2}\left[\frac{n}{n_0}\right]} \bar{m}_0\{T>n;\ \tau_{i-1}\le n<\tau_i\} + \bar{m}_0\{T>n;\ \tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]}\le n\}.$$

First we estimate the second summand of (23).

LEMMA 1.15. *For every $n>2n_0$ and for $\varepsilon_0>0$ as in Proposition 1.13*

$$\bar{m}_0\{T>n;\ \tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]}\le n\} \le \bar{m}_0(\bar{\Delta}_0)(1-\varepsilon_0)^{\frac{1}{2}\left[\frac{n}{n_0}\right]-1}.$$

PROOF. Since $T>n>\tau_{i-1}$ we have

$$\bar{m}_0\{T>n;\ \tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]}\le n\} \le \bar{m}_0\{T>\tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]}\} =$$
$$\bar{m}_0\{T>\tau_1\}\bar{m}_0\{T>\tau_2|\ T>\tau_1\}...\bar{m}_0\{T>\tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]}|\ T>\tau_{\frac{1}{2}\left[\frac{n}{n_0}\right]-1}\}.$$

Notice that $\bar{m}_0\{T>\tau_1\}\le\bar{m}_0(\bar{\Delta}_0)$. The first item of Proposition 1.13 implies that each of the other terms is less than $1-\varepsilon_0$. Substituting these into the above equation finishes the proof. □

Now, we begin estimating the first summand of (23). Start with the cases $i = 1, 2$.

LEMMA 1.16. *For $i=1,2$ and every $n>n_0$ we have*

$$\bar{m}_0\{T>\tau_{i-1};\tau_{i-1}\le n<\tau_i\}\le K\bar{\mathcal{M}}_{\left[\frac{n}{2}\right]-n_0}.$$

PROOF. For $i=1$ we take advantage of the fact that τ_1 depends only on the first coordinate. Then we have

$$\bar{m}_0\{T>0;\tau_1>n\} = m_0\{\tau_1>n\}m_0'(\Delta_0') \le m_0'(\Delta_0')m\{\hat{R}>n-n_0\}$$
$$\le m_0'(\Delta_0')\sum_{j\ge n-n_0} m_0\{R>j\} \le m_0'(\Delta_0')\bar{\mathcal{M}}_{n-n_0}.$$

which proves the statement in this case by the definition of K and using the fact that $\bar{\mathcal{M}}_k$ is monotone decreasing in k. For $i=2$ we have

$$\bar{m}_0\{T>\tau_1;\tau_1\le n<\tau_2\}\le\bar{m}_0\{\tau_2>n\}\le\bar{m}_0\{\tau_2-\tau_1\ge\frac{n}{2}\}+\bar{m}_0\{\tau_1\ge\frac{n}{2}\}.$$

From the second item of Proposition 1.13 we have

$$\bar{m}_0\{\tau_2-\tau_1\ge\frac{n}{2}\}\le K_0\bar{\mathcal{M}}_{\left[\frac{n}{2}\right]-n_0}.$$

The second item is estimated as in the case $i=1$ and so we get

$$\bar{m}_0\{T>\tau_1;\tau_1\le n<\tau_2\}\le K_0\bar{\mathcal{M}}_{\left[\frac{n}{2}\right]-n_0}+m_0'(\Delta_0')\bar{\mathcal{M}}_{\left[\frac{n}{2}\right]-n_0}\le K\bar{\mathcal{M}}_{\left[\frac{n}{2}\right]-n_0}.$$

which completes the proof in this case also. □

We now consider the general case.

Lemma 1.17. *For each* $i \geq 3$

$$\bar{m}_0\{T > n;\ \tau_{i-1} \leq n < \tau_i\} \leq \sum_{j=1}^{i} K(1-\varepsilon_0)^{i-3} \bar{\mathcal{M}}_{[\frac{n}{i}]-n_0}.$$

Proof. Since $T > n \geq \tau_{i-1}$ we have

$$\bar{m}_0\{T > n;\ \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}_0\{T > \tau_{i-1};\ n < \tau_i\}.$$

Moreover, from $\tau_i = \tau_i - \tau_{i-1} + \tau_{i-1} - \tau_{i-2} + ... + \tau_1 - \tau_0 > n$ we obtain that there is at least one $j \in [1, i]$ such that $\tau_j - \tau_{j-1} > \frac{n}{i}$ and therefore

$$\bar{m}_0\{T > n;\ \tau_{i-1} \leq n < \tau_i\} \leq \sum_{j=1}^{i} \bar{m}_0\{T > \tau_{i-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\}.$$

For each $i, j \geq 3$ we write

$$\bar{m}_0\{T > \tau_{i-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} = Y_1 \cdot Y_2 \cdot Y_3 \tag{24}$$

where

$$Y_1 := \bar{m}_0\{T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\},$$

$$Y_2 := \bar{m}_0\{T > \tau_1\}\bar{m}_0\{T > \tau_2 | T > \tau_1\} ... \bar{m}_0\{T > \tau_{j-2} | T > \tau_{j-3}\},$$

$$Y_3 := \bar{m}_0\{T > \tau_j | T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} ...\ \bar{m}_0\{T > \tau_{i-1} | T > \tau_{i-2}; \tau_j - \tau_{j-1} > \frac{n}{i}\}.$$

By the second item of Proposition 1.13 we have

$$Y_2 \leq \bar{m}_0(\bar{\Delta}_0)(1-\varepsilon_0)^{j-3}. \tag{25}$$

For the first term, note that

$$Y_1 := \bar{m}_0\{T > \tau_{j-1}; \tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\} \leq \bar{m}_0\{\tau_j - \tau_{j-1} > \frac{n}{i} | T > \tau_{j-2}\}.$$

By construction $\{T > \tau_{j-2}\}$ can be written as a union of elements of ξ_{j-1} and so, by the second item of Proposition 1.13,

$$Y_1 \leq K_0 \bar{\mathcal{M}}_{[\frac{n}{i}]-n_0}. \tag{26}$$

For the third term, since τ_j and τ_{j-1} are constant on the elements of ξ_j, if $\tau_j(\bar{x}) - \tau_{j-1}(\bar{x}) > \frac{n}{i}$ for some point $\bar{x}$, then it holds on $\xi_j(\bar{x})$. By construction, for $k \geq j$ the partition ξ_k is finer than ξ_j and $\{T > \tau_{k-1}\}$ can be written as a union of elements of ξ_k. Hence $\{T > \tau_{k-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\}$ can be covered with elements of ξ_k. Using the first item of Proposition 1.13 in each partition element gives

$$\bar{m}_0\{T > \tau_k |\ T > \tau_{k-1}; \tau_j - \tau_{j-1} > \frac{n}{i}\} \leq 1 - \varepsilon_0.$$

This immediately gives

$$Y_3 \leq (1-\varepsilon_0)^{i-j}. \tag{27}$$

Substituting (25), (26), (27) into (24) we get the assertion of Lemma 5.5. For the case $i \geq 3$ and $j < 3$ proof will be the same but only without Y_2. □

Notice that substituting the estimates in the statements of Lemmas 1.16 and 1.17 into (23) gives the statement in Proposition 1.14.

5.2. Super polynomial cases. In this subsection the we give the proof of the tail estimates for exponential and stretched exponential cases. Let

$$A(i) = \{\mathbf{k} = (k_1, ..., k_{i-1}) \in \mathbb{N}^{i-1} : \sum_j k_j \leq n,\ k_j \geq n_0, j = 1, ..., i-1\},$$

$$\bar{\mathcal{M}}(i,n) = \max_{\mathbf{k}\in A(i)} \bar{\mathcal{M}}_{n-\sum_j k_j - n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0}. \tag{28}$$

Remark 1.18. It is known fact (see for example [**65**]) that the cardinality card$A(i)$ of $A(i)$ is bounded above by $\binom{n+i-n_0}{i-1}$.

Let δ be a sufficiently small number, which will be specified later. We first prove the following technical statement from which both exponential and stretched exponential cases will follow.

Proposition 1.19. *For sufficiently large n and any $\theta' \in (0,1]$ we have*

$$\bar{m}_0\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} \binom{n+i-n_0}{i-1} K_0^i \bar{\mathcal{M}}(i,n) + \bar{m}(\bar{\Delta}_0)(1-\varepsilon_0)^{[\delta n^{\theta'}]-1}.$$

To prove Proposition 1.19 we first write, as in (23),

$$\bar{m}_0\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} \bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} + \bar{m}_0\{T > n; \tau_{[\delta n^{\theta'}]} \leq n\}. \tag{29}$$

As in polynomial case, we start by estimating the second summand of (29).

Lemma 1.20. *For sufficiently large n we have*

$$\bar{m}_0\{T > n; \tau_{[\delta n^{\theta'}]} \leq n\} \leq \bar{m}(\bar{\Delta}_0)(1-\varepsilon_0)^{[\delta n^{\theta'}]-1}.$$

Proof. The proof is identical to the proof of Lemma 1.15. □

To estimate the second summand of (29), first we fix i and prove the following

LEMMA 1.21. *For sufficiently large n, for every $i \leq [\delta n^{\theta'}]$ we have*

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}(\bar{\Delta}_0) K_0^i card A(i) \bar{\mathcal{M}}(i,n).$$

PROOF. For any $\bar{x} \in \bar{\Delta}_0$ with $\tau_{i-1}(\bar{x}) \leq n < \tau_i(\bar{x})$ there is $(k_1, ..., k_{i-1}) \in A(i)$ such that $k_j = \tau_j(\bar{x}) - \tau_{j-1}(\bar{x})$ for $j = 1, ..., i-1$ and $\tau_i(\bar{x}) - \tau_{i-1}(\bar{x}) > n - \sum_j k_j$. Now, for every $\mathbf{k} = (k_1, ..., k_{i-1}) \in A(i)$ let

$$P(\mathbf{k}, i) = \bigcap_{j=1}^{i-1} \{\bar{x} \in \bar{\Delta} : \ \tau_j(\bar{x}) - \tau_{j-1}(\bar{x}) = k_j\} \tag{30}$$

and

$$Q(\mathbf{k}, i) = P(\mathbf{k}, i) \cap \{\bar{x} \in \bar{\Delta} : \ \tau_i(\bar{x}) - \tau_{i-1}(\bar{x}) > n - \sum_j k_j\}. \tag{31}$$

Using the above observation and notations we can write

$$\bar{m}_0\{T > n; \tau_{i-1} \leq n < \tau_i\} \leq \bar{m}_0\{\tau_{i-1} \leq n < \tau_i\} = \sum_{\mathbf{k} \in A(i)} \bar{m}_0\{Q(\mathbf{k}, i)\}. \tag{32}$$

