Subexponential Decay of Correlations for Compact Group Extensions of Nonuniformly Expanding Systems

Henk Bruin
Department of Maths and Stats
University of Surrey
Guildford GU2 7XH, UK

Mark Holland
Department of Maths and Stats
University of Surrey
Guildford GU2 7XH, UK

Ian Melbourne
Department of Maths and Stats
University of Surrey
Guildford GU2 7XH, UK

23 July, 2004; Revised 8 October, 2004

Abstract

We show that the renewal theory developed by Sarig and Gouëzel in the context of nonuniformly expanding dynamical systems applies also to the study of compact group extensions of such systems. As a consequence, we obtain results on subexponential decay of correlations for equivariant Hölder observations.

1 Introduction

Suppose that $f : X \to X$ is a discrete dynamical system with ergodic measure μ. If $\phi, \psi : X \to \mathbb{R}$ lie in $L^2(X)$, we define the correlation function

$$\rho_{\phi,\psi}(n) = \int_X \phi(\psi \circ f^n)d\mu - \int_X \phi \, d\mu \int_X \psi \, d\mu.$$

The dynamical system is mixing if $\rho_{\phi,\psi}(n) \to 0$ as $n \to \infty$ for all $\phi, \psi \in L^2(X)$. For certain classes of dynamical systems and sufficiently regular observations ϕ, ψ, it is possible to estimate the speed at which $\rho_{\phi,\psi}(n) \to 0$. For Axiom A diffeomorphisms, it is known that the correlation function decays exponentially for Hölder observations (see for example [3, 20, 19]).

The early proofs in the uniformly hyperbolic case revolve around quasicompactness of a certain transfer operator. This method also applies to certain hyperbolic systems with singularities and to certain nonuniformly hyperbolic situations. Such systems can often be modelled by the tower construction of Young [23] and then exponential return asymptotics guarantee the existence of a "physical" measure μ and exponential decay of correlations.

Several methods have been developed to deal with the case where the rate of decay of correlations are slower than exponential. (For a recent survey, see Baladi [2].) These methods include Birkhoff cones [14, 22] and probabilistic coupling [24]. In particular, Young towers with subexponential return asymptotics have subexponential (stretched exponential or polynomial) decay of correlations [24]. Sarig [21] used renewal sequences to obtain lower bounds for decay of correlations and this method was generalised by Gouëzel [7]. In particular, the results of [7, 21] show that the subexponential decay rates of Young [24] are optimal. (See also [12, Theorem 4.3].)

In this paper, we are interested in group extensions $X \times G$ where G is a compact connected Lie group with Haar measure ν and $f : X \to X$ is a dynamical system of the type described above with ergodic measure μ. Given a Hölder cocycle $h : X \to G$, we define the G-extension $f_h : X \times G \to X \times G$ by $f_h(x, g) = (fx, gh(x))$. The product measure $m = \mu \times \nu$ is f_h-invariant and is ergodic/mixing under mild hypotheses on f and h (see [5] for the case when X is uniformly hyperbolic). We take mixing as given in this paper, and direct attention to the rate of mixing. For general Hölder observations $\phi, \psi : X \times G \to \mathbb{R}$, existing results are restricted to the case when X is uniformly hyperbolic and either G is semisimple or X is infranil Anosov, see Dolgopyat [4]. Nicol *et al.* [18] introduced a class of *equivariant* observations $\phi : X \times G \to \mathbb{R}^d$ of the form $\phi(x, g) = g \cdot v(x)$ where $\mathbb{R}^d$ is a representation of G and $v : X \to \mathbb{R}^d$. The statistics of such observations arise naturally in dynamical systems with Euclidean symmetry [18]. The correlation function $\rho_{\phi,\psi}(n)$ is defined as before but with ψ replaced by ψ^T, and so takes values in the space of $d \times d$ matrices.

Results on exponential decay of correlation for equivariant observations on compact group extensions were obtained by [6] in the case when X is uniformly hyperbolic. This was extended in [15] to include the general situation where the transfer operator is quasicompact for the X dynamics. In a sense, exponential decay on X leads to exponential decay on $X \times G$ for equivariant observations (more accurately, the proof of exponential decay on X by quasicompactness leads to the result on $X \times G$).

An important open problem is to obtain results on subexponential decay of correlations for sufficiently regular observations of compact group extensions. Previously, there were no such results even for equivariant observations. In this paper, we deduce subexponential decay results for equivariant observations on $X \times G$ in certain situations where subexponential decay can be proved on X. (Again, it is the proof that extends, not the result directly.) The technique of proof is perhaps unexpected. The Hilbert cones and probabilistic coupling methods mentioned above fail for equivariant observations — Hilbert cones uses positivity of the transfer operator; in coupling the observation is viewed as the density for a probability measure; neither makes sense here. Instead we use the renewal sequence method of Sarig [21] and Gouëzel [7] for obtaining upper bounds for decay rates, even though this method was introduced for obtaining lower bounds. (The fact that it can be used for obtaining upper bounds for general Hölder observations was pointed out in Gouëzel [10].) In fact, the method is much easier to apply in our context than in the nonequivariant situation, but we do not obtain lower bounds. (These two statements are related since the leading term in [7, 21] vanishes in the equivariant case.)

(a) Statement of the main result

Let (X, d) be a locally compact separable bounded metric space with Borel probability measure η and let $f : X \to X$ be a nonsingular transformation for which η is ergodic. Let $Y \subset X$ be a measurable subset with $\eta(Y) > 0$. We suppose that there is an at most countable measurable partition $\{Y_j\}$ with $\eta(Y_j) > 0$, and that there exist integers $r_j \geq 1$, and constants $\lambda > 1$; $C, D > 0$ and $\gamma \in (0, 1)$ such that for all j,

(1) $f^{r_j} : Y_j \to Y$ is a (measure-theoretic) isomorphism.

