Mixing for invertible infinite measure systems

Ian Melbourne *

24 November 2012; Revised 18 April 2014

Abstract

In a recent paper, Melbourne and Terhesiu [Operator renewal theory and mixing rates for dynamical systems with infinite measure, *Invent. Math.* **189** (2012), 61–110] obtained results on mixing and mixing rates for a large class of noninvertible maps preserving an infinite ergodic invariant measure.

Here, we are concerned with extending these results to the invertible setting. Mixing is established for a large class of infinite measure invertible maps. Assuming additional structure, in particular exponential contraction along stable manifolds, it is possible to obtain good results on mixing rates and higher order asymptotics.

1 Introduction

There is a well-developed theory of mixing and rates of mixing (decay of correlations) for large classes of dynamical systems with finite ergodic invariant measure. Uniformly hyperbolic (Axiom A) maps are topologically mixing up to a finite cycle [24], and topological mixing is equivalent to mixing for reasonable measures. Then it is natural to restrict to the mixing case and focus on mixing rates. The standard approach of Sinai, Bowen and Ruelle [6, 22, 23] is to pass via a Markov partition to symbolic dynamics. Quotienting out the stable directions leads to a uniformly expanding system and the transfer operator L restricted to Hölder functions contracts exponentially quickly to the invariant density. This leads easily to exponential decay of correlations for the uniformly expanding quotient system. An elementary approximation argument then extends this result to the original system.

A similar approach holds in the nonuniformly hyperbolic case for systems modelled by a Young tower with exponential tails [29]. (This incorporates classical examples such as dispersing billiards, Hénon-like attractors, and Lorenz-like maps.) The approach also extends to systems modelled by Young towers with subexponential decay of correlations [30]. In fact, there was an oversight in the literature where many

*Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK

authors assumed that the approximation argument for passing from noninvertible to invertible had been checked, though Young's paper [30] was entirely in the noninvertible setting. This was resolved recently by Gouëzel [10] based on ideas in [7]; an exposition can be found in [20, Appendix B].

For infinite measure systems, results on mixing and rates of mixing were recently obtained by Melbourne and Terhesiu [19] in the noninvertible context. A natural question is to extend these results to the invertible case via the above approximation argument. It turns out that this is significantly more complicated than in the finite measure case, for reasons that will become transparent shortly. In this paper, we resolve the question of mixing in the invertible setting. Also, we extend results on rates of mixing [19, 25, 26] from the noninvertible case, but these extensions are satisfactory only when exponential (or very rapid) contraction is assumed along stable manifolds.

We now discuss briefly the setting for the results in [19] and in this paper, and explain why the approximation step to pass from noninvertible to invertible is more difficult than in the finite measure case. We will denote invertible maps by $f : M \to M$ and the corresponding quotient maps by $\bar f : \bar X \to \bar X$. The corresponding ergodic invariant measures are denoted μ and $\bar\mu$.

Let L denote the (normalised) transfer operator for $\bar f$ (so $\int Lv\, w\, d\bar\mu = \int v\, w \circ \bar f\, d\bar\mu$ for $v \in L^1(\bar X)$, $w \in L^\infty(\bar X)$). In the finite measure case, we would expect that $L^n v \to \int v\, d\bar\mu$ and would study the rate of this convergence, which translates into the rate of decay of correlations. However, in the infinite case $|L^n v|_1 \to 0$ for all $v \in L^1$, so the aim is to find normalising constants a_n such that $a_n L^n v \to \int v\, d\bar\mu$ (in some sense) for suitable observables v.

An important ingredient for studying infinite measure systems, crucial also in [19], is the existence of a "good" inducing set $\bar Y \subset \bar X$ of finite nonzero measure. Define the *first return time* $\varphi : \bar Y \to \mathbb{Z}^+$ and the first return map $\bar F = \bar f^\varphi : \bar Y \to \bar Y$ by setting

$$\varphi(y) = \inf\{n \ge 1 : \bar f^n y \in \bar Y\}, \quad \bar F(y) = \bar f^{\varphi(y)}(y). \tag{1.1}$$

It is assumed that $\bar f$ is conservative so that φ is defined almost everywhere in Y, but that φ is nonintegrable.

We require further that $\bar F : \bar Y \to \bar Y$ is uniformly expanding (in a sense that will be made precise later) and that the tails of the first return time are regularly varying:

$$\bar\mu(y \in \bar Y : \varphi(y) > n) = \ell(n) n^{-\beta}, \tag{1.2}$$

where $\beta \in (0, 1]$ and ℓ is a slowly varying function[1].

Set $d_\beta = \frac{1}{\pi} \sin \beta\pi$ and define

$$a_n = \begin{cases} d_\beta^{-1} \ell(n) n^{1-\beta}, & \beta \in (0,1) \\ \sum_{j<n} \ell(j) j^{-1}, & \beta = 1 \end{cases}. \tag{1.3}$$

[1]A measurable function $\ell : (0, \infty) \to (0, \infty)$ is *slowly varying* if $\lim_{x\to\infty} \ell(\lambda x)/\ell(x) = 1$ for all $\lambda > 0$. A *regularly varying* function is one of the form $\ell(x)x^q$ where $\ell(x)$ is slowly varying

Then for $\beta \in (\frac{1}{2}, 1]$, Melbourne and Terhesiu [19] extended results of Garsia and Lamperti [9] from the scalar probability case to prove that $\lim_{n\to\infty} |a_n 1_{\bar Y} L^n v - \int v\, d\bar\mu|_\infty = 0$ for Hölder observables v supported in $\bar Y$. It follows that we obtain the mixing result

$$\lim_{n\to\infty} a_n \int v\, w \circ \bar f^n \, d\bar\mu = \int v\, d\bar\mu \int w\, d\bar\mu,$$

for all observables v, w supported in $\bar Y$ with v Hölder and w integrable. (The result is stated precisely in Theorem 2.1(a) below, and the extension to the invertible setting is Theorem 2.2(a).)

For $\beta \in (0, \frac{1}{2}]$, it is known that such a result cannot hold without further assumptions. Gouëzel [13] used ideas of Doney [8] to show that the result goes through under the additional *smooth tails* condition

$$\bar\mu(y \in \bar Y : \varphi(y) = n) \le C\ell(n) n^{-(\beta+1)}. \tag{1.4}$$

(See Theorem 2.1(b) below, and Theorem 2.2(b) for the extension to the invertible setting.)

In the first paragraph of the introduction, we mentioned that in the classical uniformly hyperbolic setting, there is an elementary approximation method that enables results on decay of correlations to be passed from a quotient noninvertible system to the underlying invertible one. There is a single fundamental and transparent reason why the approximation method to pass from noninvertible to invertible maps is not straightforward in the infinite measure setting. In the finite measure case, mixing and rates of mixing are proved for observables supported on the whole space. In the infinite case, we work with observables supported in $\bar Y$ and Y. Starting with observables $v, w : M \to \mathbb{R}$ supported in Y, approximation leads to observables $\bar v$, $\bar w$ that are not supported in $\bar Y$. Moreover, the support of $\bar v$ and $\bar w$ depends on the level of approximation. This accounts for the delicate nature of some of the arguments in this paper.

Remark 1.1 It turns out that all the problems above are associated with approximating the w observable. If it is assumed that w depends only on future coordinates so that it is necessary only to deal with the v observable, then it is relatively easy to recover mixing and mixing rates from the noninvertible case regardless of rates of contraction along stable manifolds. See Subsection 5.1.

Example 1.1 An important class of noninvertible maps amenable to the methods in [19] are one-dimensional Pomeau-Manneville intermittent maps [21]. We focus here on the family of maps $\bar f : [0, 1] \to [0, 1]$ studied by Liverani *et al.* [16]. For $\gamma > 0$, these have the form

$$\bar f(x) = \begin{cases} x(1 + 2^\gamma x^\gamma), & x \in [0, \frac{1}{2}) \\ 2x - 1, & x \in [\frac{1}{2}, 1] \end{cases}, \tag{1.5}$$

and there is a unique (up to scaling) absolutely continuous $\bar f$-invariant measure $\bar\mu$ for each γ. The measure is finite if and only if $\gamma < 1$. Taking $\bar Y = [\frac{1}{2}, 1]$, conditions (1.2) and (1.4) are satisfied with $\beta = 1/\gamma$ and ℓ asymptotically constant. Moreover, these maps satisfy the remaining technical assumption in [19] and here; namely that the first return map $\bar F : \bar Y \to \bar Y$ is a Gibbs-Markov map (see Section 2.1 below).

The analogous class of invertible maps are studied in Hu and Young [15, 14]. These "almost Anosov" diffeomorphisms are uniformly hyperbolic except at one fixed point. It is reasonable to expect that they often satisfy the required technical assumptions (it is easy to construct simplified examples of intermittent diffeomorphisms which do so, see Section 8) in which case our results on mixing apply. Moreover, the examples in [15] have one neutral expanding direction and one strictly contracting direction so that there is exponential contraction along stable manifolds, implying mixing rates similar to those in the noninvertible setting [19]. However, Hu [14] considers almost Anosov diffeomorphisms where both the contracting and expanding directions are neutral, for which our methods would yield results on mixing but with poor mixing rates. In independent work, Liverani and Terhesiu [17] have developed a different technique which often yields optimal mixing rates in such situations, and moreover avoids assumptions of a Markov nature, but which is currently restricted to the case where there is a global smooth stable foliation. See Remark 8.1 for a more detailed comparison of the work presented here and in [17].

The remainder of the paper is organised as follows. In Section 2, we recall the results on mixing of [13, 19] from the noninvertible setting, and state the corresponding result, Theorem 2.2, in the invertible setting. In Section 3, we collect some standard techniques related to Young towers. In Section 4, we prove a key estimate. Sections 5 and 6 contain the proof of Theorem 2.2. In Section 7, we discuss rates of mixing. Finally, in Section 8, we describe examples to which our results apply.

Notation We write $a_n \ll b_n$ as $n \to \infty$ if there is a constant $C > 0$ such that $a_n \le C b_n$ for all $n \ge 1$.

2 Statement of results on mixing

In this section, we describe the results of [13, 19] about mixing for nonuniformly expanding maps, and our main results about mixing for nonuniformly hyperbolic diffeomorphisms.

2.1 Mixing for nonuniformly expanding maps

The results of [13, 19] apply in particular to systems with first return maps that are Gibbs-Markov (uniformly expanding plus good distortion). This includes parabolic rational maps of the complex plane (Aaronson *et al* [2]) and Thaler's class of interval maps with indifferent fixed points [27].