Notice that for each i and each $\mathbf{k} \in A(i)$ we have

$$\begin{aligned} \bar{m}_0\{Q(\mathbf{k}, i)\} = \bar{m}_0(\bar{\Delta}_0) \bar{m}_0\{\tau_1 = k_1 | \bar{\Delta}_0\} \bar{m}_0\{\tau_2 - \tau_1 = k_2 | P(\mathbf{k}, 2)\}... \\ ...\bar{m}_0\{\tau_i - \tau_{i-1} > n - \sum k_j | P(\mathbf{k}, i)\}. \end{aligned} \tag{33}$$

Since for any j, the set $\{P(\mathbf{k}, j)\}$ is a union of elements of ξ_j, from the second item of Proposition 1.13 we get

$$\bar{m}_0\{\tau_j - \tau_{j-1} = k_j | P(\mathbf{k}, j)\} \leq K_0 \bar{\mathcal{M}}_{k_j - n_0}$$

for each $1 \leq j \leq i-1$, and

$$\bar{m}_0\{\tau_i - \tau_{i-1} > n - n_0 - \sum k_j | P(\mathbf{k}, i)\} \leq K_0 \bar{\mathcal{M}}_{n - n_0 - \sum k_j}.$$

Substituting this into (33) and using the definition of $\bar{\mathcal{M}}(i,n)$ in (28), we obtain

$$\bar{m}_0\{Q(\mathbf{k}, i)\} \leq \bar{m}_0(\bar{\Delta}_0) K_0^i \bar{\mathcal{M}}_{n - n_0 - \sum k_j} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} \leq \bar{m}_0(\bar{\Delta}_0) K_0^i \bar{\mathcal{M}}(i,n).$$

Substituting this into (32) completes the proof of Lemma 1.21. □

PROOF OF PROPOSITION 1.19. To prove the Proposition we just substitute the statements in the two Lemmas above into (29) and use the upper bound for the cardinality of $A(i)$, see Remark (1.18). □

Proof of the Theorem in the stretched exponential case. Here we consider the case $\mathcal{M}_n = \mathcal{O}(e^{-\tau n^\theta})$ for some $\tau, \theta > 0$ and we show that

$$\bar{m}_0\{T > n\} \leq Ce^{-\tau' n^{\theta'}} \tag{34}$$

for some $C_1, \tau' > 0$ and $\theta > \theta' > 0$. We start with the following basic estimate

Lemma 1.22. *For sufficiently large n, every $i \leq [\delta n^{\theta'}]$ and every $\boldsymbol{k} \in A(i)$ we have*

$$\bar{\mathcal{M}}_{n-\sum_j k_j - n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} \leq {C_1}^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)}.$$

The statement and proof of Lemma 1.22 depend only on some relatively standard but non-trivial estimates concerning the tails of sequences which decay at a stretched exponential rate. In order to simplify the exposition we give the proof at the end of the section. From Lemma 1.22 and the definition of $\bar{\mathcal{M}}(i,n)$ we obtain

$$\bar{\mathcal{M}}(i,n) \leq {C_1}^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)}. \tag{35}$$

On the other hand combining Pascal's rule with Stirling's formula we get

$$\binom{n+i-n_0}{i-1} \leq \binom{n}{[\delta n^{\theta'}]} \leq C_2 e^{\varepsilon n^{\theta'} \log n} < C_2 e^{\varepsilon n^\theta} \tag{36}$$

for $\varepsilon > 0$ such that $\varepsilon \to 0$ as $\delta \to 0$. Substituting inequalities (35) and (36) into Proposition 1.19 we get

$$\bar{m}\{T > n\} \leq \sum_{i \leq [\delta n^{\theta'}]} C_2 e^{\varepsilon n^\theta} K_0^i {C_1}^i e^{-\tau(n-in_0)^\theta} n^{i(1-\theta)} + \bar{m}(\bar{\Delta}_0)(1-\varepsilon_0)^{[\delta n^{\theta'}]-1}. \tag{37}$$

The second term of (37) is of order $e^{n^{\theta'} \delta \log(1-\varepsilon_0)}$. Hence, it remains to prove similar asymptotics for the first summand. Notice first of all that the terms in the sum are monotone increasing in i. Therefore denoting these terms by a_i we have that

$$\sum_{i \leq [\delta n^{\theta'}]} a_i \leq a_{[\delta n^{\theta'}]}[\delta n^{\theta'}] = e^{\varepsilon n^\theta} (K_0 C_1)^{[\delta n^{\theta'}]} e^{-\tau(n-[\delta n^{\theta'}]n_0)^\theta} n^{[\delta n^{\theta'}](1-\theta)}[\delta n^{\theta'}].$$

Writing the right hand side in exponential form, we have

$$a_{[\delta n^{\theta'}]}[\delta n^{\theta'}] = \exp(\varepsilon n^\theta + \delta n^{\theta'} \log(K_0 C_1) - \tau(n - n_0[\delta n^{\theta'}])^\theta + (1-\theta)[\delta n^{\theta'}] \log n + \log(\delta n^{\theta'})).$$

Factoring out $n^{\theta'}$ and using the general inequality $a^\theta - b^\theta \leq (a-b)^\theta$ the exponent is $\leq$

$$n^{\theta'} \left\{ (\varepsilon - \tau) n^{\theta - \theta'} + \delta \log(K_0 C_1) + \tau(\delta n_0)^\theta n^{(\theta-1)\theta'} + (1-\theta)\delta \log n + n^{-\theta'} \log(\delta n^{\theta'}) \right\}.$$

Since this last expression is clearly decreasing in δ and negative for $\delta > 0$ sufficiently small we get the stretched exponential bound as required. □

Proof of the Theorem in the exponential case. Here we consider the case $\mathcal{M}_n = \mathcal{O}(e^{-\tau n})$ for some $\tau > 0$ and we show that

$$\bar{m}_0\{T > n\} \le Ce^{-\tau' n} \tag{38}$$

for some $\tau' > 0$.

Lemma 1.23. *For all $i \le [\delta n]$ and every $\boldsymbol{k} \in A(i)$ we have*

$$\bar{\mathcal{M}}_{n-\sum_j k_j - n_0} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} \le C^i e^{-\tau(n - in_0)}.$$

Proof. Since $\mathcal{M}_n$ is decaying exponentially fast, there exists $C' > 0$ such that for any $j = 1, ..., i-1$ we have

$$\bar{\mathcal{M}}_{k_j - n_0} = \sum_{\kappa \ge k_j - n_0} \mathcal{M}_\kappa \le C' \sum_{\kappa \ge k_j - n_0} e^{-\tau\kappa} = \frac{C'}{1 - e^{-\tau}} e^{-\tau(k_j - n_0)}$$

and similarly

$$\bar{\mathcal{M}}_{n - n_0 - \sum k_j} \le \frac{C'}{1 - e^{-\tau}} e^{-\tau(n - n_0 - \sum k_j)}.$$

Letting $C := C'/(1 - e^{-\tau})$ and substituting these estimates into the expression in the lemma finishes the proof. $\square$

By Lemma 1.23 and the definition of $\bar{\mathcal{M}}(i, n)$ we obtain

$$\bar{\mathcal{M}}(i, n) \le C^i e^{-\tau(n - in_0)}. \tag{39}$$

On the other hand, since $i \le [\delta n]$, there exists a uniform constant C_1 such that

$$\binom{n + i - n_0}{i - 1} \le \binom{n}{[\delta n]} \le C_1 e^{\varepsilon n} \tag{40}$$

for $\varepsilon > 0$ such that $\varepsilon \to 0$ as $\delta \to 0$. Choosing $\theta' = 1$ in Proposition 1.19 and substituting inequalities (39) and (40) we obtain

$$\bar{m}_0\{T > n\} \ge \sum_{i \le [\delta n]} C_1 e^{\varepsilon n} (K_0 C)^i e^{-\tau(n - in_0)} + \bar{m}(\Delta_0)(1 - \varepsilon_0)^{[\delta n] - 1}. \tag{41}$$

Choose δ small enough so that $\tau'(\delta) = \tau(1 - \delta n_0) - \delta \log(K_0 C) - \varepsilon > 0$. This is possible, since as $\delta \to 0$ we have $\tau'(\delta) \to \tau$. Notice, that each term of the first summand in (41) can be written in an exponential form with exponent

$$\varepsilon n + i \log(K_0 C) - \tau(n - in_0) = -\tau'(\delta) n + \tau n_0 (i - \delta n) + (i - \delta n) \log(K_0 C).$$

Hence, we can write the first summand of (41) as

$$C_1 e^{-\tau'(\delta) n} \sum_{i \le [\delta n]} e^{(i - \delta n)(\log(K_0 C) + \tau n_0)} \le C_1 e^{-\tau'(\delta) n} \sum_{j \ge 1} e^{-j(\log(K_0 C) + \tau n_0)}. \tag{42}$$

Notice that the series in the right hand side of (42) is convergent. Therefore, by choosing $\tau' = \min\{\tau'(\delta), \delta\log(1-\varepsilon_0)\}$ we get (38). □

Here we prove Lemma 1.22. First we prove the following

LEMMA 1.24. *Let $\tau > 0$, $\theta \in (0,1)$. For all $n \geq (\frac{2}{\tau\theta})^{1/\theta}$ we have*

$$\sum_{k\geq n} e^{-\tau k^\theta} \leq \frac{2}{\tau\theta} e^{-\tau n^\theta} n^{1-\theta}.$$

PROOF. First of all note that we have

$$\sum_{k\geq n} e^{-\tau k^\theta} \leq \int_n^\infty e^{-\tau x^\theta} dx.$$

After change of variables $t = \tau x^\theta$ we obtain

$$\int_n^\infty e^{-\tau x^\theta} dx = \frac{1}{\theta\tau^{1/\theta}} \int_{\tau n^\theta}^\infty e^{-t} t^{1/\theta - 1} dt. \tag{43}$$

In [**51**] it was proved that for any $a, B > 0$ and $x > \frac{B(a-1)}{B-1}$

$$\int_x^\infty t^{a-1} e^{-t} dt < B x^{a-1} e^{-x}. \tag{44}$$

Substituting (44) with $a = 1/\theta$ and $B = 2$ into the right hand side of (43) finishes the proof. □

Now, we are ready to prove Lemma 1.22. By Lemma 1.24 and definition of $\bar{\mathcal{M}}_n$, for sufficiently large n we have

$$\bar{\mathcal{M}}_{k_j - n_0} \leq C' \sum_{\kappa \geq k_j - n_0} e^{-\tau\kappa^\theta} \leq \frac{2C'}{\tau\theta} e^{-\tau(k_j - n_0)^\theta} (k_j - n_0)^{1-\theta}$$

for any $0 \leq j \leq i-1$ and

$$\bar{\mathcal{M}}_{n-n_0-\sum k_j} \leq \frac{2C'}{\tau\theta} e^{-\tau(n-n_0-\sum k_j)^\theta} (n - \sum k_j - n_0)^{1-\theta}.$$

Using $a^\alpha + b^\alpha \geq (a+b)^\alpha$ for $\alpha \in (0,1)$ and $a, b \geq 0$ we obtain

$$\begin{aligned} \bar{\mathcal{M}}_{n-n_0-\sum k_j} \prod_{j=1}^{i-1} \bar{\mathcal{M}}_{k_j - n_0} &\leq \left(\frac{2C'}{\tau\theta}\right)^i e^{\tau(n-in_0)^\theta} (n - \sum_j k_j - n_0)^{1-\theta} \prod (k_j - n_0)^{1-\theta} \\ &\leq \left(\frac{2C'}{\tau\theta}\right)^i e^{\tau(n-in_0)^\theta} n^{i(1-\theta)}. \end{aligned}$$

Taking $C_1 = 2C'/\tau\theta$ we obtain the statement in the Lemma.