(2) $d(f^{r_j}x, f^{r_j}y) \geq \lambda d(x, y)$ for all $x, y \in Y_j$.

(3) $d(f^k x, f^k y) \leq C d(f^{r_j}x, f^{r_j}y)$ for all $x, y \in Y_j$, $k < r_j$.

(4) $g_j = \frac{d(\eta|_{Y_j} \circ (f^{r_j})^{-1})}{d\eta|_Y}$ satisfies $|\log g_j(x) - \log g_j(y)| \leq D d(x, y)^\gamma$ for almost all $x, y \in Y$.

(5) $\sum_j r_j \eta(Y_j) < \infty$.

We say that a dynamical system f satisfying (1)–(5) is *nonuniformly expanding.*

Define the *return time function* $r : Y \to \mathbb{N}$ by $r|_{Y_j} \equiv r_j$. Condition (5) says that $\int_Y r \, d\eta < \infty$. The map $f^Y : Y \to Y$ given by $f^Y(y) = f^{r(y)}(y)$ is the corresponding *induced map.* (By condition (2), f^Y is uniformly expanding.) It can be shown (see Young [24, Theorem 1]) that there is a unique invariant probability measure μ on X that is equivalent to η.

Let G be a compact connected Lie group with Haar measure ν. Given a measurable cocycle $h : X \to G$, we define the G-extension $f_h : X \times G \to X \times G$ with f_h invariant measure $m = \mu \times \nu$. Forward iterates are given for $n \geq 1$ by

$$f_h^n(x, g) = (f^n x, g h_n(x)), \quad \text{where } h_n(x) = h(x)h(fx) \cdots h(f^{n-1}x).$$

In particular, we obtain the induced transformation on $Y \times G$ given by $(y,g) \mapsto (f^Y y, gh^Y(y))$ where $h^Y : Y \to G$ is defined to be $h^Y(y) = h_{r(y)}(y)$.

Now let $\mathbb{R}^d$ be an orthogonal representation of G (so G can be viewed as a closed subgroup of the $d \times d$ orthogonal matrices). We say that $\phi : X \times G \to \mathbb{R}^d$ is an equivariant observation if $\phi(x,g) = g \cdot v(x)$ where $v : X \to \mathbb{R}^d$. Note that $\phi \in L^\infty(X \times G)$ if and only if $v \in L^\infty(X)$ in which case $|\phi|_\infty = |v|_\infty$. If $v : X \to \mathbb{R}^d$ is γ-Hölder, then we say that ϕ is Hölder and we define $\|\phi\|_\gamma = |v|_\infty + |v|_\gamma$ where $|v|_\gamma = \sup_{x \neq y} |v(x) - v(y)|/d(x,y)^\gamma$.

Theorem 1.1 *Let $f_h : X \times G \to X \times G$ be a weak mixing compact group extension of a nonuniformly expanding map as above, where $h : X \to G$ is a Hölder cocycle. Assume that*

$$\mu(y \in Y : r_j(y) \geq n) = O(n^{-(\beta+1)}),$$

for some $\beta > 1$. Let G act orthogonally on $\mathbb{R}^d$. Then there exists a constant $C > 0$ such that for all equivariant observations $\phi, \psi : X \times G \to \mathbb{R}^d$ with ϕ Hölder and $\psi \in L^\infty$,

$$\rho_{\phi,\psi}(n) \leq C\|\phi\|_\gamma \, |\psi|_\infty n^{-\beta}.$$

Similar results hold for more general decay rates, including stretched exponential (see Section 2(a)).

Remark 1.2 (a) Define $\mathrm{Fix}\, G = \{v \in \mathbb{R}^d : gv = v \text{ for all } g \in G\}$. We can always decompose $\mathbb{R}^d = \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ where G acts trivially on $\mathbb{R}^{d_1}$ ($\mathrm{Fix}\, G = \mathbb{R}^{d_1}$) and fixed-point freely on $\mathbb{R}^{d_2}$ ($\mathrm{Fix}\, G = \{0\}$). When G acts trivially, we are in the situation studied by [7, 21] so we focus on the case $\mathrm{Fix}\, G = \{0\}$.
(b) Suppose that $\mathrm{Fix}\, G = \{0\}$ and that ϕ is supported in $Y \times G$. Then we obtain the improved estimate $\rho_{\phi,\psi}(n) \leq C\|\phi\|_\gamma \, |\psi|_\infty n^{-(\beta+1)}$ if the cocycle h is either (i) locally constant, or (ii) supported in Y. See Remark 5.6.

Remark 1.3 A standard application of decay of correlations is to deduce statistical limit laws such as the central limit theorem or the almost sure invariance principle. However, stronger results can be obtained more easily by observing that the full map $f_h : X \times G \to X \times G$ is a discrete suspension over the induced map on $Y \times G$. By [17], the limit laws on $X \times G$ can be deduced from the corresponding laws on $Y \times G$ (see [8]). The CLT and ASIP on $X \times G$ is proved in [16] under hypotheses more general than in Theorem 1.1.

The remainder of the paper is organised as follows: In Section 2, we prove a simplified special case of the Renewal Theorem of [7, 21] which suffices for our purposes. Then we turn to the proof of Theorem 1.1. Various parts of the proof are best treated at different levels of abstraction, so in Section 3 we relate nonuniformly expanding maps to Young towers and Gibbs-Markov maps.

In Section 4, we describe the set up of inducing in the context of compact group extensions and we use the renewal theorem to relate properties of the dynamics on $X \times G$ to properties of the induced system on $Y \times G$. At this point, conditions (1)–(5) are not required.

In Section 5, we prove Theorem 1.1 for Hölder observations supported in $Y \times G$. This is done by proving the corresponding result when $f : X \to X$ is a Markov map and $f^Y : Y \to Y$ is a Gibbs-Markov map. This is certainly the case for nonuniformly expanding maps (Condition (3) is not required here.)

