Polynomial Bounds for the Decay of Correlations in Non-uniformly Expanding Maps

Huyi Hu * Sandro Vaienti †

November 8, 2015

Abstract

We give conditions under which nonuniformly expanding maps exhibit upper and lower bounds of polynomial type for the decay of correlations and for a large class of observables. We show that if the Lasota-Yorke type inequalities for the transfer operator of a first return map are satisfied in a Banach space $B$, and the absolutely continuous invariant measure obtained is weak mixing, in terms of aperiodicity, then under some renewal condition, the maps have polynomial decay of correlations for observables in $B$. We also provide some general conditions that give aperiodicity for expanding maps in higher dimensional spaces. As applications, we obtain polynomial decay, including lower bound in some cases, for piecewise expanding maps with an indifferent fixed point and for which we also allow non-Markov structure and unbounded distortion. The observables are functions that have bounded variation or satisfy quasi-Hölder conditions respectively and, in the case of polynomial lower bounds, their support is bounded away from the neutral fixed points.

Contents

0 Introduction 2

*Mathematics Department, Michigan State University, East Lansing, MI 48824, USA. e-mail: <hu@math.msu.edu>.
†UMR-7332 Centre de Physique Théorique, CNRS, Université d’Aix-Marseille, Université du Sud Toulon-Var and FRUMAM, Fédération de Recherche des Unités des Mathématiques de Marseille, CPT Luminy, Case 907, F-13288 Marseille CEDEX 9. e-mail: <vaienti@cpt.univ-mrs.fr>. HH was supported by the University of Marseille and the University of Toulon during his visits at the Center of Theoretical Physics in Luminy. SV was supported by the ANR-Project Perturbations, by the CNRS-PEPS Mathematical Methods of Climate Theory and by the PICS (Projet International de Coopération Scientifique), Propriétés statistiques des systèmes dynamiques déterministes et aléatoires, with the University of Houston, n. PICS05968. Part of this work was done while he was visiting the Centro de Modelamiento Matemático, UMI2807, in Santiago de Chile with a CNRS support (délégation).
The purpose of this paper is to study polynomial decay of correlations for invariant measures which are absolutely continuous with respect to the Lebesgue measure on compact subsets of $\mathbb{R}^n$. Typically the maps $T$ which we consider are non uniformly expanding and they may neither have a Markov partition nor exhibit bounded distortion. The main tool we use is the transfer (Perron-Frobenius) operator on induced subsystems endowed with the first return map.

We now explain in details the content of this paper. Let us consider a non uniformly expanding map $T$ defined on a compact subset $X \subset \mathbb{R}^n$, with or without discontinuities. Since we do not have necessarily bounded distortion or Markov partitions, Hölder continuous functions are not preserved under the transfer operator. Therefore we will work on Banach spaces $B$ embedded in $L^1$ with respect to the Lebesgue measure, and use a norm $\| \cdot \|_B$ stronger than the $L^1$ norm $\| \cdot \|_{L^1}$. We will give some conditions on $B$ under which the results apply, see Assumption B.

Let us now take a subset $\tilde{X} \subset X$ and define the first return map $\tilde{T}$. The first ingredient of our theorem is the Lasota-Yorke inequality for the transfer operator $\mathcal{P}$ of $\tilde{T}$ with respect to the norms $\| \cdot \|_B$ and $\| \cdot \|_{L^1}$. Hence, $\mathcal{P}$ has a fixed point $\tilde{h}$ that defines an absolutely continuous measure $\tilde{\mu}$ invariant under $\tilde{T}$. The measure $\tilde{\mu}$ can be extended to a measure $\mu$ on $X$ invariant under $T$. We may assume ergodicity for $\tilde{\mu}$, otherwise we can take an ergodic component. Then the ergodicity of $\tilde{\mu}$ gives ergodicity of $\mu$. However, we also need some mixing property for $\mu$. Therefore our second ingredient is to require that the function $\tau$ given by the first return time is aperiodic, which is equivalent to the weak mixing of $\mu$ for $T$. The third ingredient are precise tail estimates as they are required in the renewal theory approach. In this regard, let us call
\\[\|R_n\|\] the operator norm (see below) of the \(n\)-th power of the transfer operator restricted to the level sets with first return time \(\tau = n\); then we will ask that
\\[\sum_{k=n+1}^{\infty} ||R_k||\]
decays at least as \(n^{-\beta}\), with \(\beta > 1\). Such a decay gives also an estimate, through the exponent \(\beta\), of the \textit{error term} denoted by the function \(F_\beta(n)\) in the basic inequality (1.5) of Theorem A below. Whenever that error term goes to zero faster than \(\sum_{k=n+1}^{\infty} \mu(\tau > k)\), the latter sum gives a lower bound for the decay of correlations and we will refer to this situation as the optimal rate: this will be shown to hold in the situations of Section 5. In any case it follows from the general theory of renewal developed by Sarig [Sr] and successively improved by Gouëzel [Go], that the decay of correlations
\\[\text{Cov}(f, g \circ T^n) := \int f \circ T^n d\mu - \int f d\mu \int g d\mu],
is polynomial for functions \(f \in \mathcal{B}\) and \(g \in L^\infty\) (with respect to the Lebesgue measure) and \(\text{supp } f, \text{supp } g \subseteq \hat{X}\).

The proof of aperiodicity in Theorem B is particularly technical. We use some results in the theory developed in the paper [ADSZ], where aperiodicity is proved for a large class of interval maps, and some methods in [AD] for skew product rigidity. We extend the aperiodicity result to the multidimensional setting without Markov partitions thus pursuing the program started in [ADSZ], which was just oriented to treat the non-Markov cases especially for one-dimensional systems.

Our results in Part I suggest a more unified approach for the theory originally proposed by Sarig in [Sr] for nonuniformly expanding systems. In [Sr], to get upper and lower bounds of polynomial decay of correlations for a dynamical system, one needs the existence of an invariant measure, and then spectral gap and aperiodicity. Our results show that for some nonuniformly expanding maps, all of the conditions are implied by some Lasota-Yorke type of inequalities in some suitable Banach spaces. More precisely, Theorem A and B indicate that in order to get polynomial decay of correlations for some nonuniformly expanding maps \(T\) with respect to its absolutely continuous invariant and mixing measure \(\mu\), we should proceed in the following way:

(i) Choose a subset \(\hat{X} \subset X\), define a first return map \(\hat{T}\), and verify Assumption T(a) to (d) stated in Section 1.

(ii) Find a suitable Banach space \(\mathcal{B}\) that satisfies Assumption B(a) to (f) stated in Section 1.

(iii) Verify three Lasota-Yorke type of inequalities (1.3), (1.4), and (1.7); practically, if one can obtain (1.3), the others follow by similar computations.

(iv) Estimate \(\|R_n\|_{\mathcal{B}}\).

The second part of the paper is for application, in which we use the above procedure. In [Sr] and [Go], the authors took the Banach space \(\mathcal{B}\) as the set of
functions that are Hölder continuous on certain sets, and obtained polynomial decay of correlations for systems for which the induced set is Gibbs-Markov. In this paper we will study piecewise smooth expanding maps with an indifferent fixed point in one and higher dimensional spaces in which Markov properties may not hold, and obtain polynomial decay of correlations for observables in different Banach spaces.

In the one-dimensional case we use the set of bounded variation functions for the Banach space $B$, and we find that the decay rates are of order $n^{d-1}$ if near the fixed point the map has the form $T(x) \approx x + x^{1+\gamma}$, $\gamma \in (0,1)$ and $d = 1/\gamma$. Upper bounds for the decay of correlations for these kinds of maps were already given by Young [Yo2] and by Melbourne and Terhesiu, see Sect. 5.3 in [MT].

As a matter of fact, one of the main goals in our paper is to obtain polynomial decay of correlations for piecewise smooth expanding maps with an indifferent fixed point in higher dimensional spaces.

For a large class of those maps, we constructed, in a previous paper ([HV]), an acim by using the Lasota-Yorke inequality. Those maps were written in the form of (4.4) near the indifferent fixed point $p$, where the local behavior is precisely given by an isometry plus homogeneous terms and higher order terms. In the present paper we show that such maps have polynomial decay of correlations for observables in $B$. As we said above, in the estimates we must compare the decay of $\|R_n\|$ with the measure of the level sets with the first return time larger than $n$. The former will be estimated by the norms $\|DT^{-n}\|$ or the determinants $|\det DT^{-n}|$, where the latter is often of order $n^{-(m/\gamma)-1}$, with $m = \dim X$ and $\gamma$ is given in (4.4)*. If $\|R_n\|$ decreases as $|\det DT^{-n}|$, then it usually has the same order as $\mu(\tau = n)$, which approaches to 0 faster than $\mu(\tau > n)$ does. Since $\sum_{k\geq n} \mu(\tau > k)$ gives the optimal decay rates of correlations and $\sum_{k\geq n} \|R_k\|$ determines the order of the error terms $F_\beta(n)$, we can get both upper and lower estimates for decay rates.

In particular we will obtain optimal rates under the assumption that all the pre-images of some neighborhood of $p$ do not intersect discontinuities, (see Theorem E and examples in Subsection 5.2 for more details). This is satisfied for instance whenever $T$ has a Markov partition or a finite range structure (see Remark 5.1). Moreover in Example 5.5 and thereafter we show the existence and the abundance of those systems with all the pre-images of some neighborhood of $p$ not intersecting discontinuities, but without any Markov and finite range structure. Whenever $\|R_n\|$ decreases as $\|DT^{-n}\|$, we get polynomial upper bounds (Theorem D). In the special case $\dim X = 1$, $\|DT^{-n}\|$ and $|\det DT^{-n}|$ become the same order, and we can get optimal rates of decay of correlations using $\|DT^{-n}\|$ (Theorem C).

We would like to point out two main issues which make the higher dimen-

*Notice that $T^{-n}$ denotes the inverse of $T^n$ restricted to the domain of injectivity containing $p$. 

4
sional case more complicated. The first is due to unbounded distortion of the map. This is caused by different expansion rates in different directions as a point move away from the indifferent fixed point even if $DT_p = \text{id}$ at the fixed point (see Example 1, part (A) in [HV]). The second comes from the difficulty to estimate the decreasing rates of the norm $\|R_n\|$ for quasi-Hölder spaces: Theorems D and E deal with these situations under certain hypotheses. One surely needs more work to weaken those assumptions and achieve optimal decay for a much larger class of maps.

To study statistical properties for non uniformly hyperbolic or expanding systems, it is necessary, in a few approaches, to find some “good” part on which one can get bounded distortion, like Pesin’s blocks ([Ps]), elements in Young’s tower ([Yo1, Yo2]), or some neighborhood near points that have hyperbolic times ([ABV]). Another direction is to work directly on some Banach spaces, like bounded variation functions ([LY]) or quasi-Hölder functions ([Ss]), that are preserved by the transfer operator. Our paper follows the latter way which revealed to be particularly efficient and fruitful for non-uniformly expanding endomorphisms.

We also observe that renewal theory has been recently used by Melbourne and Terhesiu in the already quoted paper [MT] to give, in some situations, upper bounds under less restrictive conditions.

We conclude this introduction with a few precisions on the structure of the exposition. There will be three blocks of long assumptions called B, S and T. “B” stays for “Banach” and contains the hypothesis on the function spaces. “S” collects the basic assumptions which allowed Sarig to establish his renewal results. “T” deals with the assumptions which must be satisfied for all the maps $T$ we will work with; then “T” splits into Assumptions T’ for 1-D maps (Section 3) and Assumptions T” for higher dimensional maps. The latter maps are covered in Sections 4 and 5 and they share all the Assumptions T”, but (b) and (c) which will change according to the role played by the derivative (Sect. 4) and the determinant (Sect. 5). We would like to stress that in the lack of any convincing classification of higher-dimensional non-uniformly expanding endomorphisms, one must delimitate very carefully the domain of the maps under investigation: we are basically using here the class of transformations we introduced in [HH] which, despite a few technical requirements which explain the length of the list of assumptions, show interesting behaviors like unbounded distortion and the presence of sigma-finite components.

Part I: Conditions for Polynomial Decay Rates
1 Assumptions and statements of results

Let $X \subset \mathbb{R}^m$ be a subset with positive Lebesgue measure $\nu$. We assume $\nu(X) = 1$. Let $d$ be the (euclidean) metric induced from $\mathbb{R}^m$.

The transfer operator $\mathcal{P} = \mathcal{P}_{\nu} : L^1(X, \nu) \to L^1(X, \nu)$ is defined by $\int \psi \circ T \phi \, d\nu = \int \psi \, \mathcal{P} \phi \, d\nu \ \forall \phi \in L^1(X, \nu), \ \psi \in L^\infty(X, \nu)$, where the latter $L^p$ spaces are defined with respect to the Lebesgue measure $\nu$ on the Borel $\sigma$-algebra of $X$.

Let $\tilde{X} \subset X$ be a measurable subset of $X$ with positive Lebesgue measure.

Recall that the first return map of $T$ with respect to $\tilde{X}$ is defined by $\tilde{T}(x) = T^{\tau(x)}(x)$, where $\tau(x) = \min\{i \geq 1 : T^i x \in \tilde{X}\}$ is the return time. We put $\hat{\nu}$ the normalized Lebesgue measure on $\tilde{X}$. Then we let $\hat{\mathcal{P}} = \mathcal{P}_{\hat{\nu}}$ be the transfer operator of $\hat{T}$.

Moreover we define

$$R_n f = 1_{\tilde{X}} \cdot \mathcal{P}^n(f 1_{\{\tau = n\}}) \quad \text{and} \quad T_n f = 1_{\tilde{X}} \cdot \mathcal{P}^n(f 1_{\tilde{X}}) \quad (1.1)$$

for any function $f$ on $\tilde{X}$. For any $z \in \mathbb{C}$, denote $R(z) = \sum_{n=1}^\infty z^n R_n$. It is clear that $\hat{\mathcal{P}} = R(1) = \sum_{n=1}^\infty R_n$.

For simplicity of notation, we regard the space $L^1(\tilde{X}, \hat{\nu})$ as a subspace of $L^1(X, \nu)$ consisting of functions supported on $\tilde{X}$, and we denote it by $L^1(\hat{\nu})$ or $L^1$ sometimes and when no ambiguity arises.

Suppose that there is a seminorm $| \cdot |_B$ for functions in $L^1(\tilde{X}, \hat{\nu})$. Consider the set $B = B(\tilde{X}) = \{f \in L^1(\tilde{X}, \hat{\nu}) : |f|_B < \infty\}$. Define a norm on $B$ by

$$\|f\|_B = |f|_B + \|f\|_1$$

for $f \in B$, where $\|f\|_1$ is the $L^1$ norm. We assume that $B$ satisfies the requirements stated below; the assumptions (a) to (c) will be necessary to establish the spectral gap of the induced transfer operator, while conditions (d) to (f) will be useful to prove aperiodicity. We first define a set $U \subset \tilde{X}$ to be almost open with respect to $\hat{\nu}$ if for $\hat{\nu}$ almost every point $x \in U$, there is a neighborhood $V(x)$ such that $\hat{\nu}(V(x) \setminus U) = 0$.

Assumption B. (a) (Compactness) $B$ is a Banach space and the inclusion $B \hookrightarrow L^1(\hat{\nu})$ is compact; that is, any bounded closed set in $B$ is compact in $L^1(\hat{\nu})$.

(b) (Boundness) The inclusion $B \hookrightarrow L^\infty(\hat{\nu})$ is bounded; that is, $\exists C_b > 0$ such that $\|f\|_\infty \leq C_b \|f\|_B$ for any $f \in B$.

(c) (Algebra) $B$ is an algebra with the usual sum and product of functions, in particular there exists a constant $C_a$ such that $\|fg\|_B \leq C_a \|f\|_B \|g\|_B$ for any $f, g \in B$. 

6
(d) (Denseness) The image of the inclusion $\mathcal{B} \hookrightarrow L^1(\nu)$ is dense in $L^1(\nu)$.

(e) (Lower semicontinuity) For any sequence $\{f_n\} \subseteq \mathcal{B}$ with $\lim_{n \to \infty} f_n = f$ $\nu$-almost everywhere, $|f|_{\mathcal{B}} \leq \liminf_{n \to \infty} |f_n|_{\mathcal{B}}$.

(f) (Openness) For any nonnegative function $f \in \mathcal{B}$, the set $\{f > 0\}$ is almost open with respect to $\nu$.

Remark 1.1. Assumption B(f) means that functions in $\mathcal{B}$ are not far from continuous functions.

The possibility of computing a lower bound for the decay of correlations relies on the following important result first established by Sarig [Sr] and successively improved by Gouëzel [Go], here we take Gouëzel’s version. We will denote $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$.

**Theorem.** [Sr, Go] Let $T_n$ be bounded operators on a Banach space $\mathcal{B}$ such that $T(z) = I + \sum_{n \geq 1} z^n T_n$ converges in Hom($\mathcal{B}, \mathcal{B}$) for every $z \in \mathbb{D}$. Assume that:

(R1) (Renewal equation) for every $z \in \mathbb{D}$, $T(z) = (I - R(z))^{-1}$, where $R(z) = \sum_{n \geq 1} z^n R_n$, $R_n \in$ Hom($\mathcal{B}, \mathcal{B}$) and $\sum_{n \geq 1} \|R_n\| < +\infty$.

(R2) (Spectral gap) $1$ is a simple isolated eigenvalue of $R(1)$.

(R3) (Aperiodicity) for every $z \in \mathbb{D} \setminus \{1\}$, $I - R(z)$ is invertible.

Let $P$ be the eigenprojection of $R(1)$ at $1$. If $\sum_{k > n} \|R_k\| = O(1/n^\beta)$ for some $\beta > 1$ and $PR'(1)P \neq 0$, then for all $n$,

$$T_n = \frac{1}{\lambda} P + \frac{1}{\lambda^2} \sum_{k=n+1}^{\infty} P_k + E_n,$$

where $\lambda$ is given by $PR'(1)P = \lambda P$, $P_n = \sum_{k \geq n} PR_k P$ and $E_n \in$ Hom($\mathcal{B}, \mathcal{B}$) satisfies $\|E_n\| = O(1/n^\beta)$ if $\beta > 2$, $O(\log n/n^2)$ if $\beta = 2$, and $O(1/n^{2\beta-2})$ if $2 > \beta > 1$.

In the dynamical setting we are interested in, the operators $T_n$ and $R_n$ are defined by (1.1). In order to check points (R1) to (R3) in the statement above, one needs equivalent dynamical properties, which are summarized in the following spectral-like assumption.

**Assumption S.** Let $X \subset \mathbb{R}^m$ be compact subset with $\nu(X) = 1$ and $\widehat{X} \subset X$ be a compact subset of $X$.

Let $T : X \to X$ be a map whose first return map with respect to $\hat{X}$ is $\hat{T} = T^r$, and $\mathcal{B}$ be a Banach space satisfying Assumption B(a) to (c). We assume the following.

\[ \text{Remark 1.1.} \] Assumption B(f) means that functions in $\mathcal{B}$ are not far from continuous functions.
(S1) (Lasota-Yorke inequality) There exist constants $\eta \in (0,1)$ and $D > 0$ such that for any $f \in \mathcal{B}$,
\[
|\mathcal{P}f|_{\mathcal{B}} \leq \eta |f|_{\mathcal{B}} + D\|f\|_{1}; \quad (1.3)
\]

(S2) (Spectral radius) There exist constants $B, \hat{D} > 0$ and $\tilde{\eta} \in (0,1)$ such that for any $f \in \mathcal{B}, z \in \mathbb{T},$
\[
\|R(z)^{n}f\|_{\mathcal{B}} \leq |z^{n}|(B\tilde{\eta}^{n}\|f\|_{\mathcal{B}} + \hat{D}\|f\|_{1}); \quad (1.4)
\]

(S3) (Ergodicity) The measure $\hat{\mu}$ given by $\hat{\mu}(f) = \hat{\nu}(hf)$ is ergodic, where $\hat{h}$ is a fixed point of $\mathcal{P}.$

(S4) (Aperiodicity) The function $e^{it}$ given by the return time is aperiodic, that is, the only solution for $e^{it} = f/f \circ \hat{T}$ which holds almost everywhere with a measurable function $f : \hat{X} \to \mathbb{S},$ is provided by $f$ constant almost everywhere and $t = 0.$

(S5) (Return times tail) The $\mathcal{B}$-norm of the operator $R_{n}$ is summable and satisfies
\[
\sum_{k=n+1}^{\infty} \|R_{k}\|_{\mathcal{B}} = O(n^{-\beta}) \text{ for some } \beta > 1.
\]

Before commenting on the correspondence between assumptions (R1) - (R3) and (S1) - (S5), we observe that Assumption S allows us to get the following important theorem on decay of correlations whose proof is basically an adaptation of the analogous result in the seminal paper by Sarig [Sr], see also [Go]:

**Theorem A.** Let us suppose that Assumption (S) is satisfied; then there exists a constant $C > 0$ such that for any function $f \in \mathcal{B},$ $g \in L^{\infty}(X,\nu)$ with supp $f,$ supp $g \subset \hat{X},$
\[
\left|\text{Cov}(f, g \circ T^{n}) - \sum_{k=n+1}^{\infty} \mu(\tau > k) \right| \int f d\mu \int g d\mu \leq CF_{\beta}(n)\|g\|_{\infty}\|f\|_{\mathcal{B}}, \quad (1.5)
\]
where $F_{\beta}(n) = 1/n^{\beta}$ if $\beta > 2,$ $(\log n)/n^{2}$ if $\beta = 2,$ and $1/n^{2\beta-2}$ if $2 > \beta > 1.$

**Comments.**
1. We begin by observing that whenever for the system $(\hat{T}, \hat{\nu})$ the Lasota-Yorke’s inequality (1.3) is satisfied for any function $f \in \mathcal{B},$ and if the Banach space $\mathcal{B}$ satisfies Assumption B(a), then $\hat{\mu}$ has a fixed point $\hat{h} \in \mathcal{B}$ with $\hat{h} \geq 0$ and $\hat{T}\hat{h} = \hat{h},$ and the measure $\hat{\mu}$ defined by $\hat{\mu}(f) = \hat{\nu}(f\hat{h})$ is $\hat{T}$ invariant. Such a measure $\hat{\mu}$ can be extended to an absolutely continuous invariant measure $\mu$ on $X$ in the usual way (see e.g. [Kk]). It is well known that if $\hat{\mu}$ is ergodic, so is $\mu.$ Ergodicity of $\hat{\mu},$ as required by (S3) immediately implies (R2).