6. Proof in the non-invertible setting

In this section we give the proof of the Theorem in the non-invertible setting. We will prove each of the required properties (G1)-(G5) in separate subsections, thus completing the proof of the existence of a tower for the product system $f : M \to M$. The tail estimates obtained above in the exponential, stretched exponential and polynomial case, then complete the proof.

(G1) Markov property. In Section 3 we carried out a the construction of the collection $\bar{\eta}$ of subsets of $\bar{\Delta}_0$ which, by the first item of Proposition 1.13 forms a partition of $\bar{\Delta}_0$ mod 0. In particular we have an induced map $\bar{F}^T : \bar{\Delta}_0 \to \bar{\Delta}_0$ which, using the canonical identification between $\bar{\Delta}_0$ as the base of the tower and as a subset of the ambient product manifold M, corresponds to an induced map $\bar{f}^T : \bar{\Delta}_0 \to \bar{\Delta}_0$ where $\bar{f} : M \to M$ is the product map on M. The construction of $\bar{\eta}$ and the induced map implies that $\bar{f}^T$ satisfies the Markov property (G1).

(G2) Uniform Expansion. The uniform expansivity condition follows immediately from the fact that it holds by assumption for the individual components and that the return time T is a sum of return times for each of the individual components. Then

$$\|(D\bar{f}^T)^{-1}(x)\| \leq \max\{\|(Df^T)^{-1}(x)\|, \|(Df'^T)^{-1}(x)\|\} \leq \beta.$$

(G3) Bounded distortion. Recall first of all from the definition of Gibbs-Markov-Young tower the notion of *separation time* and let s, s' denote the separation time of points in Δ_0, Δ_0' with respect to the partitions η, η' respectively. We let D, β, D', β' be the distortion constants as in (2) for f and f' respectively and let

$$\bar{D} = \max\left\{\frac{D\beta}{1-\beta}, \frac{D'\beta'}{1-\beta'}\right\}, \quad \bar{\beta} = \max\{\beta, \beta'\}.$$

Now let $\bar{s}$ denote the separation time of points in $\bar{\Delta}_0$ with respect to the product map $\bar{f}$ and the partition $\bar{\eta}$ constructed above. Then, using the usual identification of $\bar{\Delta}_0$ with the subset of manifold, to prove the bounded distortion condition (G3) it is sufficient to show that for any $\Gamma \in \bar{\eta}$ and $\bar{x}, \bar{y} \in \Gamma$ such that $\bar{s}(\bar{x}, \bar{y}) < \infty$ we have

$$\left|\log \frac{\det D\bar{f}^T(\bar{x})}{\det D\bar{f}^T(\bar{y})}\right| \leq \bar{D}\bar{\beta}^{\bar{s}(\bar{f}^T(\bar{x}), \bar{f}^T(\bar{y}))}.$$

To prove this, note first that, by the property of Jacobian and absolute value we have

$$\left|\log \frac{\det D\bar{f}^T(\bar{x})}{\det D\bar{f}^T(\bar{y})}\right| \leq \left|\log \frac{\det f^T(x)}{\det Df^T(y)}\right| + \left|\log \frac{\det Df'^T(x')}{\det Df'^T(y')}\right| \tag{45}$$

for all $\bar{x} = (x, x')$, $\bar{y} = (y, y') \in \Gamma$. Moreover, notice that since the determinants are all calculated at return times T, we can use the identification $\bar{f}^T = \bar{F}^T$ and reformulate the above expressions in terms of the Jacobians JF^T.

For the first term, notice that the simultaneous return time T can be written as a sum of return times to Δ_0. Without loss of generality, assume $T = R_1 + ... + R_k$, for some k. For any $x \in \Delta_0$, let $x_0 = x$ and $x_j = F^{R_1+...+R_j}(x)$ where $j = 1, ..., k-1$. Then

$$JF^T(x) = JF^{R_1}(x)JF^{R_2}(x_1)JF^{R_3}(x_2)...JF^{R_k}(x_{k-1}).$$

Since $F^T(\pi\Gamma) = \Delta_0$ we have $F^{R_1}(\pi\Gamma) \subset \eta(x_1)$, ..., $F^{R_1+...+R_{k-1}}(\pi\Gamma) \subset \eta(x_{k-1})$. Hence, for the points $x, y \in \pi\Gamma$ the sequences x_j and y_j belong to the same element of η for all $j = 0, ..., k-1$, which implies

$$s(F^{R_{j+1}}(x_j), F^{R_{j+1}}(y_j)) = s(F^T(x), F^T(y)) + R_{j+2} + ... + R_k.$$

Hence,

$$\left|\log \frac{JF^T(x)}{JF^T(y)}\right| \leq \sum_{j=1}^{k} \left|\log \frac{JF^{R_j}(x)}{JF^{R_j}(y)}\right| \leq \sum_{j=1}^{k} \left|\frac{JF^{R_j}(x)}{JF^{R_j}(y)} - 1\right|$$

$$\leq \sum_{j=0}^{k} D\beta^{s(F^{R_{j+1}}(x_j),F^{R_{j+1}}(y_j))} \leq D\beta^{s(F^T(x),F^T(y))} \sum_{j=0}^{k} \beta^{R_{j+2}+...+R_k}$$

$$\leq D\beta^{s(F^T(x),F^T(y))} \sum_{j=0}^{\infty} \beta_F^j \leq D\frac{\beta}{1-\beta}\beta^{s(F^T(x),F^T(y))}.$$

The second summand is estimated similarly and we get

$$\left|\log \frac{JF'^T(x')}{JF'^T(y')}\right| \leq D'\frac{\beta'}{1-\beta'}\beta'^{s'(F'^T(x'),F'^T(y'))}.$$

Using the fact that for any $w, z \in \bar{\Delta}_0$ we have $\bar{s}(w, z) \leq \min\{s(\pi w, \pi z), s'(\pi' w, \pi' z)\}$ we obtain the required bound.

(G4) Integrability. Follows immediately from the tail estimates obtained above.

6.1. Aperiodicity. As mentioned above, conditions (G1)-(G4) imply the existence of an ergodic $\bar{f}$-invariant probability measure $\bar{\mu}$. Moreover it is known by standard results that this measure is mixing if and only if the aperiodicity conditions is satisfied. Thus it is sufficient to show that $\bar{\mu}$ is mixing. To see this, let μ, μ' be the invariant, mixing, probability measures associated to the maps f, f' as introduced in Section 3.3. Then the measure $\mu \times \mu'$ is invariant and mixing for the product map $\bar{f}$ and thus it is sufficient to show that $\bar{\mu} = \mu \times \mu'$ to imply that $\bar{\mu}$ is mixing. This follows again by standard uniqueness arguments. Indeed, both $\bar{\mu}$ and $\mu \times \mu'$ are ergodic and equivalent to the reference measure, at least on the set $\bar{\Delta}_0$. Thus, by Birkhoff's ergodic Theorem, for any integrable function, their time averages converge to the same limit and so $\int \varphi d\bar{\mu} = \int \varphi d\mu \times \mu'$ implying that $\bar{\mu} = \mu'$.

7. Proof in the invertible setting

Let $F_i : \mathcal{T}_i \circlearrowleft$ be the two towers corresponding to maps $f_i : M_i \circlearrowleft$, $i = 1, 2$ as in the statement of the Theorem. Then, from conditions (A1)-(A7) we can obtain GMY-towers by considering the system obtained by the equivalence relation $\sim$ on Λ^i, $i = 1, 2$ defined as $x \sim y$ if and only if $y \in \gamma^s(x)$. Then on $\Delta_0^i = \Lambda^i / \sim$ we have the partition $\mathcal{P}^i = \{\Delta_{0,j}\} := \{\Lambda_{0,j}/\sim\}$ and the return time function $R^i : \Delta_{0,j} \to \mathbb{Z}^+$, and the quadruples $(F_i, R^i, \mathcal{P}_i, s_i)$, $i = 1, 2$ satisfy conditions (G1)-(G5). Moreover there is natural projection $\bar{\pi}_i : \mathcal{T} \to \Delta$ that sends each stable manifold to a point. We can then define the direct product of these two "quotient" GMY-towers and, from previous construction we obtain a new GMY tower for this product. Thus, on $\Delta_0^1 \times \Delta_0^2$ we have a partition $\hat{\mathcal{P}}$, and return time $T : \Delta_0 \to \mathbb{N}$ such that for any $A \in \hat{\mathcal{P}}$ we have $(F_1 \times F_2)^T(A) = \Delta_0^1 \times \Delta_0^2$. On the other hand we know that each $A \in \hat{\mathcal{P}}$ is of form $A_1 \times A_2$ and $\bar{\pi}_i^{-1}(A_i) \subset \Lambda^i$, $i = 1, 2$. Then $\mathcal{Q} = \{\bar{\pi}_1^{-1}(A_1) \times \bar{\pi}^{-1}(A_2) | A_i \in \hat{\mathcal{P}}_i, i = 1, 2\}$ gives the desired partition of $\Lambda^1 \times \Lambda^2$. Indeed, $(f_1 \times f_2)(\bar{\pi}_1^{-1}(A_1) \times \bar{\pi}^{-1}(A_2))$ is a u-subset of $\Lambda^1 \times \Lambda^2$ because at return times we have $f_i^T = F_i^T$, $i = 1, 2$.