We recall the key definitions [1, 2]. Let $(\bar X,\bar\mu)$ be a Lebesgue space with countable measurable partition $\alpha_{\bar X}$. Let $\bar f:\bar X\to\bar X$ be an ergodic, conservative, measure-preserving, Markov map transforming each partition element bijectively onto a union of partition elements. Recall that $\bar f$ is *topologically mixing* if for all $a,b\in\alpha_{\bar X}$ there exists $N\ge 1$ such that $b\subset\bar f^n a$ for all $n\ge N$.

Let $\bar Y$ be a union of partition elements with $\bar\mu(\bar Y)\in(0,\infty)$. Define the first return time $\varphi:\bar Y\to\mathbb{R}$ and first return map $\bar F=\bar f^{\varphi}:\bar Y\to\bar Y$ as in (1.1). Let α be the partition of $\bar Y$ consisting of nonempty cylinders of the form $a\cap(\bigcap_{j=1}^{n-1}T^{-j}\xi_j)\cap T^{-n}\bar Y$ where $a,\xi_j\in\alpha_{\bar X}$, and $a\subset\bar Y$, $\xi_j\subset\bar X\setminus\bar Y$. Fix $\theta\in(0,1)$ and define $d_\theta(x,y)=\theta^{s(x,y)}$ where the *separation time* $s(x,y)$ is the greatest integer $n\ge 0$ such that $\bar F^n x$ and $\bar F^n y$ lie in the same partition element in α. It is assumed that the partition α separates orbits of $\bar F$, so $s(x,y)$ is finite for all $x\neq y$; then d_θ is a metric. Let $F_\theta(\bar Y)$ be the Banach space of d_θ-Lipschitz functions $v:\bar Y\to\mathbb{R}$ with norm $\|v\|_\theta=|v|_\infty+|v|_\theta$ where $|v|_\theta=\sup_{x\neq y}|v(x)-v(y)|/d_\theta(x,y)$.

Define the potential function $p=\log\frac{d\bar\mu}{d\bar\mu\circ\bar F}:\bar Y\to\mathbb{R}$. We require that p is uniformly piecewise Lipschitz: that is, $p|_a$ is d_θ-Lipschitz for each $a\in\alpha$ and the Lipschitz constants can be chosen independent of a. We also require the big images condition $\inf_a\bar\mu(\bar F a)>0$. Such a Markov map $\bar F:\bar X\to\bar X$ with partition α, with uniformly piecewise Lipschitz potential and satisfying the big images property, is called a *Gibbs-Markov map.*

Throughout we assume that $\bar\mu(\bar X)=\infty$ (in other words φ is nonintegrable) and that $\bar\mu$ is normalised so that $\bar\mu(\bar Y)=1$.

Theorem 2.1 *Let $\bar f:\bar X\to\bar X$ be as above with nonintegrable return time $\varphi:\bar Y\to\mathbb{Z}^+$ satisfying* (1.2) *and Gibbs-Markov return map $\bar F=\bar f^{\varphi}:\bar Y\to\bar Y$. Define a_n as in* (1.3).

(a) (Melbourne and Terhesiu [19]) *If $\beta\in(\frac12,1]$, then*

$$\lim_{n\to\infty}a_n\int_{\bar X}v\,w\circ\bar f^n\,d\bar\mu=\int_{\bar X}v\,d\bar\mu\int_{\bar X}w\,d\bar\mu,$$

for all observables $v,w:\bar X\to\mathbb{R}$ supported in $\bar Y$ with $v\in F_\theta(\bar Y)$ and $w\in L^1$.

(b) (Gouëzel [13]) *The same conclusion holds also for $\beta\in(0,\frac12]$ under the additional "smooth tails" condition* (1.4).

2.2 Mixing for nonuniformly hyperbolic diffeomorphisms

Let $f:M\to M$ be a diffeomorphism (possibly with singularities) defined on a Riemannian manifold (M,d). We assume that f is nonuniformly hyperbolic in the sense of Young [29, 30]. The precise definitions are somewhat technical; here we are content to focus on the parts necessary for understanding this paper, referring to the original papers for further details.

As part of this set up, there is an f-invariant ergodic conservative measure μ. Let $Y \subset M$ with $\mu(Y) \in (0, \infty)$ and rescale so that $\mu(Y) = 1$. Define the first return time function $\varphi : Y \to \mathbb{Z}^+$ given by $\varphi(y) = \inf\{n \geq 1 : f^n y \in Y\}$. (By conservativity, φ is finite almost everywhere on Y and we may suppose without loss that φ is finite on the whole of Y.) Also, define the first return map $F = f^\varphi : Y \to Y$. (In [29, 30], it is not required that φ is the first return to Y, but this is a crucial ingredient in [19] and in this paper.)

Let $\{Z_j\}$ denote a measurable partition of Y such that φ is constant on partition elements. Let s denote the *separation time* with respect to the map $F : Y \to Y$. That is, if $y, y' \in Y$, then $s(y, y')$ is the least integer $n \geq 0$ such that $F^n y$, $F^n y'$ lie in distinct partition elements of Y.

Let $\{W^s\}$ and $\{W^u\}$ denote two measurable partitions of Y. If $y \in Y$, the elements containing y are labelled $W^s(y)$ and $W^u(y)$.

(P1) $F(W^s(y)) \subset W^s(Fy)$ for all $y \in Y$.

(P2) There exist constants $C \geq 1$, $\gamma_0 \in (0, 1)$ such that

- (i) If $y' \in W^s(y)$, then $d(F^n y, F^n y') \leq C\gamma_0^n$ for all $n \geq 0$.
- (ii) If $y' \in W^u(y)$, then $d(F^n y, F^n y') \leq C\gamma_0^{s(y,y')-n}$ for all $0 \leq n \leq s(y, y')$.
- (iii) If $y, y' \in Y$, then $d(f^j y, f^j y') \leq Cd(Fy, Fy')$ for all $0 \leq j < \min\{\varphi(y), \varphi(y')\}$.

Hypotheses (P1) and (P2) imply that $F : Y \to Y$ is uniformly hyperbolic with stable and unstable disks W^s and W^u. We assume a local product structure, namely that each W^s intersects each W^u in precisely one point.

Let $\bar{Y} = Y/\sim$ where $y \sim y'$ if $y \in W^s(y')$ and define the partition $\{\bar{Z}_j\}$ of $\bar{Y}$. We obtain a well-defined first return time function $\varphi : \bar{Y} \to \mathbb{Z}^+$ and first return map $\bar{F} : \bar{Y} \to \bar{Y}$. Define $\bar{\pi}_*\mu = \bar{\mu}$ where $\bar{\pi} : Y \to \bar{Y}$ is the quotient map.

(P3) The map $\bar{F} : \bar{Y} \to \bar{Y}$ and partition $\{\bar{Z}_j\}$ separate points in $\bar{Y}$. (It follows that $d_\theta(y, y') = \theta^{s(y,y')}$ defines a metric on $\bar{Y}$ for each $\theta \in (0, 1)$.)

(P4) $\bar{F} : \bar{Y} \to \bar{Y}$ is a Gibbs-Markov map with respect to the partition $\alpha = \{\bar{Z}_j\}$ with ergodic invariant probability measure $\bar{\mu}$.

We omit the additional assumptions in Young [29] that guarantee that μ is a physical measure for $F : Y \to Y$. The results here do not rely on this property.

Our main result is the following.

Theorem 2.2 *Assume that $f : M \to M$ is a nonuniformly hyperbolic diffeomorphism satisfying conditions (P1)–(P4) above. Assume that the return time function $\varphi : Y \to \mathbb{Z}^+$ is nonintegrable and has regularly varying tails as in* (1.2).

Assume further either (a) that $\beta \in (\frac{1}{2}, 1]$, or (b) that $\beta \in (0, \frac{1}{2}]$ and φ has smooth tails as in (1.4).

Then

$$\lim_{n\to\infty} a_n \int_M v\, w \circ f^n \, d\mu = \int_M v \, d\mu \int_M w \, d\mu,$$

for all Hölder observables $v, w : M \to \mathbb{R}$ supported in Y.

Remark 2.1 The results in this paper are restricted to the situation where the first return quotient map $\bar{F} : \bar{Y} \to \bar{Y}$ is Gibbs-Markov. However, the results in [19], described in Theorem 2.1(a), hold in a much more general functional analytic setting which includes the important class of one-dimensional "AFN" maps studied in [31]. (These are the non-Markovian analogue of the maps studied by Thaler [27].) Results for invertible maps where the quotient map is AFN are obtained in [17].

3 Towers and approximation of observables

In this section, we recall standard material on modelling nonuniformly expanding maps and nonuniformly hyperbolic diffeomorphisms by tower maps (one-sided and two-sided respectively), and how to approximate observables to pass from two-sided to one-sided towers.

3.1 One-sided towers

Let $\bar{f} : \bar{X} \to \bar{X}$ be a nonuniformly expanding map with Gibbs-Markov first return map $\bar{F} = \bar{f}^\varphi : \bar{Y} \to \bar{Y}$ as in Subsection 2.1. Define the *tower* $\bar{\Delta} = \{(y, j) \in \bar{Y} \times \mathbb{Z} : 0 \le j < \varphi(y)\}$ and the *tower map* $\bar{f}_{\bar{\Delta}} : \bar{\Delta} \to \bar{\Delta}$ given by $\bar{f}_{\bar{\Delta}}(y, j) = (y, j+1)$ for $j \le \varphi(y) - 2$ and $\bar{f}_{\bar{\Delta}}(y, j) = (\bar{F}y, 0)$ for $j = \varphi(y) - 1$.

The base of the tower $\{(y, 0) : y \in \bar{Y}\}$ is naturally identified with $\bar{Y}$ and so we may regard $\bar{Y}$ as a subset of both $\bar{X}$ and $\bar{\Delta}$. Then $\mu_{\bar{\Delta}} = \bar{\mu} \times \text{counting}$ is an $\bar{f}_{\bar{\Delta}}$-invariant measure on $\bar{\Delta}$.

Define the projection $\pi : \bar{\Delta} \to \bar{X}$, $\pi(y, j) = \bar{f}^j y$. Then $\pi \bar{f}_{\bar{\Delta}} = \bar{f}\pi$ and $\pi_* \mu_{\bar{\Delta}} = \bar{\mu}$. Thus $\bar{f}_{\bar{\Delta}}$ is an extension of $\bar{f}$ with the same first return map $\bar{F} : \bar{Y} \to \bar{Y}$ and return time function $\varphi : \bar{Y} \to \mathbb{Z}^+$ as the original map.