2. The aperiodicity condition (R3) follows if we show that 1 is not an eigenvalue of $R(z)$ for $|z| = 1$ with $z \neq 1.$ Let us fix $0 < t < 2\pi$ and put $z = e^{it};$
if we suppose that \( R(z)f = f \) for some nonzero \( f \in \mathcal{B} \), by the arguments developed in the proof of the Lemma 6.6 in [Go], that is equivalent to the equation \( e^{-ist} f \circ \hat{T} = f \) almost everywhere. By the aperiodicity condition (S4) we conclude that \( t = 0 \) and \( f \) is a constant \( \hat{\mu} \)-almost everywhere which is a contradiction.

3. Practically, (1.4) usually can be obtained in a similar way as (1.3), for example, see the proof of Theorem D. On the other hand, since \( \hat{T} = R(1) \), (1.4) implies the Lasota-Yorke inequality for \( \hat{T}^n \) for some \( n > 0 \) with \( B\hat{\eta}^n < 1 \).

4. As we said in the Introduction, Assumption (S4) is actually equivalent to the fact that \( \mu \) is weak mixing for \( T \) (see e.g [PP]). Since decay of correlations implies mixing, we obtain that with the Lasota-Yorke inequality, weak mixing implies mixing.

Assumption (S4) is usually difficult to check. However, for piecewise expanding systems, the condition can be verified and we will give some sufficient conditions in Theorem B below.

The more general version of aperiodicity is the following. Let \( G \) be a locally compact Abelian polish group. A measurable function \( \phi : \hat{X} \rightarrow G \) is aperiodic if the only solutions for \( \gamma \circ \phi = \lambda f / f \circ T \) almost everywhere with \( \gamma \in \hat{G}, |\lambda| = 1 \) and a measurable function \( f : \hat{X} \rightarrow G \) are \( \gamma = 1, \lambda = 1 \) and \( f \) constant almost everywhere, see [ADSZ] and references therein. Here we only consider the case \( \gamma = \text{id}, \phi = e^{it}T \), and \( G \) being the smallest compact subgroup of \( S \) containing \( e^{it} \).

We denote by \( B_\varepsilon(\Gamma) \) the \( \varepsilon \) neighborhood of a set \( \Gamma \subset X \). Recall that the notion almost open is given before the statement of Assumption B. We now state a few conditions which must be satisfied by all the maps considered from now on.

**Assumption T.**

(a) (Piecewise smoothness) There are countably many disjoint sets \( U_1, U_2, \cdots \) almost open with respect to \( \nu \), with \( \hat{X} = \bigcup_{i=1}^{\infty} \bar{U}_i \), a compact set, such that for each \( i, \hat{T}_i := \hat{T}|_{\bar{U}_i} \) extends to a \( C^{1+\alpha} \) diffeomorphism from \( \bar{U}_i \) to its image, and \( \tau[\nu_i] \) is constant; we will use the symbol \( \hat{T}_i \) to denote the extension as well.

(b) (Finite images) \( \{\hat{T}U_i : i = 1, 2, \cdots\} \) is finite, and \( \nu(B_\varepsilon(\partial\hat{T}U_i)) = O(\varepsilon) \quad \forall i = 1, 2, \cdots \).

(c) (Expansion) There exists \( s \in (0, 1) \) such that \( d(\hat{T}x, \hat{T}y) \geq s^{-1}d(x, y) \quad \forall x, y \in \bar{U}_i \quad \forall i \geq 1 \).

(d) (Topological mixing) \( T : X \rightarrow X \) is topological mixing.
Remark 1.2. Conditions (b) and (c) in Assumption T correspond to conditions (F) and (U) in [ADSZ]. There is there a third assumption, (A), which is distortion and which is not necessarily guaranteed in our systems. With this precision, we could regard the systems satisfying Assumption T(a)-(c) as higher dimensional “AFU” systems. Returning to the finite image condition T (b), it is used in proof of Lemma 2.1 below, to get μ(A_{n,n_0}) \to 1 as n_0 \to \infty and this is a consequence of a “small image boundary” as explained in the first footnote of the proof of Theorem B. This could probably be replaced by some weaker conditions of the same type, still allowing to prove Lemma 2.1. However, that property is one of the conditions for “AFU” systems and it is easy to check: so we simply took it.

Remark 1.3. We mention that if T has relatively prime return times on almost all points \( x \in \hat{X} \), then Condition (d) is satisfied. The former means that for any neighborhood \( U \) of \( x \), there is a point \( y \in U \) and return times \( \tau'(x) \) and \( \tau'(y) \) such that \( T^{\tau'(x)}(x), T^{\tau'(y)}(x) \in U \) and the greatest common divisor \( (\tau'(x), \tau'(y)) = 1 \). Here \( \tau'(x) \) and \( \tau'(y) \) are not necessarily the first return time.

Take a partition \( \xi \) of \( \hat{X} \). Consider a family of skew-products of the form

\[
\hat{T} = \hat{T}_S : \hat{X} \times Y \to \hat{X} \times Y, \quad \hat{T}_S(x, y) = (\hat{T}x, S(\xi(x))(y)),
\]

(1.6)

where \( (Y, \mathcal{F}, \rho) \) is a Lebesgue probability space, \( \text{Aut}(Y) \) is the collection of its automorphisms, that is, invertible measure-preserving transformations, and \( S : \xi \to \text{Aut}(Y) \) is arbitrary.

Consider functions \( \hat{f} \in L^1(\hat{\nu} \times \rho) \) and define

\[
|\hat{f}|_{\hat{B}} = \int_Y |\hat{f}(\cdot, y)| d\rho(y), \quad \|\hat{f}\|_{\hat{B}} = |\hat{f}|_{\hat{B}} + \|\hat{f}\|_{L^1(\hat{\nu} \times \rho)}.
\]

Then we let

\[
\hat{B} = \{ \hat{f} \in L^1(\hat{\nu} \times \rho) : |\hat{f}|_{\hat{B}} < \infty \}.
\]

It is easy to see that with the norm \( \| \cdot \|_{\hat{B}}, \hat{B} \) is a Banach space.

The transfer operator \( \hat{\mathcal{P}} = \hat{\mathcal{P}}_{\hat{\nu} \times \rho} \) acting on \( L^1(\hat{\nu} \times \rho) \) is defined as the dual of the operator \( \hat{f} \to \hat{f} \circ \hat{T} \) from \( L^\infty(\hat{\nu} \times \rho) \) to itself. Note that if \( Y \) is a space consisting of a single point, then we can identify \( \hat{X} \times Y, \hat{T} \) and \( \hat{\mathcal{P}} \) with \( \hat{X}, \hat{T} \) and \( \mathcal{P} \) respectively.

**Theorem B.** Suppose \( \hat{T} \) satisfies Assumption T(a) to (d) and \( \mathcal{B} \) satisfies Assumption B(d) to (f), and \( \hat{\mathcal{P}} \) satisfies the Lasota-Yorke inequality

\[
|\hat{\mathcal{P}}\hat{f}|_{\hat{B}} \leq \hat{\eta} \|\hat{f}\|_{\hat{B}} + \hat{D} \|\hat{f}\|_{L^1(\hat{\nu} \times \rho)}
\]

(1.7)

for some \( \hat{\eta} \in (0,1) \) and \( \hat{D} > 0 \). Then the absolutely continuous invariant measure \( \hat{\mu} \) obtained from the Lasota-Yorke inequality (1.3) is ergodic and \( e^{i\pi r} \) is aperiodic. Therefore Assumptions (S3) and (S4) follow.
Remark 1.4. The theorem is for ergodicity and aperiodicity of \( \mu \). As we mentioned in Comment (4) above, aperiodicity of \( \mu \) is equivalent to weak mixing for \( \mu \) with respect to \( T \). So practically, if we know that \( \mu \) is mixing or weak mixing for \( T \), then we do not need to use the theorem.

Remark 1.5. Same as for (1.4), the inequality (1.7) may be obtained in a similar way as (1.3). This is because any \( S(\xi(x)) \) is a measure preserving transformation, and therefore \( \mathcal{F} \) and \( \mathcal{F} \) have the same potential function, see the proof of Theorem D.

Remark 1.6. It is well known that for \( C^{1+\alpha}, \alpha > 1 \), uniformly expanding maps or uniformly hyperbolic diffeomorphisms, the absolutely continuous invariant measures or the SRB measure \( \mu \) are ergodic if the maps are topological mixing, see e.g. [Bo] for invertible case; the noninvertible case can be obtained similarly.

However, it is not the case if the conditions on \( C^{1+\alpha} \) or uniformity of hyperbolicity fails. In [Qu] the author gives an example of \( C^1 \) uniformly expanding maps of the unit circle, and in [HPT] the authors provide an example of \( C^\infty \) diffeomorphisms, where the Lebesgue measure is preserved and topological mixing does not give ergodicity. In the proof of the theorem we in fact give some additional conditions under which topological mixing implies ergodicity (see Lemma 2.2).

2 Aperiodicity

The proof of Theorem B is based on a result in [ADSZ]. We briefly mention the terminology used there.

A fibered system is a quintuple \((X, A, \nu, T; \xi)\), where \((X, A, \nu, T)\) is a non-singular transformation on a \( \sigma \)-finite measure space and \( \xi \subset A \) is a finite or countable partition (mod \( \nu \)) such that:

1. \( \xi_\infty = \bigvee_{i=0}^{\infty} T^{-i} \xi \) generates \( A \);
2. every \( A \in \xi \) has positive measure;
3. for every \( A \in \xi, T|_A : A \to TA \) is bimeasurable invertible with nonsingular inverse.

The transformation given in (1.6) is called the skew products over \( \xi \). We will denote with \( \xi_n \) the \( n \)-join \( \xi_n = \bigvee_{i=0}^{n-1} T^{-i} \xi \), and with \( \xi_n(x) \) the element (cylinder) of the partition \( \xi_n \) containing the point \( x \). Consider the corresponding transfer operator of \( \mathcal{F} = \mathcal{F}_{\nu \times \nu} \). A fibred system \((X, A, \nu, T, \xi)\) with \( \nu \) finite is called skew-product rigid if for every invariant function \( h(x, y) \) of \( \mathcal{F} \) of an arbitrary skew product \( T_S \), the set \( \{h(\cdot, y) > 0\} \) is almost open (mod \( \nu \)) for almost every \( y \in Y \). In [ADSZ], a set \( U \) being almost open (mod \( \nu \)) means that for \( \nu \) almost every \( x \in U \), there is a positive integer \( n \) such that \( \nu(\xi_n(x) \setminus U) = 0 \). Since
the partition $\xi$ we are interested in satisfies $\nu(\partial A) = 0$ for any $A \in \xi_n$ and $\hat{T}$ is piecewise smooth, the fact that $\xi_\infty$ generates $\mathcal{A}$ implies that the definition given there is the same as we defined for Assumption B(f).

A set that can be expressed in the form $\hat{T}^n\xi_n(x)$, $n \geq 1$ and $x \in \hat{X}$, is called an image set. A cylinder $C$ of length $n_0$ is called a cylinder of full returns, if for almost all $x \in C$ there exist $n_k \to \infty$ such that $\hat{T}^{n_k}\xi_{n_k+n_0}(x) = C$. In this case we say that $\hat{T}^{n_k}(C)$ is a recurrent image set.

Our proof of Theorem B is based on a result given in Theorem 2 in [ADSZ]:

**Theorem.** Let $(X, \mathcal{A}, \mu, T, \xi)$ be a skew-product rigid measure preserving fibered system whose image sets are almost open. Let $G$ be a locally compact Abelian polish group. If $\gamma \circ \phi = \lambda f/f \circ T$ holds almost everywhere, where $\phi : X \to \hat{G}$, $\xi$ measurable, $\gamma \in \hat{G}$, $\lambda \in \mathbb{S}$, then $f$ is constant on every recurrent image set.

**Warning.** In the proof of Theorem B and the lemmas below we will work exclusively on the induced space $\hat{X}$ and with measures $\hat{\nu}$ and $\hat{\mu}$ and density $h$. So we will drop the hat on those notations.

**Proof of Theorem B.** Recall that $\mu$ is an $\hat{T}$ invariant measure with density $h$, where $h$ is the fixed point of $\hat{\mathcal{P}}$ in $\mathcal{B}$. By Lemma 2.2 we know that $\mu$ is ergodic. So we only need to prove that $e^{it\sigma}$ is aperiodic.

Denote by $\mathcal{A}$ the Borel $\sigma$-algebra inherited from $\mathbb{R}^m$. Take a countable partition $\xi$ of $\hat{X}$ into $\{U_i\}$ or finer. We also require that each $A \in \xi$ is almost open, and $\nu(B, \partial T^\xi)) = O(\varepsilon)$, where $\partial T^\xi = \cup_{A \in \xi} \partial (TA)$. The latter is possible because we can take smooth surfaces as the boundary of the elements of $\xi$, in addition to Assumption T(b) \footnote{This assumption is in fact used to get the measure of an $\epsilon$-neighborhood of the boundary of $T\xi$ of order $\epsilon$.}. Since $\hat{T}$ is uniformly expanding by Assumption T(c), we know that each element of $\xi_\infty = \bigvee_{i=0}^{\infty} \hat{T}^{-i}\xi$ contains at most one point.\footnote{In fact, if $x, y \in \xi_n$, then for any $i > 0$, $\hat{T}^ix$ and $\hat{T}^iy$ are always in the same elements of $\xi$, and hence in the same $U_n$, for some $n_i > 0$. On the other hand, by Assumption T(c) we have $d(\hat{T}^ix, \hat{T}^iy) \geq s^{-i}d(x, y)$. If $d(x, y) \neq 0$, then $d(\hat{T}^ix, \hat{T}^iy) \to \infty$, contradicting the facts that $\hat{X}$ is compact. We in fact recall that in Lebesgue spaces a necessary and sufficient condition for $\xi_n \to \mathcal{A}$ is that there exists a set of zero measure $N \subset \hat{X}$ such that for $x, y \in \hat{X}/N$ (with $x \neq y$) there exists $n \geq 1$ and $U \in \xi_n$ such that $x \in U$ but $y \notin U$.}

So $\xi_\infty$ generates $\mathcal{A}$. We may regard that each $A \in \xi$ has positive measure, otherwise we can use $\hat{X} \setminus A$ to replace $\hat{X}$. Also, for every $A \in \xi$, $\hat{T}|_A : A \to \hat{TA}$ is a diffeomorphism, and therefore $\hat{T}|_A$ is bimeasurable invertible with nonsingular inverse. So the quintuple $(\hat{X}, \mathcal{A}, \mu, \hat{T}, \xi)$ is a measure preserving fibered system.

The construction of $\xi$ implies that $\mu(\partial \xi) = \nu(\partial \xi) = 0$. Hence, $\mu(\partial \xi_n) = \nu(\partial \xi_n) = 0$ for any $n \geq 1$. Note that the intersection of finite number of almost open sets is still almost open. Differentiability of $\hat{T}$ on each $U_i$ implies that all elements $\xi_n(x)$ of $\xi_n$ are almost open, and therefore all image sets $\hat{T}^n\xi_n(x)$ are almost open with respect to $\mu$.

To get skew product rigidity, let us consider the skew product $\hat{T}_S$ defined in (1.6) for any $(Y, \mathcal{F}, \rho)$. Let $\mathcal{P} = \mathcal{P}_{\nu \times \rho}$ be the transfer operator and $\hat{h} an
invariant function, that is, \( \widehat{\mathcal{D}} h = h \). By Proposition 2.3 below we know that \( h \in \widehat{B} \). Hence, for \( \rho \)-almost every \( y \in Y \), \( \widehat{h}(\cdot, y) \in B \). By Assumption B(f), \( \{ \widehat{h}(\cdot, y) > 0 \} \) is almost open mod \( \nu \). This gives the skew product rigidity.

So far we have verified all conditions in the theorem of [ADSZ] stated above. Applying the theorem to the equation \( e^{itT} = f / f \circ \widehat{T} \) almost everywhere, where \( f : \widehat{X} \to \mathcal{S} \) is a measurable function, we get that \( f \) is constant on every recurrent image sets \( J \).

Now we prove aperiodicity, by following similar arguments in [Go]. Assume the equation \( e^{itT} = f / f \circ \widehat{T} \) holds almost everywhere for some real number \( t \) and a measurable function \( f : \widehat{X} \to \mathcal{S} \). By Lemma 2.1 below we get that \( \widehat{X} \) contains a recurrent image set \( J \) with \( \mu(J) > 0 \). By the theorem above, we know that \( f \) is constant, say \( c \), almost everywhere on \( J \). By the absolute continuity of \( \mu \) and the fact that \( \{ h > 0 \} \) is \( \nu \)-almost open, we can find an open set \( J' \subset J \) of positive \( \mu \)-measure. By Assumption T(d), \( T \) is topological mixing. Therefore for all sufficiently large \( n \), we have \( T^{-n}J' \cap J' \neq \emptyset \). Since the intersection is open\(^5\), we get that \( \mu(T^{-n}J' \cap J') > 0 \). So for any typical point \( x \) in \( T^{-n}J' \cap J' \), there is \( k > 0 \) such that \( T^nx = \widehat{T}^kx \), and \( n = \sum_{i=0}^{k-1} T^i(x) \). Since \( e^{itT} = f / f \circ \widehat{T} \) along the orbit of \( x \), we have

\[
e^{int} = e^{it \sum_{i=0}^{k-1} \tau^i(x)} = f(x) \frac{f(\widehat{T}x)}{f(\widehat{T}x)} \cdots \frac{f(\widehat{T}^{k-1}x)}{f(\widehat{T}^kx)} = \frac{f(x)}{f(\widehat{T}^kx)} = \frac{c}{c} = 1.
\]

Since this is true for all large \( n \), by replacing \( n \) by \( n + 1 \) we get that \( e^{it} = 1 \). It follows that \( t = 0 \) and \( f = f \circ \widehat{T} \) almost everywhere which implies that \( f \) must be a constant almost everywhere since \( \mu \) is ergodic.

To prove Lemma 2.1, we need a result from Lemma 2 in Section 4 in [ADSZ]. We state it as the next lemma. The setting for the lemma is a conservative fibered system. So it can be applied directly to our case.

**Lemma.** A cylinder \( C \subset \eta_{n_0} \) is a cylinder of full return if and only if there exists a set \( K \) of positive measure such that for almost every \( x \in K \), there are \( n_i \to \infty \) with \( \widehat{T}^{n_i} \eta_{n_i+n_0} = C \).

**Lemma 2.1.** There is a recurrent image set \( J \) contained in \( \widehat{X} \) with \( \mu(J) > 0 \).

**Proof.** Recall that \( s \) is given in Assumption T(e). Take \( C_\xi > 0 \) such that \( \text{diam} D \leq C_\xi \) for all \( D \in \xi \). Set

\[
A'_{k,n_0} = \{ x \in \widehat{X} : x \not\in B_{C_{\xi} e^{k+n_0}(\partial \widehat{T} \xi)} \},
\]

\[
A_{n,n_0} = \bigcap_{k=0}^{n-1} \widehat{T}^{-k} A'_{k,n_0}.
\]
By the construction of $\xi$, there is $C' > 0$ such that
\[ \nu(A'_{k,n_0}) \geq 1 - C'C_s^{k+n_0}. \]
By Assumption B(b), $\|h\|_{\infty} < \infty$. So if we take $C = C'C_s/\|h\|_{\infty}/(1 - s)$, then
\[ \mu(A'_{k,n_0}) \geq 1 - C'C_s/\|h\|_{\infty}^{k+n_0} = 1 - C(1 - s)^{k+n_0}. \]
Since $\mu$ is an invariant measure, $\mu(A_{n_0},n_0) \geq 1 - C(1 - s)^{n_0}$. Hence, $\mu(A_{n_0},n_0)$ is less than
\[ C \left( 1 - C(1 - s)^{n_0} \right) \]
for any given $n_0$. Then $\mu(A_{n_0},n_0)$ has the measure bounded below by a positive number for all $n_0 > 0$, and the bound can be chosen arbitrarily close to 1 by taking $n_0$ sufficiently large.

Note that $\xi_n$ is a partition with at most countably many elements. For each $n_0 > 0$, let $B'_{n_0}$ be the union of finite elements of $\xi_{n_0}$ such that $\mu(B'_{n_0}) > 1 - Cs^{n_0}/2$. Then set $B_{n_0} = B'_{n_0} \cap T^{-n}B'_{n_0}$. Clearly, $\mu(B_{n_0}) \geq 1 - Cs^{n_0}$. Denote $C_{n_0} = A_{n_0} \cap B_{n_0}$. We have $\mu(C_{n_0}) \geq 1 - 2Cs^{n_0}$. Hence, $\sup_{n_0} \mu(C_{n_0}) = \infty$ for all large $n_0$. A generalized Borel-Cantelli Lemma by Kochen and Stone (see [Yaj]) gives that for any given $n_0 > 0$, the set of points that belong to infinitely many $C_{n_0} \cap n$ has the measure bounded below by
\[ \limsup_{n_0} \frac{\sum 1_{i \leq i \leq n} \mu(C_{i,n_0}) \mu(C_{k,n_0})}{\sum 1_{i \leq k \leq n} \mu(C_{i,n_0} \cap C_{k,n_0})}. \]
Note that if $n_0 \to \infty$, then $\mu(C_{i,n_0}) \to 1$ as $n_0 \to \infty$ and uniformly in $i$ by the previous lower bound on $\mu(C_{n_0})$. Hence the upper limit goes to 1 as $n_0 \to \infty$.