All that is left is to check the properties (A1)-(A7). Those that refer to the combinatorial structure follow immediately from the discussion above, others follow immediately from the corresponding properties of the quotient tower. The only new property to check here is the second item in (A3). This follows easily by noticing that from the definition of T we have $s_T(x, y) \leq \min\{s_{R^1}(x_1, y_1), s_{R^2}(x_2, y_2)\}$, where s_T, s_{R^1}, s_{R^2} denote separation times with respect to return times T, R^1, R^2 and $x = (x_1, x_2)$, $y = (y_1, y_2)$ and therefore, by the definition of product metric, we have

$$\begin{aligned} \operatorname{dist}_M((f^T)^n(x), (f^T)^n(y)) &= \max_{i=1,2}\{\operatorname{dist}_i((f_i)^T(x_i), (f_i)^T(y_i))\} \\ &\leq C\beta^{\min\{s_{R^1}(x_1,y_1), s_{R^2}(x_2,y_2)\}} \leq C\beta^{s_T(x,y)}. \end{aligned}$$

Tail estimates transfer directly, since we need to estimate $m_\gamma\{T > n\}$ for $\gamma \in \Gamma^u$.

CHAPTER 2

Decay of Correlations for Continuous observables

In this chapter we study the problem of decay of correlations for continuous observables. Here we present two results and some counterexamples, which shows there is no specific rate of decay of correlation if the classes of observables are too big. First show that if an invertible system admits a Young tower then the rate of decay of correlations can be related to the decay rate of tail and the modulus of continuity of observables. The second is an abstract result that shows if the rate of decay of correlations is known for smooth observables and the cofactor has special form then the rate of decay of correlations for continuous observables can be estimated in terms of modulus of continuity.

1. Decay of correlation for invertible systems

Below we give our main technical result and its applications to Hénon maps and Solenoid maps with intermittency.

Let $f : M \to M$ be a $C^{1+\varepsilon}$ diffeomorphism a map with Young tower $F : \mathcal{T} \to \mathcal{T}$ (see section 2 for definition). Let Λ denote its base and $\mathcal{P}$ its partition. Define a sequence of partitions:

$$\mathcal{P}_0 = \mathcal{P} \quad \text{and} \quad \mathcal{P}_n = \bigvee_{i=0}^{n-1} F^{-i}\mathcal{P} \quad \text{for} \quad n \geq 1.$$

By definition of the elements of $\mathcal{P}_n$ contain stable manifolds of fixed size and shrink in the unstable direction as n goes to infinity. We would like to keep track of the speed of shrinking. In this purpose introduce the sequence

$$\delta_n = \sup\{\text{diam}\pi(F^n(P)) : P \in \mathcal{P}_{2n}\}.$$

The main result of this section is the following

THEOREM 2.1. *Let $f : M \to M$ be a diffeomorphism of Riemannian manifold M, which admits a Young tower with return time R. Then for any $\varphi, \psi \in \mathcal{C}(M)$ we have*

$$C_n(\varphi, \psi; \mu) \leq 2(\|\varphi\|_\infty + \|\psi\|_\infty)\mathcal{R}_{\varphi,\psi}(\delta_n) + u_n, \tag{46}$$

where u_n is a sequence of positive numbers defined as follows:

(i) *If* $m\{R > n\} \leq C\theta^n$ *for some* $C > 0$ *and* $\theta \in (0,1)$, *then there exist* $\theta' \in (0,1)$ *and* $C' > 0$ *such that* $u_n \leq C'\theta'^n$.

(ii) *If* $m\{R > n\} \leq Ce^{-cn^\eta}$, *for some* $C, c > 0$ *and* $\eta \in (0,1)$, *then there are* $C', c' > 0$ *such that* $u_n \leq C'e^{-c'n^\eta}$.

(iii) *If* $m\{R > n\} \leq Cn^{-\alpha}$ *for some* $C > 0$ *and* $\alpha > 1$ *then there exists* $C' > 0$ *such that* $u_n \leq C'n^{1-\alpha}$.

Next two theorems show that if the rates of decay of correlations are known for smooth observables and moreover cofactor in (0.1) depends on the observables in an explicit way, namely if there exists constant C that depends only on f such that

$$C_n(\varphi, \psi; \mu) \leq C\|\varphi\|_{\mathcal{B}_1}\|\psi\|_{\mathcal{B}_2}\gamma_n. \tag{47}$$

Then the rates of decay of correlations for continuous observables can be obtained in term of modulus of continuity and the rates of decay of correlations for smooth observables.

We state two separate theorems depending on whether f is invertible or non-invertible. Reason of this distinction is fact that to obtain the rates of decay of correlations for invertible systems we need to require that both of the observables have some regularity, whereas for non-invertible systems it is sufficient if one of the observables is regular see section 6 for examples.

THEOREM 2.2. *Let* $f : M \to M$ *be an invertible map defined on smooth manifold* M. *Suppose that* (47) *holds for any* $\tilde{\varphi}, \tilde{\psi} \in C^r(M)$ *and for some* $r \geq 1$. *Then for any* $\varphi, \psi \in \mathcal{C}(M)$ *and* $\varepsilon > 0$ *we have*

$$C_n(\varphi, \psi; \mu) \leq 2\mathcal{R}_\varphi(\varepsilon)\|\psi\|_\infty + 2\mathcal{R}_\psi(\varepsilon)\|\varphi\|_\infty + C\varepsilon^{-2r}\mathcal{R}_\varphi(\varepsilon)\mathcal{R}_\psi(\varepsilon)a_n.$$

REMARK 2.3. Notice that Theorem 2.2 does not imply Theorem 2.1, since the existence of tower does not give information about cofactor $C(\varphi, \psi)$.

Now, we state a similar theorem for non-invertible systems.

THEOREM 2.4. *Let* $f : M \to M$ *be a map defined on of smooth manifold* M. *Suppose that* (47) *holds for any* $\tilde{\varphi} \in L^\infty(M, \mu)$ *and* $\tilde{\psi} \in C^r(M)$ *for some* $r \geq 1$. *Then for any* $\varphi \in L^\infty(M, \mu), \psi \in \mathcal{C}^0(M)$ *and* $\varepsilon > 0$ *we have*

$$C_n(\varphi, \psi; \mu) \leq 2\mathcal{R}_\psi(\varepsilon)\|\varphi\|_\infty + C\varepsilon^{-r}\|\varphi\|_\infty\mathcal{R}_\psi(\varepsilon)a_n.$$

The parameter ε in Theorems 2.2 and 2.4 is free and we can choose it in a convenient way to us. Below we give some applications and corollaries of the above theorems.

2. Applications

We start with the applications of Theorem 2.1 to Hénon maps and Solenoid maps with intermittent behavior.

2.1. Hénon maps. Let $T_{a,b} : \mathbb{R} \to \mathbb{R}$ denote the Hénon map i.e.

$$T_{a,b}(x, y) = (1 - ax^2 + y, bx).$$

In [**11, 12, 13**] it was shown that there exists positive measure set $\mathcal{A}$ of parameters (a, b) such that for any $(a, b) \in \mathcal{A}$ corresponding $T_{a,b}$ admits unique SRB-measure which is mixing and the speed of mixing is exponential for Hölder continuous observables. It was shown by constructing Young towers with exponential tails. Here we assume that $(a, b) \in \mathcal{A}$, and investigate the problem of decay of correlations for continuous observables.

Theorem 2.5. *Let $f = T_{a,b}$ and $(a, b) \in \mathcal{A}$. Then there exists $\theta \in (0, 1)$ such that for any $\varphi, \psi \in \mathcal{C}(\mathbb{R}^2)$*

$$C_n(\varphi, \psi; \mu) \le 2(\|\varphi\|_\infty + \|\psi\|_\infty)\mathcal{R}_{\varphi,\psi}(\theta^n) + C\theta^n, \tag{48}$$

where C depends on φ, ψ and f.

The following corollary is a direct application of the above theorem.

Corollary 2.6. *Let $f = T_{a,b}$ and $(a, b) \in \mathcal{A}$.*

(i) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim e^{\alpha \log \varepsilon}$ i.e. if the observables φ, ψ are Hölder continuous with exponent α then $C_n(\varphi, \psi; \mu) \lesssim e^{-\alpha' n}$, $\alpha' = \alpha|\log \theta|$.*

(ii) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim e^{-|\log \varepsilon|^\alpha}$, $\alpha \in (0, 1)$ then $C_n(\varphi, \psi; \mu) \lesssim e^{-n^\alpha}$.*

(iii) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim |\log \varepsilon|^{-\alpha}$, for $\alpha > 0$ then $C_n(\varphi, \psi; \mu) \lesssim n^{-\alpha}$.*[1]

Notice that the upper bound in (i) is the same as in [**12**], which is natural to expect. The estimates in (ii)-(iii) are new.

2.2. Solenoid with intermittency. The second class of maps we consider is a solenoid map with an indifferent fixed point [**6**], which is defined as follows. Let $M = \mathbb{S}^1 \times D^2$, where D^2 is a unit disk in $\mathbb{R}^2$. For $(x, y, z) \in M$ let

$$g(x, y, z) = \left(f(x), \frac{1}{10}y + \frac{1}{2}\cos x, \frac{1}{10}z + \frac{1}{2}\sin x\right),$$

[1] Throughout the paper we use the notation $a \lesssim b$ if there exists a constant C independent of ε such that $a \le Cb$.

where $f : \mathbb{S}^1 \to \mathbb{S}^1$ is a map of degree $d \geq 2$ with the properties: f is C^2 on $\mathbb{S}^1 \setminus \{0\}$; f is C^1 on $\mathbb{S}^1$ and $f' > 1$ on $\mathbb{S}^1 \setminus \{0\}$; $f(0) = 0$, $f'(0) = 1$ and there exists $\gamma \in (0,1)$ such that $-xf'' \approx |x|^\gamma$ for all $x \neq 0$.

Theorem 2.7. *Let g be the map described above. Assume that $\gamma < 1$. Then for any $\varphi, \psi \in \mathcal{C}(M)$*

$$C_n(\varphi, \psi; \mu) \leq 2(\|\varphi\|_\infty + \|\psi\|_\infty)\mathcal{R}_{\varphi,\psi}(n^{-1/\gamma}) + Cn^{1-1/\gamma} \tag{49}$$

where constant C depends on φ and ψ.

Direct application of the theorem for specific classes of observables gives.