In Section 6, we prove Theorem 1.1 for Hölder observations supported in $X \times G$. This is done by proving the corresponding results for Young towers. The result for nonuniformly expanding maps follows immediately from the result for Young towers.

2 A simplified renewal theorem

Let $\mathrm{Hom}(\mathcal{B})$ denote the space of bounded linear operators on a Banach space $\mathcal{B}$. Let $R_n \in \mathrm{Hom}(\mathcal{B})$. We assume that

(H1) $\sum_{j=n}^{\infty} \|R_j\| = O(n^{-\beta})$, where $\beta \geq 0$.

Set $R(\omega) = \sum_{n=1}^{\infty} R_n e^{in\omega}$. It follows from (H1) that $R : S^1 \to \mathrm{Hom}(\mathcal{B})$ is a well-defined map. Next, we assume that

(H2) The spectrum of $R(\omega)$ does not contain 1 for all $\omega \in S^1$.

Define $T(\omega) = (I - R(\omega))^{-1}$. Hypothesis (H2) guarantees that $T : S^1 \to \mathrm{Hom}(\mathcal{B})$ is well-defined.

Theorem 2.1 *Assume that (H1) and (H2) are valid with $\beta > 0$ not an integer. Then the Fourier coefficients T_n of T satisfy $\|T_n\| = O(n^{-\beta})$.*

Corollary 2.2 *Assume that (H1) and (H2) are valid with $\beta > 0$ and $\beta \neq 1$. Then the Fourier coefficients T_n of T satisfy $\|T_n\| = O(n^{-\beta})$.*

Proof By Theorem 2.1, we may assume that $\beta \in \{2, 3, \ldots\}$. In particular, the Fourier coefficients $\|T_n\|$ are summable. Now apply [7, Lemma 4.5]. ∎

Remark 2.3 (1) Much of the difficulty, and hence depth, of the work in [7, 21] stems from the fact that 1 is automatically an eigenvalue of $R(0)$ in their context. Hence (H2) is violated, necessitating a combination of complex analytic techniques

and Fourier analysis. In our work, Fourier analysis alone suffices.
(2) As the proof of Corollary 2.2 shows, it suffices to show that Theorem 2.1 holds for $\beta \in (0,1)$ and that $\|T_n\|$ is summable for $\beta > 1$. However the details for $\beta > 1$ not an integer are elementary and so are included. In addition, the estimates in Theorem 2.1 are explicit (this is clear in the proof) which may be of use for estimating mixing times (though we do not pursue that issue here).

(a) Stretched exponential rates and convolutive sequences

Corollary 2.2 can be generalised significantly using ideas in the Ph. D. thesis of Gouëzel [10, Définition 2.2.10]. Adapting the definitions there, we say that a sequence of positive numbers $u_n > 0$ is *convolutive* if $\sum_{n=1}^{\infty} u_n < \infty$ and there exists a constant $C > 0$ such that

(i) $(u * u)_n \le C u_n$ for all $n \ge 1$,

(ii) $u_m \le C u_n$ for all $m \ge n$,

(iii) $(1-\epsilon)^n = O(u_n)$ for all $\epsilon > 0$,

(iv) for all $\epsilon > 0$, there exists $N \ge 1$ such that the sequence $v_n = 1_{n \ge N} u_n$ satisfies $(v * v)_n \le \epsilon u_n$ for all $n \ge 1$.

For u_n convolutive, it follows from Gouëzel [10] that if $\sum_{j=n}^{\infty} \|R_j\| = O(u_n)$ and (H2) is valid, then $\|T_n\| = O(u_n)$.

It is easily verified that $u_n = n^{-\beta}$ is convolutive for $\beta > 1$, so we recover Corollary 2.2 in this case. Also, stretched exponential sequences $u_n = \tau^{n^\gamma}$ with $\tau, \gamma \in (0,1)$ are convolutive, so we obtain a stretched exponential version of Corollary 2.2.

The above definition of convolutive sequence is simpler than in [10], so it is worth mentioning where the various conditions are used. The method in [10] is to apply a generalised version of the Wiener lemma from harmonic analysis (if $f \in L^1(S^1, \mathbb{C})$ is nonvanishing and has summable Fourier coefficients, then f^{-1} has summable Fourier coefficients) [13, p. 202]. Conditions (i) and (ii) guarantee the existence of a suitable Banach algebra. Conditions (iii) and (iv) ensure that the Wiener lemma holds in this Banach algebra. The remaining conditions in [10] are not required here due to the simplification mentioned in Remark 2.3(1).

(b) Proof of Theorem 2.1

In the remainder of this section, we prove Theorem 2.1. Throughout, if $\beta \ge 0$ then we write $\beta = k + \alpha$ where $k = [\beta]$ and $\alpha \in [0,1)$.

Lemma 2.4 *Assume that* $R(\omega) = \sum_{n=1}^{\infty} R_n e^{in\omega}$ *with* $\sum_{j=n}^{\infty} \|R_j\| = O(n^{-\beta})$ *where* $\beta > 0$ *is not an integer. Then* $R : S^1 \to \mathbb{R}$ *is* C^β.

Proof First suppose that $\beta = \alpha \in [0,1)$. We follow [7, 21]. Let $\omega_1, \omega_2 \in S^1$ and fix $N \geq 1$. Write $S_n = \sum_{j=n}^{\infty} \|R_j\|$. Then

$$\begin{aligned}\|R(\omega_1) - R(\omega_2)\| &\leq \sum_{n=1}^{N-1} |e^{in\omega_1} - e^{in\omega_2}| \|R_n\| + 2\sum_{n=N}^{\infty} \|R_n\| \\ &\leq |\omega_1 - \omega_2| \sum_{n=1}^{N-1} n(S_n - S_{n+1}) + 2\sum_{n=N}^{\infty}(S_n - S_{n+1}) \\ &\leq |\omega_1 - \omega_2| \sum_{n=1}^{N-1} S_n + 2S_N \leq C|\omega_1 - \omega_2| \sum_{n=1}^{N-1} n^{-\alpha} + 2CN^{-\alpha} \\ &\leq C(1-\alpha)^{-1} N^{-\alpha}\Big\{N|\omega_1 - \omega_2| + 2(1-\alpha)\Big\}.\end{aligned}$$

Let $N = 1/|\omega_1 - \omega_2| + c$ where $0 \leq c < 1$. Then $N^{-\alpha} \leq |\omega_1 - \omega_2|^\alpha$ and $N|\omega_1 - \omega_2| \leq 1 + 2\pi$ so that $\|R(\omega_1) - R(\omega_2)\| \leq C(1-\alpha)^{-1}(3 + 2\pi - 2\alpha)|\omega_1 - \omega_2|^\alpha$, as required.