Denote
\[ \Gamma_{n_0} = \{ x \in \hat{X} : x \in C_{n_0} \text{ infinitely often} \}. \]
The above argument gives $\mu(\Gamma_{n_0}) \to 1$ as $n_0 \to \infty$.

Note that for a one-to-one map $T, T(A \cap T^{-1}B) = B$ if and only if $B \subset TA$. Since $\xi_n(x) = \xi(x) \cap T^{-1}(\xi_{n-1}(Tx))$, and $T$ is a local diffeomorphism, we know that $\hat{T} \xi_n(x) = \xi_{n-1}(\hat{T}x)$ if and only if $\xi_{n-1}(\hat{T}x) \subset \hat{T} \xi(x)$. Inductively, $\hat{T}^n \xi_{n+1}(x) = \xi_{n}(\hat{T}^n x)$ if and only if $\xi_{n+i+n}(T^i x) \subset \hat{T}^i \xi(\hat{T}^{i-1}x)$ for $i = 1, \cdots, n$. If $x \in A_{n_0,n_0}$ for some $n_0 > 0$, then $\hat{T}^{i-1} x \not\in B_{C_s^{i+n_0}}(\hat{B} \xi \hat{T}^i x)$ for all $i = 1, \cdots, n$. Since the diameter of each member of $\xi$ is less than $C_s$, by Assumption T(c), $\text{diam} \xi_n(x) \leq C_s^{n_0}$ for any $x \in \hat{X}$ and $n_0 \geq 0$. We get $\xi_{n+i+n}(T^i x) \subset \hat{T}^i \xi(\hat{T}^{i-1}x)$ and therefore $\hat{T}^n \xi_{n+i+n}(x) = \xi_{n}(\hat{T}^n x)$.

Consequently, if $x \in \Gamma_{n_0},$ then $x \in A_{n_0,n_0} = A_{n_0,n_0} \cap B_{n_0,n_0}$ for infinitely many $n_0$. Hence, $\hat{T}^n \xi_{n_0+i+n}(x) = \xi_{n_0}(\hat{T}^n x)$ and $\hat{T}^n x \in B_{n_0}$ for infinitely many $n_0$.

Take $n_0 > 0$ such that $\mu(\Gamma_{n_0}) > 0$. Since $B_{n_0}$ consists of only finitely many elements in $\xi_{n_0},$ we know that there is an element $C \in \xi_{n_0}$ with $C \subset B_{n_0}$ such that
\[ \mu \{ x: \hat{T}^n \xi_{n_0+i+n}(x) = \xi_{n_0}(\hat{T}^n x) = C \text{ infinitely often} \} > 0. \]
By the above lemma from [ADSZ], $C$ is a cylinder of full returns. Hence, $\hat{T}^n C$ is a recurrent image set. Since $\mu$ is an invariant measure, $\mu(\hat{T}^n C) > 0$ and therefore $\mu(J) > 0$.

\textbf{Lemma 2.2.} Suppose $T$ and $B$ satisfies Assumption T(d) and B(f) respectively. Then there is only one absolutely continuous invariant measure $\mu$ which is ergodic.
Suppose By Assumption B(d), suppose that of Krengel ([Kr]), there exists a nonnegative function \( f \) that lim inf satisfies Lasota-Yorke inequality (1.7). Then any \( L \) more general cases. To simplify the notation we will write from now on \( h \) ections are This means \( e \) where \( e \) and \( e \) and \( e \) respectively. Hence, \( e \) \( b \) \( T \) is topological mixing, there is \( e \) \( b \) \( T \) \( e \) \( b \), hence \( e \) \( b \) \( T \) \( e \). Since \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \). Hence \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \). Therefore, \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \) \( b \) \( T \) \( e \) \( b \). which is a contradiction.

We are left with the proof that any fixed point \( \tilde{h} \) of \( \tilde{\mathcal{P}} \) belongs to \( \mathcal{B} \). The result was proved for Gibbs-Markov maps in [AD]. We show that it holds in more general cases. To simplify the notation we will write from now on \( L^1(\nu \times \rho) \) instead of \( L^1(\tilde{X} \times Y; \nu \times \rho) \).

**Proposition 2.3.** Suppose that \( \mathcal{B} \) satisfies Assumption B(d) and (e), and \( \tilde{\mathcal{P}} \) satisfies Lasota-Yorke inequality (1.7). Then any \( L^1(\nu \times \rho) \) function \( h \) on \( \tilde{X} \times Y \) that satisfies \( \tilde{\mathcal{P}}_{\nu \times \rho} h = \tilde{h} \) belongs to \( \mathcal{B} \).

**Proof.** By Assumption B(d), \( \mathcal{B} \) is dense in \( L^1(\tilde{X}, \nu) \). It is easy to see that \( \tilde{\mathcal{B}} \) is dense in \( L^1(\nu \times \rho) \). Hence, for any \( \varepsilon > 0 \) we can find a nonnegative function \( f_\varepsilon \in \tilde{\mathcal{B}} \) such that \( \| f_\varepsilon - \tilde{h} \|_{L^1(\nu \times \rho)} < \varepsilon \). By the stochastic ergodic theorem of Krenkel ([Kr]), there exists a nonnegative function \( \tilde{h}_\varepsilon \in L^1(\nu \times \rho) \) and a subsequence \( \{n_k\} \) such that

\[
\lim_{k \to \infty} \frac{1}{n_k} \sum_{\ell=0}^{n_k-1} \tilde{\mathcal{P}}^\ell f_\varepsilon = \tilde{h}_\varepsilon \quad \nu \times \rho\text{-a.e.} \tag{2.2}
\]

and \( \tilde{\mathcal{P}} \tilde{h}_\varepsilon = \tilde{h}_\varepsilon \).

Note that Lasota-Yorke inequality (1.7) implies that for any \( \tilde{f} \in \tilde{\mathcal{B}} \), \( \ell \geq 1 \),

\[
|\tilde{\mathcal{P}}^\ell \tilde{f}|_\tilde{\mathcal{B}} \leq \tilde{\eta}^\ell |\tilde{f}|_\tilde{\mathcal{B}} + \tilde{D}^\ell \| \tilde{f} \|_{L^1(\nu \times \rho)} \leq \tilde{D}_2 \| \tilde{f} \|_\tilde{\mathcal{B}}, \tag{2.3}
\]

where \( \tilde{D}^* = \tilde{D} \tilde{\eta} / (1 - \tilde{\eta}) \geq \tilde{D}(\tilde{\eta} + \cdots + \tilde{\eta}^{\ell-1}) \) and \( \tilde{D}_2 = 1 + \tilde{D}^* \). Denote \( \psi_k = \frac{1}{n_k} \sum_{\ell=0}^{n_k-1} \tilde{\mathcal{P}}^\ell f_\varepsilon \). By (2.3) \( \psi_k \leq \tilde{D}_2 \| \tilde{f} \|_\tilde{\mathcal{B}} \). On the other hand (2.2) implies that \( \liminf_{k \to \infty} \psi_k(x, y) = \tilde{h}_\varepsilon(x, y) \) for \( \nu \)-a.e. \( x \in \tilde{X} \), \( \rho \)-a.e. \( y \in Y \). Hence, by Assumption B(e) and Fatou’s lemma we obtain

\[
|\tilde{h}_\varepsilon|_\tilde{\mathcal{B}} = \int_Y \lim_{k \to \infty} |\psi_k(\cdot, y)|_\mathcal{B} d\rho(y) \leq \int_Y \liminf_{k \to \infty} |\psi_k(\cdot, y)|_\mathcal{B} d\rho(y) \leq \int_Y \liminf_{k \to \infty} |\psi_k(\cdot, y)|_\mathcal{B} d\rho(y) \leq \tilde{D}_2 \| \tilde{f}_\varepsilon \|_\tilde{\mathcal{B}}. \tag{2.4}
\]

This means \( \tilde{h}_\varepsilon \in \tilde{\mathcal{B}} \).
By Fatou’s Lemma and the fact that $\mathcal{P}$ is a contraction on $L^1(\nu \times \rho)$, it follows immediately that (2.2) and $\mathcal{P}\hat{h} = \hat{h}$ imply

$$
\|\hat{h} - \hat{h}_\varepsilon\|_{L^1(\nu \times \rho)} \leq \liminf_{k \to \infty} \frac{1}{n_k} \sum_{l=0}^{n_k-1} \|\mathcal{P}^l(\hat{h} - \hat{f}_\varepsilon)\|_{L^1(\nu \times \rho)} \leq \|\hat{h} - \hat{f}_\varepsilon\|_{L^1(\nu \times \rho)} \leq \varepsilon.
$$

By the first inequality of (2.3) we know that for any $n \geq 1$,

$$
\|\hat{h}_\varepsilon\|_{\bar{B}} = \|\mathcal{P}^n\hat{h}_\varepsilon\|_{\bar{B}} \leq \tilde{D}^n\|\hat{h}_\varepsilon\|_{\bar{B}} + \tilde{D}^*\|\hat{h}_\varepsilon\|_{L^1(\nu \times \rho)}.
$$

Sending $n$ to infinity we get $\|\hat{h}_\varepsilon\|_{\bar{B}} \leq \tilde{D}^*\|\hat{h}_\varepsilon\|_{L^1(\nu \times \rho)} \leq \tilde{D}^*\|\hat{h}\|_{L^1(\nu \times \rho)} + \varepsilon$.

Replace now $\varepsilon$ with a decreasing sequence $c_n \to 0$ as $n \to \infty$. Since $\hat{h}_{c_n}$ converges in $L^1(\nu \times \rho)$ to $\hat{h}$, there is a subsequence $c_i$ such that $\lim_{i \to \infty} \hat{h}_{c_i} = \hat{h}$, $\nu \times \rho$-a.e.. Then by the same arguments as for (2.4), we get

$$
|\hat{h} - \hat{h}_{c_i}|_{\bar{B}} \leq \liminf_{i \to \infty} |\hat{h}_{c_i} - \hat{h}_{c_i}|_{\bar{B}} \leq 2 \sup_{0 \leq \xi \leq 1} \|\hat{h}_\varepsilon\|_{\bar{B}} \leq 2\tilde{D}(\|\hat{h}\|_{L^1(\nu \times \rho)} + 1).
$$

So we obtain $\hat{h} - \hat{h}_{c_i} \in \bar{B}$.

Therefore $\hat{h} = (\hat{h} - \hat{h}_{c_i}) + \hat{h}_{c_i} \in \bar{B}$ and this completes the proof.  \qed

**Part II: Applications to Non-Markov Maps**

We now apply our results to piecewise expanding non-Markov maps with an indifferent fixed point. We use different Banach spaces for maps in one and higher dimensional spaces.

**3 Systems on the interval**

The object of this section is twofold: to give an example of a Banach space which fits our assumptions, and to provide the lower bound for the decay of correlations. Moreover, we will use a large space of observables, bounded variation function instead of Hölder continuous functions.

Let $X = I = [0,1]$ and $\nu$ be the Lebesgue measure on $X$.

Recall that for a map $T : X \to X$ and a subset $\tilde{X} \subset X$, the corresponding first return map is denoted by $\tilde{T} : \tilde{X} \to \tilde{X}$; $\hat{\nu}$ will denote again the normalized Lebesgue measure over $\tilde{X}$.

Assume that $T : X \to X$ is a map satisfying the following assumptions.

**Assumption T'.** (a) (Piecewise smoothness) There are points $0 = a_0 < a_1 < \cdots < a_K = 1$ such that for each $j$, $T_j = T|_{I_j}$ is a $C^2$ diffeomorphism on its image, where $I_j = (a_{j-1}, a_j)$.

(b) (Fixed point) $T(0) = 0$. 

16
(c) (Expansion) There exists \( z \in I_1 \) such that \( T(z) \in I_1 \) and \( \Delta := \inf_{x \in \mathcal{X}} |T'(x)| > 2 \) for any \( x \in \mathcal{X} \), where \( \mathcal{X} = [z, 1] \).

(d) (Distortion) \( \Gamma := \sup_{x \in [z, 1]} |\mathcal{T}'(x)|/|\mathcal{T}'(x)| < \infty \).

(e) (Topological mixing) \( T : I \to I \) is topological mixing.

Denote \( J = [0, z) \) and \( \mathcal{X} = \mathcal{X}_J = X \setminus J \). \( I_0 = TJ \setminus J \subset I_1 \). We also denote the first return map \( \mathcal{T} = \mathcal{T}_J \) by \( \mathcal{T}_{ij} \) if \( \mathcal{T} = \mathcal{T}^{-1}_i \mathcal{T}_j \). Further, we denote \( I_{0i} = I_1 \setminus J \), \( I_{0j} = I_j \mathcal{T}^{-1}_j I_0 \) if \( j > 1 \), and \( I_{ij} = \mathcal{T}^{-1}_j I_0 \) for \( i > 0 \). Hence, \( \{I_{ij} : i = 0, 1, 2, \ldots \} \) form a partition of \( I_1 = (a_j, b_j) \) for \( j = 2, \ldots, K \). Also, we denote \( I_{ij} = [a_{ij}, b_{ij}] \) for any \( i = 0, 1, 2, \ldots \) and \( j = 1, \ldots, K \).

Recall that the variation of a real or complex valued function \( f \) on \([a, b]\) is defined by
\[
V^b_a(f) = V_{[a, b]}(f) = \sup_{\xi \in \Xi} \sum_{i=1}^n |f(x^{(i)}) - f(x^{(i-1)})|,
\]
where \( \xi \) is a finite partition of \([a, b]\) into \( a = x^{(0)} < x^{(1)} < \ldots < x^{(n)} = b \) and \( \Xi \) is the set of all such partitions. A function \( f \in L^1([a, b], \nu) \), where \( \nu \) denotes the Lebesgue measure, is of bounded variation if \( V_{[a, b]}(f) = \inf_g V_{[a, b]}(g) < \infty \), where the infimum is taken over all the functions \( g = f \nu\)-a.e.. Let \( \mathcal{B} \) be the set of functions \( f \in L^1(\mathcal{X}, \nu) \), \( f : \mathcal{X} \to \mathbb{R} \) with \( V(f) := V_{\mathcal{X}}(f) < \infty \). For \( f \in \mathcal{B} \), denote by \( \|f\|_{\mathcal{B}} = V(f) \), the total variation of \( f \). Then we define \( \|f\|_{\mathcal{B}} = \|f\|_1 + |f|_{\mathcal{B}} \), where the \( L^1 \) norm is intended with respect to \( \nu \). It is well known that \( \| \cdot \|_{\mathcal{B}} \) is a norm, and with the norm, \( \mathcal{B} \) becomes a Banach space.

To obtain the decay rates, we also assume that there are constants \( 0 < \gamma < 1, \gamma' > \gamma \) and \( \bar{C} > 0 \) such that in a neighborhood of the indifferent fixed point \( p = 0 \),
\[
T(x) = x + \bar{C}x^{1+\gamma} + O(x^{1+\gamma'}),
T'(x) = 1 + \bar{C}(1 + \gamma)x^{\gamma} + O(x^{\gamma'}),
T''(x) = \bar{C}\gamma(1 + \gamma)x^{\gamma-1} + O(x^{\gamma'-1}).
\]

For any sequences of numbers \( \{a_n\} \) and \( \{b_n\} \), we write \( a_n \approx b_n \) if \( \lim_{n \to \infty} a_n/b_n = 1 \), and \( a_n \sim b_n \) if \( c_1b_n \leq a_n \leq c_2b_n \) for some constants \( c_2 > c_1 > 0 \).

Denote
\[
d_{ij} = \sup\{\mathcal{T}'_{ij}(x)|x : x \in I_{ij}\}, \quad d_n = \max\{d_{n,j} : 2 \leq j \leq K\}.
\]

**Theorem C.** Let \( \mathcal{X} \), \( \mathcal{T} \) and \( \mathcal{B} \) are defined as above. Suppose \( T \) satisfies Assumption T’ (a) to (e). Then Assumption B(a) to (f) and assumptions (S1) to (S4) are satisfied and \( \|R_n\| = O(d_n) \). Hence, if \( d_n = O(n^{-\beta+1}) \) for some \( \beta > 1 \), then there exists \( \bar{C} > 0 \) such that for any functions \( f \in \mathcal{B}, g \in L^\infty(X, \nu) \) with \( \text{supp}\ f, \text{supp}\ g \subset \mathcal{X} \), (1.5) holds.
In particular, if $T$ satisfies (3.1) near 0, then $\sum_{k=n+1}^{\infty} \mu(\tau > k)$ has the order $n^{-\left(\frac{1}{\gamma} - 1\right)}$ and $d_n$ has the order $O(n^{-\left(\frac{1}{\gamma} + 1\right)})$, with $\beta = \frac{1}{\gamma}$. Since $\frac{1}{\gamma} - 1 < \frac{1}{\gamma}$ and $\frac{1}{\gamma} - 1 < 2(\frac{1}{\gamma} - 1)$ we have

$$\text{Cov}(f, g \circ T^n) \approx \sum_{k=n+1}^{\infty} \mu(\tau > k) \int f \mu \int g \mu \sim \frac{1}{n^{\left(\frac{1}{\gamma} - 1\right)}}.$$ 

It is well known that if the map $T$ allows a Markov partition, then the decay of correlations is of order $O(n^{-\left(\frac{1}{\gamma} - 1\right)})$ (see e.g. [Hu], [Sr],[LSV], [PY]). For non-Markov case, the upper bound estimate is given in [Yo2], in [MT], and [Sr] for observables with some Hölder property. With the methods in [Sr], the lower bound can be obtained by estimating the lower bound of the decay rate of the tower. Since our methods do not require Markov properties, the decay rates can be obtained directly from the size of the sets $\{\tau \geq k\}$; we also stress that our observables are functions with bounded variation.

**Proof of Theorem C.** By Lemma 3.1 below, $B$ satisfies Assumption B(a) to (f). By Lemma 3.2, we know that conditions $S(1)$ and $S(2)$ are satisfied. Notice that all requirements of Assumption T are satisfied, since part (a), (c) and (d) follow from Assumption $T'(a), (c)$ and (e) directly, and part (b) follows from the definition of $\tilde{T}$. Moreover Lemma 3.2 (iii) gives (1.7). Hence Theorem B can be applied and therefore conditions $S(3)$ and $S(4)$ are satisfied.

The estimate $\|R_n\| = O(d_n)$ follows from Lemma 3.3: we have thus proved the decay of correlations (1.5).

Suppose that $T$ also satisfies (3.1). Denote by $z_n \in I_1$ the point such that $T^n(z_n) = z$. It is well known that $z_n \approx (\gamma n)^{-1/\gamma}$ (see e.g. Lemma 3.1 in [HV]), and then we can obtain $(T^{-n}_1)'(x) = O(n^{-\frac{1}{\gamma} - 1})$. It follows that $d_n = O(n^{-\frac{1}{\gamma} - 1})$, since the density function $h$ is bounded on $\tilde{X}$, $\mu(\tau > k) \leq C_1 \nu(\tau > k) \leq C_2 z_k$ for some $C_1, C_2 > 0$. Hence $\sum_{k=n+1}^{\infty} \mu(\tau > k) = O(n^{-\frac{1}{\gamma} + 1})$.

**Lemma 3.1.** $B$ is a Banach space satisfying Assumption B(a) to (f) with $C_a = C_b = 1$.

**Proof.** These are standard facts, see for instance the proofs in Chapter 1 in [Br].

**Lemma 3.2.** There exist constants $\eta \in (0, 1)$ and $D, \tilde{D} > 0$ satisfying

(i) for any $f \in B$, $|\tilde{F}f|_{\tilde{B}} \leq \eta \|f\|_B + D \|f\|_{L^1(\rho)}$;

(ii) for any $f \in B$, $\|R(z)f\|_B \leq |z| \eta \|f\|_B + \tilde{D} \|f\|_{L^1(\rho)}$; and

(iii) for any $f \in \tilde{B}$, $|\tilde{F}\tilde{f}|_{\tilde{B}} \leq \eta \|\tilde{f}\|_{\tilde{B}} + D \|\tilde{f}\|_{L^1(\rho \times \rho)}$.