Corollary 2.8. *Let g be a map as in the theorem above. Then*

(i) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim e^{\alpha \log \varepsilon}$ i.e. the observables φ, ψ are Hölder continuous with exponent α then $C_n(\varphi, \psi; \mu) \lesssim e^{-\min\{\alpha/\gamma, 1/\gamma - 1\} \log n}$,*

(ii) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim e^{-|\log \varepsilon|^\alpha}$, $\alpha \in (0,1)$ then $C_n(\varphi, \psi; \mu) \lesssim e^{-(\log n)^\alpha}$.*

(iii) *If $\mathcal{R}_{\varphi,\psi}(\varepsilon) \lesssim |\log \varepsilon|^{-\alpha}$, for $\alpha > 0$ then $C_n(\varphi, \psi; \mu) \lesssim |\log n|^{-\alpha}$.*

The estimate in the first item coincides with the one given in [**6**] the remaining two are new.

Below we give applications of Theorem 2.4 for classes of observables considered in [**46**]. We choose this classes of observables not only to compare the results, but it also shows the dependence of bounds on the regularity of observables and hyperbolic quality of the underlying system.

For $\gamma \in (0,1)$ and $\delta > 1$ the following classes were defined in [**46**].

- $(R1, \gamma) := \{\psi : \mathcal{R}_\psi(\varepsilon) = \mathcal{O}(e^{\gamma \log \varepsilon})\}$.
- $(R2, \gamma) := \{\psi : \mathcal{R}_\psi(\varepsilon) = \mathcal{O}(e^{-|\log \varepsilon|^\gamma})\}$.
- $(R3, \delta) := \{\psi : \mathcal{R}_\psi(\varepsilon) = \mathcal{O}(e^{-(\log|\log \varepsilon|)^\delta})\}$.
- $(R4, \delta) := \{\psi : \mathcal{R}_\psi(\varepsilon) = \mathcal{O}(|\log \varepsilon|^{-\delta}\}$.

Notice that the first class corresponds to the class of Hölder continuous functions with exponent γ.

Corollary 2.9. (1) *Assume that $\gamma_n = \theta^n$ for some $\theta \in (0,1)$. Then*

(i) *If $\psi \in (R1, \gamma)$ then $C_n(\varphi, \psi; \mu) \lesssim \theta^{(n\gamma)/r}$.*

(ii) *If $\psi \in (R2, \gamma)$ then $C_n(\varphi, \psi; \mu) \lesssim e^{-\tau n^\gamma}$, where $\tau = \frac{|\log \theta|^\gamma}{r^\gamma}$.*

(iii) *If $\psi(R3, \delta)$ then $C_n(\varphi, \psi; \mu) \lesssim e^{-(\log n)^\delta}$.*

(iv) *If $\psi \in (R4, \delta)$ then $C_n(\varphi, \psi; \mu) \lesssim n^{-\delta}$.*

(2) *Assume that* $a_n = n^{-\alpha}$ *for some* $\alpha > 0$. *Then*

(i) *If* $\psi \in (R1, \gamma)$ *then* $C_n(\varphi, \psi; \mu) \lesssim n^{-\alpha/r}$.

(ii) *If* $\psi \in (R2, \gamma)$ *then* $C_n(\varphi, \psi; \mu) \lesssim e^{-|\log n|^{\gamma}}$.

(iii) *If* $\psi(R3, \delta)$ *then* $C_n(\varphi, \psi; \mu) \lesssim e^{-(\log|\log n|)^{-\delta}}$.

(iv) *If* $\psi \in (R4, \delta)$ *then* $C_n(\varphi, \psi; \mu) \lesssim (\log n)^{-\delta}$.

To obtain the Corollary we just choose $\varepsilon \cdot a_n = 1$. We recover all the results of [**46**] and [**25**] in the exponential case. But as formulas show that in the polynomial case for the observables of low regularity our estimates are not sufficiently good. This is because of the approximation we use.

Notice that there is an analogous corollary of Theorem 2.1, with the only difference, that we require both of the observables to have some regularity which recovers all the results of [**71**].

3. Proof of Theorem 2.1

In this section we reduce the system to a non-invertible system, as in [**68**]. We start by defining a special measure on Λ.

3.1. The natural measures on the unstable manifolds. We fix $\hat{\gamma} \in \Gamma^u$. For any $\gamma \in \Gamma^u$ and $x \in \gamma \cap \Lambda$ let $\hat{x}$ be the point $\gamma^s(x) \cap \hat{\gamma}$. Define for $x \in \gamma \cap \Lambda$

$$\hat{u}(x) = \prod_{i=0}^{\infty} \frac{\det Df^u(f^i(x))}{\det Df^u(f^i(\hat{x}))}.$$

By item (b) of assumption (A4) $\hat{u}$ satisfies the bounded distortion property. For each $\gamma \in \Gamma^u$ define the measure m_γ as

$$\frac{dm_\gamma}{d\mathrm{Led}_\gamma} = \hat{u}\mathbf{1}_{\gamma\cap\Lambda},$$

where $\mathbf{1}_{\gamma\cap\Lambda}$ is the characteristic function of $\gamma \cap \Lambda$. The proof of the following lemma is fairly standard and can be found in [**6**].

Lemma 2.10. (i) *Let* Θ *be the map defined in* (*A4*). *Then*

$$\Theta_* m_\gamma = m_{\gamma'}$$

for any $\gamma, \gamma' \in \Gamma^u$.

(ii) *Let* $\gamma, \gamma' \in \Gamma^u$ *be such that* $f^R(\gamma \cap \Lambda_i) \subset \gamma'$, *and let* $Jf^R(x)$ *denote the Jacobian of* f^R *with respect to the measures* m_γ *and* $m_{\gamma'}$. *Then* $Jf^R(x)$ *is* *constant on the stable manifolds and there is $C > 0$ such that for every $x, y \in \gamma \cap \Lambda$*

$$\left|\frac{Jf^R(x)}{Jf^R(y)} - 1\right| \le C\beta^{s(f^R(x), f^R(y))}.$$

3.2. Quotient tower. Let $\bar{\Lambda} = \Lambda/\sim$, where $x \sim y$ if and only if $y \in \gamma^s(x)$. This equivalence relation gives rise to a *quotient tower*

$$\Delta = \mathcal{T}/\sim$$

with $\Delta_\ell = \mathcal{T}_\ell/\sim$ and its partition into $\Delta_{0,i} = \mathcal{T}_{0,i}/\sim$ which we denote by $\bar{\mathcal{P}}$. There is a natural projection $\bar{\pi} : \mathcal{T} \to \Delta$. Since f^R preserves stable leaves and R is constant on them. Hence, return time $\bar{R}$ and separation time $\bar{s}$ are well defined by R and s. Moreover, we can define a tower map $\bar{F} : \Delta \to \Delta$. Let m be a measure whose restriction onto unstable manifolds is m_γ. Lemma 2.10 implies that there is a measure $\bar{m}$ on Δ whose restriction to each $\gamma \in \Gamma^u$ is m_γ. We let $J\bar{F}$ denote the Jacobian of $\bar{F}$ with respect to $\bar{m}$. It is clear that $\bar{F} : \Delta \to \Delta$ is an expanding tower. The mixing properties of $\bar{F}$ was studied firstly in [**68**]. Several papers appeared improving or extending the results given in [**68**], for example [**30, 32, 46**]. Here we combine the results from [**6**] and [**69**]. To state the theorem we introduce the space of Hölder continuous functions on Δ as

$$\mathcal{F}_\beta = \{\varphi : \Delta \to \mathbb{R} : \exists C_\varphi \text{ such that } |\varphi(x) - \varphi(y)| \le C_\varphi \beta^{\bar{s}(x,y)} \ \forall x, y \in \Delta\}.$$

$$\mathcal{F}_\beta^+ = \Big\{\varphi \in \mathcal{F} : \exists C_\varphi \text{ such that on each } \Delta_{\ell,i}, \text{ either } \varphi \equiv 0, \text{ or } \varphi > 0, \left|\frac{\varphi(x)}{\varphi(y)} - 1\right| \le C_\varphi \beta^{\bar{s}(x,y)} \ \forall x, y \in \Delta_{\ell,i}\Big\}.$$

THEOREM 2.11. [**6, 68**]

(i) *$\bar{F}$ admits unique mixing acid $\bar{\nu}$; $d\bar{\nu}/d\bar{m} \in \mathcal{F}^+$ and $d\bar{\nu}/d\bar{m} > c > 0$.*

(ii) *Let λ be a probability measure with $\varphi = d\lambda/d\bar{m} \in \mathcal{F}_\beta^+$.*

(1) *If $\bar{m}\{\bar{R} > n\} \le C\theta^n$ for some $C > 0$ and $\theta \in (0,1)$ then there exists $C' > 0$ and $\theta' \in (0,1)$ such that $|\bar{F}_*^n\lambda - \bar{\nu}| \le C'\theta'^n$.*

(2) *If $\bar{m}\{\bar{R} > n\} \le Ce^{-cn^\eta}$ for some $C, c > 0$ and $\eta \in (0,1]$ then there exists $C', c' > 0$ such that $|\bar{F}_*^n\lambda - \bar{\nu}| \le C'e^{-c'n^\eta}$. Moreover c' does not depend on φ, C' depends only on C_φ.*

(3) *If $\bar{m}\{\bar{R} > n\} \le Cn^{-\alpha}$ for some $C > 0$ and $\alpha > 1$ then there exists $C' > 0$ such that $|\bar{F}_*^n\lambda - \bar{\nu}| \le C'n^{1-\alpha}$.*

3.3. Approximation of correlations. Here we establish the relation between the original problem and problem of estimating the decay rates of correlations on the quotient tower, and then apply Theorem 2.11.