3.2 Two-sided towers

Now let $f : M \to M$ be a nonuniformly hyperbolic diffeomorphism, as in Subsection 2.2, with first return map $F = f^\varphi : Y \to Y$ that quotients to a Gibbs-Markov map $\bar{F} : \bar{Y} \to \bar{Y}$. Form the quotient tower $\bar{\Delta}$ as in Subsection 3.1, Similarly, starting from $F : Y \to Y$ and $\varphi : Y \to \mathbb{Z}^+$, we can form a tower Δ and tower map $f_\Delta : \Delta \to \Delta$ such that $F = f^\varphi : Y \to Y$ is the first return map for f_Δ. Again, $\mu_\Delta = \mu \times$counting is an f_Δ-invariant measure on Δ. Define the semiconjugacy $\pi : \Delta \to M$, $\pi(y, \ell) = f^\ell y$. Then we can take $\mu = \pi_*\mu_\Delta$ to be the measure on M mentioned in Section 2.2.

We use the separation time for $F : Y \to Y$ to define a separation time on Δ: define $s((y, \ell), (y', \ell')) = s(y, y')$ if $\ell = \ell'$ and 0 otherwise. This drops down to separation times s on $\bar{\Delta}$ and $\bar{Y}$. (The separation time on $\bar{Y}$ coincides with the one defined in Subsection 2.1.)

For $q = (y, \ell), q' = (y', \ell) \in \Delta$, we write $q' \in W^s(q)$ if $y' \in W^s(y)$ and $q' \in W^u(q)$ if $y' \in W^u(y)$. Conditions (P2) translate as follows.

(P2′) There exist constants $C \ge 1$, $\gamma_0 \in (0, 1)$ such that for all $q, q' \in \Delta$, $n \ge 1$,

- (i) If $q' \in W^s(q)$, then $d(\pi f_\Delta^n q, \pi f_\Delta^n q') \le C\gamma_0^{\psi_n(q)}$, and
- (ii) If $q' \in W^u(q)$, then $d(\pi f_\Delta^n q, \pi f_\Delta^n q') \le C\gamma_0^{s(q,q')-\psi_n(q)}$,

where $\psi_n(q) = \#\{j = 0, \dots, n - 1 : f^j q \in Y\}$ is the number of visits of q to Y by time n.

Proposition 3.1 $d(\pi f_\Delta^n q, \pi f_\Delta^n q') \le C\gamma_0^{\min\{\psi_n(q), s(q,q')-\psi_n(q)\}}$ *for all* $q, q' \in \Delta$, $n \ge 1$.

Proof This is immediate from conditions (P2′) and the product structure on Y (cf. [18, Corollary 5.3]). ∎

Given $\theta \in (0, 1)$, we define the symbolic metric d_θ on $\bar{\Delta}$ by setting $d_\theta(q, q') = \theta^{s(q,q')}$. This restricts to the metric d_θ defined on $\bar{Y}$ in Section 2.1.

3.3 Approximation of observables

Throughout, we work with Hölder observables $v_0, w_0 : M \to \mathbb{R}$ with fixed Hölder exponent $\eta \in (0, 1]$. Let $v = v_0 \circ \pi : \Delta \to \mathbb{R}$ be the lift of such a C^η observable $v_0 : M \to \mathbb{R}$. For each $k \ge 1$, define $v_k : \Delta \to \mathbb{R}$,

$$v_k(q) = \inf\{v(f_\Delta^k q') : s(q, q') \ge 2\psi_k(q)\}.$$

We list some standard properties of v_k. Let L denote the transfer operator corresponding to $\bar{f}_{\bar{\Delta}} : \bar{\Delta} \to \bar{\Delta}$, defined by $\int_{\bar{\Delta}} Lv\, w\, d\mu_{\bar{\Delta}} = \int_{\bar{\Delta}} v\, w \circ \bar{f}_{\bar{\Delta}}\, d\mu_{\bar{\Delta}}$ for $v \in L^1$, $w \in L^\infty$.

Proposition 3.2 *The function* v_k *lies in* $L^\infty(\Delta)$ *and projects down to a Lipschitz observable* $\bar{v}_k : \bar{\Delta} \to \mathbb{R}$. *Moreover, setting* $\gamma = \gamma_0^\eta$ *and* $\theta = \gamma^{\frac{1}{2}}$,

(a) $|\bar{v}_k|_\infty = |v_k|_\infty \le |v_0|_\infty$.

(b) $|v \circ f_\Delta^k(q) - v_k(q)| \le C\|v_0\|_{C^\eta}\gamma^{\psi_k(q)}$ *for* $q \in \Delta$.

(c) $|\gamma^{\psi_k}|_1 \to 0$ *as* $k \to \infty$.

(d) $\|L^k\bar v_k\|_\theta \le C\|v_0\|_{C^\eta}$.

Proof Again, this is standard (see for example [20, Proposition B.5]). If $s(q,q') \ge 2\psi_k(q)$, then $v_k(q) = v_k(q')$. It follows that v_k is piecewise constant on a measurable partition of Δ, and hence is measurable, and that $\bar v_k$ is well-defined. Part (a) is immediate.

Recall that $v = v_0 \circ \pi$ where $v_0 : M \to \mathbb{R}$ is C^η. Let $q \in \Delta$. By Proposition 3.1 and the definition of v_k,

$$\begin{aligned}|v \circ f_\Delta^k(q) - v_k(q)| &= |v_0(\pi f_\Delta^k q) - v_0(\pi f_\Delta^k q')| \le \|v_0\|_{C^\eta} d(\pi f_\Delta^k q, \pi f_\Delta^k q')^\eta \\ &\le C^*\gamma^{\min\{\psi_k(q), s(q,q')-\psi_k(q)\}},\end{aligned}$$

where q' is such that $s(q,q') \ge 2\psi_k(q)$. In particular, $s(q,q') - \psi_k(q) \ge \psi_k(q)$, so we obtain part (b). Part (c) is immediate by the bounded convergence theorem since $\psi_k \to \infty$ as $k \to \infty$ almost everywhere.

The proof of part (d) is a bit longer and we refer to [20, Proposition B.5(c)]. ∎

We require some further less standard properties that stem from the special role of Y in infinite ergodic theory.

Proposition 3.3 *Let* $k \ge 1$.

(a) If $\operatorname{supp} v \subset Y$, *then* $\operatorname{supp} v_k \subset f_\Delta^{-k}Y$.

(b) If $\operatorname{supp} v \subset Y$, *then* $\operatorname{supp} L^k\bar v_k \subset \bar Y$.

Proof Assume that $\operatorname{supp} v \subset Y$. Suppose $q \notin f_\Delta^{-k}(Y)$, that is $f_\Delta^k q \notin Y$. Then $f_\Delta^k q' \notin Y$ for all q' with $s(q,q') \ge k$, and hence $v(f_\Delta^k q') = 0$ for such points. In particular, $v_k(q) = 0$ proving (a).

Combining part (a) with the first statement in Proposition 3.2, we obtain that $\operatorname{supp} \bar v_k \subset \bar f_\Delta^{-k}\bar Y$. Now $(L^k\bar v_k)(q) = \sum_{\bar f_\Delta^k q' = q} e^{p_k}(q')\bar v_k(q')$ where $p_k = p + p\circ f + \cdots + p \circ f^{k-1}$. Suppose $q \notin \bar Y$ and that $\bar f_\Delta^k q' = q$. We have $\bar v_k(q') = v(f_\Delta^k q'')$ where $s(q'',q') \ge 2k$ and hence $s(f_\Delta^k q'', q) \ge k$. In particular, $f_\Delta^k q'' \notin Y$ so $\bar v_k(q') = v(f_\Delta^k q'') = 0$. It follows that each term in $(L^k\bar v_k)(q)$ is zero, proving (b). ∎

4 A key estimate

In this section, we prove the following key estimate. Let $\bar Y_k = \bar f_\Delta^{-k}\bar Y$.

Lemma 4.1 *There exists a sequence $c_{k,n}$ with the property that $\lim_{n\to\infty} c_{k,n} = 0$ for all k such that*

$$\Big|\int_{\bar\Delta}\Big(a_n L^{n-k}v - \int_{\bar\Delta} v\, d\bar\mu_{\bar\Delta}\Big) w\, d\bar\mu_{\bar\Delta}\Big| \le c_{k,n}\|v\||w|_\infty,$$

for all $v \in F_\theta(\bar Y)$, $w \in L^\infty(\bar Y_k)$, $k, n \ge 0$.

The proof of this result is entirely at the level of the quotient tower and we suppress the bars. Also, we write f and μ, omitting the subscript Δ.

We require some preliminary notation. Define $T_n, P : L^1(Y) \to L^1(Y)$ to be the operators $T_n v = 1_Y L^n(1_Y v)$ and $Pv = 1_Y \int_Y v\, d\mu$.