18
Proof. (i) Denote \( x_{ij} = \hat{T}_{ij}^{-1}(x) \), and \( \hat{g}(x_{ij}) = |\hat{T}_{ij}'(x_{ij})|^{-1} \). By the definition, we have
\[
\hat{\mathcal{F}}(x) = \sum_{j=1}^{\infty} \sum_{i=0}^{K} f(\hat{T}_{ij}^{-1}x)\hat{g}(\hat{T}_{ij}^{-1}x)1_{\hat{T}_{ij}}(x).
\]

Take a partition \( \xi \) of \( \hat{T}_{ij} \) into \( \hat{T}_{ij}a_{ij} = x^{(0)} < x^{(1)} < \cdots < x^{(k_{ij})} = \hat{T}_{ij}b_{ij} \), where we assume \( \hat{T}_{ij}a_{ij} < \hat{T}_{ij}b_{ij} \) without loss of generality. Whenever \( \hat{T}_{ij} \) intersects more than one intervals \( I_k = (a_k, b_k) \) in the case \( i = 0 \), then we put the endpoints \( a_k \) and \( b_k \) into the partition. Denote \( x^{(i)}_{ij} = \hat{T}_{ij}^{-1}x^{(i)} \). We have
\[
\sum_{\ell=1}^{k_{ij}} \left| f(x^{(i)}_{ij})\hat{g}(x^{(i)}_{ij}) - f(x^{(i-1)}_{ij})\hat{g}(x^{(i-1)}_{ij}) \right| \leq \sum_{\ell=1}^{k_{ij}} \hat{g}(x^{(i)}_{ij})|f(x^{(i)}_{ij}) - f(x^{(i-1)}_{ij})| + \sum_{\ell=1}^{k_{ij}} |f(x^{(i-1)}_{ij})| |\hat{g}(x^{(i)}_{ij}) - \hat{g}(x^{(i-1)}_{ij})|.
\]

By (3.2), \( \hat{g}(x^{(i)}_{ij}) \leq d_{ij} \). By definition, \( \sum_{\ell=1}^{k_{ij}} |f(x^{(i-1)}_{ij}) - f(x^{(i)}_{ij})| \leq V_{ij}(f) \). Also, by the mean value theorem and Assumption \( T'(d) \),
\[
\frac{|\hat{g}(\hat{c}^{(i)}_{ij}) - \hat{g}(c^{(i)}_{ij})|}{x^{(i)}_{ij} - x^{(i-1)}_{ij}} \leq |\hat{g}'(c^{(i)}_{ij})| = |\hat{T}'(c^{(i)}_{ij})|/|\hat{T}'(c^{(i)}_{ij})|^2 \leq \Gamma,
\]
where \( c^{(i)}_{ij} \in [x^{(i-1)}_{ij}, x^{(i)}_{ij}] \). Using the fact that
\[
\lim_{\max(|x^{(i)}_{ij} - x^{(i-1)}_{ij}|) \to 0} \sum_{\ell=1}^{k_{ij}} |f(x^{(i-1)}_{ij})|(x^{(i)}_{ij} - x^{(i-1)}_{ij}) = \int_{a_{ij}}^{b_{ij}} |f|d\nu,
\]
we get from (3.3) that
\[
V_{\hat{T}_{ij}}((f \cdot \hat{g}) \circ \hat{T}_{ij}^{-1}) \leq d_{ij}V_{ij}(f) + \Gamma \int_{\hat{T}_{ij}} |f|d\nu.
\]

Denote \( c = \min\{\nu(\hat{T}_{ij}) : i = 1, 2, \cdots, 1 \leq j \leq K\} \), where \( c > 0 \) because there is only finite number of images \( \hat{T}_{ij} \). It can be shown that (see e.g. [Br], Ch. 3)
\[
V(\hat{\mathcal{F}}) \leq 2 \sum_{j=1}^{K} V_{\hat{T}_{ij}}((f \cdot \hat{g}) \circ \hat{T}_{ij}^{-1}) + 2c^{-1}\|f\|_1.
\]

By Assumption \( T'(c) \), \( d_{ij} \leq \Delta^{-1} \) for all \( i = 1, 2, \cdots \) and \( j = 1, \cdots, K \). Hence
\[
|\hat{\mathcal{F}}f|_B = V(\hat{\mathcal{F}}) \leq 2\Delta^{-1}V(f) + 2\Gamma \int |f|d\nu + 2c^{-1}\|f\|_1 = \eta |f|_B + D\|f\|_1,
\]
where \( \eta = 2\Delta^{-1} < 1 \) and \( D = 2\Gamma + 2c^{-1} > 0 \).

Part (ii) and (iii) can be proved in a similar way for the proof of corresponding part of Lemma 4.2. \(\Box\)
Lemma 3.3. There exists a constant $C_R > 0$ such that $\|R_n\|_{\mathcal{B}} \leq C_R d_n$ for all $n > 0$.

Proof. For $f \in \mathcal{B}$, denote

$$R_{ij}f = 1_{\tilde{X}} \cdot \mathcal{P}^i (f1_{I_{ij}})(x).$$

(3.5)

Hence $R_i = \sum_{j=1}^{K} R_{ij}$ and $\mathcal{P} = \sum_{i=0}^{\infty} \sum_{j=1}^{K} R_{ij}$ by definition and linearity of $\mathcal{P}$.

Assume $i > 0$; since $\tilde{T}_{ij} [a_{ij}, b_{ij}] = I_0 \subset I$, by (3.2), $\hat{\nu}(I_{ij}) \leq d_{ij} \hat{\nu}(I_0) < d_{ij}$.

Hence, by Assumption B(b),

$$\int_{I_{ij}} |f| d\hat{\nu} \leq \|f\|_{\infty} \hat{\nu}(I_{ij}) \leq C_b \|f\|_{\mathcal{B}} \cdot d_{ij} \hat{\nu}(I_0) \leq C_b d_{ij} \|f\|_{\mathcal{B}}.$$  

(3.6)

Note that $V_{I_{ij}} (f) \leq V(f) = |f|_{\mathcal{B}} < \|f\|_{\mathcal{B}}$. By (3.4),

$$V_{\tilde{T}_{ij}}((f \cdot \tilde{\gamma}) \circ \tilde{T}_{ij}^{-1}) \leq d_{ij} \|f\|_{\mathcal{B}} + \Gamma C_b d_{ij} \|f\|_{\mathcal{B}} = (1 + \Gamma C_b) d_{ij} \|f\|_{\mathcal{B}}.$$  

(3.7)

Since $R_{ij}f(x) = 1_{\tilde{X}}(x) \cdot (f \cdot \tilde{\gamma}) \circ \tilde{T}_{ij}^{-1}(x)$, we have

$$|R_{ij}f|_{\mathcal{B}} \leq 2V_{\tilde{T}_{ij}}((f \cdot \tilde{\gamma}) \circ \tilde{T}_{ij}^{-1}) + 2 \frac{1}{\hat{\nu}(I_0)} \int_{I_{ij}} |f| d\hat{\nu}.$$  

By (3.6) and (3.7),

$$|R_{ij}f|_{\mathcal{B}} \leq 2(1 + \Gamma C_b) d_{ij} \|f\|_{\mathcal{B}} + 2C_b d_{ij} \|f\|_{\mathcal{B}}.$$

On the other hand, by (3.5) and (3.6), we have

$$\|R_{ij}f\|_{L^1} = \int_{\tilde{X}} \mathcal{P}^{i+1} (f1_{I_{ij}}) d\hat{\nu} = \int_{I_{ij}} f d\hat{\nu} \leq \int_{I_{ij}} |f| d\hat{\nu} \leq C_b d_{ij} \|f\|_{\mathcal{B}}.$$  

Hence, we get

$$\|R_{ij}f\|_{\mathcal{B}} = |R_{ij}f|_{\mathcal{B}} + \|R_{ij}f\|_{L^1} \leq 2(1 + \Gamma C_b) + 3C_b d_{ij} \|f\|_{\mathcal{B}}.$$

By the definition of $R_{ij}$ and $d_n$, we get

$$\|R_n f\|_{\mathcal{B}} \leq \sum_{j=1}^{K} \|R_{n-1,j} f\|_{\mathcal{B}} \leq K'(2 + 2\Gamma C_b + 3C_b) d_n,$$

where $K' < K$ is the number of preimages of $I_0$ that are not in $I_1$. So the result follows with $C_R = K'(2 + 2\Gamma C_b + 3C_b)$.

□
4 Systems on multidimensional spaces: generalities and the role of the derivative

The main difficulty to investigate the statistical properties for systems with an indifferent fixed point \( p \) in higher dimensional space is that near \( p \) the system could have unbounded distortion in the following sense: there are uncountably many points \( z \) near \( p \) such that for any neighborhood \( V \) of \( z \), we can find \( \tilde{z} \in V \) with the ratio

\[
| \det DT_1^{-n}(z) | / | \det DT_1^{-n}(\tilde{z}) |
\]

unbounded as \( n \to \infty \) (see Example in Section 2 in [HV]). For this reason we need a more extensive analysis of the expanding features around the neutral fixed point. This has been accomplished in the previous quoted paper and in order to construct an absolutely continuous invariant measure by adding the Assumption T” below, which, together with (4.4), will also be used to get the rate of mixing.

4.1 Setting and statement of results.

Let \( X \subset \mathbb{R}^m \), \( m \geq 1 \), be again a compact subset with \( \text{int} X = X \), \( d \) the Euclidean distance, and \( \nu \) the Lebesgue measure on \( X \) with \( \nu(X) = 1 \).

Assume that \( T : X \to X \) is a map satisfying the following assumptions.

**Assumption T”.** (a) (Piecewise smoothness) There are finitely many disjoint open sets \( U_1, \ldots, U_K \) with piecewise smooth boundary such that \( X = \bigcup_{i=1}^{K} \overline{U_i} \) and for each \( i \), \( T_i : T|_{U_i} \) can be extended to a \( C^{1+\alpha} \) diffeomorphism \( T_i : \overline{U_i} \to B_{\varepsilon_i}(T_i U_i) \), where \( \overline{U_i} \supset U_i, \alpha \in (0, 1] \) and \( \varepsilon_i > 0 \).

(b) (Fixed point) There is a fixed point \( p \in U_1 \) such that \( T^{-1} p \notin \partial U_j \) for any \( j = 1, \ldots, K \).

(c) (Topological mixing)

\[ T : X \to X \text{ is topologically mixing}. \]

**Remark 4.1.** Assumption T”(b) will allow us to get a good structure for the first return map around any pre-images of \( p \) different from \( p \) itself. In particular there will be an open neighborhood for each of those pre-images which is partitioned in level sets ordered with increasing first return time starting from 2 and with the same (large) image for the induced map. As we will see, this induction scheme reveals to be particularly useful when we will consider the transfer operator on the quasi-Hölder function space; in this regard we also defer to our previous paper [HV].

Before continuing with the list of assumptions we need to introduce a few more quantities and notations.
For any \( \varepsilon_0 > 0 \), denote

\[
G_U(x, \varepsilon, \varepsilon_0) = 2 \sum_{j=1}^K \frac{\nu(T_j^{-1} B_\varepsilon(\partial TU_j) \cap B_{1-s\varepsilon_0}(x))}{\nu(B_{1-s\varepsilon_0}(x))}.
\]

From now on we assume that the indifferent fixed point \( p = 0 \).

For any \( x \in U_i \), we define \( s(x) \) as the inverse of the slowest expansion near \( x \), that is,

\[
s(x) = \min \{ s : d(x, y) \leq sd(Tx, Ty), y \in U_i, d(x, y) \leq \min \{ \varepsilon_1, 0.1|x| \} \}.
\]

where the factor 0.1 forces the points \( y \) to stay in a ball around \( x \) which does not intersect the origin, though any other small factor would work as well.

Take an open neighborhood \( Q \) of \( p \) such that \( TQ \subset U_1 \), then let

\[
s = s(Q) = \max \{ s(x) : x \in X \setminus Q \}.
\]

(4.1)

Let \( \hat{T} = \hat{T}_Q \) be the first return map with respect to \( \hat{X} = \hat{X}_Q = X \setminus Q \). Then for any \( x \in U_j \), we have \( \hat{T}(x) = T_j(x) \) if \( T_j(x) \notin Q \), and \( \hat{T}(x) = T_1 T_j(x) \) for some \( i > 0 \) if \( T_j(x) \in Q \). Denote \( \hat{T}_{ij} = T_i T_j \) for \( i > 0 \).

Further, we take \( Q_0 = TQ \setminus Q \). Then we denote \( U_{01} = U_1 \setminus Q \), \( U_{0j} = U_j \setminus T_j^{-1} Q \) if \( j > 1 \), and \( U_{ij} = T_j^{-1} Q_0 \) for \( i > 0 \). Hence, \( \{ U_{ij} : i = 0, 1, 2, \cdots \} \) form a partition of \( U_j \) for \( j = 2, 3, \cdots, K \).

For \( 0 < \varepsilon \leq \varepsilon_0 \), we denote

\[
G_Q(x, \varepsilon, \varepsilon_0) = 2 \sum_{j=1}^K \sum_{i=0}^\infty \frac{\nu(\hat{T}_{ij}^{-1} B_\varepsilon(\partial Q_0) \cap B_{1-s\varepsilon_0}(x))}{\nu(B_{1-s\varepsilon_0}(x))},
\]

and

\[
G(x, \varepsilon, \varepsilon_0) = G_U(x, \varepsilon, \varepsilon_0) + G_Q(x, \varepsilon, \varepsilon_0), \quad G(\varepsilon, \varepsilon_0) = \sup_{x \in \hat{X}} G(x, \varepsilon, \varepsilon_0).
\]

(4.2)

**Assumption T". (continued)**

(d) (Expansion) \( T \) satisfies \( 0 < s(x) < 1 \ \forall x \in X \setminus \{ p \} \).

Moreover, there exists an open region \( Q \) with \( p \in Q \subset \overline{Q} \subset TQ \subset \overline{TQ} \subset U_1 \) and constants \( \alpha \in (0, \alpha], \eta \in (0,1) \), such that for all \( \varepsilon_0 \) small,

\[
s^\alpha + \lambda \leq \eta < 1,
\]

where \( s \) is defined in (4.1) and

\[
\lambda = 2 \sup_{0 < \varepsilon \leq \varepsilon_0} \frac{G(\varepsilon, \varepsilon_0)}{\varepsilon^\alpha} \varepsilon_0^\alpha.
\]

(4.3)
(e) (Distortion) For any \( b > 0 \), there exist \( \zeta > 0 \) such that for any small \( \varepsilon_0 \) and \( \varepsilon \in (0, \varepsilon_0) \), we can find \( 0 < N = N(\varepsilon) \leq \infty \) with

\[
\frac{|\det DT_1^{-n}(y)|}{|\det DT_1^{-n}(x)|} \leq 1 + \zeta \varepsilon \quad \forall y \in B_\varepsilon(x), \ x \in B_{\varepsilon_0}(Q_0), \ n \in (0, N],
\]

and

\[
\sum_{n=N}^{\infty} \sup_{y \in B_\varepsilon(x)} |\det DT_1^{-n}(y)| \leq b^{m+\alpha} \quad \forall x \in B_{\varepsilon_0}(Q_0),
\]

where \( \alpha \) is given in part (d) and \( m \) is the dimension of the ambient space.

For sake of simplicity of notations, we may assume \( \hat{\alpha} = \alpha \).

Remark 4.2. We stress that the measure \( \nu(T_1^{-1}B_\varepsilon(\partial TU_j)) \) usually plays an important role in the study of statistical properties of systems with discontinuities. Here \( G_U(x, \varepsilon, \varepsilon_0) \) gives a quantitative measurement of the competition between the expansion and the accumulation of discontinuities near \( x \). We refer to [Ss], Section 2, for more details about its geometric meaning. Furthermore it is proved, still in [Ss] Lemma 2.1, that if the boundary of \( U_i \) consists of piecewise \( C^1 \) codimension one embedded compact submanifolds, then

\[
G_U(\varepsilon, \varepsilon_0) \leq 2N_U \frac{\gamma_m^{-1} s \epsilon}{\gamma_m \gamma_m (1 - s)} \varepsilon_0(1 + o(1)),
\]

where \( N_U \) is the maximal number of smooth components of the boundary of all \( U_i \) that meet in one point and \( \gamma_m \) is the volume of the unit ball in \( \mathbb{R}^m \).

Remark 4.3. If \( T^{-1}TQ \cap \partial U_j = \emptyset \) for any \( j \), then for any small \( \varepsilon_0 \), either \( G_Q(x, \varepsilon, \varepsilon_0) = 0 \) or \( G_U(x, \varepsilon, \varepsilon_0) = 0 \), and therefore we have \( G(x, \varepsilon, \varepsilon_0) = \max\{G_U(x, \varepsilon, \varepsilon_0), G_Q(x, \varepsilon, \varepsilon_0)\} \).

Remark 4.4. If \( T \) has bounded distortion (see below), then \( G_Q \) is roughly equal to the ratio between the volume of \( B_{\varepsilon_0}(\partial Q_0) \) and the volume of \( Q_0 \). Therefore if \( \varepsilon_0 \) is small enough, then \( \sup_{x \in X} \{G_Q(x, \varepsilon, \varepsilon_0)\} \) is bounded by \( \sup_{x \in X} \{G_U(x, \varepsilon, \varepsilon_0)\} \).

Remark 4.5. We include Assumption \( T''(e) \) since near the fixed point distortion for \( DT_1 \) is unbounded in general. It requires that either distortion of \( DT_1^{-n} \) is small, or \( |\det DT_1^{-n}| \) itself is small.

Remark 4.6. There are some sufficient conditions under which Assumption \( T''(d) \) and (e) could be easily verified. We refer [HV] for more details, see in particular Theorems B and C in that paper.

If near \( p \) distortion is bounded, then Assumption \( T''(e) \) is automatically satisfied and it will be stated as follows (it could be regarded as the case \( N(\varepsilon) = \infty \) for any \( \varepsilon \in (0, \varepsilon_0) \)):

Assumption \( T'' \). (variant)
There exist $J > 0$ such that for any small $\varepsilon_0$ and $\varepsilon \in (0, \varepsilon_0)$,
\[
\frac{|\det DT^{-n}(y)|}{|\det DT^{-n}(x)|} \leq 1 + Je^\alpha \quad \forall y \in B_\varepsilon(x), \ x \in B_{\varepsilon_0}(Q_0), \ n \geq 0.
\]

**Remark 4.7.** It is well known that if $\dim X = m = 1$, any system that has the form given by (4.4) below near the fixed point satisfies Assumption $T''(e')$. The systems given in Example 4.1 satisfy it too.

To estimate the decay rates, we often consider the following special cases: There are constants $\gamma' > \gamma > 0$, $C_i, C'_i > 0$, $i = 0, 1, 2$, such that in a neighborhood of the indifferent fixed point $p = 0$,
\[
\begin{align*}
|y|(1 - C_0|y|^{\gamma} + O(|y|^{\gamma'})) &\leq |DT^{-1}(y)| \leq |y|(1 - C_0|y|^{\gamma} + O(|y|^{\gamma'})), \\
1 - C_1|y| + O(|y|^{\gamma'}) &\leq DDT^{-1}(y) \leq 1 - C_1|y| + O(|y|^{\gamma'}), \\
C_2|y|^{-\gamma} + O(|y|^{-\gamma'-1}) &\leq D^2T^{-1}(y) \leq C_2|y|^{-\gamma} + O(|y|^{-\gamma'-1}).
\end{align*}
\]
where $||DT^{-1}||, ||DT||$ etc., denote the operator norm.

We will now define the space of functions particularly adapted to study the action of the transfer operator on the class of maps just introduced. If $\Omega$ is a Borel subset of $\hat{X}$, we define the oscillation of $f$ over $\Omega$ by the difference of essential supremum and essential infimum of $f$ over $\Omega$:
\[
\text{osc}(f, \Omega) = \text{Essup}_\Omega f - \text{Einf}_\Omega f.
\]
We notice that the function $x \to \text{osc}(f, B_\varepsilon(x))$ is measurable.

For $0 < \alpha < 1$ and $\varepsilon_0 > 0$, we define the quasi-Hölder seminorm of $f$ with $\text{supp} f \subset \hat{X}$ as
\[
|f|_B = \sup_{0 < r \leq \varepsilon_0} \varepsilon^{-\alpha} \int_{\hat{X}} \text{osc}(f, B_r(x)) d\nu(x),
\]
where $\nu$ is the normalized Lebesgue measure on $\hat{X}$, and take the space of the functions as
\[
\mathcal{B} = \left\{ f \in L^1(\hat{X}, \nu) : |f|_B < \infty \right\},
\]
and then equip it with the norm
\[
\|f\|_B = \|f\|_{L^1(\hat{X}, \nu)} + |f|_B.
\]
Clearly, the space $\mathcal{B}$ does not depend on the choice of $\varepsilon_0$, though $|f|_B$ does.

Let $s_{ij} = \sup \{ ||DT^{-1}_{ij}(x)|| : x \in B_{\varepsilon_0}(Q_0) \}$, and $s_n = \max \{ s_{n-1,j} : j = 2, \cdots, K \}$.

*Since the boundary of $\hat{X}$ is piecewise smooth, we could define the space of the function directly on $\hat{X}$ instead of $\mathbb{R}^m$ as it was done in [Ss].
Theorem D. Let $\hat{X}, \hat{T}$ and $B$ be defined as above. Suppose $T$ satisfies Assumption $T''(a)$ to $(e)$. Then there exist $\varepsilon_0 \geq \varepsilon_1 > 0$ such that Assumption $B(a)$ to $(f)$ and conditions $S(1)$ to $S(4)$ are satisfied and $\|R_n\| = O(s_n^\beta)$. Hence, if $\sum_{k=n+1}^\infty s_n^\beta = O(n^{-\beta})$ for some $\beta > 1$, then there exists $C > 0$ such that for any functions $f \in B$, $g \in L^\infty(X, \nu)$ with $\text{supp} f, \text{supp} g \subset \hat{X}$, (1.5) holds.

Remark 4.8. For Lipschitz observables, the rates of decay of correlation are given by the rates of decay of $\mu(\tau > n)$ if the systems have Markov partitions and bounded distortion. It is generally believed that for Hölder observables, the decay rates may be slower if the Hölder exponents become smaller. It is unclear to the authors whether the rates we get are optimal. In the next section, we will put stronger conditions on the systems so that we can get optimal rates for Hölder observables with the Hölder exponents larger than or equal to $\alpha$.

Remark 4.9. For one dimensional systems the rates given in the theorem are optimal, since the decreasing rates given by the norm of derivatives are the same as those given by determinants (see the discussion in the Introduction or Section 4 for more details). So the theorem provides the same decay rates as Theorem C does, but for different sets of observables, since functions with bounded variation are not necessary quasi-Hölder functions and vice versa.

Before giving the proof, we present an example.

Example 4.1. Assume that $T$ satisfies Assumption $T''(a)$ to $(d)$, and near the fixed point $p = 0$, the map $T$ satisfies

$$T(z) = z(1 + |z|^\gamma + O(|z|^\gamma')), $$

where $z \in X \subset R^m$ and $\gamma' > \gamma$.

Denote $z_n = T_1^{-n}z$. We showed in Lemma 3.1 in [HV] that $|z_n| = \frac{1}{(\gamma n)^{1/\gamma}} + O\left(\frac{1}{n^{1/\gamma}}\right)$, where $\tilde{\gamma} < \gamma$. Using this fact we can check that $T$ satisfies also Assumption $T''(e')$. Hence, the theorem can be applied.