Let $\pi : \mathcal{T} \to M$, $\bar{\pi} : \mathcal{T} \to \Delta$ be the tower projections, then we have $\bar{\nu} = \bar{\pi}_* \nu$ and $\mu = \pi_* \nu$. Given $\varphi, \psi \in \mathcal{C}^0(M)$ define $\tilde{\varphi} = \varphi \circ \pi$ and $\tilde{\psi} = \psi \circ \pi$. By definition

$$\int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu = \int (\tilde{\varphi} \circ F^n)\tilde{\psi} d\nu - \int \tilde{\varphi} d\nu \int \tilde{\psi} d\nu,$$

which shows it is sufficient to obtain estimates for the lifted observables. This will be done by approximating the lifted observables with piecewise constant observables on tower. For $k \le n/4$ define $\bar{\varphi}_k$ as follows

$$\bar{\varphi}_k|_P = \inf\{\tilde{\varphi} \circ F^k(x) | \ x \in P\}, \text{ where } P \in \mathcal{P}_{2k}.$$

Define $\bar{\psi}_k$ in a similar way, and note that $\bar{\varphi}_k$ and $\bar{\psi}_k$ are constant on the stable leaves. Hence, we can consider them as a function defined on quotient tower Δ. The main result of this section is the following

PROPOSITION 2.12.

$$|\mathcal{C}_n(\tilde{\varphi}, \tilde{\psi}; \nu) - \mathcal{C}_n(\bar{\varphi}_k, \bar{\psi}_k; \bar{\nu})| \le 2(\|\varphi\|_\infty + \|\psi\|_\infty)\mathcal{R}_{\psi,\varphi}(\delta_k).$$

PROOF. The proof consists of several steps and follows the argument in [**6**]. First we claim

$$|C_n(\tilde{\varphi}, \tilde{\psi}; \nu) - C_{n-k}(\bar{\varphi}_k, \tilde{\psi}; \nu)| \le 2\|\psi\|_\infty \mathcal{R}_\varphi(\delta_k). \tag{50}$$

Indeed, using the fact $C_n(\tilde{\varphi}, \tilde{\psi}; \nu) = C_{n-k}(\tilde{\varphi} \circ F^k, \tilde{\psi}; \nu)$ the left hand side of 50 can be written as

$$\left| \int (\tilde{\varphi} \circ F^k - \bar{\varphi}_k)\tilde{\psi} d\nu + \int (\tilde{\varphi} \circ F^k - \bar{\varphi}_k) d\nu \int \tilde{\psi} d\nu \right|$$
$$\le 2\|\tilde{\psi}\|_\infty \int |\tilde{\varphi} \circ F^k - \bar{\varphi}_k| d\nu.$$

By definition of $\bar{\varphi}_k$ for $x \in P$ we have

$$|\tilde{\varphi} \circ F^k(x) - \bar{\varphi}_k| \le \sup_{x,y \in P} |\tilde{\varphi}(F^k(x)) - \tilde{\varphi}(F^k(y))| \le \mathcal{R}_\varphi(\delta_k).$$

Combining last two inequalities implies desired conclusion. Now, let $\bar{\psi}_k \nu$ be the measure whose density with respect to ν is $\bar{\psi}_k$ and let $\tilde{\psi}_k = dF^k_*(\bar{\psi}_k \nu)/d\nu$. Then

$$|C_{n-k}(\bar{\varphi}_k, \tilde{\psi}; \nu) - C_{n-k}(\bar{\varphi}_k, \tilde{\psi}_k; \nu)| \le 2\|\varphi\|_\infty \mathcal{R}_\psi(\delta_k). \tag{51}$$

After substituting and simplifying the expression we obtain

$$|C_{n-k}(\bar{\varphi}_k, \tilde{\psi}; \nu) - C_{n-k}(\bar{\varphi}_k, \tilde{\psi}_k; \nu)| \le 2\|\varphi\|_\infty \left| \int (\tilde{\psi} - \tilde{\psi}_k) d\nu \right|.$$

First observe that $F^k_*((\tilde{\psi}\circ F^k)\nu) = \tilde{\psi}\nu$. Letting $|\cdot|$ denote the variational norm for measures we have

$$\left|\int(\tilde{\psi}-\tilde{\psi}_k)d\nu\right| = |F^k_*((\tilde{\psi}\circ F^k)\nu) - F^k_*(\bar{\psi}_k\nu)|$$
$$\leq |(\tilde{\psi}\circ F^k - \bar{\psi}_k)\nu| = \int |\psi\circ F^k - \bar{\psi}_k| d\nu.$$

As in the proof of (50) we have $|\psi\circ F^k - \bar{\psi}_k| \leq \mathcal{R}_\psi(\delta_k)$ which implies relation 51. Combining the inequalities (50), (51) and the equality $C_{n-k}(\bar{\varphi}_k, \tilde{\psi}_k;\nu) = C_n(\bar{\varphi}_k,\bar{\psi}_k;\bar{\nu})$ from [**6**] finishes the proof. □

It remains to prove decay of correlations for observables $\bar{\varphi}_k$ and $\bar{\psi}_k$ on the quotient tower. We start with the usual transformations that simplify the correlation function. Without lost of generality assume that $\bar{\psi}_k$ is not identically zero. Let $b_k = \left(\int(\bar{\psi}_k + 2\|\bar{\psi}_k\|_\infty)d\bar{\nu}\right)^{-1}$ and $\hat{\psi}_k = b_k(\bar{\psi}_k + 2\|\bar{\psi}_k\|_\infty)$, then we have $\int \hat{\psi}_k d\bar{\nu} = 1$, $\|\bar{\psi}_k\|_\infty \leq b_k^{-1} \leq 3\|\bar{\psi}_k\|_\infty$ and $1 \leq \hat{\psi}_k \leq 3$. Moreover, $\hat{\psi}_k$ is constant on the elements of $\mathcal{P}_{2k}$. Thus

$$C_n(\bar{\varphi}_k,\bar{\psi}_k;\bar{\nu}) = \left|\int(\bar{\varphi}\circ\bar{F}^n)\bar{\psi}_k d\bar{\nu} - \int\bar{\varphi}_k d\bar{\nu}\int\bar{\psi}_k d\bar{\nu}\right| =$$
$$\frac{1}{b_k}\left|\int(\varphi\circ\bar{F}^n)\hat{\psi}_k d\bar{\nu} - \int\bar{\varphi}_k d\bar{\nu}\right| \leq \tag{52}$$
$$\frac{1}{b_k}\|\bar{\varphi}_k\|_\infty\int\left|\frac{d\bar{F}^n_*(\hat{\psi}_k\bar{\nu})}{d\bar{m}} - \frac{d\bar{\nu}}{d\bar{m}}\right|d\bar{m}.$$

Now, letting $\hat{\lambda}_k = \bar{F}^{2k}_*(\hat{\psi}_k\bar{\nu})$ we conclude

$$C_n(\bar{\varphi}_k,\bar{\psi}_k;\bar{\nu}) \leq \frac{1}{b_k}\|\varphi\|_\infty\left|F^{n-2k}_*\hat{\lambda}_k - \bar{\nu}\right|. \tag{53}$$

Note that the density of $\hat{\lambda}_k$ belongs to the class $\mathcal{F}^+$ (see Lemma 4.1, [**6**]). Hence, we can apply Theorem 2.11 to $\left|F^{n-2k}_*\hat{\lambda}_k - \bar{\nu}\right|$ and obtain estimates for $C_n(\bar{\varphi}_k,\bar{\psi}_k;\bar{\nu})$. Substituting them into Proposition 2.12 finishes the proof of The Main Technical Theorem.

4. Proofs of Theorems 2.5 and 2.7

We start this section with the following auxiliary construction. Consider the sequence of stopping times for the points in Λ defined as follows:

$$S_0 = 0,\ S_1 = R \text{ and } S_{i+1} = S_i + R\circ f^{S_i},\ \text{for } i \geq 1. \tag{54}$$

Let $\mathcal{Q}_0$ be the partition of Λ into Λ_i's. Define the sequence of partitions $\mathcal{Q}_k$ as: $x, y \in \Lambda$ belong to the same element of $\mathcal{Q}_k$ if the following conditions hold.

(i) $f^R(x)$ and $f^R(y)$ have the same stopping times up to time $k-1$.

(ii) $f^{S_i}(f^R(x))$ and $f^{S_i}(f^R(y))$ belong to the same element of $\mathcal{Q}_0$ for each $0 \le i \le k-1$.

(iii) $f^{S_k(Q)}$ is a u-subset.

For $Q \in \mathcal{Q}_0$ let $R(Q)$ denote its return time. Let $k \ge 1$ be arbitrary integer and define a sequence

$$\bar{\delta}_k = \sup_{Q \in \mathcal{Q}_0} \bar{\delta}_k(Q),$$

where $\bar{\delta}_k(Q)$ is defined as follows:

(1) For $k > R(Q) - 1$, let

$$\bar{\delta}_k := \sup_{0 \le \ell \le R(Q)-1} \{\mathrm{diam}(f^\ell(A \cap \gamma)) : \gamma \in \Gamma^u, A \in \mathcal{Q}_{k-R(Q)+1+\ell}, A \subset Q\}.$$

(2) For $k \le R(Q) - 1$, let

$$\bar{\delta}_k^0(Q) := \sup_{0 \le \ell < R(Q)-k} \{\mathrm{diam}(f^\ell(Q \cap \gamma)) : \gamma \in \Gamma^u\},$$

$$\bar{\delta}_k^+ := \sup_{R(Q)-k \le \ell \le R(Q)-1} \{\mathrm{diam}(f^\ell(A \cap \gamma)) : \gamma \in \Gamma^u, A \in \mathcal{Q}_{k-R(Q)+1+\ell}, A \subset Q\}$$

and define

$$\bar{\delta}_k(Q) = \sup\{\bar{\delta}_k^0(Q), \bar{\delta}_k^+(Q)\}.$$

From Lemma 3.2 in [**6**] we have

$$\mathrm{diam}(\pi(F^k(P)) \le C \max\{\beta^k, \bar{\delta}_k\} \tag{55}$$

for any $P \in \mathcal{P}_{2k}$, $k \ge 0$, and some $C > 0$. This is the main estimate we use to prove Theorems 2.5 and 2.7.

In [**6**] it was proven that $\bar{\delta}_k \lesssim k^{-1/\gamma}$ and $m\{R > k\} \le k^{-1/\gamma}$ for the Solenoid map with intermittent fixed point. Substituting this into (55) and applying item (ii) of the Main Technical Theorem finishes the proof of Theorem 2.7.

In [**13**] it was shown that for any $(a, b) \in \mathcal{A}$ the corresponding Hénon map admits a Young tower for which the tail of the return time decays exponentially. Therefore to complete the proof of Theorem 2.6 it is sufficient to show that $\bar{\delta}_k$ decays exponentially. We will show this in the following lemma and complete the proof in this case as well.

In [**68**] (see [**13**] for the details of the construction) it was shown that Hénon maps satisfy backward contraction on the unstable leaves, that is there exits $C > 0$ such that for all $x, y \in \Lambda_i$ with $y \in \gamma^u(x)$ and $0 \le n \le R_i$

$$\mathrm{dist}(f^n(x), f^n(y)) \le C\beta^{R_i - n}. \tag{56}$$

LEMMA 2.13. *$\exists C > 0$ and $\beta' \in (0,1)$ such that $\bar{\delta}_k \le C\beta'^k$.*

PROOF. A) We start with the case $k \le R(P) - 1$ and $0 \le \ell < R(P) - k$. By (56) for any $x \in P$

$$\mathrm{diam}(f^\ell(P \cap \gamma^u(x))) \le C\beta^{R(P)-\ell} \le C\beta^k.$$

This implies $\bar{\delta}_k^0 \lesssim \beta^k$.