Following [11] (see also [12, 18, 20] for example) we define a sequence of operators $A_n : L^1(Y) \to L^1(\Delta)$ that behave like L^n but summing only over preimages that lie in Y and do not return to Y in n iterates of f. This is analogous to $T_n = 1_Y L^n 1_Y$ which is like L^n but summing only over preimages that lie in Y and have returned to Y at time n. In symbols,

$$(T_n v)(y) = \sum_{\substack{f^n z = y\\ y,z\in Y}} e^{p_n}(z)v(z), \quad (A_n v)(q) = \sum_{\substack{f^n z = q, z\in Y\\ f^j z\notin Y, j=1,\ldots,n}} e^{p_n}(z)v(z),$$

and we make the convention that $T_0 = A_0 = I$. Then $L^n 1_Y$ is the convolution

$$L^n 1_Y = \sum_{j=0}^{n} A_j T_{n-j}. \tag{4.1}$$

It is easy to check that $(A_j v)(y,\ell) = \delta_{j,\ell} v(y)$ where $\delta_{j,\ell}$ is the Kronecker delta. In particular,

$$1_\Delta = \sum_{j=0}^{\infty} A_j 1_Y. \tag{4.2}$$

Proposition 4.1 $|1_{Y_k} A_j 1_Y|_1 \le \mu(\varphi > j) - \mu(\varphi > j+k)$ *for all j, k,*

Proof Since $1_{Y_k} A_j 1_Y$ is supported on the j'th level of the tower, $|1_{Y_k} A_j 1_Y|_1 = \mu_\Delta(\{\ell = j\} \cap Y_k)$. But $\{\ell = j\} \cap Y_k \subset \{(y,\ell) : \ell = j,\ j < \varphi(y) \le k+j\}$ and so $\mu_\Delta(\{\ell = j\} \cap Y_k) \le \mu(\varphi > j) - \mu(\varphi > j+k)$. ∎

Using (4.1) and (4.2), we obtain

$$\begin{aligned}\int_\Delta\Big(a_n L^{n-k}v - \int_\Delta v\, d\mu\Big) w\, d\mu &= \int_\Delta\Big(a_n \sum_{j=0}^{n-k} A_j T_{n-k-j} v - \Big(\int_\Delta v\, d\mu\Big)\sum_{j=0}^{\infty} A_j 1_Y\Big) 1_{Y_k} w\, d\mu \\ &= \int_\Delta w \sum_{j=0}^{n-k} 1_{Y_k} A_j 1_Y \Big(a_n T_{n-k-j} v - Pv\Big)\, d\mu - (Pv)\int_\Delta w \sum_{j=n-k+1}^{\infty} 1_{Y_k} A_j 1_Y\, d\mu.\end{aligned}$$

Define $U_n : F_\theta(Y) \to F_\theta(Y)$, $U_n v = a_n T_n v$. Then $\int_\Delta \Big(a_n L^{n-k} v - \int_\Delta v\, d\mu \Big) w\, d\mu = E_1(k,n) + E_2(k,n) - E_3(k,n)$, where

$$E_1(k,n) = \int_\Delta w \sum_{j=0}^{n-k} 1_{Y_k} A_j 1_Y \Big(U_{n-k-j} v - Pv \Big)\, d\mu$$

$$E_2(k,n) = \int_\Delta w \sum_{j=0}^{n-k} 1_{Y_k} A_j 1_Y (a_n - a_{n-k-j}) T_{n-k-j} v\, d\mu$$

$$= \int_\Delta w \sum_{j=0}^{n-k} 1_{Y_k} A_j 1_Y \Big(\frac{a_n}{a_{n-k-j}} - 1 \Big) U_{n-k-j} v\, d\mu$$

$$E_3(k,n) = (Pv) \int_\Delta w \sum_{j=n-k+1}^{\infty} 1_{Y_k} A_j 1_Y\, d\mu.$$

We make use of the fact, proved in [19], that

$$\lim_{n\to\infty} \|U_n - P\| = 0. \tag{4.3}$$

To estimate $E_1(k,n)$, it is useful to recall the following "discrete dominated convergence theorem".

Proposition 4.2 *Let $b_{n,j} \in \mathbb{R}$, $n, j \ge 0$. Suppose that $\lim_{n\to\infty} b_{n.j} = 0$ for each j and there is a sequence $c_j > 0$ such that $|b_{n,j}| \le c_j$ for all j, n and $\sum_{j=0}^\infty c_j < \infty$. Then $\lim_{n\to\infty} \sum_{j=0}^\infty b_{n,j} = 0$.* ∎

Proposition 4.3 $\lim_{n\to\infty} E_1(k,n) = 0$ *for each fixed k.*

Proof We keep k fixed throughout. By Proposition 4.1,

$$|E_1(k,n)| \le |w|_\infty \sum_{j=0}^{n-k} |1_{Y_k} A_j 1_Y|_1 |U_{n-k-j} v - Pv|_\infty \le \|v\| |w|_\infty \sum_{j=0}^{n-k} b_{n,j}$$

where $b_{n,j} = (\mu(\varphi > j) - \mu(\varphi > j + k)) \|U_{n-k-j} - P\|$. By (4.3), $\lim_{n\to\infty} b_{n,j} = 0$ for each j, and there is a constant C such that $b_{n,j} \le c_j = C(\mu(\varphi > j) - \mu(\varphi > j + k))$. Moreover, $\sum_{j=0}^\infty c_j = C \sum_{j=0}^{k-1} \mu(\varphi > j) < \infty$. Now apply Proposition 4.2. ∎

Next, it follows from (4.3) that $\|U_n\|$ is bounded, so

$$|E_2(k,n)| \le |w|_\infty \sum_{j=0}^{n-k} |1_{Y_k} A_j 1_Y|_1 |U_{n-k-j} v|_\infty \Big| \frac{a_n}{a_{n-k-j}} - 1 \Big|$$

$$\ll \|v\| |w|_\infty \sum_{j=0}^{n-k} |1_{Y_k} A_j 1_Y|_1 \Big| \frac{a_n}{a_{n-k-j}} - 1 \Big|.$$

It is convenient to split the sum into two parts, defining

$$E_2'(k,n)=\sum_{0\le j\le n/2}|1_{Y_k}A_j1_Y|_1\Big|\frac{a_n}{a_{n-k-j}}-1\Big|,$$
$$E_2''(k,n)=\sum_{n/2<j\le n-k}|1_{Y_k}A_j1_Y|_1\Big|\frac{a_n}{a_{n-k-j}}-1\Big|.$$

Proposition 4.4 $\lim_{n\to\infty}E_2'(k,n)=0$ *for each fixed* k.

Proof We keep k fixed throughout. By Proposition 4.1, $E_2'(k,n)\le\sum_{0\le j\le n/2}b_{n,j}$ where $b_{n,j}=(\mu(\varphi>j)-\mu(\varphi>j+k))|\frac{a_n}{a_{n-k-j}}-1|$. Recalling the definition in (1.3), we note that a_n is regularly varying (this is immediate for $\beta<1$ and is a consequence of Karamata for $\beta=1$). By Potter's bounds, a_n/a_{n-k-j} is bounded for j,n with $j\le n/2$. Hence $b_{n,j}\ll c_j=\mu(\varphi>j)-\mu(\varphi>j+k)$ which is summable as in Proposition 4.3. Also $\lim_{n\to\infty}a_n/a_{n-k-j}=1$ for each fixed j. Hence we can again apply Proposition 4.2. ∎

Define $\tilde\ell(n)=\sum_{j<n}\ell(j)j^{-1}$.

Proposition 4.5 $E_2''(k,n)\ll\begin{cases}k\ell(n)^2n^{-(2\beta-1)} & \beta<1\\ k\tilde\ell(n)\ell(n)n^{-1} & \beta=1\end{cases}$ *for all* k,n.

Proof We have $a_n/a_{n-k-j}\ll a_n$, so by Proposition 4.1,

$$|E_2''(k,n)|\ll a_n\sum_{n/2<j\le n-k}(\mu(\varphi>j)-\mu(\varphi>j+k))\le a_nk\mu(\varphi>n/2)\ll a_nk\mu(\varphi>n),$$

and the result follows by definition of a_n. ∎

Proposition 4.6 *For* $k\le n/2$, $|E_3(k,n)|\ll\Big|\int_\Delta v\,d\mu\Big||w|_\infty k\ell(n)n^{-\beta}$.

Proof By Proposition 4.1, suppressing the factor $|\int_\Delta v\,d\mu||w|_\infty$,

$$|E_3(k,n)|\le\sum_{j=n-k+1}^{\infty}|1_{Y_k}A_j1_Y|_1\le\sum_{j=n-k+1}^{\infty}\mu(\varphi>j)-\mu(\varphi>j+k)$$
$$=\sum_{j=n-k+1}^{n}\mu(\varphi>j)\le k\mu(\varphi>n-k+1).$$

Since $\mu(\varphi>n)$ is regularly varying, it follows from Potter's bounds that $\mu(\varphi>n-k+1)/\mu(\varphi>n)$ is bounded for k,n with $k\le n/2$. Hence $|E_3(k,n)|\ll k\mu(\varphi>n)$ proving the result. ∎

Proof of Lemma 4.1 Since $\int_\Delta\Big(a_nL^{n-k}v-\int_\Delta v\,d\mu\Big)w\,d\mu=E_1(k,n)+E_2(k,n)-E_3(k,n)$, the result follows from Propositions 4.3–4.6. ∎

5 Proof of Theorem 2.2(a)

In this section, we prove part (a) of our main result, Theorem 2.2, establishing mixing for infinite measure nonuniformly hyperbolic diffeomorphisms with $\beta > \frac{1}{2}$.

Given C^η observables $v_0, w_0 : M \to \mathbb{R}$, supported in Y, let $v = v_0 \circ \pi, w = w_0 \circ \pi : \Delta \to \mathbb{R}$ be the lifted observables to the tower. (Since we are interested in observables on Y and $F = f^\varphi$ is a first return map, $\pi|_Y = \mathrm{id}_Y$ and we could simply write $v_0 = v$ and $w_0 = w$, but this does not seem helpful.) Since $\pi : \Delta \to M$ is a semiconjugacy and $\mu = \pi_* \mu_\Delta$, to prove Theorem 2.2 it is suffices to work entirely in the context of the tower map $f_\Delta : \Delta \to \Delta$ and the one-sided quotient $\bar{f}_{\bar{\Delta}} : \bar{\Delta} \to \bar{\Delta}$. For notational simplicity, from now on we write f, $\bar{f}$, μ and $\bar{\mu}$ instead of f_Δ, $\bar{f}_{\bar{\Delta}}$, μ_Δ and $\bar{\mu}_{\bar{\Delta}}$. Thus the result we want to prove becomes

$$\lim_{n\to\infty} a_n \int_\Delta v\, w \circ f^n \, d\mu = \int_\Delta v\, d\mu \int_\Delta w\, d\mu,$$

for $v, w : Y \to \mathbb{R}$ Hölder. Again for notational simplicity, from now on we write $\|v\|$ instead of $\|v\|_{C^\eta}$ and similarly for w.

Since the arguments for dealing with v and w are very different, it is convenient to give the arguments separately in Subsections 5.1 and 5.2. The combined argument is given in Subsection 5.3.

5.1 The v observable

In this subsection, we show how to deal with the v observable: we assume the set up of Theorem 2.2(a), but we suppose that w depends only on future coordinates (that is, w is constant along stable disks and hence projects to an observable on $\bar{\Delta}$).

Lemma 5.1 *There exists a constant $C > 0$ such that*

$$\Big| a_n \int_\Delta v\, w \circ f^n \, d\mu - \int_\Delta v\, d\mu \int_\Delta w\, d\mu \Big| \le C \|U_n - P\| \|v\| |w|_1.$$

for all $v, w : Y \to \mathbb{R}$ where v is Hölder and w is L^∞ such that w depends only on future coordinates, and all $n \ge 0$.