If the dimension $m = 1$, then $T^n$ maps the interval $[z_{n+1}, z_n] = [z_{n+1}, T(z_{n+1})]$ to its image $[z_1, z_0]$ bijectively. It follows that $\|DT_1^{-n}\|$ is roughly proportional to $|z_n|^{1+\gamma}/(|z_0| - |z_1|)$, since the length of the interval $[z_{n+1}, T(z_{n+1})]$ is roughly equal to $|T(z_{n+1}) - z_{n+1}| \approx |z_0|^{1+\gamma}$, see also Lemma 3.1 and Lemma 3.2 in [HV] for a more formal derivation. So $s_n = O\left(\frac{1}{n^{1+1/\gamma}}\right)$ and $\sum_{k=n+1}^\infty s_k^\alpha = O\left(\frac{1}{n^{\alpha-1}}\right)$. If $\gamma \in (0, 1)$ is such that $\alpha(1/\gamma + 1) > 1$, the series is convergent. Also, as stated in Theorem C in the last section, $\sum_{k=n+1}^\infty \mu(\tau > k) = O\left(\frac{1}{n^{\frac{1}{\gamma} - 1}}\right)$.
So if $\alpha(1/\gamma + 1) > 1/\gamma$, the sum involving $s_k^\alpha$ decreases faster. We get that the decay rate is given by

$$\left| \text{Cov}(f, g \circ T^n) \right| = O\left( \sum_{k=n+1}^{\infty} \mu(\tau > k) \right) = O\left( \frac{1}{n^{1/2-\beta}} \right),$$

for $f \in B, g \in L^\infty(X, \nu)$ with $\text{supp} f, \text{supp} g \subset \hat{X}$ and with $\beta = \frac{1}{2}$. This gives the same results as in Theorem C for quasi Hölder test functions instead for functions of bounded variations.

On the other hand, if $m \geq 2$, then $T_1^{-n}$ maps a sphere about the fixed point of radius $|z|$ to a sphere of radius $|z_n|$, if higher order terms are ignored. Hence, $DT_1^{-n}$ contracts vectors in the tangent space of the sphere at the rate of order $|z_n|$. To see the contracting rates along the radial direction, i.e., the direction orthogonal to the tangent space of the spheres, we note that restricted to each ray the map has the form $T(r) = r(1 + r^\gamma + O(r^\gamma))$. Hence, by the above arguments for one dimensional case, $DT_1^{-n}$ contracts vectors in the radial direction at the rates of order $|z_n|^{1+\gamma}$. So the norm $\|DT_1^{-n}\|$ is roughly proportional to $|z_n|$, and $s_n = O\left( \frac{1}{n^{1+\gamma/2}} \right)$ and $\sum_{k=n+1}^{\infty} s_k^\alpha = O\left( \frac{1}{n^{(\alpha/\gamma)-1}} \right)$. If $\gamma \in (0, 1/2)$ is such that $\alpha/\gamma > 1$, the series is convergent. By defining $\beta := \frac{2}{\gamma} - 1$ we can now consider the three cases $\beta > 2, 1 < \beta < 2, \beta = 2$ in order to determine the error term $F_\beta(n)$. Let us take, for instance, $\beta > 2$, which requires $\alpha/\gamma > 3$.

Note that $\nu(\tau > n)$ is of the same order as $|z_n|^m$, and therefore $\mu(\tau > n) = O\left( \frac{1}{n^{m/\gamma}} \right)$. It follows that $\sum_{k=n+1}^{\infty} \mu(\tau > k) = O\left( \frac{1}{n^{(m/\gamma)-1}} \right)$. Since the order is higher, by (1.5), we get

$$\left| \text{Cov}(f, g \circ T^n) \right| \leq C/n^\beta.$$

### 4.2 Proof of Theorem D

The proof of Theorem D requires a few preparatory lemmas.

First of all and in order to deduce the spectral properties of $\hat{P}$ from the Lasota-Yorke inequality, one needs to verify Assumption B on the space of functions $B$.

**Lemma 4.1.** $B$ is a Banach space satisfying Assumptions B(a) to (f) with $C_a = 2C_b = 2 \gamma_m^{-1} \epsilon_0^{-m}$, where $\gamma_m$ is the volume of the unit ball in $\mathbb{R}^m$.

**Proof.** Parts (a), (b) and (c) are stated in Propositions 3.3 and 3.4 in [Ss] with $C_b = \max\{1, \epsilon \alpha \} / \gamma_m \epsilon_0^m$ and $C_a = 2 \max\{1, \epsilon \alpha \} / \gamma_m \epsilon_0^m$. Part (d) follows from the fact that Hölder continuous functions with compact support in $\hat{X}$ are dense in $L^1(\hat{X}, \hat{\nu})$. 

---

26
Let us now assume $f(u) = \lim_{n \to \infty} f_n(u)$ for $\nu$-a.e. $u \in \mathbb{R}^m$. Take $x \in \mathbb{R}^m$, and $\varepsilon \in (0, \varepsilon_0)$. It is easy to see that for almost every pair of $y, z \in B_\varepsilon(x)$, we have

$$|f(y) - f(z)| \leq \lim_{n \to \infty} |f_n(y) - f_n(z)| \leq \liminf_{n \to \infty} \text{osc}(f_n, B_\varepsilon(x)).$$

Hence, $\text{osc}(f, B_\varepsilon(x)) \leq \liminf_{n \to \infty} \text{osc}(f_n, B_\varepsilon(x))$. By Fatou’s lemma, we have

$$\int \text{osc}(f, B_\varepsilon(x))d\nu \leq \liminf_{n \to \infty} \int \text{osc}(f_n, B_\varepsilon(x))d\nu.$$

This implies $|f|_B \leq \liminf_{n \to \infty} |f_n|_B$. We get part (e).

It leaves to show part (f). For a function $f \in B$, denote

$$D_n(f) = \left\{ x \in \mathbb{R}^m : \liminf_{\varepsilon \to 0} \text{osc}(f, B_\varepsilon(x)) > \frac{1}{n} \right\}, \quad D(f) = \bigcup_{n=1}^{\infty} D_n(f).$$

Clearly $D(f)$ is the set of discontinuity points of $f$. If $\nu(D(f)) > 0$, then there exists $N > 0$ such that $\text{Leb}(D_N(f)) > \epsilon > 0$. Notice that $D_N(f) = \bigcup_{k \geq 1} S_k$, where $S_k = \bigcap_{n \geq k} \{ x : \text{osc}(f, B_\frac{1}{n}(x)) > \frac{1}{k} \}$ is an increasing sequence of measurable sets.

For $k$ big enough we still have $\nu(S_k) > \epsilon$ and therefore, for such a $k$:

$$|f|_B \geq \sup_{\varepsilon > 0} \varepsilon^{-a} \int_{D_n(f)} \text{osc}(f, B_\varepsilon(x))d\nu(x) \geq \sup_{\varepsilon > 0} \varepsilon^{-a} \int_{S_k} \text{osc}(f, B_\varepsilon(x))d\nu(x) = \infty.$$ 

This means $f \notin B$; in other words, any $f \in B$ satisfies $\nu(D(f)) = 0$.

Take any $f \in B$ with $f \geq 0$ almost everywhere. If $f(x) = 2c > 0$ for some $x \notin D(f)$, then there is $\varepsilon > 0$ such that $\text{osc}(f, B_\varepsilon(x)) \leq c$. Hence, $f(x') \geq c > 0$ for almost every point $x' \in B_\varepsilon(x)$. So $B_\varepsilon(x) \setminus \{ f > 0 \}$ has Lebesgue measure zero. This implies that $\{ f > 0 \}$ is almost open and therefore part (f) follows.

Before stating the next lemma, we recall that the space $B$ depends on the exponent $\alpha$ and the value of the seminorms on $\varepsilon_0$: as we did above, we will not index $B$ with these two parameters. Moreover all the integrals in the next proof will be performed over $\tilde{X}$.

Lemma 4.2. There exists $\varepsilon_* > 0$ such that for any $\varepsilon_0 \in (0, \varepsilon_*)$, we can find constants $\eta \in (0, 1)$ and $D, \hat{D} > 0$ satisfying

(i) for any $f \in B$, $|\tilde{\mathcal{P}}f|_B \leq \eta |f|_B + D\|f\|_{L^1(\nu)}$;

(ii) for any $f \in B$, $\|R(z)f\|_B \leq \eta (\|f\|_B + \hat{D}\|f\|_{L^1(\nu)})$; and

(iii) for any $\tilde{f} \in \tilde{B}$, $\|\tilde{\mathcal{P}}f\|_{\tilde{B}} \leq \eta \|\tilde{f}\|_{\tilde{B}} + D\|\tilde{f}\|_{L^1(\tilde{\nu} \times \rho)}$.
Proof. By Assumption T'' (d), \( s^\alpha + \lambda < 1 \). Therefore if we first choose \( b \) small enough, we obtain \( \zeta \) according to Assumption T''(e), and then we can take \( \varepsilon_0 \) small enough in order to get

\[
\eta := (1 + \zeta \varepsilon_0^m)(s^\alpha + \lambda) + 2\gamma_m^{-1}bK' < 1, \tag{4.8}
\]

where \( K' \) is the number of \( j \) such that \( U_{ij} \neq \emptyset \). Clearly, \( \eta \) is decreasing with \( \varepsilon_0 \).

Let us define:

\[
D := 2\zeta + 2(1 + \zeta \varepsilon_0^m)\lambda / \varepsilon_0^m + 2\gamma_m^{-1}bK' > 0. \tag{4.9}
\]

For any \( x \in \tilde{X} \), let us denote \( x_{ij} = \tilde{T}^{-1}_{ij}x, \tilde{g}_{ij}(x) = |\det D\tilde{T}_{ij}(x)|^{-1} \) and for \( f \in B \):

\[
R_{ij}f = 1_{\tilde{X}} \cdot \mathcal{P}^i(f1_{U_{ij}})(x). \tag{4.10}
\]

Clearly,

\[
R_{ij}f(x) = f(x_{ij})\tilde{g}(x_{ij})1_{U_{ij}}(x_{ij}). \tag{4.11}
\]

Hence \( R_i = \sum_{j=1}^K R_{ij} \) and \( \mathcal{P} = \sum_{i=0}^{\infty} \sum_{j=1}^K R_{ij} \) by definition and the linearity of \( \mathcal{P} \). We also define

\[
G_{ij}(x, \varepsilon, \varepsilon_0) = \frac{\nu(\tilde{T}^{-1}_{ij}B_{\epsilon}(\partial \tilde{T}U_{ij}) \cap B_{\epsilon}(1-s)\varepsilon_0(x))}{\nu(B_{\epsilon}(1-s)\varepsilon_0(x))}. \tag{4.12}
\]

Clearly, \( G(x, \varepsilon, \varepsilon_0) = 2 \sum_{i=0}^{\infty} \sum_{j=1}^K G_{ij}(G(x, \varepsilon, \varepsilon_0)). \)

For any \( \varepsilon \in (0, \varepsilon_0] \), take \( N = N(\varepsilon) > 0 \) as in Assumption T''(e).

For \( i \leq N(\varepsilon) \) and by the proof of Proposition 6.2 in [HV], we know that

\[
\text{osc}(R_{ij}f, B_{\varepsilon}(x)) = \text{osc}((f\tilde{g}) \circ \tilde{T}^{-1}_{ij}B_{\varepsilon}(x)).
\]

The computation in that proof also gives

\[
\text{osc}(f\tilde{g}, \tilde{T}^{-1}_{ij}B_{\varepsilon}(x) \cap U_{ij}) \leq (1 + \zeta e^\alpha)\text{osc}(f, B_{\varepsilon}(x_{ij}) \cap U_{ij})\tilde{g}(x_{ij}) + 2\zeta e^\alpha |f(x_{ij})|\tilde{g}(x_{ij}).
\]

Notice that \( \text{osc}(f, B_{\varepsilon}(x_{ij}) \cap U_{ij}) \leq \text{osc}(f, B_{\varepsilon}(x_{ij})). \) By integrating and using (4.11) we get

\[
\int \text{osc}((f\tilde{g}) \circ \tilde{T}^{-1}_{ij}, B_{\varepsilon}())1_{\tilde{T}U_{ij}} \, d\nu \leq \int [(1 + \zeta e^\alpha)R_{ij}\text{osc}(f, B_{\varepsilon}()) + 2\zeta e^\alpha R_{ij}|f|] \, d\nu. \tag{4.13}
\]

On the other hand, by the same arguments as in Section 4 of [SS], we get

\[
\int 2\text{Esup}_{B_{\varepsilon}(x)}(f\tilde{g}) \circ \tilde{T}^{-1}_{ij}1_{B_{\varepsilon}(\partial \tilde{T}U_{ij})} \, d\nu 
\leq 2(1 + \zeta e^\alpha) \int_{\tilde{X}} G_{ij}(x, \varepsilon, \varepsilon_0) [f(x) + \text{osc}(f, B_{\varepsilon_0}(x))] \, d\nu. \tag{4.14}
\]

28
Therefore by (4.12), (4.13) and (4.14),

\[ |R_{ij}|_{B} = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \text{osc}(R_{ij} f, B_{\varepsilon}(\cdot)) d\nu \]

\[ \leq \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int [(1 + \zeta \varepsilon^\alpha)R_{ij} \text{osc}(f, B_{\varepsilon}(\cdot)) + 2\zeta \varepsilon^\alpha R_{ij}|f|] d\nu \]

\[ + \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} 2(1 + \zeta \varepsilon^\alpha) \int G_{ij}(x, \varepsilon, \varepsilon_0)[|f|(x) + \text{osc}(f, B_{\varepsilon_0}(x))] d\nu. \]  

(4.15)

For \( i > N(\varepsilon) \), by the definition of oscillation we obtain directly that

\[ \text{osc}(R_{ij} f, B_{\varepsilon}(x)) \leq 2\|f\|_{\infty} \sup_{T_{ij}^{-1}B_{\varepsilon}(x)} \hat{g}. \]

Hence, by Assumption B(b) with \( C_{b} = \gamma_{m}^{-1} \varepsilon_{0}^{-m} \), we have

\[ |R_{ij}|_{B} = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \text{osc}(R_{ij} f, B_{\varepsilon}(\cdot)) d\nu \]

\[ \leq 2\|f\|_{\infty} \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \sup_{T_{ij}^{-1}B_{\varepsilon}(x)} \hat{g} d\nu \]

\[ \leq 2(\gamma_{m}^{-1} \varepsilon_{0}^{-m})^{-1}(\|f\|_{B} + \|f\|_{1}) \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \sup_{T_{ij}^{-1}B_{\varepsilon}(x)} \hat{g} d\nu. \]  

(4.16)

(i) We first note that for all \( 0 < \varepsilon \leq \varepsilon_0 \),

\[ \varepsilon^{-\alpha} \sum_{i=0}^{N(\varepsilon)} \sum_{j=1}^{K} R_{ij} \text{osc}(f, B_{\varepsilon}(\cdot)) d\nu \leq \varepsilon^{-\alpha} \int \mathcal{F} \text{osc}(f, B_{\varepsilon}(\cdot)) d\nu \]

\[ \leq s^\alpha (s\varepsilon)^{-\alpha} \int \text{osc}(f, B_{\varepsilon}(\cdot)) d\nu \leq s^\alpha |f|_{B}, \]  

(4.17)

\[ \varepsilon^{-\alpha} \sum_{i=0}^{N(\varepsilon)} \sum_{j=1}^{K} 2(1 + \zeta \varepsilon^\alpha)G_{ij}(\cdot, \varepsilon, \varepsilon_0)[|f| + \text{osc}(f, B_{\varepsilon}(\cdot))] d\nu \]

\[ \leq \varepsilon^{-\alpha} 2(1 + \zeta \varepsilon^\alpha)G(\varepsilon, \varepsilon_0) \int [\|f\| + \text{osc}(f, B_{\varepsilon}(\cdot))] d\nu \]

\[ \leq (1 + \zeta \varepsilon^\alpha)\lambda \varepsilon_{0}^{-\alpha} \|f\|_{1} + |f|_{B}, \]

(4.18)

where we used (4.2) and (4.3). Also, by Assumption T''(e) and Assumption B(b) with \( C_{b} = \gamma_{m}^{-1} \varepsilon_{0}^{-m+\alpha} \), we have that for all \( 0 < \varepsilon \leq \varepsilon_0 \):

\[ \varepsilon^{-\alpha} \|f\|_{\infty} \int \sum_{N(\varepsilon)}^{\infty} \sum_{j=1}^{K'} \sup_{T_{ij}^{-1}B_{\varepsilon}(x)} \hat{g} d\nu \leq \varepsilon^{-\alpha} \|f\|_{\infty} \cdot bK' \varepsilon_{m+\alpha} \leq \gamma_{m}^{-1} bK' \|f\|_{B}. \]  

(4.19)
Since \( \tilde{\mathcal{D}} f(x) = \sum_{i=0}^{\infty} \sum_{j=1}^{K} R_{ij} f(x) \), by (4.15) and (4.16), and using (4.17) to (4.19), we obtain that \( |\tilde{\mathcal{D}} f|_{S} \) is bounded by

\[
\sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \left[ \int \sum_{i=0}^{K} \sum_{j=1}^{K} \text{osc}(R_{ij} f, B_{\varepsilon}(x)) d\nu + \int \sum_{i=0}^{K} \sum_{j=1}^{K} \text{osc}(R_{ij} f, B_{\varepsilon}(x)) d\nu \right]
\]

\[
\leq (1 + \varepsilon_0^\alpha) s_0^\varepsilon |f|_S + 2|f|_1 + (1 + \varepsilon_0^\alpha) \lambda \varepsilon_0^{-\alpha} |f|_1 + |f|_S + 2\gamma^{-1} b K' |f|_S
\]

\[
\leq [(1 + \varepsilon_0^\alpha) (s_0^\alpha + \lambda) + 2\gamma^{-1} b K'] |f|_S + [2\zeta + 2(1 + \varepsilon_0^\alpha) \lambda \varepsilon_0^{-\alpha} + 2\gamma^{-1} b K'] |f|_1.
\]

By definition of \( \eta \) in (4.8) and \( D \) in (4.9) we get the desired inequality.

(ii) Note that for any real valued function \( f \) and \( z \in \mathbb{C} \), we have \( \text{osc}(z f, B_{\varepsilon}(x)) = |z| \text{osc}(f, B_{\varepsilon}(x)) \). Also, note that if \( \{a_n\} \) is a sequence of positive numbers and \( z \in \mathbb{R} \), \( |\sum_{n=1}^{\infty} z^n a_n| \leq |z| \sum_{n=1}^{\infty} a_n \). Hence we have

\[
|R(z)f|_S \leq |z| \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \sum_{i=0}^{K} \sum_{j=1}^{K} \int \text{osc}(R_{ij} f, B_{\varepsilon}(x)) d\nu \leq |z| |\tilde{\mathcal{D}} f|_S.
\]

By part (i), the inequality becomes

\[
|R(z)f|_S \leq |z| |\eta f|_S + D |f|_1.
\]

Also, since \( \tilde{\mathcal{D}} \) and \( R_n \) are positive operators,

\[
\|R(z)f\|_1 \leq \sum_{n=1}^{\infty} \|z^n R_n f\|_1 \leq |z| \sum_{n=1}^{\infty} \|R_n f\|_1 = |z| \|\tilde{\mathcal{D}} f\|_1 = |z| \|f\|_1.
\]

It follows that

\[
\|R(z)f\|_S \leq |z| |\eta f|_S + (D + 1) |f|_1.
\]

We finally get the expected result with \( \hat{D} = D + 1 \).