Next suppose that $\beta = k + \alpha \geq 1$. Repeatedly differentiating the power series for R yields $R^{(k)}(\omega) = \sum_{n=1}^{\infty} n^k R_n e^{in\omega}$. Let $E_n = \sum_{j\geq n} j^k \|R_j\|$. We claim that $E_n = O(n^{-\alpha})$. It follows that $R^{(k)}$ is C^α as required.

It remains to prove the claim. We have $S_n = \sum_{j\geq n} \|R_j\| \leq Cn^{-(k+\alpha)}$. Compute that

$$E_n = \sum_{j\geq n} j^k \|R_j\| = \sum_{j\geq n} j^k (S_j - S_{j+1}) = n^k S_n + \sum_{j\geq n+1} (j^k - (j-1)^k) S_j$$

Using the identity $x^k - y^k = (x-y)(x^{k-1} + x^{k-2}y + \cdots + y^{k-1})$, we see that $j^k - (j-1)^k \leq kj^{k-1}$. Hence $E_n \leq Cn^{-\alpha} + Ck\sum_{j\geq n+1} j^{-(\alpha+1)} \leq C(1 + k/\alpha)n^{-\alpha}$. ∎

Lemma 2.5 *Suppose that* $T : S^1 \to \mathrm{Hom}(\mathcal{B})$ *is* C^β *where* $\beta \geq 0$ *and let* $T_n \in \mathrm{Hom}(\mathcal{B})$ *denote the Fourier coefficients of* T. *Then* $\|T_n\| = O(n^{-\beta})$.

Proof First suppose that $\beta = \alpha \in [0,1)$. The estimate is a standard result for $\mathcal{B} = \mathbb{C}$ (see [13, p. 25]) which easily generalises as in Sarig [21].

Next suppose that $\beta = k + \alpha \geq 1$. Integrating by parts k times yields

$$T_n = \frac{1}{2\pi} \int_0^{2\pi} T^{(k)}(\omega) \frac{e^{-in\omega}}{(in)^k} d\omega.$$

Now $E_n = (in)^k T_n$ are the Fourier coefficients of the C^α function $T^{(k)}$ and so $\|E_n\| = O(n^{-\alpha})$. Hence $\|T_n\| = O(n^{-\beta})$ as required. ∎

Proof of Theorem 2.1 By Lemma 2.4, R is C^β, and it follows from (H2) that $T = (I-R)^{-1}$ is C^β. The result follows from Lemma 2.5. ∎

3 Background on Young towers and Gibbs-Markov maps

As mentioned earlier, certain parts of this paper are best formulated at various different levels of abstraction. Eventually, we prove Theorem 1.1 for G-extensions of a Young tower which is an effective model for nonuniformly expanding maps. The result for nonuniformly expanding maps follows from the result for towers, but the tower has an additional Markov structure which simplifies the proofs. The base of the tower has a Gibbs-Markov structure. We recall background material about Young towers in Subsection (a) and about Gibbs-Markov maps in Subsection (b).

(a) Young towers

Let (Δ_0, m_0) be a probability space and let $F_0 : \Delta_0 \to \Delta_0$ be a measure-preserving transformation. Given an L^1 function $r : \Delta_0 \to \mathbb{N}$ with $r > 0$ almost everywhere (such an r is called a roof function), we define the suspension

$$\Delta = \{(x, \ell) \in \Delta_0 \times \mathbb{N} : 0 \le \ell \le r(x)\}/\sim$$

where $(x, r(x)) \sim (F_0 x, 0)$. Define the suspension map $F : \Delta \to \Delta$ by $F(x, \ell) = (x, \ell + 1)$ computed subject to identifications. An F-invariant probability measure is given by $m = m_0 \times \xi / \int_{\Delta_0} r \, dm_0$ where ξ is counting measure.

Let $\{\Delta_{j,0}\}$ be an at most countable measurable partition of Δ_0 that separates points in Δ_0, such that $r_j = r|_{\Delta_{j,0}}$ is constant, and such that $F_0 : \Delta_{j,0} \to \Delta_0$ is a measurable isomorphism. For each j and $0 \le \ell < r_j$, let $\Delta_{j,\ell} = \Delta_{j,0} \times \{\ell\}$. This defines a partition $\{\Delta_{j,\ell}\}$ of Δ.

Define a *separation time* function $s : \Delta \times \Delta \to \mathbb{N}$ as follows: If x, y lie in distinct partition elements, then $s(x, y) = 0$. If $x, y \in \Delta_{j,0}$ for some j, then $s(x, y)$ is the greatest integer $n \ge 0$ such that $F_0^n x$ and $F_0^n y$ lie in the same partition element of Δ_0. Finally, if $x, y \in \Delta_{j,\ell}$ then write $x = F^\ell x_0$, $y = F^\ell y_0$ where $x_0, y_0 \in \Delta_{j,0}$ and define $s(x, y) = s(x_0, y_0)$. Since $\{\Delta_{j,0}\}$ separates points, $s(x, y) = \infty$ if and only if $x = y$. Hence $d_\theta(x, y) = \theta^{s(x,y)}$ defines a metric on Δ for any choice of $\theta \in (0, 1)$.