Proof Write

$$\begin{aligned} a_n \int_\Delta v\, w \circ f^n \, d\mu - \int_\Delta v\, d\mu \int_\Delta w\, d\mu \\ = a_n \int_\Delta v \circ f^\ell\, w \circ f^{n+\ell} \, d\mu - \int_\Delta v \circ f^\ell \, d\mu \int_\Delta w\, d\mu = I_1(\ell, n) + I_2(\ell, n) - I_3(\ell, n), \end{aligned}$$

where

$$I_1(\ell,n)=a_n\int_\Delta v_\ell\, w\circ f^{n+\ell}\,d\mu-\int_\Delta v_\ell\,d\mu\int_\Delta w\,d\mu,$$

$$I_2(\ell,n)=a_n\int_\Delta (v\circ f^\ell-v_\ell)\, w\circ f^{n+\ell}\,d\mu,\quad I_3(\ell,n)=\int_\Delta (v\circ f^\ell-v_\ell)\,d\mu\int_\Delta w\,d\mu.$$

By Proposition 3.2 and the assumption on w, both v_ℓ and w depend only on future coordinates. Hence we can write

$$\int_\Delta v_\ell\, w\circ f^{n+\ell}\,d\mu=\int_{\bar\Delta}\bar v_\ell\, w\circ \bar f^{n+\ell}\,d\bar\mu=\int_{\bar\Delta}L^nL^\ell\bar v_\ell\, w\,d\bar\mu$$

Moreover, $\operatorname{supp} L^\ell\bar v_\ell\subset\bar Y$ by Proposition 3.3(b) and $\operatorname{supp} w\subset\bar Y$ by assumption, so $L^nL^\ell\bar v_\ell\, w=T_nL^\ell\bar v_\ell\, w$. Hence

$$I_1(\ell,n)=a_n\int_{\bar\Delta}T_nL^\ell\bar v_\ell\, w\,d\bar\mu-\int_{\bar\Delta}L^\ell\bar v_\ell\,d\bar\mu\int_{\bar\Delta}w\,d\bar\mu=\int_{\bar\Delta}\Big(a_nT_nL^\ell\bar v_\ell-PL^\ell\bar v_\ell\Big)w\,d\bar\mu,$$

and so

$$|I_1(\ell,n)|\le|(a_nT_n-P)L^\ell\bar v_\ell|_\infty\,|w|_1\le\|U_n-P\|\|L^\ell\bar v_\ell\|_\theta|w|_1\ll\|U_n-P\|\|v\||w|_1,$$

where we have used Proposition 3.2(d).

Next, by Proposition 3.2(b),

$$|I_2(\ell,n)|\ll a_n\|v\||w|_\infty|\gamma^{\psi_\ell}|_1,\quad |I_3(\ell,n)|\ll\|v\||w|_1|\gamma^{\psi_\ell}|_1.$$

By Proposition 3.2(c), $\lim_{\ell\to\infty}I_j(\ell,n)=0$ for all n for $j=2,3$. The result follows. ∎

5.2 The w observable

In this subsection, we show how to deal with the w observable: we assume the set up of Theorem 2.2(a), but we suppose that v depends only on future coordinates.

Lemma 5.2 *Let $c_{k,n}$ be as in Lemma 4.1. There exists a constant $C>0$ such that*

$$\Big|a_n\int_\Delta v\, w\circ f^n\,d\mu-\int_\Delta v\,d\mu\int_\Delta w\,d\mu\Big|\le C(c_{k,n}+|\gamma^{\psi_k}|_1)\|v\|\|w\|,$$

for all $v,w:Y\to\mathbb{R}$ Hölder such that v depends only on future coordinates, and all $0\le k\le n$.

Proof Write

$$\begin{aligned}&a_n\int_\Delta v\, w\circ f^n\,d\mu-\int_\Delta v\,d\mu\int_\Delta w\,d\mu \qquad (5.1)\\ &\quad=a_n\int_\Delta v\circ f^k\, w\circ f^{n+k}\,d\mu-\int_\Delta v\,d\mu\int_\Delta w\circ f^k\,d\mu=I_1(k,n)+I_2(k,n)-I_3(k,n),\end{aligned}$$

where

$$
\begin{aligned}
I_1(k,n) &= a_n \int_\Delta v\circ f^k\, w_k\circ f^n\, d\mu - \int_\Delta v\, d\mu \int_\Delta w_k\, d\mu, \\
I_2(k,n) &= a_n \int_\Delta v\circ f^k\, (w\circ f^k - w_k)\circ f^n\, d\mu, \\
I_3(k,n) &= \int_\Delta v\, d\mu \int_\Delta (w\circ f^k - w_k)\, d\mu.
\end{aligned}
$$

By the same argument used for $I_1(\ell,n)$ in Subsection 5.1, $\int_\Delta v\circ f^k\, w_k\circ f^n\, d\mu = \int_{\bar\Delta} v\circ \bar f^k\, \bar w_k\circ \bar f^n\, d\bar\mu$. Hence for all $k\le n$,

$$
I_1(k,n) = a_n \int_{\bar\Delta} v\circ \bar f^k\, \bar w_k\circ \bar f^n\, d\bar\mu - \int_{\bar\Delta} v\, d\bar\mu \int_{\bar\Delta} \bar w_k\, d\bar\mu = \int_{\bar\Delta} \Big(a_n L^{n-k} v - \int_{\bar\Delta} v\, d\bar\mu\Big) \bar w_k\, d\bar\mu.
$$

By Lemma 4.1 and Proposition 3.2(a), $|I_1(k,n)| \le c_{k,n}\|v\||w|_\infty$.

Next, note that

$$
\begin{aligned}
I_2(k,n) &= a_n \int_\Delta (1_Y v)\circ f^k [1_{f^{-k}Y}(w\circ f^k - w_k)]\circ f^n\, d\mu \\
&= a_n \int_\Delta (1_Y\, 1_Y\circ f^n)\circ f^k\, v\circ f^k\, (w\circ f^k - w_k)\circ f^n\, d\mu,
\end{aligned}
$$

so by Proposition 3.2(b), for $n\ge k$,

$$
\begin{aligned}
|I_2(k,n)| &\ll |v|_\infty \|w\| a_n \int_\Delta (1_Y\, 1_Y\circ f^n)\circ f^k\, \gamma^{\psi_k}\circ f^n\, d\mu \\
&= |v|_\infty \|w\| a_n \int_{\bar\Delta} 1_{\bar Y}\, (1_{\bar Y}\circ \bar f^k\, \gamma^{\psi_k})\circ \bar f^{n-k}\, d\bar\mu \\
&= |v|_\infty \|w\| a_n \int_{\bar\Delta} (L^{n-k} 1_{\bar Y})\, 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu = |v|_\infty \|w\| (I_2'(k,n) + I_2''(k,n)),
\end{aligned}
$$

where

$$
I_2'(k,n) = \int_{\bar\Delta} \Big(a_n L^{n-k} 1_{\bar Y} - \int_{\bar\Delta} 1_{\bar Y}\, d\bar\mu\Big)\, 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu, \quad I_2''(k,n) = \int_{\bar\Delta} 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu.
$$

It follows from Lemma 4.1 that $|I_2'(k,n)| \le c_{k,n}\|1_{\bar Y}\||1_{\bar Y_k}\gamma^{\psi_k}|_\infty \le c_{k,n}$. Also $|I_2''(k,n)| \le |\gamma^{\psi_k}|_1$.

Finally, by Proposition 3.2(b), $|I_3(k,n)| \ll |v|_1 \|w\| |\gamma^{\psi_k}|_1$. ∎

5.3 The v and w observables

Finally, we consider the general case where v and w depend on the past and future.

Lemma 5.3 *Let $c_{k,n}$ be as in Lemma 4.1. There exists a constant $C > 0$ such that*

$$\Big|a_n \int_\Delta v\, w \circ f^n \, d\mu - \int_\Delta v \, d\mu \int_\Delta w \, d\mu\Big| \le C(c_{k,n} + |\gamma^{\psi_k}|_1)\|v\|\|w\|,$$

for all $v, w : Y \to \mathbb{R}$ Hölder, and all $0 \le k \le n$.

Proof We combine the methods used to prove Lemmas 5.1 and 5.2. Write

$$\begin{aligned}
a_n \int_\Delta v\, w \circ f^n \, d\mu &- \int_\Delta v \, d\mu \int_\Delta w \, d\mu \\
&= a_n \int_\Delta v \circ f^{\ell+k}\, w \circ f^{k+n+\ell} \, d\mu - \int_\Delta v \circ f^\ell \, d\mu \int_\Delta w \circ f^k \, d\mu \\
&= I_1(k,\ell,n) + I_2(k,\ell,n) + I_3(k,\ell,n) - I_4(k,\ell,n) - I_5(k,\ell,n),
\end{aligned}$$

where

$$\begin{aligned}
I_1(k,\ell,n) &= a_n \int_\Delta v_\ell \circ f^k \, w_k \circ f^{n+\ell} \, d\mu - \int_\Delta v_\ell \, d\mu \int_\Delta w_k \, d\mu, \\
I_2(k,\ell,n) &= a_n \int_\Delta v_\ell \circ f^k \, (w \circ f^k - w_k) \circ f^{n+\ell} \, d\mu, \\
I_3(k,\ell,n) &= a_n \int_\Delta (v \circ f^\ell - v_\ell) \circ f^k \, w \circ f^{k+n+\ell} \, d\mu, \\
I_4(k,\ell,n) &= \int_\Delta v_\ell \, d\mu \int_\Delta (w \circ f^k - w_k) \, d\mu, \\
I_5(k,\ell,n) &= \int_\Delta (v \circ f^\ell - v_\ell) \, d\mu \int_\Delta w \circ f^k \, d\mu.
\end{aligned}$$

Now

$$\int_\Delta v_\ell \circ f^k \, w_k \circ f^{n+\ell} \, d\mu = \int_{\bar\Delta} \bar v_\ell \, \bar w_k \circ \bar f^{n-k+\ell} \, d\bar\mu = \int_{\bar\Delta} L^{n-k} L^\ell \bar v_\ell \, \bar w_k \, d\bar\mu.$$

Hence

$$\begin{aligned}
I_1(k,\ell,n) &= a_n \int_{\bar\Delta} L^{n-k} L^\ell \bar v_\ell \, \bar w_k \, d\bar\mu - \int_{\bar\Delta} L^\ell \bar v_\ell \, d\bar\mu \int_{\bar\Delta} \bar w_k \, d\bar\mu \\
&= \int_{\bar\Delta} \Big(a_n L^{n-k} L^\ell \bar v_\ell - \int_{\bar\Delta} L^\ell \bar v_\ell \, d\bar\mu\Big) \bar w_k \, d\bar\mu.
\end{aligned}$$