B) Now, consider the case $k \le R(P)-1$, and $R(P)-k < \ell \le R(P)-1$. Notice that for any $Q \subset P$, $Q \in \mathcal{P}_{k-R(P)+\ell+1}$ the stopping times $S_1, ..., S_{\ell'}$, $\ell' = k - R(P) + \ell$, are constant on Q and $f^{S_i}(Q) \subset P_i$, for some $P_i \in \mathcal{Q}_0$, $i = 1, ..., \ell'$. Let $r_1, ..., r_{\ell'-1}$ be the return times of these elements. By (56) we have

$$\mathrm{diam}(Q \cap \gamma^u) \le C\beta^{R(P)+r_1+...+r_{\ell'-1}}$$

Since $R(P) + r_1 + ... + r_{\ell'-1} \ge l + k$ using again the inequality (56)

$$\mathrm{diam}(f^\ell(Q \cap \gamma^u)) \le C\beta^{R(P)+r_1+...+r_{\ell'-1}} \le C\beta^k.$$

This implies $\bar{\delta}_k^+ \lesssim \beta^k$, which finishes the proof when $k \le R(P)$. The case $k > R(P)$ is treated as B). ☐

In this chapter we give a unified approach to obtain the rates of decay of correlations for continuous observables for the systems defined on smooth manifolds. Our argument is similar to the ones that were used in [**27, 40**] to show that exponential mixing for $C^r, r \ge 1$ observables implies exponential mixing for Hölder observables. Here we show the argument works in more general situation. More precisely, if decay rates for $C^r, R \ge 1$ observables are known, then it is possible to obtain decay rates of correlations for observables with lower regularity. This shows that in fact, it is sufficient to obtain decay rates for smooth observables.

5. Proof of Theorem 2.2

Here we prove the Theorem 2.2. The proof of Theorem 2.4 is analogous. Since the approximation we are going to use is local we can directly work on $\mathbb{R}^m$.

We start by introducing the mollifiers. Let $k = (k_1, ..., k_m)$ be a multiplex with length $|k| = k_1 + ... + k_m$. Denote by D^k a differential operator of order $|k|$. Let $\rho \in C^\infty(\mathbb{R}^m)$ be such that[2]

i) $\rho \ge 0$, $\sup(\rho) \subset B(0,1)$, $\int_{\mathbb{R}^m} \rho d\mu = 1$;

[2] where $B(0,1)$ denotes the unit ball centered at origin

ii) there exists constant C such that $|D_x^k \rho(x)| \leq C$, for any k with $|k| \in \{1, ..., r\}$, where $r > 1$ is some fixed integer.

Given $\varepsilon > 0$ define $\rho_\varepsilon(x) = \varepsilon^{-m}\rho(\varepsilon^{-1}x)$. Then $\sup(\rho_\varepsilon) \subset B(0, \varepsilon)$ and $\int_{\mathbb{R}^m} \rho_\varepsilon d\mu = 1$. For $\phi \in \mathcal{C}^0(\mathbb{R}^m)$ define smoothed version $\phi_\varepsilon \in C^\infty(\mathbb{R}^m)$ as

$$\phi_\varepsilon(x) = \int_{\mathbb{R}^m} \phi(y)\rho_\varepsilon(x - y)dy. \tag{57}$$

The following lemma gives estimates on the quality of the approximation and bound for its derivatives.

Lemma 2.14. *For any $x \in \mathbb{R}^m$ we have*

$$|\phi(x) - \phi_\varepsilon(x)| \leq \mathcal{R}_\phi(\varepsilon) \quad \textit{and} \quad \|\phi_\varepsilon(x)\|_{C^r} < \varepsilon^{-r} C\mathcal{R}_\phi(\varepsilon). \tag{58}$$

Proof. Since $\int \rho_\varepsilon(y)dy = 1$, using the definition of $\mathcal{R}_\phi(\varepsilon)$ we have

$$\begin{aligned} |\phi_\varepsilon(x) - \phi(x)| &\leq |\int \phi(y)\rho_\varepsilon(x-y)dy - \int \phi(x)\rho_\varepsilon(x-y)dy| \\ &\leq \int |\phi(y) - \phi(x)|\rho_\varepsilon(x-y)dy \leq \mathcal{R}_\phi(\varepsilon). \end{aligned}$$

To prove the second item, it is sufficient to show that for any $k = 1, 2, ..., r$ one has $|D^k \phi_\varepsilon(x)| \lesssim \varepsilon^{-k}\mathcal{R}_\phi(\varepsilon)$.

Direct calculation shows that $D^k\phi_\varepsilon(x) = \int D^k\rho_\varepsilon(x-y)\phi(y)dy$. Noting that $\int D^k\rho(x-y)dy = 0$ we can write

$$\begin{aligned} |D^k\phi_\varepsilon(x)| &\leq \int |D^k\rho_\varepsilon(x-y)(\phi(y) - \phi(x))|dy \\ &\leq \varepsilon^{-m-k}\mathcal{R}_\phi(\varepsilon) Vol B(0, \varepsilon) \leq \varepsilon^{-k}\mathcal{R}_\phi(\varepsilon)C. \end{aligned}$$

□

Now, for given observables $\varphi, \psi \in \mathcal{C}^0(\mathbb{R}^n)$ we approximate their correlations by the correlations of mollified observables $\varphi_\varepsilon, \psi_\varepsilon \in C^\infty(\mathbb{R}^n)$ as follows.

$$\begin{aligned}
&|\int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu| \\
&\leq |\int \varphi \circ f^n(\psi - \psi_\varepsilon) d\mu| + |\int (\varphi \circ f^n - \varphi_\varepsilon \circ f^n)\psi_\varepsilon d\mu| \\
&+ |\int \varphi_\varepsilon \circ f^n \psi_\varepsilon d\mu - \int \varphi_\varepsilon d\mu \int \psi_\varepsilon d\mu| \\
&+ |\int (\varphi_\varepsilon - \varphi) d\mu \int \psi_\varepsilon d\mu| + |\int \varphi d\mu \int (\psi - \psi_\varepsilon) d\mu| \\
&\leq 2\mathcal{R}_\varphi(\varepsilon) \int \psi d\mu + 2\mathcal{R}_\psi(\varepsilon) \int \varphi d\mu + C\|\varphi_\varepsilon\|_{C^r} \|\psi_\varepsilon\|_{C^r} a_n.
\end{aligned}$$

To obtain the last inequality we used the assumption on decay rates for smooth observables and the first item of (58). Now, using the second item of (58) we obtain

$$\begin{aligned}
&|\int (\varphi \circ f^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu| \leq \\
&\leq 2\mathcal{R}_\varphi(\varepsilon)\|\psi\|_\infty + 2\mathcal{R}_\psi(\varepsilon)\|\varphi\|_\infty + C\varepsilon^{-2r}\mathcal{R}_\varphi(\varepsilon)\mathcal{R}_\psi(\varepsilon) a_n.
\end{aligned}$$

6. No decay for too big classes of observables

Here we give some counterexamples to show non-existence of decay rates. First example concerns the general (invertible or non-invertible) map and the second example is only for invertible maps.

6.1. General case.

Example 1. Let $f : M \to M$ be map defined on a probability space (M, μ) and preserving μ. Suppose that μ is mixing non-atomic probability measure. *For any sequence of positive numbers γ_n such that $\lim_{n\to\infty} \gamma_n = 0$ there exists $\varphi, \psi \in L^2(M, \mu)$ such that $C_n(\varphi, \psi; \mu)$ decays slower than γ_n.*

Proof. Let us fix a function $\varphi \in L^2(\mu)$, with $\int \varphi d\mu = 0$ and $\|\varphi\|_{L^2} = 1$. Define a sequence of linear functionals $\Phi_n : L^2(M, \mu) \to \mathbb{R}$

$$\Phi_n := C_n(\varphi, \psi; \mu) = \int \psi(\varphi \circ f^n) d\mu.$$

Assume by contradiction that there exists a sequence $\gamma_n \to 0$ such that for every $\varphi \in L^2(\mu)$

$$\sup_n |\Phi_n/\gamma_n| < \infty.$$

Hence, by the uniform boundedness principle [**35**],

$$\|\Phi_n\|/\gamma_n < \infty.$$

On the other hand for any n, we have $\psi = \varphi \circ f^n \in L^2(M,\mu)$ and

$$\Phi_n(\psi) = \int (\varphi \circ f^n)^2 = 1.$$

Since we can choose arbitrary n and $\gamma_n \to 0$ as $n \to \infty$ we get a contradiction. □

6.2. Invertible case.

EXAMPLE 2. Now, let $f : M \to M$ be an invertible system with non-atomic mixing measure μ. *If $\varphi \in L^2(M,\mu)$ then (f,μ) does not exhibit decay of correlations with respect to any Banach space $\mathcal{B} \subset L^2(M,\mu)$.*

PROOF. Assume by contradiction there exits γ_n $\gamma_n \to 0$ and a constant $C := C(\varphi,\psi)$ such that

$$C_n(\varphi,\psi;\mu) \le C\gamma_n.$$

Choose $\psi \in \mathcal{B}$ be such that $\int \psi d\mu = 0$, and $\|\psi\|_{L^2} \neq 0$. Then we have

$$\begin{aligned}\|\psi\|_{L^2} = \|\psi \circ f^{-n}\|_{L^2} &= \sup_{\|\varphi\|_{L^2}=1} \int \varphi\psi \circ f^{-n} d\mu \\ &= \sup_{\|\varphi\|_{L^2}=1} \int \varphi \circ f^n \psi d\mu \le C_n(\varphi,\psi,\mu) \le C\gamma_n,\end{aligned}$$

which is a contradiction, since the left hand side is fixed positive and right hand side converges to 0. □

Bibliography

[1] J.F. Alves, C. Dias, S. Luzzatto, V. Pinheiro, SRB measures for partially hyperbolic systems whose central direction is weakly expanding. Preprint.

[2] J.F. Alves, C. Dias, S. Luzzatto, Geometry of expanding absolutely continuous invariant measures and the liftability problem, Ann. Inst. Henri Poincaré, Analyse Non Linéaire, 30, 101-120, 2013.

[3] J.F.Alves, J.M.Freitas, S.Luzzatto, S.Vaienti, From rates of mixing to recurrence times via large deviations, Advances in Mathematics, 228, 1203-1236, 2011.