Definition 3.1 Let g_j be the Jacobian of $F_0|_{\Delta_{j,0}} : \Delta_{j,0} \to \Delta_0$. The suspension $f : \Delta \to \Delta$ is a called a *Young tower* if there exist constants $\theta \in (0,1)$ and $C > 0$ such that for each j, $|\log g_j(x) - \log g_j(y)| \le C d_\theta(x,y)$ for all $x, y \in \Delta_{j,0}$.

Now suppose that $f : X \to X$ is a nonuniformly expanding map with induced uniformly expanding map $f^Y : Y \to Y$ as described in the introduction. In particular, suppose that $\lambda > 1$ and $\gamma \in (0,1)$ are the constants appearing in conditions (2) and (4). Let $\Delta = \{(y, \ell) : y \in Y, \ \ell = 0, \dots, r(y) - 1\}$, so Δ is the disjoint union of r_j copies of each Y_j. Define a measure μ on Δ by setting $\mu|_{Y_j \times \{\ell\}} = m|_{Y_j}$. Define $F : \Delta \to \Delta$ by setting $F(y, \ell) = (y, \ell + 1)$ for $0 \le \ell < r(y) - 1$ and $F(y, r(y) - 1) = (f^Y y, 0)$.

Proposition 3.2 *Let $\theta = 1/\lambda^\gamma$. Then the map $F : \Delta \to \Delta$ is a tower with $\Delta_0 = Y$, $F_0 = f^Y$ and $\Delta_{j,0} = Y_j$. Moreover, there is a constant $C' \ge 1$ such that $d(x,y) \le C'[d_\theta(x,y)]^{1/\gamma}$ for all $x, y \in \Delta$.*

Proof It follows from condition (2) for nonuniform expansion that $d(x,y) \le \operatorname{diam}(Y)/\lambda^{s(x,y)}$ for all $(x,y) \in \Delta$. Hence the partition $\Delta_{j,0}$ of Δ_0 separates points and the required distortion condition on g_j is immediate, so Δ is a tower.

If x, y lie in the same partition element of Δ_ℓ, then write $x = f^\ell x_0$, $y = f^\ell y_0$ so $d(f^Y x_0, f^Y y_0) \le \operatorname{diam}(Y)/\lambda^{s(x,y)}$. By condition (3) for nonuniform expansion,

$$d(x,y) \le C d(f^Y x_0, f^Y y_0) \le C \operatorname{diam}(Y)/\lambda^{s(x,y)} = C \operatorname{diam}(Y)[d_\theta(x,y)]^{1/\gamma}.$$

The last statement of the proposition follows. ∎

Define the projection $\pi : \Delta \to X$ by $\pi(y, \ell) = f^\ell(y)$. Then π is a measure-preserving isomorphism and it follows as in the proof of Proposition 3.2 that

$$d(\pi(x), \pi(y))^\gamma \le C'' d_\theta(x,y).$$

for all $(x,y) \in \Delta$. In particular, if $v : (X, d) \to \mathbb{R}$ is globally γ-Hölder, then $v \circ \pi : (\Delta, d_\theta) \to \mathbb{R}$ is globally Lipschitz.

We conclude that Theorem 1.1 for G-extensions of nonuniformly expanding maps is equivalent to the corresponding theorem for tower maps (with Hölder observations replaced by Lipschitz observations).

(b) Markov maps and Gibbs-Markov maps

A background reference for Markov maps is [1, Section 4.7]. We follow the presentation in [7]. Let (X, μ) be a Lebesgue space. Recall that a measure-preserving transformation $f : X \to X$ is a *Markov map* if there is a measurable partition α

of X such that if $a \in \alpha$ with $\mu(a) > 0$, then fa is a union of elements of α and $f|_a$ is injective. Moreover, it is assumed that the partition α separates points. If $a_0, \dots, a_{n-1} \in \alpha$, we define the cylinder $[a_0, \dots, a_{n-1}] = \cap_{i=0}^{n-1} f^{-i} a_i$.

Suppose that $Y = \cup_{a \in \widetilde{\alpha}} a$ is a union of elements of α, and define the induced map $f^Y : Y \to Y$. This is a measure-preserving transformation with respect to $\mu^Y = \mu|_Y$. Moreover, f^Y is a Markov map with respect to the partition β consisting of all cylinders for $f : X \to X$ of the form $b = [a, \xi_1, \dots, \xi_{n-1}, Y]$ where $a \in \widetilde{\alpha}, \xi_1, \dots, \xi_{n-1} \notin \widetilde{\alpha}$.

If $b_0, \dots, b_{n-1} \in \beta$, we define the n-cylinder $[b_0, \dots, b_{n-1}]_Y = \cap_{i=0}^{n-1} (f^Y)^{-i} b_i$. These cylinders can be used to define a metric on Y in terms of separation times. Fix $\theta \in (0,1)$, and define $d_\theta(x,y) = \theta^{s(x,y)}$ where $s(x,y)$ is the greatest integer $n \geq 0$ such that x, y lie in the same n-cylinder.

We shall suppose that the induced map is additionally a *Gibbs-Markov* map satisfying the following properties:

(i) *Big images property:* There exists $c > 0$ such that $\mu(f^Y b) \geq c$ for all $b \in \beta$.

(ii) *Distortion:* There exists $\theta \in (0,1)$ such that that $p|_b : b \to \mathbb{R}$ is Lipschitz with respect to d_θ for all $b \in \beta'$ where $p(x) = p^Y(x) = \log \dfrac{d\mu^Y}{d\mu^Y \circ f^Y}$ and β' denotes the smallest partition of Y such that $f^Y b$ is a union of atoms in β' for all $b \in \beta$.