By Proposition 3.3(b), $\operatorname{supp} L^\ell \bar v_\ell \subset \bar Y$. Hence by Lemma 4.1 and Proposition 3.3(a,d),

$$|I_1(k,\ell,n)| \le c_{k,n} \|L^\ell \bar v_\ell\| |\bar w_k|_\infty \ll c_{k,n} \|v\| |w|_\infty.$$

Arguing as for $I_2(k,n)$ in the proof of Lemma 5.2,

$$\begin{aligned}
|I_2(k,\ell,n)| &\ll |v|_\infty \|w\| a_n \int_\Delta 1_Y \circ f^{k+\ell}\, 1_Y \circ f^{k+n+\ell}\, \gamma^{\psi_k} \circ f^{n+\ell}\, d\mu \\
&= |v|_\infty \|w\| a_n \int_{\bar\Delta} 1_{\bar Y}\, (1_{\bar Y} \circ \bar f^k\, \gamma^{\psi_k}) \circ \bar f^{n-k}\, d\bar\mu \\
&= |v|_\infty \|w\| a_n \int_{\bar\Delta} (L^{n-k} 1_{\bar Y})\, 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu \\
&= |v|_\infty \|w\| \Big\{ \int_{\bar\Delta} (a_n L^{n-k} 1_{\bar Y} - \int_{\bar\Delta} 1_{\bar Y}\, d\bar\mu)\, 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu + \int_{\bar\Delta} 1_{\bar Y_k}\, \gamma^{\psi_k}\, d\bar\mu \Big\} \\
&\le |v|_\infty \|w\| (c_{k,n} + |\gamma^{\psi_k}|_1).
\end{aligned}$$

By Proposition 3.2(b),

$$\begin{gathered}
|I_3(k,\ell,n)| \ll a_n \|v\| |w|_\infty |\gamma^{\psi_\ell}|_1, \quad |I_4(k,\ell,n)| \ll |v|_\infty \|w\| |\gamma^{\psi_k}|_1, \\
|I_5(k,\ell,n)| \ll \|v\| |w|_1 |\gamma^{\psi_\ell}|_1.
\end{gathered}$$

Combining these estimates, letting $\ell \to \infty$, and using Proposition 3.2(c), we obtain the required result. ∎

Proof of Theorem 2.2(a) Recall that $\lim_{n\to\infty} c_{k,n} = 0$ for each k. Hence it follows from Lemma 5.3 that $\limsup_{n\to\infty} |a_n \int_\Delta v\, w \circ f^n\, d\mu - \int_\Delta v\, d\mu \int_\Delta w\, d\mu| \ll |\gamma^{\psi_k}|_1$. The result follows from Proposition 3.2(c). ∎

6 Proof of Theorem 2.2(b)

In this section, we prove part (b) of our main theorem, dealing with the case $\beta \in (0, \frac12]$. Recall that it is necessary to assume an additional condition even for the noninvertible result, namely the smooth tails condition $\mu(\varphi = n) \le C\ell(n) n^{-(\beta+1)}$. The proof of Theorem 2.2(b) is almost identical to the proof of part (a). The only estimate that requires modification is the one for $E_2''(k,n)$ in Proposition 4.5.

In fact, the milder smooth tail condition

$$\mu(\varphi = n) \le C\ell(n) n^{-(\beta+q)}, \tag{6.1}$$

suffices, provided that $q > 1 - \frac12\beta$.

Remark 6.1 Let $q \in [0,1]$. Given that $\mu(\varphi > n) = \ell(n) n^{-\beta}$, condition (6.1) is equivalent to the condition that

$$|\ell(n+1) - \ell(n)| \ll \ell(n) n^{-q}.$$

Recall that $a_n = d_\beta^{-1} \ell(n) n^{1-\beta}$. We list some elementary consequences of smooth tails.

Proposition 6.1 *Assume smooth tails with $q \in (1-\beta, 1]$. Then*

(i) $\mu(\varphi > j) - \mu(\varphi > j+k) \ll k\ell(j) j^{-(\beta+q)}$.

(ii) $\left|\dfrac{a_{n+j}}{a_n} - 1\right| \ll j n^{-q}$. ∎

Proposition 6.2 *Assume the smooth tails condition with $q \in (1-\beta, 1)$. Then $E_2''(k,n) \ll k\ell(n) n^{-(\beta+2q-2)}$. If $q = 1$, then $E_2''(k,n) \ll k\ell(n) n^{-\beta} \log n$.*

Proof Recall that $E_2''(k,n) = \sum_{n/2<j\le n-k} |1_{Y_k} A_j 1_Y|_1 \left|\frac{a_n}{a_{n-k-j}} - 1\right|$. By Propositions 4.1 and 6.1,

$$E_2''(k,n) \ll k \sum_{j=n/2}^{n-k} \ell(j) j^{-(\beta+q)} (k+j)(n-k-j)^{-q} \ll k\ell(n) n^{1-\beta-q} \sum_{j=n/2}^{n-k} (n-k-j)^{-q}$$

and the result follows. ∎

In particular, if $q > 1 - \frac{1}{2}\beta$, then $\lim_{k\to\infty} \lim_{n\to\infty} E_2''(k,n) = 0$. All the other terms were covered in the proof of Theorem 2.2(a) and hence the proof of Theorem 2.2(b) is complete.

7 Mixing rates

As mentioned in the introduction, we obtain some results on mixing rates and higher order asymptotics, but in general they are fairly weak. However, if we assume sufficiently rapid contraction along stable manifolds and smooth tails with q large enough, then we obtain essentially optimal results. For simplicity, throughout this section we assume exponential contraction along stable manifolds and the smooth tails condition (1.4) with $q = 1$.

7.1 Some more estimates

Lemma 7.1 *Assume the smooth tails condition* (1.4)*. If the mixing rate in the non-invertible case is given by $\|a_n T_n - P\| \ll (\log n)^c n^{-\tau}$, where $\tau \in (0,1)$, $c \ge 0$, then $|E_1(k,n)| \ll \|v\| |w|_\infty k (\log n)^c n^{-\tau}$ for all k, n.*

Proof Recall that $|E_1(k,n)| \le \|v\| |w|_\infty \sum_{j=0}^{n-k} |1_{Y_k} A_j 1_Y|_1 \|U_{n-k-j} - P\|$. By Propositions 4.1 and 6.1(i),

$$|E_1(k,n)| \le \|v\| |w|_\infty k \sum_{j=0}^{n-k} \ell(j) j^{-(\beta+1)} (\log(n-k-j))^c (n-k-j)^{-\tau},$$

which gives the desired result since the convolution of $\ell(n)n^{-(\beta+1)}$ and $(\log n)^c n^{-\tau}$ is $O((\log n)^c n^{-\tau})$. ∎

Lemma 7.2 *Assume the smooth tails condition* (1.4). *Then* $|E_2(k,n)| \ll \|v\||w|_\infty(k^2n^{-1} + k\ell(n)n^{-\beta}\log n)$ *for* $k \le n/3$.

Proof We already obtained this estimate for $E_2''(k,n)$ in Proposition 6.2 so it remains to handle $E_2'(k,n)$. The following very rough estimate suffices.

Recall that $E_2'(k,n) = \sum_{0\le j\le n/2} |1_{Y_k}A_j 1_Y|_1 \Big|\frac{a_n}{a_{n-k-j}} - 1\Big|$. By Propositions 4.1 and 6.1,

$$
\begin{aligned}
|E_2'(k,n)| &\ll k\sum_{j=1}^{n/2} \ell(j)j^{-(\beta+1)}(k+j)(n-k-j)^{-1} \ll kn^{-1}\sum_{j=1}^{n/2}\ell(j)j^{-(\beta+1)}(k+j) \\
&= k^2n^{-1}\sum_{j=1}^{n/2}\ell(j)j^{-(\beta+1)} + kn^{-1}\sum_{j=1}^{n/2}\ell(j)j^{-\beta} \ll k^2n^{-1} + k\ell(n)n^{-\beta}.
\end{aligned}
$$

∎

Corollary 7.1 *Assume the smooth tails condition* (1.4). *If the mixing rate in the noninvertible case is given by* $\|a_nT_n - P\| \ll (\log n)^c n^{-\tau}$, *where* $\tau \in (0,1)$, $c \ge 0$, *then*

$$
\begin{aligned}
&\Big| a_n \int_M v\, w\circ f^n\, d\mu - \int_M v\, d\mu \int_M w\, d\mu \Big| \\
&\qquad \ll \|v\|\|w\|\{k(\log n)^c n^{-\tau} + k^2n^{-1} + k\ell(n)n^{-\beta}\log n + |\gamma^{\psi_k}|_1\},
\end{aligned}
$$

for all Hölder $v,w : M\to\mathbb{R}$ *supported in* Y, *and all* $0\le k\le n/3$.

Proof This combines the various estimates obtained throughout this paper. We claim that these combine to give the estimate

$$
\|v\|\|w\|\{k(\log n)^c n^{-\tau} + k^2n^{-1} + k\ell(n)n^{-\beta}\log n + |\gamma^{\psi_k}|_1 + a_n|\gamma^{\psi_\ell}|_1\}.
$$

Since ℓ is arbitrary the result follows from Proposition 3.2(c).

To prove the claim it is required to estimate the five terms $I_j(k,\ell,n)$, $j = 1,\dots,5$, that appear in Section 5.3. By Proposition 3.2(b), $|I_3(k,\ell,n)| \ll a_n|\gamma^{\psi_\ell}|_1$, $|I_4(k,\ell,n)| \ll |\gamma^{\psi_k}|_1$, $|I_5(k,\ell,n)| \ll |\gamma^{\psi_\ell}|_1$.

The calculations in Section 5.3 show that

$$
|I_1(k,\ell,n)| \ll c_{k,n}, \quad |I_2(k,\ell,n)| \ll (c_{k,n} + |\psi_k|_1),
$$

where $c_{k,n}$ is as in Lemma 4.1. Moreover, by the proof of Lemma 4.1, we can obtain an estimate for $c_{k,n}$ by estimating the expressions $E_j(k,\ell,n)$, $j=1,2,3$. This is done using Lemma 7.1, Lemma 7.2 and Proposition 4.6 respectively. ∎

7.2 Exponential contraction

Suppose that there is exponential contraction along stable manifolds. This corresponds to strengthening condition (P2)(i) in Section 2.2.