[4] J. F. Alves, X. Li, Gibbs-Markov-Young structure with (stretched) exponential recurrence times for partially hyperbolic attractors, Preprint.

[5] J.F.Alves, S. Luzzatto, V. Pinheiro, Markov structures and decay of correlations for non-uniformly expanding dynamical systems, Ann. Inst. Henri Poincaré, Analyse Non Linéaire 22, 817-839, 2005.

[6] J.F. Alves, V. Pinheiro, Slow rates of mixing for dynamical systems with hyperbolic structures, J. Stat. Phys., 131, 505-534, 2008.

[7] J.F. Alves, D. Schnellmann, Ergodic properties of Viana-like maps with singularities in the base dynamics. Proc. AMS. 141, 3493–3955, 2013

[8] A. Avez, Propriétés ergodiques des endomorphisms dilatants des variétés compactes, C.R. Acad. Sci. Paris Sér. A-B 266, A610-A612, 1968.

[9] V. Baladi, Positive transfer operators and decay of correlations, World Scientific, 2000.

[10] V. Baladi, S. Gouëzel, Stretched exponential bounds for the correlations of the Viana-Alves skew products. Second Workshop on Dynamics and Randomness, Universidad de Chile, 2002.

[11] M. Benedicks, L. Carleson, The dynamics of the Hénon map. Ann. Math., 122:1-25, 1985.

[12] M. Benedicks, L.-S. Young, Sinai-Bowen-Ruelle measures for certain Hénon maps, Invent. Math, 112, 541-576, 1993.

[13] M. Benedicks, L.-S. Young, Markov extensions and decay of correlations for certain Hénon maps. Géométrie complexe et systémes dynamiques (Orsay, 1995). Astérisque No. 261, 13-56, 2000.

[14] V.I.Bogachev, Measure theory v. 1. Springer, 2006, 500 p.

[15] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, Springer Lecture Notes in Math., 75 p., 1975.

[16] X. Bressaud, R. Fernández, A. Galves, Decay of correlations for non-Hölderian dynamics. A coupling approach, Electron. J. Probab. 4, no. 3, 19 p. (electronic) 1999.

[17] H. Bruin, S. Luzzatto, S. van Strien, Decay of correlations in one-dimensional dynamics, Annales de l'ENS, 36, 621-646, 2003.

[18] J. Buzzi, V. Maume-Deschamps, Decay of correlations on towers with non-Hölder Jacobian and non-exponential return time, Discrete and Continuous Dynamical Systems, 12, no. 4, 639-656, 2005.

[19] J. Buzzi, O. Sester, M. Tsujii, Weakly expanding skew-products of quadratic maps, Ergod. Th. Dynam. Syst.23, 1401-1414, 2003.

[20] N. Chernov, Statistical properties of piecewise smooth hyperbolic systems in high dimensions. Disc. and Cont. Dyn. Syst. 5, 425–448, 1999

[21] N. Chernov, Decay of correlations and dispersing billiards. J. Stat. Phys. 94, 513-556, 1999.

[22] N. Chernov, R. Markarian, Chaotic billiards. Volume 127 of Mathematical Surveys and Monographs. Providence, RI: Amer. Math. Soc., 2006.

[23] P. Collet, A remark about uniform de-correlations prefactor, Preprint.

[24] I. Cornfeld, S. Fomin, Y. Sinai, Ergodic theory, Springer-Verlag, p. 486, 1982.

[25] K. Diaz-Ordaz, Decay of correlations for non-Hölder observables for one-dimensional expanding Lorenz-Like maps, Discrete and continuous dynamical systems, Vol. 15, no. 1, 159-176, 2006.

[26] K. Diaz-Ordaz, M.P. Holland, S. Luzzatto, Statistical properties of one-dimensional maps with critical points and singularities, Stochastics and Dynamics, Vol. 6, No. 4, 423-458, 2006.

[27] D. Dolgopyat, On decay of correlations in Anosov flows, Annals of Mathematics 147: 2 357-390, 1998.

[28] D. Dolgopyat, On dynamics of mostly contracting diffeomorphisms. Comm. Math. Phys, 213:181-201, 2000.

[29] A. Fisher, A. Lopes, Exact bounds for the polynomial decay of correlation, $1/f$ noise and the CLT for the equilibrium state of a non-Hölder potential,Nonlinearity,14, 1071-1104 2001.

[30] S. Gouëzel, Decay of correlations for nonuniformly expanding systems, Bull. Soc. Math. France, Vol. 134, no. 1, 1-31, 2006.

[31] F. Hofbauer, G. Keller Ergodic properties of invariant measures for piecewise monotonic transformations. Math. Z. 180:119-140, 1982.

[32] M. Holland Slowly mixing systems and intermittency maps, Ergodic theory and Dynamical Systems, Vol. 25, no. 1, 133-159, 2004.

[33] H.Hu, Decay of correlations for piecewise smooth maps with indifferent fixed points, Ergodic theory and Dynamical Systems, Vol 24, no. 2, 495–524, 2004.

[34] S. Isola, On the systems with finite ergodic degree, Far East Journal of Dynamical Systems, 5, 1-62, 2003.

[35] T. Kato, Perturbation theory for linear operators, Springer-Verlag, 1980.

[36] G. Keller, T. Nowicki, Spectral theory, zeta functions and the distributions of points for Collet-Eckman maps, Comm. Math. Phys. 149, 31-69, 1992.

[37] A. Kondah, V. Maume, B. Schmitt, Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non Höldiériennes, Ann. Inst. H. Poincaré Probab. Statist., 33, no. 6, 675-695 1997.

[38] A. Lasota, J.A. Yorke, On the existence of invariant measures for piecewise monotonic transformations, Trans. of the AMS, Vol. 186, 1973.

[39] C. Liverani, Decay of correlations. Annals Math. 142 (2), 239-301, 1995.

[40] C. Liverani, On contact Anosov flows, Annals of Mathematics, 159:3, 1275-1312, 2004.

[41] C. Liverani, Multidimensional expanding maps with singularities: a pedestrian approach Ergod. Th. Dyn. Syst. 33, 168-182, 2013.

[42] A. Lopes, Entropy and large deviations, Nonlinearity 3, 527-546, 1990.

[43] S. Luzzatto, Stochastic-like behaviour in non-uniformly expanding maps, in Handbook of Dynamical Systems, Vol 1B. 265-326, Elsevier, 2006.

[44] S. Luzzatto, I. Melbourne, Statistical properties and decay of correlations for interval maps with critical points and singularities, Commun. Math. Phys. 320, 21-35, 2013.

[45] V. Lynch, Non-uniformly expanding dynamical systems and decay of correlations for non-Hölder continuous observables, PhD Thesis, University of Warwick, 2003.

[46] V. Lynch, Decay of correlations for non-Hölder observables, Discrete and Continuous Dynamical Systems, 16, no. 1, 19-46, 2006.
[47] I. Melbourne, M. Nicol, Large deviations for nonuniformly hyperbolic systems, Trans. Amer. Math. Soc. 360, 6661-6676, 2008.
[48] V. Maume-Deschamps, Projective metrics and mixing properties on towers, Trans. Amer. Math. Soc.,353, 3371-3389, 2001.
[49] I. Melbourne, M. Nicol, Large deviations for nonuniformly hyperbolic systems, Trans. Amer. Math. Soc. 360, 6661-6676, 2008.
[50] I. Melbourne, M. Nicol, Almost sure invariance principle for nonuniformly hyperbolic systems, Commun. Math. Phys., 260, 131-146, 2005.
[51] P. Natalini, B. Palumbo, Inequalities for the Incomplete Gamma function, Mathematical Inequalities & Applications, 3 (1), 69-77, 2000.
[52] T. Nowicki, S. van Strien, Absolutely continuous invariant measures for C^2 unimodal maps satisfying the Collet-Eckmann conditions, Invent. Math. 93, 619-635, 1988.
[53] T. Nowicki, D. Sands, Non-uniform Hyperbolicity and Universal Bounds for S-unimodal Maps, Invent. Math. 132, no. 3, 633-680, 1988.
[54] V. I. Oseledec, A multiplicative ergodic theorem: Lyapunov characteristic numbers of dynamical systems, Trans. Moscow Math. Soc. 19, 197-231, 1968.
[55] Ya. B. Pesin, Characteristic Lyapunov exponents and smooth ergodic theory, Russ. Math. Surveys, 32, 55-114, 1977.
[56] Ya. B. Pesin, Families of invariant manifolds corresponding to non-zero characteristic exponents, Math. of the USRR, Izvestija 10, 1261-1305, 1978.
[57] Ya. B. Pesin, Ya. G. Sinai, Gibbs measures for partially hyperbolic attractors, Erg. Th. & Dynam. Syst.2, 417-438, 1982.
[58] V. Pinheiro, Expanding Measures Ann. I. H. Poincare – AN 28, 889-939, 2011.
[59] M. Pollicott, Rates of mixing for potentials of summable variation, Trans. Amer. Math. Soc, 352, 843-853, 2000.
[60] M. Pollicott, M. Yuri, Statistical properties of maps with indifferent periodic points, Commun. Math. Phys. 217, 503-520, 2001.
[61] D. Ruelle, A measure associated with Axiom A attractors. Amer. J. Math. 98:619-654, 1976.
[62] M. Ruziboev, Decay of correlations for invertible maps with non-Hölder observables, Dynamical Systems
[63] Y. Sinai, Gibbs measures in ergodic theory, Russ. Math. Surveys, 27(4) 1972.
[64] Y. Sinai, Dynamical systems with elastic reflections, Ergodic properties of dispersing billiards. Uspekhi. Mat. Nauk 25, 141-192, 1970.
[65] T.Tao, V.H.Vu Additive combinatorics, Cambridge studies in advanced mathematics, v. 105, 2006.
[66] D. Thomine, A spectral gap for transfer operators of piecewise expanding maps, Discrete and continuous time Dynamical systems, Vol. 30, n. 3, 917-944, 2011.
[67] L.-S. Young , Decay of correlations for certain quadratic maps, Comm. Math. Phys. Vol. 146, No. 1, 123-138, 1992.
[68] L.-S. Young, Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. 147, 585-650, 1998.
[69] L.-S. Young, Recurrence times and rates of mixing, Israel J. Math. 110, 153-188, 1999.
[70] L.-S. Young, What are SRB measures, and which dynamical systems have them? J. Stat. Phys. Vol. 108, no. 5-6, 733-754, 2002.
[71] H.-K. Zhang, Decay of correlations on non-Hölder observables, Int. J. Nonlin. Sci. 10:3 375-385, 2010.