Remark 3.3 Clearly, conditions (1), (2) and (4) for nonuniformly expanding maps guarantee that the induced map $f^Y : Y \to Y$ (or the map $F_0 : \Delta \to \Delta_0$ for the tower map in Subsection (a)) is Gibbs-Markov with respect to the partition $\{Y_j\}$ with $\theta = 1/\lambda^\gamma$. By condition (1), the big images property is trivially satisfied since each Y_j is mapped onto the whole of Y. Condition (2) guarantees that the partition separates points. Condition (4) gives the distortion, indeed $\beta' = \{Y\}$, and p is globally Lipschitz.

It follows in a standard manner from assumptions (i) and (ii) that there exists a constant $D > 0$ such that for all $x, y \in [b_0, \dots, b_{k-1}]_Y$,

$$\left| \frac{e^{p_k(x)}}{e^{p_k(y)}} - 1 \right| \leq D\theta^{-k} d_\theta(x,y) \quad \text{and} \quad D^{-1} \leq \frac{\mu[b_0, \dots, b_{k-1}]_Y}{e^{p_k(x)}} \leq D, \tag{3.1}$$

where $p_k(x) = p(x) + p(f^Y x) + \cdots + p((f^Y)^{k-1} x)$.

4 Compact group extensions

In this section, we begin by recalling the formalism of inducing for discrete dynamical systems in the context of compact group extensions and equivariant observations.

Then we apply the renewal theorem (Corollary 2.2) to this situation. We also prove a partial result towards the verification of hypothesis (H2).

(a) Inducing and compact group extensions

Let (X, μ) be a probability space, $f : X \to X$ a measure preserving transformation, and $Y \subset X$ a measurable subset with $\mu(Y) > 0$. By Poincaré recurrence, for almost every point y there is an integer $n \geq 1$ such that $f^n y \in Y$. Let Z_n consist of those $y \in Y$ for which n is the least integer such that $f^n y \in Y$. Then we have the measurable partition $Y = \cup_{n\geq 1} Z_n$ and we define the induced transformation $f^Y : Y \to Y$ by setting $f^Y y = f^n y$ for $y \in Z_n$. (To make the connection with nonuniformly expanding maps, Z_n is the union of those subsets Y_j with $r_j = n$.)

Given a cocycle $h : X \to G$ with corresponding G-extension $f_h : X \times G \to X \times G$, we define (as in the Introduction) the induced G-extension $(f_h)^{Y\times G} : Y \times G \to Y \times G$. Note that $(f_h)^{Y\times G} = (f^Y)_{h^Y}$ where

$$h^Y = h_n = h \cdot h \circ f \cdots h \circ f^{n-1} \text{ on } Z_n.$$

We consider equivariant observations $\phi : X \times G \to \mathbb{R}^d$ of the form $\phi(x, g) = gv(x)$ where $v \in L^1(X, \mathbb{R}^d)$. The standing assumption $\operatorname{Fix} G = \{0\}$ has the consequence that $\int_{X\times G} \phi \, dm = 0$ for all equivariant observations ϕ (since $\int_G gv \, d\nu = 0$ for all $v \in \mathbb{R}^d$).

Corresponding to the map $f : X \to X$, we define as usual the transfer (or Perron-Frobenius) operator L on $L^1(X)$: if $v \in L^1(X)$, then Lv is the unique element of $L^1(X)$ such that $\int_X Lv \, w \, d\mu = \int_X v \, w \circ f \, d\mu$ for all $w \in L^\infty(X)$. The operator L defines (componentwise) an operator $L : L^1(X, \mathbb{C}^d) \to L^1(X, \mathbb{C}^d)$ for all $d \geq 1$. Similarly, we obtain a transfer operator $\widehat{L} : L^1(X \times G, \mathbb{C}^d) \to L^1(X \times G, \mathbb{C}^d)$ corresponding to $f_h : X \times G \to X \times G$. We define the twisted transfer operator $L_h : L^1(X, \mathbb{C}^d) \to L^1(X, \mathbb{C}^d)$ by $L_h v = L(h^{-1}v)$ (taking the complexified action of G on $\mathbb{C}^d$).

The analogous definitions apply to the induced transformation $f^Y : Y \to Y$. We have the transfer operators $R : L^1(Y, \mathbb{C}^d) \to L^1(Y, \mathbb{C}^d)$ and $\widehat{R} : L^1(Y \times G, \mathbb{C}^d) \to L^1(Y \times G, \mathbb{C}^d)$, and and the twisted transfer operator $R_h : L^1(Y, \mathbb{C}^d) \to L^1(Y, \mathbb{C}^d)$ defined by $R_h v = R((h^Y)^{-1}v)$. (We avoid the more natural, but cumbersome, notation R_{h^Y}.)

Proposition 4.1 *Let $\phi(x, g) = gv(x)$ be an equivariant observation with $v \in L^1(X, \mathbb{C}^d)$ or $v \in L^1(Y, \mathbb{C}^d)$. Then $(\widehat{L}\phi)(x, g) = g(L_h v)(x)$ and $(\widehat{R}\phi)(y, g) = g(R_h v)(y)$.*

Proof This is standard, see for example [6]. ∎

We define the following bounded linear operators on $L^1(X, \mathbb{C}^d)$:

$$T_n v = L_h^n(1_Y v)1_Y, \qquad R_n v = L_h^n(1_{Z_n} v)1_Y.$$

Proposition 4.2 *(a)* $T_n = \sum_{i_1+\cdots+i_k=n} R_{i_k} \cdots R_{i_2} R_{i_1}$.

(b) Restricting to $L^1(Y)$, we have

$$R_h(e^{ir\omega}v) = \sum_{n=1}^{\infty} R_n v e^{in\omega}$$

where $r : Y \to \mathbb{N}$ is given by $r|_{Z_n} \equiv n$.