(iii) The transfer operator \( \tilde{\mathcal{D}} \) has the form (see also [ADSZ])

\[
(\tilde{\mathcal{D}} f)(x, y) = \sum_{n=0}^{\infty} \sum_{j=1}^{K} \tilde{f}(\tilde{T}_{ij}^{-1} x, \Pi_{ij}) g(\tilde{T}_{ij}^{-1} x) J_{ij} \tilde{f}(x, y),
\]

for any \( \tilde{f} \in \mathcal{B} \), where \( S(U_{ij}) : Y \to Y \) are automorphisms. Let us denote:

\[
(\tilde{R}_{ij} \tilde{f})(x, y) = \tilde{f}(\tilde{T}_{ij}^{-1} x, S(U_{ij})^{-1}(y)) g(\tilde{T}_{ij}^{-1} x) J_{ij} \tilde{f}(x, y).
\]

Following the same computations as above, we get formulas similar to (4.15) and (4.16) but with \( R_n \) and \( \tilde{T}_{ij} \) replaced by \( R_{ij} \) and \( \tilde{T}_{ij} \) respectively, and \( f(\cdot) \) replaced by \( \tilde{f}(\cdot, y) \). Denote \( y_1 = S(U_{ij})^{-1}(y) \). Instead of (4.15) and (4.16), we get that for \( i < N(\varepsilon) \),

\[
|\tilde{R}_{ij} \tilde{f}(\cdot, y)|_S = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \text{osc}(\tilde{R}_{ij} \tilde{f}(\cdot, y_1), B_{\varepsilon}(\cdot)) d\nu
\]

\[
\leq \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \left[ (1 + \varepsilon_0^\alpha) \tilde{R}_{ij} \text{osc}(\tilde{f}(\cdot, y_1), B_{\varepsilon}(\cdot)) + 2\varepsilon_0^{-\alpha} \tilde{R}_{ij} \tilde{f}(\cdot, y_1) \right] d\nu + 2G_{ij}(x, \varepsilon, \varepsilon_0) (1 + \varepsilon_0^\alpha) (\text{osc}(\tilde{f}(\cdot, y_1), B_{\varepsilon}(\cdot)) + |\tilde{f}(\cdot, y_1)|)
\]

30
and for \( i \geq N(\varepsilon) \),
\[
|\tilde{R}_{ij} \tilde{f}(\cdot, y)|_{\mathcal{B}} = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \text{osc}(\tilde{R}_{ij} \tilde{f}(\cdot, y_1), B_{\varepsilon}(\cdot)) d\tilde{\nu} \\
\leq 2(\gamma_0 \varepsilon_0^{\alpha})^{-1} (|\tilde{f}(\cdot, y_1)|_{\mathcal{B}} + \|\tilde{f}(\cdot, y_1)\|_{L^1(\nu)}) \varepsilon^{-\alpha} \sup_{0 < \varepsilon \leq \varepsilon_0} \int_{\tilde{T}_{ij}^{-1} B_{\varepsilon}(x)} \tilde{g} d\tilde{\nu}.
\]

We observe that for any \( x, S(U_{ij}) : Y \to Y \) preserves the measure \( \rho \). We set
\[
\tilde{f}(x) = \int_{S} \tilde{f}(x, y_1) d\rho(y), \quad \text{osc}(\tilde{f}(\cdot), B_{\varepsilon}(\cdot)) = \int_{S} \text{osc}(\tilde{f}(\cdot, y_1), B_{\varepsilon}(\cdot)) d\rho(y).
\]
Integrating with respect to \( \tilde{\nu} \) and using Fubini's theorem, we get
\[
|\tilde{R}_{ij} \tilde{f}|_{\mathcal{B}} \leq \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \left[ (1 + \varepsilon \alpha) \tilde{R}_{ij} \text{osc}(\tilde{f}(\cdot), B_{\varepsilon}(\cdot)) + 2\varepsilon \alpha \tilde{R}_{ij} |\tilde{f}(\cdot)| \right] + 2G_{ij}(x_{ij}, \varepsilon, \varepsilon_0)(1 + \varepsilon \alpha) \left( \text{osc}(\tilde{f}(\cdot), B_{\varepsilon}(\cdot)) + |\tilde{f}(\cdot)| \right) d\tilde{\nu}
\]
and
\[
|\tilde{R}_{ij} \tilde{f}|_{\mathcal{B}} \leq 2(\gamma_0 \varepsilon_0^{\alpha})^{-1} (|\tilde{f}|_{\mathcal{B}} + \|\tilde{f}\|_{L^1(\tilde{\nu} \times \rho)}) \varepsilon^{-\alpha} \sup_{0 < \varepsilon \leq \varepsilon_0} \int \sup_{T_{ij}^{-1} B_{\varepsilon}(x)} \tilde{g} d\tilde{\nu}.
\]

By Fubini's theorem, we have also
\[
|\tilde{f}|_{\mathcal{B}} = \sup_{0 < \varepsilon \leq \varepsilon_0} \varepsilon^{-\alpha} \int \text{osc}(\tilde{f}(\cdot), B_{\varepsilon}(\cdot)) d\tilde{\nu},
\]
and
\[
|\tilde{f}|_{L^1(\tilde{\nu} \times \rho)} = \int |\tilde{f}(\cdot)| d\tilde{\nu}. \quad \text{Using the same arguments as in the proof of part (i) we get}
\]
\[
|\tilde{P} \tilde{f}(\cdot, y)|_{\mathcal{B}} \leq \sum_{n=0}^{\infty} \sum_{j=1}^{K} \tilde{R}_{ij} |\tilde{f}|_{\mathcal{B}} \leq (1 + \varepsilon \alpha) |\tilde{f}|_{\mathcal{B}} + 2\varepsilon \|\tilde{f}\|_{L^1(\tilde{\nu} \times \rho)}
\]
\[
+ (1 + \varepsilon \alpha) \lambda(|\tilde{f}|_{\mathcal{B}} + \varepsilon^{-\alpha} \|\tilde{f}\|_{L^1(\tilde{\nu} \times \rho)}) + 2\gamma_0^{-1} bK'(|\tilde{f}|_{\mathcal{B}} + \|\tilde{f}\|_{L^1(\tilde{\nu} \times \rho)}),
\]
and therefore the result of part (iii) with the same \( \eta \) and \( D \) giving in (4.8) and (4.9) respectively. \( \square \)

**Lemma 4.3.** There exists a constant \( C_R > 0 \) such that \( |R_n|_{\mathcal{B}} \leq C_R \varepsilon_0^{\alpha} \) for all \( n > 0 \).

**Proof.** Since \( R_i = \sum_j R_{ij} \), we only need to prove the results for \( R_{ij} \).

Take \( \varepsilon \in (0, \varepsilon_0) \). Choose any \( b > 0 \) and let \( N(\varepsilon) \) be given by Assumption T\(^{(\alpha)}\)(\( \varepsilon \)).

We first consider the case \( n = i + 1 \leq N(\varepsilon) \).

By the definition of \( R_{ij} \) given in (4.10), we have for any \( f \in \mathcal{B} \),
\[
\int R_{ij} f d\tilde{\nu} = \int_\mathcal{X} 1_{\tilde{X}} : \mathcal{D}^{i+1}(f 1_{U_{ij}}) d\tilde{\nu} = \int_\mathcal{X} f 1_{U_{ij}} d\tilde{\nu} = \int_{U_{ij}} f d\tilde{\nu}. \quad (4.20)
\]
We now denote $d_{ij} = \sup \{|\det D\tilde{T}_{ij}^{-1}(x)| : x \in B_r(Q_0)\}$. Since for any $x$, $|\det D\tilde{T}_{ij}^{-1}(x)| \leq ||D\tilde{T}_{ij}^{-1}(x)||$, we have $d_{ij} \leq s_{ij}$. Since $\tilde{T}U_{ij} = Q_0$, 

$$\nu(U_{ij}) \leq d_{ij}\nu(Q_0) \leq s_{ij}\nu(Q_0).$$

(4.21)

Hence by Assumption B(b),

$$\int R_{ij}f d\nu \leq ||f||_{L^\infty(\nu)}\nu(U_{ij}) \leq C_0\nu(Q_0)s_{ij}||f||_B.$$  

(4.22)

By similar arguments as for (4.20), we have

$$\int R_{ij}\text{osc}(f, B_{s_{ij}\varepsilon(\cdot)})d\nu \leq \int \text{osc}(f, B_{s_{ij}\varepsilon(\cdot)})1_{U_{ij}}d\nu \leq s_{ij}\varepsilon\alpha/|f||_B.$$  

(4.23)

We note that for each $j$, $\tilde{T}U_{ij} = Q_0$ and the “thickness” of $\tilde{T}_{ij}^{-1}B_r(\partial Q_0)$ is of order $s_{ij}\varepsilon$, since $\partial Q_0$ consists of piecewise smooth surfaces. So $G_{ij}(\varepsilon, \varepsilon_0) \leq C_G\varepsilon \alpha s_{ij}$ for some $C_G$ independent of $i$ and $j$. Therefore we have

$$\int \varepsilon^{-\alpha}2(1 + \varepsilon\alpha)G_{ij}(\varepsilon, \varepsilon_0)[|f| + \text{osc}(f, B_{s_{ij}\varepsilon(\cdot)})]d\nu$$

$$\leq 2(1 + \varepsilon\alpha)C_G\varepsilon^{1-\alpha}s_{ij}[||f||_{L^1(\varepsilon)} + \varepsilon_0/|f||_B].$$

Hence by (4.15) we get that

$$|R_{ij}||f||_B \leq C_R'\varepsilon_{ij}^{\alpha}[||f||_{L^1(\varepsilon)} + |f||_B] = C_R's_{ij}||f||_B$$

(4.24)

for $C_R' = (1 + \varepsilon_0/\varepsilon_0)(1 + 2C_G\varepsilon_0^{1-\alpha}) + 2C_0\varepsilon_0(Q_0)$.

We now consider the case $n = i + 1 > N(\varepsilon)$. As we mentioned in Remark 4.7, in this case $m \geq 2$. By definition, there is $C_\varepsilon > 0$ such that $\hat{g}(x_{ij}) \leq C_\varepsilon s_{ij}^{\alpha}$ for any $x_{ij} \in \tilde{T}_{ij}^{-1}B_r(Q_0)$ with $j = 2, \cdots, K$. By Assumption T''(\varepsilon) we know that for any $x \in B_r(Q_0)$, 

$$\left(\sup_{\tilde{T}_{ij}^{-1}B_r(x)}\hat{g}\right)^{1/2} \leq \left(\sum_{l=N(\varepsilon)}^\infty\sup_{\tilde{T}_{ij}^{-1}B_{l/2}(x)}\hat{g}\right)^{1/2} \leq \sqrt{b}\varepsilon^{(m+\alpha)/2} \leq \sqrt{b}\varepsilon^\alpha.$$  

Therefore we obtain

$$\sup_{\tilde{T}_{ij}^{-1}B_r(x)}\hat{g} = \left(\sup_{\tilde{T}_{ij}^{-1}B_{l/2}(x)}\hat{g}\right)^{1/2} \leq C_\varepsilon s_{ij}\sqrt{b}\varepsilon^\alpha$$

and substitute in (4.16) to get ($\alpha \leq 1$):

$$|R_{ij}||f||_B \leq C_R''s_{ij}||f||_B \leq C_R''s_{ij}||f||_B$$

for $C_R'' = 2(\gamma_0\varepsilon_0^{\alpha})^{-1}\sqrt{b}C_\varepsilon$.  

32
Finally, by (4.22), we have
\[ \|R_{ij}f\|_1 \leq \int R_{ij}|f|d\nu \leq C_B \nu(Q_0)s_{ij}\|f\|_B. \]

Thus we have \( \|R_{ij}f\|_B = (C'_R + C''_R + C_B \nu(Q_0))s_{ij}\|f\|_B \). The result of the lemma then follows. \( \square \)

We are finally ready to give the proof of Theorem D.

Proof of Theorem D. Choose \( \varepsilon_0 > 0 \) as in Lemma 4.2, and define \( B \) correspondingly by using this \( \varepsilon_0 \). By Proposition 3.3 in [Ss], \( B \) is complete, and hence is a Banach space. Then Assumption B(a) to (f) follow from Lemma 4.1.

By Lemma 4.2 we know that conditions (S1) and (S2) are satisfied. Assumption T''(a), (d) and (c) imply Assumption T (a), (c) and (d) respectively. Assumption T(b) is implied by the construction of the first return map. Lemma 4.2(iii) gives (1.7). So all conditions for Theorem B are satisfied. Hence we obtain conditions (S3) and (S4). The fact \( \|R_n\| = O(s^n) \) follows from Lemma 4.3. \( \square \)

5 Systems on multidimensional spaces: the role of the determinant in getting an optimal bound

In this section we put additional conditions on the map \( T \) that we studied in the previous chapter in order to get optimal estimates for the decay of correlations. As we anticipated in the introduction if \( \|R_n\| \) decreases, in some Banach space with norm \( \|\cdot\| \), as \( |\det DT^{-n}| \), then it usually has the same order as \( \mu(\tau = n) \), which approaches to 0 faster than \( \mu(\tau > n) \) does. Since \( \sum_{k\geq n} \mu(\tau > k) \) gives the optimal decay rates of correlations and \( \sum_{k\geq n} \|R_k\| \) determines the order of the error terms \( F_T(n) \), we can get both upper and lower estimates for decay rates.

5.1 Assumptions and statement of the results.

Let us suppose \( T \) satisfies Assumption T''(a), (d) and (c) in the last section. We replace part (b) and (c) by the following

Assumption T''. (b') (Fixed point and a neighborhood) There is a fixed point \( p \in U_1 \) and a neighborhood \( V \) of \( p \) such that \( T^{-n}(V) \cap \partial U_j = \emptyset \) for any \( j = 1, \ldots, K \) and for any \( n \geq 0 \).

(c') (Topological exactness) \( T : X \to X \) is topologically exact, that is, for any \( x \in X, \varepsilon > 0 \), there is an \( \tilde{N} = \tilde{N}(x, \varepsilon) > 0 \) such that \( T^{\tilde{N}}B_x(x) = X \).
Remark 5.1. We will introduce in Example 5.5 below class of maps in one and higher dimensions satisfying Assumption $T''(b')$, and which do not possess necessarily a Markov partition or a finite range structure (FRS). Whenever the latter conditions are satisfied, that assumption is in fact easy to verify as long as $p$ is not on the boundary of the elements of the partition in the Markov case and not on the boundary of the images in the FRS case.\footnote{Consider a partition $\mathcal{U} = \{U_i\}_i$ into domains of local injectivity of the map $T$ that have piecewise smooth boundary, and take the join $\mathcal{U}^{(n)} := \bigvee_{i=0}^{n} T^{-i} \mathcal{U}$. Let $U_{j_1, \ldots, j_n}$ the set $U_{j_1} \cap T^{-1} U_{j_2} \cap \cdots \cap T^{-(n-1)} U_{j_n} \in \mathcal{U}^{(n)}$, and call it a cylinder of rank $n$. Then the FRS says that there exists a collection of finite number of subsets $\{W_0, W_1, \ldots, W_N\}$ of $X$ of positive measure such that the $T^n$ images of any cylinder of rank $n$ for all $n > 0$ is equal to some $W_i$ in the collection. Suppose that our indifferent fixed point $p$ is not on the boundary of those subsets. Then there exists a neighborhood $V$ of $p$ disjoint with the boundary of those subsets. Now for all $n > 0$, the FRS property and the choice of $V$ give $T^n(\partial U_{j_1, \ldots, j_n}) \cap V = \emptyset$, and hence $\partial(U_{j_1, \ldots, j_n}) \cap T^{-n}(V) = \emptyset$. But the boundary of the initial partition $\mathcal{U}$ must be contained in the boundary of some $U_{j_1, \ldots, j_n}$, which implies the assumption.}

Remark 5.2. Assumption $T''(b')$ will allow us to get a better estimate for $\|R_{ij}\|_B$ which in turn will give us optimal bounds. To understand the difference with the results of Section 4, we recall that there, starting from (4.23), we got the estimate in (4.24) $|R_{ij}f|_B \leq C_R s_{ij}^2 \|f\|_B$ for some constant $C_R > 0$, and hence $\|R_{ij}f\|_B$ decreases as the speed of $s_{ij}^2$ does. This was precisely the statement of Lemma 4.3, where $s_{ij}$ was given by the norm $\|D^2 T_{ij}^{-1}\|$ of the derivatives. With Assumption $T''(b')$ and by considering a different and smaller Banach space we can get the new estimate (5.10), which lead to the upper bound $|R_{ij}f|_Q \leq C_2 \|d_{ij}\|_B \|f\|_B$ in (5.12), where $d_{ij}$ is given by the determinant $|\det D^2 T_{ij}^{-1}|$. On the other hand, estimates of the norm $|R_{ij}f|_H$ can be obtained and decreases with the same order. Other explications and details will be given in the proof.

Since we want to reserve the symbol $B$ for the function space upon which we want to get the renewal type results leading to the bounds of the decay of correlations, we begin to rename the seminorm and the Banach space defined in (4.6) and (4.7) with $Q$, instead of $B$, since it will be changed in a moment. We remind that such a seminorm will depend on $\alpha$ and on $\epsilon_0$, the latter dependence affecting only the value of the seminorms. Then (4.7) will be now written as:

$$\|f\|_Q = \|f\|_{L^1(\mu)} + |f|_Q.$$

Recall that $V$ is a neighborhood of $p$ given in Assumption $T''(b')$. We denote the preimages $T^{i_k} \cdots T^{i_1} V$ by $V_{i_k \cdots i_1}$. We also denote with $T$ the set of all possible words $i_1 \cdots i_k$ such that $T_{i_k} \cdots T_{i_1} V$ is well defined, where $i_k \in \{1, \ldots, K\}$ and $k > 0$.

For an open set $O$, let $\mathcal{H} := \mathcal{H} \alpha_{\epsilon_1} = \mathcal{H} \alpha_\alpha (O, H)$ be the set of Hölder functions $f$ on $O$ that satisfies $|f(x) - f(y)| \leq H d(x, y)^\alpha$ for any $x, y \in O$ with $d(x, y) \leq \epsilon_1$.

Let $\hat{h}$ be a fixed point of the transfer operator $\hat{T}$, which will be unique under...
the assumptions of the theorem below. We now define \( \mathcal{B} \) by
\[
\mathcal{B} := \mathcal{B}^0_{\epsilon_0, \epsilon_1} = \left\{ f \in \mathcal{Q} : \exists H > 0 \text{ s.t. } (f/\hat{h})|_{V_I} \in \mathcal{H}^0_{\epsilon_1}(V_I, H) \forall I \in \hat{T} \right\}, \tag{5.1}
\]
and for any \( f \in \mathcal{B} \), let
\[
|f|_\mathcal{B} := |f|_{\mathcal{H}^0_{\epsilon_1}} = \inf \{ H : (f/\hat{h})|_{V_I} \in \mathcal{H}^0_{\epsilon_1}(V_I, H) \forall I \in \hat{T} \}.
\]
Sublemmas 5.3 and 5.4 below imply that \( \hat{h} > 0 \) on all \( V_I \), and therefore the definition makes sense. Then we take \( |\cdot|_\mathcal{Q} + |\cdot|_\mathcal{H} \) as a seminorm for \( f \in \mathcal{B} \) and define the norm in \( \mathcal{B} \) by
\[
\| \cdot \|_\mathcal{B} = \| \cdot \|_1 + |\cdot|_\mathcal{Q} + |\cdot|_\mathcal{H}. \tag{5.2}
\]
Clearly, \( \mathcal{B} \subseteq \mathcal{Q} \) and \( \|f\|_\mathcal{B} \geq \|f\|_\mathcal{Q} \) if \( f \in \mathcal{B} \).

Recall that for any sequences of numbers \( \{a_n\} \) and \( \{b_n\} \), we use \( a_n \approx b_n \) if \( \lim_{n \to \infty} \frac{a_n}{b_n} = 1 \), and \( a_n \sim b_n \) if \( c_1 b_n \leq a_n \leq c_2 b_n \) for some constants \( c_2 \geq c_1 > 0 \).

Let \( d_{ij} = \sup \{ |\det D\tilde{T}_{ij}^{-1}(x)| : x \in B_{\epsilon_0}(Q_0) \} \), and \( d_n = \max \{ d_{n-1,j} : j = 2, \cdots, K \} \).

**Theorem E.** Let \( \tilde{X}, \tilde{T} \) and \( \mathcal{B} \) be defined as above. Suppose \( T \) satisfies Assumption \( T''(a), (b'), (c'), (d) \) and \( (e) \). Then there exist \( \epsilon_0 \geq \epsilon_1 > 0 \) such that Assumption \( B(a) \) to \( (f) \) and conditions \( S(1) \) to \( S(4) \) are satisfied and \( \|R_n\|_\mathcal{B} = O(d_n^{\alpha/(m+\alpha)}) \). Hence, if \( \sum_{k=n+1}^\infty d_n^{m/(m+\alpha)} = O(n^{-\beta}) \) for some \( \beta > 1 \), then there exists \( C > 0 \) such that for any functions \( f \in \mathcal{B}, g \in L^\infty(X, \nu) \) with supp \( f \), supp \( g \subset \tilde{X}, (1.5) \) holds.

Moreover, if \( T \) satisfies \( (4.4) \) near \( p = 0 \), then
\[
\sum_{k=n+1}^\infty \mu(\tau > k) \sim n^{-\left(\frac{n}{n-1}\right)}.
\]
In this case, if \( d_n = O(n^{-\beta'}) \) for some \( \beta' > 1 \) and if
\[
\beta = \beta' \cdot \frac{m}{m+\alpha} - 1 > \max\{2, \frac{m}{\gamma} - 1\}, \tag{5.3}
\]
then
\[
\text{Cov}(f, g \circ T^n) \approx \sum_{k=n+1}^\infty \mu(\tau > k) \int f \, d\mu \int g \, d\mu \sim \frac{1}{n^{\frac{n}{n-1}}}. \tag{5.4}
\]
In particular, if Assumption \( T''(e') \) in Section 4.1 stating bounded distortion also holds, then the above statements remain true if we replace \( m/(m+\alpha) \) in (5.3) by 1.

**Remark 5.3.** For the case that \( T \) satisfies \( (4.4) \) near \( p \), Assumption \( T''(e') \) implies \( h \) is bounded away from 0 on the sets \( \{ \tau > n \} \). Hence \( \mu(\tau > n) \) and \( \nu(\tau > n) \) have the same order, and \( \sum_{k=n+1}^\infty \mu(\tau > k) \sim n^{-\left(\frac{n}{n-1}\right)} \). This is the case in Example 5.1, 5.2 and 5.4 below.
On the other hand, if Assumption $T''(c')$ only holds for an invariant subset of $X$ like in Example 7.3, then $h$ may be only supported on part of the sets $\{\tau > n\}$, and therefore $\mu(\tau > n)$ may decrease faster. In this case, $\sum_{k=n+1}^{\infty} \mu(\tau > k) = o(n^{-(\frac{m}{2} - 1)})$.

### 5.2 Examples

Before giving the proof, we present a few examples. The first four examples are concerning various decay rates, where we will always assume that $T$ satisfies Assumption $T''(a)$, $(b')$, $(c')$ and $(d)$. Example 5.5 and thereinafter are for maps satisfying Assumption $T''(b')$.

**Example 5.1.** Assume $m = 3$, and near the fixed point $p = (0, 0, 0)$, the map $T$ has the form

$$T(w) = (x(1 + |w|^2 + O(|w|^3)), y(1 + |w|^2 + O(|w|^3)), z(1 + 2|w|^2 + O(|w|^3)))$$

where $w = (x, y, z)$ and $|w| = \sqrt{x^2 + y^2 + z^2}$.

This map is very similar to that in Example 1 in [HV], although it is now in a three dimensional space. We can still use the same arguments to show that Assumption $T''(c)$ is satisfied.

Denote $w_n = T^{-n}w$; clearly, $|w| + |w|^3 + O(|w|^4) \leq |T(w)| \leq |w| + 2|w|^3 + O(|w|^4)$. By standard arguments we know that

$$\frac{1}{\sqrt{4n}} + O\left(\frac{1}{\sqrt{n^2}}\right) \leq |w_n| \leq \frac{1}{\sqrt{2n}} + O\left(\frac{1}{\sqrt{n^3}}\right)$$

(see also Lemma 3.1 in [HV]). Since we are in a three dimensional space, we now have $\nu(\tau > k) \sim \frac{1}{k^{3/2}}$, and therefore $\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{1/2}}$.