(E) If $y' \in W^s(y)$, then $d(f^n y, f^n y') \le C\gamma_0^n$ for all $n \ge 1$.

In this case, γ^{ψ_k} is replaced by γ^k.

Theorem 7.1 *Assume that $f : M \to M$ is a nonuniformly hyperbolic diffeomorphism satisfying conditions (P1)–(P4) and the exponential contraction condition (E). Suppose that φ satisfies conditions (1.2) and (1.4).*

If the mixing rate in the noninvertible case is given by $\|a_n T_n - P\| \ll (\log n)^c n^{-\tau}$, where $\tau \in (0,1)$, $c \ge 0$, then

$$\Big| a_n \int_M v\, w \circ f^n \, d\mu - \int_M v\, d\mu \int_M w\, d\mu \Big| \ll \|v\| \|w\| \{(\log n)^{c+1} n^{-\tau} + \ell(n) n^{-\beta} (\log n)^2\},$$

for all $v, w : Y \to \mathbb{R}$ Hölder, and all $n \ge 0$.

Proof Take $k = q \log n$ with q large in Corollary 7.1 ∎

The following refinement of Theorem 7.1 is useful for passing results on higher order asymptotics in the noninvertible context over to invertible systems.

Theorem 7.2 *Assume that $f : M \to M$ is a nonuniformly hyperbolic diffeomorphism satisfying conditions (P1)–(P4) and the exponential contraction condition (E). Suppose that φ satisfies conditions (1.2) and (1.4).*

Let b_n be a bounded sequence such that $b_{n+1} - b_n \ll n^{-1}$. If the mixing rate in the noninvertible case is given by $\|a_n T_n - b_n P\| \ll (\log n)^c n^{-\tau}$, where $\tau \in (0,1)$, $c \ge 0$, then

$$\Big| a_n \int_M v\, w \circ f^n \, d\mu - b_n \int_M v\, d\mu \int_M w\, d\mu \Big| \ll \|v\| \|w\| \{(\log n)^{c+1} n^{-\tau} + \ell(n) n^{-\beta} (\log n)^2\},$$

for all $v, w : Y \to \mathbb{R}$ Hölder, and all $n \ge 0$.

Proof In the notation of Lemma 5.3, we obtain $a_n \int_\Delta v\, w \circ f^n \, d\mu - b_n \int_\Delta v\, d\mu \int_\Delta w\, d\mu = J_1(k,\ell,n) + I_2(k,\ell,n) + I_3(k,\ell,n) - b_n I_4(k,\ell,n) - b_n I_5(k,\ell,n)$, where

$$\begin{aligned} J_1(k,\ell,n) &= a_n \int_\Delta v_\ell \circ f^k\, w_k \circ f^{n+\ell} \, d\mu - b_n \int_\Delta v_\ell \, d\mu \int_\Delta w_k \, d\mu \\ &= \int_{\bar\Delta} \Big(a_n L^{n-k} L^\ell \bar v_\ell - b_n \int_{\bar\Delta} L^\ell \bar v_\ell \, d\bar\mu \Big) \bar w_k \, d\bar\mu. \end{aligned}$$

Since b_n is bounded, it suffices to verify that $J_1(k,\ell,n)$ is bounded by the terms listed in Corollary 7.1.

As in the proof of Lemma 4.1, we reduce to considering three terms $F_j(k,n)$, $j=1,2,3$, given by $F_2(k,n)=E_2(k,n)$, $F_3(k,n)=b_nF_3(k,n)$ and

$$F_1(k,n)=\int_\Delta w\sum_{j=0}^{n-k}1_{Y_k}A_j1_Y\Big(U_{n-k-j}v-b_nPv\Big)\,d\mu=F_1'(k,n)+F_1''(k,n)+F_1'''(k,n),$$

where

$$\begin{aligned}F_1'(k,n)&=\int_\Delta w\sum_{j=0}^{n-k}1_{Y_k}A_j1_Y\Big(U_{n-k-j}v-b_{n-k-j}Pv\Big)\,d\mu,\\ F_1''(k,n)&=\int_\Delta w\sum_{0\le j\le n/2}1_{Y_k}A_j1_Y(b_{n-k-j}-b_n)Pv\,d\mu,\\ F_1'''(k,n)&=\int_\Delta w\sum_{n/2<j\le n-k}1_{Y_k}A_j1_Y(b_{n-k-j}-b_n)Pv\,d\mu.\end{aligned}$$

The argument for $E_1(k,n)$ in Lemma 7.1 goes through unchanged for $F_1'(k,n)$. Also, by Propositions 4.1 and 6.1(i),

$$|F_1'''(k,n)|\ll|v|_1|w|_1\sum_{n/2<j\le n-k}|1_{Y_k}A_j1_Y|_1\ll k\mu(\varphi>n/2)\ll k\ell(n)n^{-\beta}.$$

So far, we used only that b_n is bounded. Finally, using Propositions 4.1 and 6.1(i) and the extra condition on b_n, for $k\le n/3$,

$$\begin{aligned}|F_1''(k,n)|&\le|v|_1|w|_1\sum_{0\le j\le n/2}|1_{Y_k}A_j1_Y|_1|b_{n-k-j}-b_n|\\ &\ll kn^{-1}\sum_{0\le j\le n/2}\ell(j)j^{-(\beta+1)}(k+j)\ll k^2n^{-1}+k\ell(n)n^{-\beta}.\end{aligned}$$

This completes the proof. ∎

Example 7.1 For the maps (1.5) studied in [16], conditions (1.2) and (1.4) are satisfied and $\ell(n)$ is asymptotically constant. For $\beta\in(\frac12,1)$, Melbourne and Terhesiu [19] obtained a mixing rate in the noninvertible case of the form

$$\|T_n-\sum_{j\ge1}d_ja_n^{-j}P\|\ll c_n,\tag{7.1}$$

where $c_n=n^{-\frac12}$, $d_1=1$ and the remaining constants d_j are typically nonzero. For such expressions we make the convention that the sum is finite, disregarding the terms for which $a_n^{-j}\ll c_n$.

if we assume exponential contraction along stable manifolds then setting $c = 0$, $\tau = \beta - \frac{1}{2}$, and $b_n = \sum_{j\ge1} d_j a_n^{-(j-1)}$ in Theorem 7.2, we obtain the almost identical mixing rate

$$\Big|\int_M v\,w\circ f^n\,d\mu - \sum_{j\ge1} d_j a_n^{-j}\int_M v\,d\mu\int_M w\,d\mu\Big| \ll \|v\|\|w\|c_n', \tag{7.2}$$

where $c_n' = (\log n)n^{-\frac{1}{2}}$. As in [19], this is optimal for $\beta \ge \frac{3}{4}$ and we obtain second order asymptotics for $\beta > \frac{3}{4}$.

Using more refined methods, Terhesiu [25, 26] has obtained much stronger results for various classes of systems based on the properties of φ. All of these results apply in particular to the maps (1.5). In [25], it is still assumed that $\beta > \frac{1}{2}$, but with the improved mixing rate $c_n = n^{-\beta}$ in (7.1), so for invertible systems with exponential contraction, we obtain (7.2) with $c_n' = (\log n)n^{-\beta}$. As in [25], this is optimal for $\beta \ge \frac{2}{3}$ and we obtain second order asymptotics for $\beta > \frac{2}{3}$.

Finally, Terhesiu [26] drops the restriction $\beta > \frac{1}{2}$ and obtains mixing rates for all $\beta \in (0,1)$. The assumptions on φ are yet more restrictive but still include the maps (1.5) and yield $c_n = (\log n)n^{-1}$ in (7.1). Accordingly, for invertible systems with exponential contraction, we obtain $c_n' = (\log n)^2 n^{-1}$ in (7.2). As in [25], this is optimal for $\beta \ge \frac{1}{2}$ and we obtain second order asymptotics for $\beta > \frac{1}{2}$.

7.3 Subexponential contraction

It is possible to relax hypothesis (E) to allow uniform but subexponential decay along stable manifolds [3]. In the stretched exponential case $d(f^n y, f^n y') \le C\gamma_0^{n^\delta}$, our results on mixing rates and higher order asymptotics are essentially unchanged (the factor $(\log n)^{c+1}$ is changed to $(\log n)^{c+1/\delta}$). More generally, we obtain essentially optimal mixing rates and higher order asymptotics if $d(f^n y, f^n y') \le Cn^{-q}$ with q sufficiently large.

7.4 Slow contraction

The original assumption (P2)(i) assumes that the contraction is as slow as the expansion. That is, expansion or contraction occurs only at returns to the set Y. This is natural for many examples, such as billiards, where there are long periods of time during which there is no contraction or expansion.

In such situations we define $\psi = 1_Y$ and note that $\psi_k = \sum_{j=0}^{k-1}\psi\circ f^j$. By the methods in [10, 20], it can be shown that $|\gamma^{\psi_k}|_1 \ll \mu(\varphi > k)$. Combined with smooth tails conditions, this can be used to obtain estimates on mixing rates. The results are too poor to write down here. (We note that a slight simplification is possible since the $k^2 n^{-1}$ term in Lemma 7.2 can be removed with some extra effort.)

8 Examples

In this section, we describe some examples in which our hypotheses can be verified. All of our results apply to Examples 8.1–8.3. In Example 8.4, Theorem 2.2 holds but hypothesis (E) fails.

Example 8.1 We begin with a particularly elementary example. Define $f : X \to X$, $X = [0,1]^2$, by setting

$$f(x_1, x_2) = \begin{cases} (x_1(1 + 2^\gamma x_1^\gamma)\,,\, \frac{1}{2}x_2) & 0 \le x_1 < \frac{1}{2} \\ (2x_1 - 1\,,\, \frac{1}{2}(x_2 + 1)) & \frac{1}{2} \le x_1 \le 1 \end{cases}.$$

We identify the first component of f with a map $\bar{f} : [0,1] \to [0,1]$ of the form in (1.5) with $\gamma \ge 1$. Recall from the introduction that $\bar{f}$ has a unique (up to scaling) absolutely continuous invariant measure $\bar{\mu}$ and this is an infinite measure. Let $\bar{Y} = [\frac{1}{2}, 1]$ and define $\varphi : \bar{Y} \to \mathbb{Z}^+$, $\bar{F} = \bar{f}^\varphi : \bar{Y} \to \bar{Y}$ to be the first return time and first return map. Then φ satisfies conditions (1.2) and (1.4) relative to $\bar{\mu}$ with $\beta = 1/\gamma$ and ℓ asymptotically constant. Moreover $\bar{F}$ is a Gibbs-Markov map.