Proof Define the sequences of bounded operators $\widehat{T}_n$, $\widehat{R}_n$ on $L^1(X \times G)$ by

$$\widehat{T}_n \phi = \widehat{L}^n(1_{Y\times G}\phi)1_{Y\times G}, \qquad \widehat{R}_n \phi = \widehat{L}^n(1_{Z_n\times G}\phi)1_{Y\times G}. \tag{4.1}$$

It follows from Proposition A.1(a) that

$$\widehat{T}_n = \sum_{i_1+\cdots+i_k=n} \widehat{R}_{i_k} \cdots \widehat{R}_{i_2} \widehat{R}_{i_1}. \tag{4.2}$$

Let $\phi(x, g) = gv(x)$. Using Proposition 4.1 and the definitions in (4.1), we compute that

$$(\widehat{T}_n\phi)(x, g) = (\widehat{L}^n(1_{Y\times G}\phi)1_{Y\times G})(x, g) = g[(L_h^n 1_Y v)1_Y](x) = g(T_n v)(x).$$

Similarly, $(\widehat{R}_n\phi)(x, g) = g(R_n v)(x)$. Substituting into (4.2) yields part (a).

Next, define $\widehat{R}_\omega \phi = \widehat{R}(e^{ir\omega}\phi)$. It follows from Proposition A.1(b) that

$$\widehat{R}_\omega = \sum_{n\geq 1} \widehat{R}_n e^{in\omega} \tag{4.3}$$

By Proposition 4.1, $(\widehat{R}_\omega\phi)(x, g) = g(R_h(e^{ir\omega}v))(x)$ and again $(\widehat{R}_n\phi)(x, g) = g(R_n v)(x)$. Substituting these into (4.3) yields part (b). ∎

(b) The renewal theorem applied to equivariant observations

We continue to assume that $f : X \to X$ is a measure-preserving transformation with induced map $f^Y : Y \to Y$, that $f_h : X \times G \to X \times G$ is a compact group extension, and that G acts orthogonally on $\mathbb{R}^d$ with $\operatorname{Fix} G = 0$.

Theorem 4.3 *Let $\mathcal{B} \subset L^1(Y, \mathbb{C}^d)$ be a Banach space with norm $\| \ \|$ satisfying $|v|_1 \le \|v\|$ for all $v \in \mathcal{B}$, such that the operators T_n, R_n restrict to bounded linear operators on $\mathcal{B}$. Assume that hypotheses (H1) and (H2) from Theorem 2.1 are satisfied with $\beta > 1$.*

Then there is a constant $C > 0$ such that for all equivariant observations $\phi(x,g) = gv(x)$, $\psi(x,g) = gw(x)$ where $v \in \mathcal{B}$ and $w \in L^\infty(Y, \mathbb{R}^d)$,

$$\Big| \int_{X \times G} \phi \, \psi^T \circ f_h^n \, dm \Big| \le C \|v\| \, |w|_\infty n^{-\beta}$$

for all $n \ge 1$.

Proof View the operators T_n, R_n as lying in $\mathrm{Hom}(\mathcal{B})$ and let

$$R(\omega) = \sum_{n \ge 1} R_n e^{in\omega}, \qquad T(\omega) = (I - R(\omega))^{-1}.$$

Hypotheses (H1) and (H2) guarantee that the maps $R, T : S^1 \to \mathrm{Hom}(\mathcal{B})$ are well-defined and by Lemma 2.4 they are C^β. In particular, the series definition of $R(\omega)$ is absolutely convergent to $R(\omega)$ and since $\beta > 1$, $T(\omega)$ has an absolutely convergent Fourier series $T(\omega) = I + \sum_{n \ge 1} \widetilde{T}_n e^{in\omega}$. By Corollary 2.2, there is a constant $C > 0$ such that $\|\widetilde{T}_n\| \le C n^{-\beta}$. Equating coefficients in $T(I - R) = I$ yields $\widetilde{T}_n = \sum_{i_1 + \cdots + i_k = n} R_{i_k} \cdots R_{i_1}$ so it follows from Proposition 4.2(a) that $T_n = \widetilde{T}_n$. Hence $\|T_n\| \le C n^{-\beta}$.

The remainder of the proof is a straightforward computation using Proposition 4.1:

$$\rho(n) = \int_{X \times G} \phi \, \psi^T \circ f_h^n \, dm = \int_{X \times G} \widehat{L}^n \phi \, \psi^T \, dm = \int_{X \times G} g(L_h^n v) \, w^T g^{-1} \, dm,$$

so that $|\rho(n)| \le | \int_X L_h^n v \, w^T d\mu |$. Since v and w are supported in Y, we can write

$$|\rho(n)| \le |\textstyle\int_X T_n v \, w^T d\mu| \le |T_n v|_1 \, |w|_\infty.$$

But $|T_n v|_1 \le \|T_n v\| \le C n^{-\beta} \|v\|$ as required. ∎

(c) Ruling out eigenvalues on the unit circle

The next result is a step towards verifying hypothesis (H2). Recall that $R(\omega) = \sum_{n \ge 1} R_n e^{in\omega}$.

Proposition 4.4 *Suppose that* $\mathrm{Fix}\, G = \{0\}$ *and that* $f_h : X \times G \to X \times G$ *is weak mixing. Then for all* $\omega \in \mathbb{R}$ *the cohomological equation*

$$R(\omega)v = v,$$

has no nonzero L^2 *solutions* $v : Y \to \mathbb{C}^d$.

Proof By Proposition 4.2(b), it is equivalent to rule out solutions to the cohomological equation $R(e^{ir\omega}(h^Y)^{-1}v) = v$. The proof is standard for $\omega \neq 0$ (cf. Pollicott & Parry [19, Proposition 6.2]), and the case $\omega = 0$ follows as in [6]. The details are provided for completeness.

Let $U : L^2(Y, \mathbb{C}^d) \to L^2(Y, \mathbb{C}^d)$ denote the isometry $Uv = e^{-ir\omega} h^Y\, v \circ f^Y$ with adjoint $U^*v = R(e^{ir\omega}(h^Y)^{-1}v)$ satisfying $U^*U = I$. It is easy to see that $Uv = v$ is equivalent to $U^*v = v$ (see for example [15, Section 2]). Hence it suffices to show that $Uv = v$ has no nonzero solutions in $L^2(Y, \mathbb{C}^d)$.