It is easy to see that $\det DT(w) = 1 + 6x^2 + 6y^2 + 8z^2 + O(|w|^3)$. So we have $|\det DT^{-1}(w)| \leq 1 - 6|w|^2 + O(|w|^3)$. By Lemma 3.2 in [HV] with $r(t) = 1 - 6t^2 + O(t^3)$, $\gamma = 2$, $C' = 6$ and $C = 1$, we get that $|\det DT^{-1}(x)| = O(1/n^3)$. Hence we have $\beta' = 3$ and $\beta = 3m/(m + \alpha) - 1 > 3 \cdot 3/(3 + 1) - 1 = 5/4$. Since $m/\gamma - 1 = 1/2$, $(5.3)$ holds, and therefore we have (5.4) with the decay rate of order $1/n^2$; contrarily to Example 4.1, we now got an optimal bound.

**Example 5.2.** Assume $m = 2$, and near the fixed point $p = (0, 0)$, the map $T$ has the form

$$T(z) = (x(1 + |z|^\gamma + O(|z|^{\gamma'})), y(1 + 2|z|^\gamma + O(|z|^{\gamma'})))$$

where $z = (x, y)$, $|z| = \sqrt{x^2 + y^2}$, $\gamma \in (0, 1)$ and $\gamma' > \gamma$.
By methods similar to Example 1 in [HV] we can check that Assumption $T''(e)$ is satisfied. Denote $z_n = T_{1-n} z$. Since $|z| + |z|^{1+\gamma} + O(|z|^\gamma) \leq |T(z)| \leq |z| + 2|z|^\gamma + O(|z|^\gamma)$, we have
\[
\frac{1}{(2\gamma n)^{1/\gamma}} + O\left(\frac{1}{n^3}\right) \leq |z_n| \leq \frac{1}{(\gamma n)^{1/\gamma}} + O\left(\frac{1}{n^\delta}\right)
\]
for some $\delta > 1/\gamma$. So $\nu(\tau > k) \sim \frac{1}{k^{2/\gamma}}$, and therefore $\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{\delta-1}}$.

It is possible to show that $|\det DT(z)| = 1 + \frac{(3 + \gamma) x^2 + (3 + 2 \gamma) y^2}{|z|^{2-\gamma}} + O(|z|^{\gamma})$. Therefore $|\det DT_{1-n}^{-1}(z)| \leq 1 - (3 + \gamma)|z|^{\gamma} + O(|z|^\gamma)$, and $|\det DT_{1-n}^{-1}(z)| = O(1/n^{1+3/\gamma})$. Hence $\beta' = 1 + \gamma/3$ and $\beta = (1 + 3/\gamma) \cdot 2/(2 + \alpha) - 1 > (1 + 3/\gamma) \cdot 2/3 - 1 = 2/\gamma - 1/3 > 2/\gamma - 1$. It means (5.3) holds, and the decay rates is of order $1/n^{\delta-1}$.

**Example 5.3.** Assume $m = 2$, and take the same map as in Example 1 in [HV], namely, near the fixed point $p = (0,0)$, the map $T$ has the form
\[
T(x, y) = (x(1 + x^2 + y^2), y(1 + x^2 + y^2)^2).
\]

The map allows an infinite absolutely continuous invariant measure. However, the map can be arranged in such a way that there is an invariant component that supports a finite absolutely continuous invariant measure $\mu$. Near the fixed point, the region of this component has the form
\[
\{z = (x, y) : |y| < x^2\}.
\]

We may regard $X$ as this component, and $T : X \to X$ satisfies the assumptions.

We can check that the map has bounded distortion near the fixed point restricted to this region. Hence, the map satisfies Assumption $T''(e')$.

Since $|z_n| = O(1/\sqrt{n})$ and for $z = (x, y)$, $|y| \leq x^2$, we get $\nu(\tau > k) \sim \frac{1}{k^{3/2}}$, and $\sum_{k=n+1}^{\infty} \nu(\tau > k) \sim \frac{1}{n^{1/2}}$.

On the other hand, $|\det DT(z)| = 1 + 5x^2 + 7y^2 + O(|z|^4)$. Since $|y| \leq x^2, |z| = |x| + O(|z|^2)$; thus $|\det DT(z)| = 1 + 5|z|^2 + O(|z|^4)$, and therefore $|\det DT_{1-n}^{-1}(z)| = O(1/n^{5/2})$. So $\beta' = 5/2$ and $\beta = 3/2$. We obtain that the decay rate is of order $1/n^{1/2}$.

**Example 5.4.** Assume $m \geq 3$ and near the fixed point $p = (0,0,0)$, the map $T$ has the form
\[
T(z) = z(1 + |z|^\gamma + O(|z|^\gamma+1)),
\]
where $m > \gamma > 0$. 

37
These examples are comparable with those in Example 4.1, except for the stronger topological assumptions which we now put on the maps. We know that these maps satisfy Assumption, $T''(e')$.

Denote $z_n = T^{-n}_1 z$. We have $|z_n| = 1/(n\gamma)^{1/\gamma} + O(1/(n\gamma)^{1+1})$ and $|\det DT(z)| = 1+(m+\gamma)|z|^{-\gamma} + O(|z|^{-\gamma+1})$. Hence, we get that $|\det DT^{-n}_1| \sim 1/n^{m+1}$. (For the relative computations see Lemma 3.1 and 3.2 in [HV]). Therefore $\beta' = \frac{m}{\gamma} + 1$ and $\beta = m/\gamma$.

On the other hand, we see that $\nu(\tau > k) = O(1/k^{m/\gamma})$, and then $\sum_{k=n+1}^\infty \nu(\tau > k) \sim \frac{1}{n^{m/\gamma}-1}$. Since $m > \gamma$, the invariant measure $\mu$ is finite and $\beta > 1$. We get that the decay rate is of order $1/n^{m/\gamma}-1$.

Now we construct examples

**Example 5.5.** Take $X = [-100, 100]$. Take a partition $\xi = \{U_0, U_i^+, U_i^- : i = 1, \ldots, 9\}$ of $X$ into 19 subintervals such that $U_0 = [-10,10]$, $U_i^+ = [-10i - 10, -10i]$ and $U_i^- = [10i, 10i + 10]$. Denote $\partial \xi = \cup_{U \in \xi} \partial U$.

Define a piecewise smooth expanding map $T : X \rightarrow X$ with an indifferent fixed point $p = 0$ as following:

(i) $T(\text{int} U_i^\pm) = \text{int} X$ for $i \neq -8, 8$ and $|T'_i(x)| \geq 10$ for all $x \notin [-3,3] \cup \partial X$;

(ii) $T(x) = x + 4|x|^{1.5}$ for $x \in [-3,3]$;

(iii) $T$ is increasing on $U_i^\pm$ and maps $\text{int} U_i^\pm$ to $\text{int} X$ linearly, that is, $T(x) = 20(x - 95)$ on $U_i^+$ and $T(x) = 20(x + 95)$ on $U_i^-$;

(iv) $T(U_8^-) = [-100, e_{\pm})$ and $T(U_8^+) = (e_{-}, 100]$, where $e_\pm \in E_{\pm}$, and $E_{\pm} = \{x \in U_0^\pm : T^n(x) \in U_0^+ \cup U_0^- \forall n \geq 0\}$.

It is clear that $T$ satisfies Assumption $T''(a)$, (b), (c'), (d) and (e'). By the choice of $E_{\pm}$, the orbits $\{T^n(e_{\pm}) : n > 0\}$ are contained in $E_{\pm} \cap E_{\pm}$, and therefore in $U_i^0 \cup U_0^\pm$. Note that all possible image sets $\{T^n(U) : U \in \cup_{i=0}^{\infty} T^{-i}(\xi)\}$ have the form $[-100,100]$, $[-100, T^n(e_{\pm}))$, $[T^n(e_{\pm}), 100]$ or $[T^n(e_{\pm}), T^n(e_{\pm})]$ up to the endpoints. So if we take $V = [-2,2]$, then $V \cap T^k(\partial U) = \emptyset$ for any $U \in \xi$ and $k \geq 0$. It follows that $T^{-k}(V) \cap \partial U = \emptyset$ for any $U \in \xi$ and $k \geq 0$. Hence, Assumption $T''(b')$ holds.

Note that $E_{+} \cup E_{-}$ is a Cantor set and therefore both sets $E_{+}$ and $E_{-}$ contain uncountably many real numbers. If we choose irrational numbers as $e_{+}$ and $e_{-}$, then both orbits $\{T^n(e_{+}) : n \geq 0\}$ and $\{T^n(e_{-}) : n \geq 0\}$ contains infinitely many points. Since for each $n$, there must be an element $U \in \cup_{i=0}^{n} T^{-i}(\xi)$ such that the end points of $T^n(U)$ contains $T^n(e)$, $e = e_{+}$ or $e_{-}$, then the set of images $\{T^n(U) : U \in \cup_{i=0}^{n} T^{-i}(\xi), n \geq 0\}$ is infinite. It means that the $T$ has neither finite Markov partition, nor finite range structure.
Remark 5.4. We mention here that $T|_{U^\pm}$ do not have to be linear. Also, the role of $U^\pm_i$ and $U^\pm_j$ can be replaced by any pairs $U^\pm_i$ and $U^\pm_j$ for $i,j \neq 0$ and $i \neq j$.

The same idea can be used to generate example of maps in higher dimensional spaces. For example, on the plane, we can take $X = [-100,100] \times [-100,100]$, and partition $X$ in to squares $U^\pm_{ij}$ of size $10 \times 10$, except for $U_0 = [-10,10] \times [-10,10]$. Near the origin we can define $T(x,y) = (x(1 + x^2 + y^2), y(1 + x^2 + y^2)^2)$ as in Example 5.3. Then we let $U^\pm_i$ and $U^\pm_j$, or $U^\pm_{ij}$, and for any connected component $i; j = \pm 0, \pm 1, \cdots, \pm 9$, to play the same role as $U^\pm_9$ and $U^\pm_8$ in the above example. That is, the map can be arranged in a way that under $T^n$, the images of boundaries of all sets in the partition are contained in the region $\{(x,y) \in X : 90 \leq |y| \leq 100\}$ or $\{(x,y) \in X : 90 \leq |x| \leq 100\}$, or both. By this way, we can construct a map $T$ that satisfies all conditions given by Assumption $T''(a), (b'), (c'), (d)$ and (e).

In fact, systems satisfying Assumption $T''(a), (b'), (c')$, and (d) are dense in the set of the systems satisfying Assumption $T''(a), (b), (c')$ and (d) in the $C^1$ topology. That is, for any system satisfying Assumption $T''(a), (b), (c')$, and (d), we can make an arbitrarily small $C^1$ perturbation to get a map $\bar{T}$ such that there exists a small neighborhood $V$ of $p$ with $\bar{T}^{-n}(V) \cap \partial U_j = \emptyset$ for any $j = 1, \ldots, K$ and for any $n \geq 0$. To see this, we first note that for any fixed $n_0$, we can get that $\bar{T}^{-n_0}(p) \cap \partial U_j = \emptyset$ for any $0 < n \leq n_0$ by using small perturbation, and then get that $\bar{T}^{-n}V \cap \partial U_j = \emptyset$ for any $0 < n \leq n_0$ by taking $V$ small enough. Further, for any connected component $V^{(n)}_i$ of $\bar{T}^{-n}V$, we require that $d(V^{(n)}_i, \partial U_j) \geq \text{diam } V^{(n)}_i$ for any $j = 1, \ldots, K$. Now we consider the case $n > n_0$. If $V^{(n)}_i \cap \partial U_j \neq \emptyset$, then we can use a small perturbation $\phi^{(n)}_i$ with both $d(\phi^{(n)}_i, \text{id})$ and $\|D\phi^{(n)}_i\|$ small such that $d(V^{(n)}_i, \partial U_j) \geq \text{diam } V^{(n)}_i$. Note that Assumption $T''(d)$ implies $s < 1/4$. It is easy to see that if $V^{(n_2)}_{i_2}$ intersects $(2 \text{diam } V^{(n_2)}_{i_1})$-neighborhood of some $V^{(n_2)}_{i_1}$ with $n_2 > n_1$, then $d(V^{(n_2)}_{i_2}, \partial U_j) < (1/4) \text{diam } V^{(n_2)}_{i_2}$. Hence, we can require $d(\phi^{(n)}_i, \text{id})$ and $\|D\phi^{(n)}_i\|$ decrease with $n$ at least by a fact 1/4 at each step. Then after a sequence of perturbations we still have $d(V^{(n)}_i, \partial U_j) \geq (1/2) \text{diam } V^{(n)}_i$ for any $n > 0$ and the $C^1$ norm of the composition of the sequence of perturbations are still small. Hence the resulting map $\bar{T}$ satisfies Assumption $T''(b')$, and obviously satisfies Assumption $T''(a), (c')$, and (d) as well. We leave the details to the reader.

5.3 Proof of Theorem E

Proof of Theorem E. Take $\varepsilon_0 > 0$ satisfying Lemma 4.2 in the previous section, and then choose $\varepsilon_1 \in (0, \varepsilon_0]$ as in Lemma 5.2 below. We reduce $\varepsilon_1$ further if necessary such that $\eta' := \eta + D_H(\varepsilon_0)\eta_0^2 < 1$, where $\eta < 1$ is given in Lemma 4.2.
and $D_{\beta}(\varepsilon_0) > 0$ is given in Lemma 5.2. Then we take $B := B_{\varepsilon_0, \varepsilon_1}$, as in (5.1). With the norm given in (5.2), $B$ satisfies Assumption B(a) to (f) by Lemma 5.1.

By Lemma 4.2 and 5.2, condition $S(1)$ is satisfied with constants $\eta$ and $D$ replaced by $\eta'$ defined as above and $D + D_{\beta}(\varepsilon_0)\varepsilon_0^2$ respectively, where $D$ is the number given in Lemma 4.2. Condition $S(2)$ can be obtained in a similar way. Assumption $T''(a)$, (d) and (e') imply Assumption $T'$ (a), (c) and (d) respectively. Assumption $T'(b)$ follows from the construction of the first return map. Lemma 4.2(iii) and 5.2(iii) give (1.7). Therefore all the conditions for Theorem B are satisfied. Hence we obtain conditions $S(3)$ and $S(4)$.

The facts $\|R_n\|_B = O(d_n^{m/(m+\alpha)})$, and $\|R_n\|_B = O(d_n)$ if Assumption $T''(e')$ is satisfied, follow from Lemma 5.5. Therefore we have established the decay of correlations (1.5).

If $T$ also satisfies (4.4), then we know that for any $z$ close to $p$, $|T_t^{-n}z|$ is of order $n^{-\gamma}$. Hence $\tilde{\nu}\{\tau > k\}$ has the order $k^{-m/\gamma}$, and $\sum_{k=n+1}^{\infty} k^{-\frac{m}{\gamma}} = O(n^{-\frac{m}{\gamma}+1})$. Then the rest of the theorem is clear. \hfill $\square$

**Lemma 5.1.** $B$ is a Banach space satisfying Assumption B(a) to (f) with $C_a = 2\gamma_m^{-1}\gamma_0^{-m+\alpha}$, where $\gamma_m$ is the volume of the unit ball in $R^n$.

**Proof.** We already know that $Q$ is a Banach space, and the proof of the completeness of $B$ follows from standard arguments. So $B$ is a Banach space.

Now we verify Assumption B(a) to (f).

By Lemma 4.1, the unit ball of $Q$ is compact in $L^1(\tilde{X}, \tilde{\nu})$. Since $\|f\|_B \geq \|f\|_Q$ for any $f \in B \subset Q$, the unit ball of $B$ is contained in the unit ball of $Q$. Since $B$ is closed in $Q$, the unit ball of $B$ is also compact. This is Assumption B(a).

Moreover, for any $f \in Q$, $\|f\|_{L^\infty} \leq C_b \|f\|_Q \leq C_b \|f\|_B$ with $C_b = \gamma_m^{-1}\gamma_0^{-m+\alpha}$. We have thus got Assumption B(b).

Invoking again Lemma 4.1, we have for any $f, g \in Q$, $\|fg\|_Q \leq C_t \|f\|_Q \|g\|_Q$, where $C_t = 2\gamma_m^{-1}\gamma_0^{-m+\alpha} = 2C_b$. It is easy to check that

$$\|fg\|_B = \|fg\|_Q + \|fg\|_H \leq C_b \|f\|_Q \|g\|_Q \|g\|_H + C_b \|g\|_Q \|f\|_H \leq C_b \|f\|_Q \|g\|_H + C_b \|g\|_Q \|f\|_H.$$

Hence,

$$\|fg\|_B \leq \|f\|_Q \|g\|_H + \|g\|_Q \|f\|_H \leq C_b \|f\|_Q \|g\|_H + C_b \|g\|_Q \|f\|_H \leq C_t \|f\|_Q \|g\|_Q \|g\|_H + C_b \|g\|_Q \|f\|_H.$$

Therefore Assumption B(c) follows with $C_a = 2\gamma_m^{-1}\gamma_0^{-m+\alpha} = 2C_b$.

Similarly, part (d) of Assumption B follows from the fact that $B$ contains all Hölder functions, and Hölder functions are dense in $L^1(\tilde{X}, \tilde{\nu})$.

Assume $f(x) = \lim_{n \to \infty} f_n(x)$ for $\tilde{\nu}$-a.e. $x \in \tilde{X}$. By the proof of Lemma 4.1 we have $|f|_Q \leq \liminf_{n \to \infty} |f_n|_Q$. For any $y, z \in V_r$, where $I \in \mathcal{I},$

$$\frac{|f(y) - f(z)|}{d(y, z)^\alpha} \leq \lim_{n \to \infty} \frac{|f_n(y) - f_n(z)|}{d(y, z)^\alpha} \leq \liminf_{n \to \infty} |f_n|_H.$$
It gives \(|f|_{\mathcal{H}} \leq \liminf_{n \to \infty} |f_n|_{\mathcal{H}}\). Since \(|f|_{\mathcal{B}} = |f|_{\mathcal{Q}} + |f|_{\mathcal{H}}\), we get part (e).

Since \(\mathcal{B} \subset \mathcal{Q}\), part (f) is directly from the fact that \(\mathcal{Q}\) satisfies Assumption B(f). 

\[\square\]

**Lemma 5.2.** Let \(\varepsilon_0\) be as in Lemma 4.2. There exists \(D_{\mathcal{H}} = D_{\mathcal{H}}(\varepsilon_0), D_{\mathcal{H}} = D_{\mathcal{H}}(\varepsilon_0) > 0\) and \(\varepsilon_- \in (0, \varepsilon_0]\) such that for any \(\varepsilon \in (0, \varepsilon_-]\), and by using the notation for the Banach space introduced in (5.1):

(i) for any \(f \in B_{\mathcal{H}_0, \varepsilon_1}^\alpha\), \(\overline{\mathcal{P}} f|_{\mathcal{H}_1} \leq s^\alpha |f|_{\mathcal{H}_1} + D_{\mathcal{H}} \varepsilon_1^\alpha \|f\|_{\mathcal{Q}_0}\);

(ii) for any \(f \in B_{\mathcal{H}_0, \varepsilon_1}^\alpha\), \(|R(z)|f|_{\mathcal{H}_1} \leq |z|(|s^\alpha |f|_{\mathcal{H}_1} + D_{\mathcal{H}} \varepsilon_1^\alpha \|f\|_{\mathcal{Q}_0})|;

(iii) and for any \(f \in B_{\mathcal{H}_0, \varepsilon_1}^\alpha\), \(\overline{\mathcal{P}} f|_{\mathcal{H}_1} \leq s^\alpha \|f\|_{\mathcal{H}_1} + D_{\mathcal{H}} \varepsilon_1^\alpha \|f\|_{\mathcal{Q}_0}\).