Turning to the full map f, we set $Y = [\frac{1}{2}, 1] \times [0,1]$ and define $\varphi : Y \to \mathbb{Z}^+$, $F = f^\varphi : Y \to Y$ to be the first return time and first return map. Clearly $\varphi(y_1, y_2) = \varphi(y_1)$ and F has the form $F(y_1, y_2) = (\bar{F}(y_1), G(y_1, y_2))$ where $|\partial G/\partial y_2|_\infty \le \frac{1}{2}$.

We now verify the hypotheses from Section 2.2. Let $\{Z_j\} = \{a \times [0,1] : a \in \alpha\}$ where α is the partition of $\bar{Y}$ for the Gibbs-Markov map $\bar{F}$ and set $\varphi_j = \varphi|_{Z_j}$. Define the family of stable disks $\{W^s\} = \{\{y_1\} \times [0,1] : y_1 \in \bar{Y}\}$ and the family of unstable disks $\{W^u\} = \{\bar{Y} \times \{x_2\} : x_2 \in [0,1]\}$. Then (P1) and (P2)(i) are immediate with $\gamma_0 = \frac{1}{2}$. Also hypotheses (P2)(ii) and (P2)(iii) are standard for $\bar{f}$ and $\bar{F}$, and are clearly inherited by f and F. Hypotheses (P3) and (P4) are immediate. Moreover, hypothesis (E) holds (also with $\gamma_0 = \frac{1}{2}$).

Example 8.2 More generally, we consider invertible maps $f : X \to X$ of the form $f(x_1, x_2) = (\bar{f}(x_1), g(x_1, x_2))$ with $\bar{f}$, $\bar{Y}$, φ and $\bar{F}$ as before, and it is assumed that $g : X \to [0,1]$ is C^2. Then φ and Y are also unchanged and $F = f^\varphi : Y \to Y$ has the form $F(y_1, y_2) = (\bar{F}(y_1), G(y_1, y_2))$. We suppose that $|\partial G/\partial y_2|_\infty \le \gamma_0$ where $\gamma_0 \in (0,1)$. The stable disks $\{W^s\}$ and the partition $\{Z_j\}$ are unchanged from Example 8.1, and the existence of unstable disks $\{W^u\}$ is a standard consequence of the uniform hyperbolicity of $F|_{Z_j}$ for each j. Hypotheses (P1)–(P4) follow. If we assume further that $|\partial g/\partial x_2|_\infty \le \gamma_0$ then hypothesis (E) holds.

Example 8.3 A rather different set of examples can be constructed along the lines of the Smale-Williams solenoids [24, 28]. First modify the map $\bar{f}$ near $\frac{1}{2}$ so that $\bar{f}$ is C^2 on $(0,1)$. Let $X = [0,1] \times \mathbb{D}$ where $\mathbb{D}$ is the closed unit disk in $\mathbb{R}^{n-1}$ and define $f_0(x_1, x_2) = (\bar{f}(x_1), g(x_2))$ for $(x_1, x_2) \in [0,1] \times \mathbb{D}$ where $g : \mathbb{D} \to \mathbb{D}$ is C^2 and

$|Dg|_\infty < 1$. Now perturb to obtain a C^2 embedding $f : X \to X$ such that $f = f_0$ on $[0, \frac{1}{2}] \times \mathbb{D}$. Note that f has a unique neutral fixed point inside $\{0\} \times \mathbb{D}$.

Let $Y = [\frac{1}{2}, 1] \times \mathbb{D}$ and define the first return time $\varphi : Y \to \mathbb{Z}^+$ and first return map $F = f^\varphi : Y \to Y$. The combinatorics in this example are almost identical to those in the previous ones, the only difference being in the first iterate of f, so it is easy to check the conditions (1.2) and (1.4). Also, F is uniformly hyperbolic if f is close enough to f_0 and conditions (P1)–(P3) as well as (E) are easily verified. Moreover, (P4) is satisfied provided the perturbation f is chosen so that $\bar{F}$ is Markov.

Example 8.4 Finally, we mention a class of examples for which hypothesis (E) fails. Proceed exactly as in Example 8.3 except that $g : \mathbb{D} \to \mathbb{D}$ is replaced by a C^2 map $g : X \to \mathbb{D}$ with derivative $Jg = D_{x_2} g$ in the $\mathbb{D}$ directions such that $|Jg|_\infty \le 1$, $|Jg|_{[\frac{1}{2},1]\times\mathbb{D}}|_\infty < 1$, but $Jg(0, x_2) \equiv I$. This means that the invariant set $\{0\} \times \mathbb{D}$ is neutral in all directions. However, writing $F = (F_1, F_2)$, we have that $|D_{x_2} F_2|_\infty \le |Jg|_{[\frac{1}{2},1]\times\mathbb{D}}|_\infty < 1$ and hypotheses (P1)–(P4) hold as before.

Remark 8.1 Independently, using different techniques the bypass the quotienting step altogether, Liverani and Terhesiu [17] have obtained similar results to the ones presented in this paper. They work with distributional function spaces of the type pioneered by Liverani and coworkers starting with [5].

The class of examples in Example 8.2 is essentially the same as those treated by Liverani and Terhesiu [17, Example 1]. Currently the function spaces in [17] are restricted to situations where there is a global smooth stable foliation. In contrast, the method presented in this paper is not dependent on the smoothness of the stable foliation. In particular, our methods apply to sufficiently small smooth perturbations, preserving the Markov structure, of the maps in Example 8.2, and their higher dimensional analogues. For such maps, generally the stable foliations have only Hölder regularity (even though the individual leaves are smooth).

Similarly, the methods in [17] currently do not generally apply to Examples 8.3 and 8.4. On the other hand, the non-Markov examples of [17, Example 2] are not covered by the method presented in this paper.

It will be interesting to see whether the the methods in [17] can be developed to prove mixing rates and higher order asymptotics in situations such as in Example 8.4 when there is not a global smooth stable foliation. In many important situations such as billiards, hyperbolicity breaks down due to long periods where there is no contraction or expansion, so that (P1)–(P4) hold but not (E). Such examples are studied for instance in [4].

Acknowledgements This research was supported in part by the CNRS during a three month appointment as chercheur en mathématique at CIRM, Marseille, Autumn 2011 where this research was begun. The research was continued at the University of Surrey and completed at the University of Warwick, partially supported by the

European Advanced Grant *StochExtHomog* (ERC AdG 320977). I am very grateful to Dalia Terhesiu for numerous helpful and encouraging discussions at various stages of the work.

References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory.* Math. Surveys and Monographs **50**, Amer. Math. Soc., 1997.

[2] J. Aaronson, M. Denker and M. Urbański. Ergodic theory for Markov fibred systems and parabolic rational maps. *Trans. Amer. Math. Soc.* **337** (1993) 495–548.

[3] J. F. Alves and D. Azevedo. Statistical properties of diffeomorphisms with weak invariant manifolds. Preprint, 2013.

[4] A. Arbieto, R. Markarian, M. J. Pacifico and R. Soares. Scaling rate for semi-dispersing billiards with non-compact cusps. *Ergodic Theory Dynam. Systems* **32** (2012) 1836–1861.

[5] M. Blank, G. Keller and C. Liverani. Ruelle-Perron-Frobenius spectrum for Anosov maps. *Nonlinearity* **15** (2002) 1905–1973.

[6] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Lecture Notes in Math. **470**, Springer, Berlin, 1975.

[7] J.-R. Chazottes and S. Gouëzel. Optimal concentration inequalities for dynamical systems. *Commun. Math. Phys.* **316** (2012) 843–889.

[8] R. A. Doney. One-sided local large deviation and renewal theorems in the case of infinite mean. *Probab. Theory Related Fields* **107** (1997) 451–465.

[9] A. Garsia and J. Lamperti. A discrete renewal theorem with infinite mean. *Comment. Math. Helv.* **37** (1962/1963) 221–234.

[10] S. Gouëzel. Private communication.

[11] S. Gouëzel. Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes. Ph. D. Thesis, Ecole Normale Supérieure, 2004.

[12] S. Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 997–1024.

[13] S. Gouëzel. Correlation asymptotics from large deviations in dynamical systems with infinite measure. *Colloq. Math.* **125** (2011) 193–212.

[14] H. Hu. Conditions for the existence of SBR measures for "almost Anosov" diffeomorphisms. *Trans. Amer. Math. Soc.* **352** (2000) 2331–2367.

[15] H. Hu and L.-S. Young. Nonexistence of SBR measures for some diffeomorphisms that are "almost Anosov". *Ergodic Theory Dynam. Systems* **15** (1995) 67–76.

[16] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999) 671–685.

[17] C. Liverani and D. Terhesiu. Mixing for some non-uniformly hyperbolic systems. Preprint.

[18] I. Melbourne. Rapid decay of correlations for nonuniformly hyperbolic flows. *Trans. Amer. Math. Soc.* **359** (2007) 2421–2441.

[19] I. Melbourne and D. Terhesiu. Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Invent. Math.* **189** (2012) 61–110.

[20] I. Melbourne and D. Terhesiu. Decay of correlations for nonuniformly expanding systems with general return times. *Ergodic Theory Dynam. Systems* **34** (2014) 893–918.

[21] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197.

[22] D. Ruelle. *Thermodynamic Formalism.* Encyclopedia of Math. and its Applications **5**, Addison Wesley, Massachusetts, 1978.

[23] Y. G. Sinaĭ. Gibbs measures in ergodic theory. *Russ. Math. Surv.* **27** (1972) 21–70.

[24] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967) 747–817.

[25] D. Terhesiu. Improved mixing rates for infinite measure preserving systems. *Ergodic Theory Dynam. Systems.* DOI:10.1017/etds.2013.59.

[26] D. Terhesiu. Mixing rates for intermittent maps of high exponent. Preprint, 2012.

[27] M. Thaler. A limit theorem for the Perron-Frobenius operator of transformations on $[0, 1]$ with indifferent fixed points. *Israel J. Math.* **91** (1995) 111–127.

[28] R. F. Williams. One-dimensional non-wandering sets. *Topology* **6** (1967) 473–487.

[29] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

[30] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188.

[31] R. Zweimüller. Ergodic properties of infinite measure-preserving interval maps with indifferent fixed points. *Ergodic Theory Dynam. Systems* **20** (2000) 1519–1549.