**Proof.** (i) Let \(\varepsilon \in (0, \varepsilon_0]\), \(J_{\mathcal{H}} > 0\) as in the proof of Sublemma 5.4 below. Suppose \(\varepsilon \in (0, \varepsilon_*]\), and \(|f|_{\mathcal{H}_1} = H\) for some \(f\). Take \(x, y \in V_I\) for some \(I \in I\) with \(d(x, y) = \varepsilon \leq \varepsilon_*\). Then by Assumption T''(e), we can take \(\zeta > 0\), \(N = N(\varepsilon) > 0\) for \(b = 1\). Note that

\[
\frac{\overline{\mathcal{P}} f(x)}{h(x)} - \frac{\overline{\mathcal{P}} f(y)}{h(y)} = \sum_{j=1}^{J} \sum_{i=1}^{\infty} \frac{\hat{g}(x_{ij})}{h(x_{ij})} \left(\frac{f(x_{ij})}{h(x_{ij})} - \frac{f(y_{ij})}{h(y_{ij})}\right) + \frac{f(y_{ij})}{h(y_{ij})} \left(\frac{\hat{g}(x_{ij})}{h(x_{ij})} - \frac{\hat{g}(y_{ij})}{h(y_{ij})}\right)
\]

\[
+ \sum_{j=1}^{J} \sum_{i=1}^{\infty} \frac{f(y_{ij})}{h(y_{ij})} \left(\frac{\hat{g}(x_{ij})}{h(x_{ij})} - \frac{\hat{g}(y_{ij})}{h(y_{ij})}\right)
\]

\[
= \sum_{j=1}^{J} \sum_{i=1}^{\infty} \frac{\hat{g}(x_{ij})}{h(x_{ij})} \left(\frac{f(x_{ij})}{h(x_{ij})} - \frac{f(y_{ij})}{h(y_{ij})}\right) + \frac{f(y_{ij})}{h(y_{ij})} \left(\frac{\hat{g}(x_{ij})}{h(x_{ij})} - \frac{\hat{g}(y_{ij})}{h(y_{ij})}\right).\]

Since \(|f|_{\mathcal{H}} = H\), we have \(|f(x_{ij})/h(x_{ij}) - f(y_{ij})/h(y_{ij})| \leq H d(x_{ij}, y_{ij})^{\alpha} \leq s^\alpha H d(x_{ij}, y_{ij})^{\alpha}\). Now, \(\overline{\mathcal{P}} h = h\) implies

\[
\sum_{j=1}^{J} \sum_{i=1}^{\infty} \frac{\hat{g}(x_{ij})}{h(x_{ij})} = 1,
\]

Thus the first sum in (5.5) is bounded by \(s^\alpha H d(x, y)\) by Assumption T''(e), \(\hat{g}(y)/\hat{h}(y) \leq e^{s^\alpha H d(x, y)^{\alpha}}\), and by Assumption T''(e), \(\hat{g}(y)/\hat{h}(y) \leq e^{s^\alpha H d(x, y)^{\alpha}}\)

Note that by our assumption, \(V_I\) does not intersect discontinuities.** By Sublemma 5.4, \(h(y)/h(x) \leq e^{s^\alpha H d(x, y)^{\alpha}}\), and by Assumption T''(e), \(\hat{g}(y)/\hat{h}(x) \leq e^{s^\alpha H d(x, y)^{\alpha}}\) if \(i \leq N(\varepsilon)\). So \([\hat{g}(y_{ij})/\hat{h}(y_{ij})/\hat{h}(y_{ij})/\hat{h}(x_{ij})/\hat{h}(x_{ij})] \leq e^{s^\alpha H d(x, y)^{\alpha}}\)

**This implies that the potential \(\hat{g}_{ij}\) of the transfer operator is continuous. Such a potential has in fact the form \(\hat{g}_{ij}(x) = |\det DT_{ij}(x)|^{-1}\), where \(T_{ij} = T_j T_i\), being \(T_i\) and \(T_j\) different determinations of the map \(T\). In the computation of the transfer operator, \(\hat{g}\) is computed in the point \(T_{ij}^{-1}x\), where \(x\) belongs to the sets of Hölder continuity \(V_I\) which are in turn the preimages of \(V\). The continuity of the potential is necessary to get the invariance of the new Banach space under the action of \(\overline{\mathcal{P}}\).
some $\zeta' > 0$. We take $\varepsilon_\cdot \in (0, \varepsilon_\cdot]$ small enough such that $e^{\varepsilon_\cdot} - 1 \leq 2\zeta' e^\varepsilon$ for any $\varepsilon_1 \leq (0, \varepsilon_-]$. Then for $d(x, y) = \varepsilon \leq \varepsilon_1$, we have

$$
\left| \frac{\hat{g}(x_{ij})\hat{h}(y_{ij})}{h(x)} - \frac{\hat{g}(y_{ij})\hat{h}(x_{ij})}{h(y)} \right| \leq 2\zeta' \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{h(x)} \cdot d(x, y)^\alpha.
$$

(5.7)

Therefore by (5.6), the second sum in (5.5) is bounded by

$$
\sum_{j=1}^K \sum_{i=1}^N f(y_{ij}) \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{h(x)} \cdot 2\zeta' d(x, y)^\alpha \leq 2\zeta' h_*^{-1} \|f\|_{\infty} d(x, y)^\alpha,
$$

where $h_*$ is the essential lower bound of $\hat{h}$ given by Sublemma 5.3.

By Assumption $T''(c')$, the third sum in (5.5) is bounded by

$$
\sum_{j=1}^K \sum_{i=1}^{\infty} f(y_{ij}) \frac{\hat{g}(x_{ij})\hat{h}(x_{ij})}{h(x)} \leq \hat{h}_-^{-2} \|\hat{h}\|_{\infty} \|f\|_{\infty} \cdot K' b' e^m + a
$$

$$
= \hat{h}_-^{-2} \|\hat{h}\|_{\infty} C_0 \|g\| \cdot K' b' e^m d(x, y)^\alpha = C_b K' b' e^m \hat{h}_-^{-2} \|\hat{h}\|_{\infty} \|f\|_{\infty} d(x, y)^\alpha,
$$

where $C_0$ is given in Lemma 4.1 which depends on $\varepsilon_0$.

Hence the result of part (1) holds with $D_R = C_b \hat{h}_-^{-1} (2\zeta' + K' b' e^m \hat{h}_-^{-1} \|\hat{h}\|_{\infty})$.

Part (ii) and (iii) can be proved by using the same estimates with the same adjustments as in the proof of Lemma 4.2. \hfill \square

**Sublemma 5.3.** There is a $\hat{h}_* > 0$ such that $\hat{h}(x) \geq \hat{h}_*$ for $\nu$-a.e. $x \in \tilde{X}$.

**Proof.** By Lemma 3.1 in [Ss], there is a ball $B_{\varepsilon}(z) \subset \tilde{X}$ such that $\text{Einf}_{\hat{h}} \hat{h} \geq \hat{h}_-$ for some constant $\hat{h}_- > 0$. By Assumption $T''(c')$, there is $\hat{N} > 0$ such that $T^{\hat{N}} B_{\varepsilon}(z) \supset X$. Then for any $x \in \tilde{X}$, there is $y_0 \in B_{\varepsilon}(z)$ such that $T^{\hat{N}} y_0 = x$.

Since $|\text{det } DT|$ is bounded above, we have $g_* := \inf \{g(y) : y \in X\} > 0$. Hence, for $\nu$-almost every $x$,

$$
\hat{h}(x) = (\mathcal{G}^{\hat{N}} \hat{h})(x) = \sum_{T^{\hat{N}} y = x} \hat{h}(y) \prod_{i=0}^{\hat{N}-1} g(T^i y) \geq \hat{h}(y_0) \prod_{i=0}^{\hat{N}-1} g(T^i y_0) \geq \hat{h}_- \hat{g}_{\hat{N}}^\circ.
$$

The result follows with $\hat{h}_* = \hat{h}_- g_{\hat{N}}^\circ$. \hfill \square

**Sublemma 5.4.** Let $\varepsilon_0$ be as in Lemma 4.2. Then there exists $J_\varepsilon > 0$ and $\varepsilon_* \in (0, \varepsilon_0]$ such that for any $x, y \in V_I$ with $d(x, y) \leq \varepsilon_*$, $I \in \mathcal{I}$,

$$
\frac{\hat{h}(x)}{\hat{h}(y)} \leq e^{J_\varepsilon d(x, y)^\alpha}.
$$

42
Proof. Since \( \hat{h} \) is the unique fixed point of \( \bar{\mathcal{P}} \), we know that \( \hat{h} = \lim_{n \to \infty} \bar{\mathcal{P}}^n 1_{\bar{X}} \), where the convergence is in \( L^1(\bar{\mu}) \). Now we consider the sequence \( f_n := \bar{\mathcal{P}}^n 1_{\bar{X}} \).

We will prove that there is \( J_h > 0 \) and \( \varepsilon_* \in (0, \varepsilon_0] \) such that for any \( n \geq 0 \), for any \( x,y \in V_f \), \( l \in \mathbb{Z} \), with \( d(x,y) \leq \varepsilon_* \),

\[
\frac{f_n(y)}{f_n(x)} \leq e^J_h d(x,y)^\alpha.
\] (5.8)

Clearly (5.8) is true for \( n = 0 \) since \( f_0(x) = 1 \) for any \( x \). We assume that it is true up to \( f_{n-1} \). Consider \( f_n \).

Note that \( f_n/\hat{h} = (1/\hat{h}) \bar{\mathcal{P}}^n (h \cdot 1_{\bar{X}}/\hat{h}) = \hat{\mathcal{L}}^n (1_{\bar{X}}/\hat{h}) \), where \( \hat{\mathcal{L}} \) is a normalized transfer operator defined by \( \hat{\mathcal{L}}(f) = (1/\hat{h}) \bar{\mathcal{P}}(hf) \). Then there are \( f_s \geq \hat{h},/s^* \) and \( f^* \leq \hat{h},/s_* \) such that \( f_s \leq f_n(x) \leq f^* \) for every \( x \in \bar{X} \) and \( n \geq 0 \), where \( \hat{h}^*, \hat{h}_* \) are the essential upper and lower bound of \( h \) respectively. Let also set: \( g_s = \inf_x f_s(x) = \inf_x \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(x_{ij}) \).

Let us set again \( b = 1 \). Then put \( \zeta > 0 \) as in Assumption T"sm(e). Let us take \( J_h > 2\zeta s^*/(1-s^*) \) so that we have \( (J_h + \zeta) s^* \leq J_h (1 + s^*) / 2 \). Then we take \( \varepsilon_* \in (0, \varepsilon_0] \) small enough such that for any \( \varepsilon \in [0, \varepsilon_*] \),

\[
e^J_h (1 + s^*) \varepsilon /2 + f^* K' b e^{m+\alpha} = f_s (g_s - K' b e^{m+\alpha}) \leq e^J_h \varepsilon^*.
\]

For any \( x, y \) in the same \( V_f \) with \( d(x, y) =: \varepsilon \leq \varepsilon_* \), we choose \( N = N(\varepsilon) \) as in Assumption T"sm(e). Let us denote with \( [f_n](N) = \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(x_{ij}) f_{n-1}(x_{ij}) \) and \{f_n\} \( N(x) = f_n(x) - [f_n](N(x) = \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(x_{ij}) f_{n-1}(x_{ij}) \). We have

\[
\frac{[f_n](N(y))}{[f_n](N(x))} \leq \sup_{1 \leq j \leq K; 0 < i \leq N} e^{\zeta d(x_{ij}, y_{ij}) \alpha} \frac{f_s (y_{ij})}{f_s (y_{ij})} \leq e^{(\zeta + J_h) s^* d(x,y)^\alpha} \leq e^J_h (1 + s^*) \varepsilon^*/2.
\]

We also get

\[
\{f_n\} \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(y_{ij}) f_{n-1}(y_{ij}) \leq f^* \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(y_{ij}) \leq f^* K' b e^{m+\alpha}.
\]

On the other hand,

\[
[f_n](N(x)) = \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(y_{ij}) f_{n-1}(y_{ij}) \geq f_s \sum_{j=1}^K \sum_{i=0}^\infty \hat{g}(y_{ij}) \geq f_s (g_s - K' b e^{m+\alpha}).
\]

By the choice of \( \varepsilon_* \), we obtain

\[
\frac{f_n(y)}{f_n(x)} \leq \frac{[f_n](N(y)) + [f_n](N(y))}{[f_n](N(x))} \leq e^J_h (1 + s^*) \varepsilon^*/2 + f^* K' b e^{m+\alpha} \leq e^J_h \varepsilon^*.
\]

This means (5.8) for \( n \) since we have set \( \varepsilon = d(x,y) \).
Lemma 5.5. There exists a constant $C_R > 0$ such that $\|R_n\|_B \leq C_R d_n^{\alpha/(m+\alpha)}$ for all $n > 0$.

If, moreover, $T$ satisfies Assumption $T''(\varepsilon')$, then $\|R_n\|_B \leq C_R d_n$ for all $n > 0$.

Proof. Since $R_i = \sum_j R_{ij}$, we only need to prove the results for $R_{ij}$.

Let $s_{ij}(x)$ be the norm of $\|D\tilde{T}^{-1}_{ij}(x)\|$, and $s_{ij} = \max\{s_{ij}(x) : x \in B_{1\varepsilon}(Q_0)\}$. Note that $\{\tau > i\} \subset T^{-1}V$ for all large $i$. We may suppose that $i$ is sufficiently large so that $B_{1\varepsilon} x_1(U_{ij}) \subset \tilde{T}^{-1}_{ij}V$.

Take $f \in B$ with $\|f\|_B = 1$.

By using (4.20) and (4.21), we apply arguments similar to (4.22) and get

$$\|R_{ij}f\|_1 = \int_{U_{ij}} |f| d\nu \leq \|f\|_\infty \hat{\nu}(U_{ij}) \leq C_0 \nu(Q_0) d_{ij} \|f\|_B. \quad (5.9)$$

Next, we consider $|R_{ij}f|_B$. Note that for any $I \in T$, $f|_{V_I} \in \mathcal{H}^\alpha(V_I, H)$ for some $H \leq \|f\|_B$. So $\text{osc}(f/h, B_{x\varepsilon}(\cdot)) \leq 2^\alpha s^\alpha \varepsilon^\alpha H \leq 2^\alpha s^\alpha \varepsilon^\alpha \|f\|_B$. Note that Sublemma 5.4 implies $\hat{\nu}(h, B_{x\varepsilon}(\cdot)) \leq 2^\alpha J_\varepsilon^\alpha \varepsilon^\alpha$ for all $x$ with $B_{x\varepsilon}(x) \in V_I$ for some $J_\varepsilon \geq J_\varepsilon > 0$. By Proposition 3.2(3) in [Ss], $\text{osc}(f, B_{x\varepsilon}(\cdot)) \leq \text{osc}(f/h, B_{x\varepsilon}(\cdot)) \hat{h} + \text{osc}(\hat{h}, B_{x\varepsilon}(\cdot)) \|f\|_\infty/\hat{h} \leq b_1 \varepsilon^\alpha \|f\|_B$, where $b_1 = 2^\alpha (H + J_\varepsilon^\alpha C_{1 \varepsilon^{-1}})$. By arguments similar to (4.20) and (4.23),

$$\int_{U_{ij}} \text{osc}(f, B_{x\varepsilon}(\cdot)) d\nu = \int_{U_{ij}} \text{osc}(f, B_{x\varepsilon}(\cdot)) d\nu \leq b_1 \varepsilon^\alpha \|f\|_B \nu(U_{ij}) \leq b_1 \varepsilon^\alpha d_{ij} \nu(Q_0) \|f\|_B \leq a_1 \varepsilon^\alpha d_{ij} \|f\|_B, \quad (5.10)$$

where $a_1 = b_1 \nu(Q_0)$.\footnote{The estimate (5.10) shows the difference with the analogous bound (4.23) and justifies the introduction of the new Banach space. In fact we can now use the local H"older property for $f$ to get an upper bound of the integral of the oscillation simultaneously in terms of the volume of $U_{ij}$, of $\varepsilon$ and of the norm of $f$. The change of variable sending $U_{ij}$ to $Q_0$, will finally produce the determinant $d_{ij}$ which will give a better upper bound for $\|R_n\|$.

$$G_{ij}(x, \varepsilon, \varepsilon_0) = 2d_{ij} \cdot b_2 \varepsilon/\nu(B(1-\varepsilon_0) \varepsilon_0(x)) \leq a_2 d_{ij} \varepsilon, \quad (5.11)$$

where $a_2 = 2b_2 \nu(B(1-\varepsilon_0) \varepsilon_0(x))$. Note that $\int \text{osc}(f, B_{x\varepsilon}(x)) d\nu \leq \varepsilon_0 \|f\|_Q$, and $\|f\|_1 + \varepsilon_0 \|f\|_Q \leq \|f\|_Q \leq \|f\|_B$. So for any $\varepsilon \in (0, \varepsilon_0]$ and $i < N(\varepsilon)$, we use (4.15), (5.10), (5.9) and (5.11) to get

$$|R_{ij} f|_Q \leq (1 + \varepsilon^\alpha) a_2 + 2\varepsilon C_{ij} \nu(Q_0) + 2(1 + \varepsilon^\alpha) a_2 \varepsilon^{1-\alpha} d_{ij} \|f\|_B \leq C_0 d_{ij} \|f\|_B, \quad (5.12)$$

$$\|R_{ij} f\|_Q \leq (1 + \varepsilon^\alpha) a_2 + 2\varepsilon C_{ij} \nu(Q_0) + 2(1 + \varepsilon^\alpha) a_2 \varepsilon^{1-\alpha} d_{ij} \|f\|_B \leq C_0 d_{ij} \|f\|_B.$$}
where $C'_2 = (1 + \zeta \varepsilon) a_1 + 2 \zeta C_0 \nu (Q_0) + 2 (1 + \zeta \varepsilon) a_2 \varepsilon^{1-\alpha}$.

For $\varepsilon \in (0, \varepsilon_0]$ and $i > N (\varepsilon)$, by Assumption $T''(\varepsilon)$ we have $d_{ij} \leq b \varepsilon^{m+\alpha}$. Hence, $\varepsilon^{-\alpha} \leq (b^{-1} d_{ij})^{-\alpha/(m+\alpha)}$. So by (4.16), we have

$$|R_{ij}f|_B \leq 2(\gamma a m_0^{m})^{-1} \cdot ||f||_Q \cdot \varepsilon^{-\alpha} \cdot d_{ij}$$

$$\leq 2(\gamma a m_0^{m})^{-1} ||f||_Q \cdot \varepsilon^{-\alpha} \cdot d_{ij}$$

(5.13)

where $C''_2 = 2(\gamma a m_0^{m})^{-1} \varepsilon^{\alpha/(m+\alpha)}$. Therefore we get that $|R_{ij}f|_Q \leq C d_{ij}^{m/(m+\alpha)}$, where $C_2 = \max \{C_2, C''_2\}$.

Now we consider $|R_{ij}f|_H$. As in the proof of Lemma 5.2, for any $x, y \in U_{ij}$,

$$|R_{ij}f(x) - R_{ij}f(y)| \leq \frac{1}{H(x)} \left| \frac{\hat{g}(x)}{h(x)} - \frac{\hat{g}(y)}{h(y)} \right|$$

$$= \frac{1}{H(x)} \left| \frac{\hat{g}(x) h(x)}{h(x)} - \frac{\hat{g}(y) h(x)}{h(y)} \right|$$

$$+ \frac{1}{H(y)} \left| \frac{\hat{g}(x) h(y)}{h(x)} - \frac{\hat{g}(y) h(y)}{h(y)} \right|$$

(5.14)

Note that $\left| \frac{\hat{g}(x) h(x)}{h(x)} - \frac{\hat{g}(y) h(y)}{h(y)} \right| \leq ||f||_H d_{ij}(x, y)^\alpha \leq ||f||_B s_{ij}^{\alpha} d_{ij}(x, y)^\alpha$.

Also, $\hat{g}(x_{ij}) \hat{h}(x_{ij})/\hat{h}(x_{ij}) \leq (\hat{h}^* / \hat{h}_*) d_{ij}$. Then the first term in the right hand side of (5.14) is bounded by $a_3 d_{ij} ||f||_B d_{ij}(x, y)^\alpha$, where $a_3 = (\hat{h}^* / \hat{h}_*) s_{ij}^{\alpha}$.

Let us take $\varepsilon = d(x, y)$; if $i \leq N (\varepsilon)$, then by (5.7),

$$|\hat{g}(x_{ij}) \hat{h}(x_{ij})/\hat{h}(x_{ij}) - \hat{g}(y_{ij}) \hat{h}(y_{ij})/\hat{h}(y_{ij})| \leq 2 \varepsilon (\hat{h}^* / \hat{h}_*) d_{ij}(x, y)^\alpha$$

Since $f(y_{ij}) / \hat{h}(y_{ij}) \leq ||f||_H / \hat{h}_* \leq C_0 \hat{h}_*^{-1} ||f||_B$, the last term in (5.14) is bounded by $a_4 d_{ij} ||f||_B d_{ij}(x, y)^\alpha$, where $a_4 = 2C_0 f'' \hat{h}_*^2$. Therefore we obtain $|R_{ij}f|_H \leq C'_3 d_{ij} ||f||_B$, where $C'_3 = b_1 + b_2$.

If $i \geq N (\varepsilon)$, then by the first inequality of (5.14), the left side of the inequality is bounded by $\max \{\hat{g}(x_{ij}) f(x_{ij})/h(x), \hat{g}(y_{ij}) h(y_{ij})/h(y)\} \leq d_{ij} ||f||_H / \hat{h}_*$. By the same arguments as for (5.13) we can get that

$$|R_{ij}f|_H \leq \varepsilon^{-\alpha} d_{ij} ||f||_H / \hat{h}_* \leq C_0 \hat{h}_*^{-1} \varepsilon^{\alpha/(m+\alpha)} d_{ij}^{m/(m+\alpha)} ||f||_B$$

$$= C''_3 d_{ij}^{m/(m+\alpha)} ||f||_B$$

where $C''_3 = C_0 \hat{h}_*^{-1} \varepsilon^{\alpha/(m+\alpha)} ||f||_B$. Then we conclude that $|R_{ij}f|_H \leq C d_{ij}^{m/(m+\alpha)} ||f||_B$, where $C_3 = \max \{C_3, C''_3\}$.

The conclusion of the first part follows by setting $C_R = C_1 + C_2 + C_3$.

If $T$ satisfies Assumption $T''(\varepsilon)$, then we can regard $N (\varepsilon) = \infty$ for any $\varepsilon > 0$. Hence we get $||R_{ij}f||_B \leq C_R d_{ij} ||f||_B$ with $C_R = C_1 + C_2 + C_3$, □

**Acknowledgments**

We wish to thank Ian Melbourne for his interest in our maps, for his helpful comments and advices.

45
We would like to thank also Romain Aimino who was trapped in interminable
discussions on aperiodicity and function spaces. We finally thank the anony-

mous referee for a very careful reading of the paper and whose comments and
suggestions helped us to improve our article.

References


of cocycles and conditional local limit theorems, *Stoch. & Dynam.*, 4
(2004), 31–62

[ABV] J.F. Alves, C. Bonatti and M. Viana, SRB measures for partially hy-
perbolic systems whose central direction in mostly expanding, *Invent.

Dynamical Systems in One Dimension*, Probability and its Applications,
Birkhauser, 1997

[Bo] R. Bowen *Equilibrium states and the ergodic theory of Anosov diffeo-

[Br] A. Broise, *Transformations dilatantes de l’intervalle et théorèmes lim-
ites*, 5-110, Astérisque, 238 (1996)

[Go] S. Gouëzel, Sharp polynomial estimates for the decay of correlations,

[He] H. Hennion, Sur un théorème spectral et son application aux noyaux

[HH] H. Hennion and L. Hervé, *Limit theorems for Markov chains and
Stochastic Properties of Dynamical Systems by Quasicompactness*, Lect.
Notes Math., 1766, Springer-Verlag, 2001

[Hu] H. Hu, Decay of correlations for piecewise smooth maps with indifferent

[HPT] H. Hu, Ya. Pesin and A. Talitskaya, A Volume Preserving Diffeomor-
phism with Essential Coexistence of Zero and Nonzero Lyapunov Expo-

46


[Qu] A. Quas, Non-ergodicity for $C^1$ expanding maps and $g$-measures (English summary), *Ergodic Theory Dynam. Systems*, 16 (1996), 531–543


47