Suppose for contradiction that $v \in L^2(Y, \mathbb{C}^d)$ is nonzero and $Uv = v$. Writing $\phi(y, g) = gv(y)$ we have

$$\phi \circ f^Y_{h^Y} = e^{ir\omega}\phi. \tag{4.4}$$

We can view $X \times G$ as a discrete suspension over $Y \times G$ by writing

$$X \times G = \{(y, g, j) \in Y \times G \times \mathbb{N} : 0 \leq j \leq r(y)\}/\sim,$$

where $(y, g, r(y)) \sim (f^Y y, gh^Y(y), 0)$. Then $f_h : X \times G \to X \times G$ is simply given by $f_h(y, g, j) = (y, g, j+1)$ computed modulo identifications. Define $\psi : Y \times G \times \mathbb{N} \to \mathbb{C}^d$ by setting $\psi(y, g, j) = e^{ij\omega}\phi(y, g)$. Condition (4.4) guarantees that ψ is well-defined as a map $\psi : X \times G \to \mathbb{C}^d$. Moreover, it is immediate that

$$\psi \circ f_h = e^{i\omega}\psi.$$

If $\omega \neq 0$, then this contradicts the assumption that f_h is weak mixing. If $\omega = 0$ then it follows from ergodicity of f_h that ψ is constant. Writing $\psi(x, g) = gw(x)$ (where $w(y, j) = e^{ij\omega}v(y)$), we obtain that $w(x) \in \mathrm{Fix}\, G = \{0\}$ for all $x \in X$ contradicting the assumption that v is nonzero. ∎

5 Observations supported in $Y \times G$

In this section, we prove a version of Theorem 1.1 for observations supported in $Y \times G$ under the assumption that the underlying dynamical system $f : X \to X$ is Markov and the induced dynamical system $f^Y : Y \to Y$ is Gibbs-Markov. As discussed in

Section 3(b), conditions (1), (2) and (4) in the definition of nonuniformly expanding map guarantee that $f^Y : Y \to Y$ is Gibbs-Markov. We assume the material from Section 3(b). In addition, we have $Y = \cup Z_n$ where the subsets Z_n are defined as in Section 4. Note that each Z_n is a union of elements of the partition β of Y defined in Section 3(b).

Let $|v|_\theta$ denote the Lipschitz constant for v (with respect to the metric d_θ) and let $\mathcal{B}$ denote the Banach space of functions $v : Y \to \mathbb{C}^d$ with norm $\|v\| = |v|_\infty + |v|_\theta < \infty$. The transfer operator $R : L^1(Y, \mathbb{C}^d) \to L^1(Y, \mathbb{C}^d)$ associated to the Gibbs-Markov map $f^Y : Y \to Y$ restricts to an operator $R : \mathcal{B} \to \mathcal{B}$ given by

$$(Rv)(x) = \sum_{f^Y y = x} e^{p(y)} v(y),$$

Let $h : X \to G$ be a measurable cocycle, and recall that $h^Y(x) = h_n(x) = h(x)h(fx) \cdots h(f^{n-1}x)$ for $x \in Z_n$. Throughout, h^Y_k means $(h^Y)_k$. Since $h(x)$ acts orthogonally on $\mathbb{R}^d$ for all $x \in X$, $|h^Y_k|_\infty = 1$ for all k. Define $|h^Y|_\theta$ as above, viewing $h^Y : Y \to G \subset \mathbb{R}^{d^2}$. Note that $|(h^Y)^{-1}|_\theta = |h^Y|_\theta$.

Let $\bar{b} = [b_0, \ldots, b_{k-1}]_Y$ be a k-cylinder for the induced map f^Y. If $x \in Y$, write $\bar{b}x = b_0 \cdots b_{k-1}x$.

Proposition 5.1 *Let $x, y \in Y$. Then $|h^Y_k(\bar{b}x) - h^Y_k(\bar{b}y)| \le \theta(1-\theta)^{-1}|h^Y|_\theta d_\theta(x, y)$.*

Proof Compute that

$$|h^Y_k(u) - h^Y_k(v)| \le \sum_{j=0}^{k-1} |h^Y|_\infty^{k-1} |h^Y((f^Y)^j u) - h^Y((f^Y)^j v)| \le \sum_{j=0}^{k-1} |h^Y|_\theta \theta^{-j} d_\theta(u, v).$$

The result follows since $d_\theta(\bar{b}x, \bar{b}y) = \theta^k d_\theta(x, y)$. ∎

For $x \in Y$, define

$$(M_{\bar{b}}v)(x) = e^{p_k(\bar{b}x)}(h^Y_k)^{-1}(\bar{b}x)v(\bar{b}x)$$

if the point $\bar{b}x = b_0 \cdots b_{k-1}x$ is defined, and zero otherwise. It is immediate from (3.1) that $|M_{\bar{b}}|_\infty \le D\mu(\bar{b})$.

Lemma 5.2 *There exists a constant $E > 0$ such that*

$$\|M_{\bar{b}}v\| \le E\mu(\bar{b})\{(1 + |h^Y|_\theta)|v|_\infty + \theta^k |v|_\theta\},$$

for all k-cylinders $\bar{b} = [b_0, \ldots, b_{k-1}]_Y$, $v \in \mathcal{B}$.

Proof See [21] for a proof in the absence of h. To simplify the notation, we suppress superscript and subscript Y's throughout the proof. Also, we write b instead of $\overline{b}$.

Let $x, y \in Y$ and compute that

$$\begin{aligned}|(M_b v)(x) - (M_b v)(y)| &= |e^{p_k(bx)} h_k^{-1}(bx) v(bx) - e^{p_k(by)} h_k^{-1}(by) v(by)| \\ &\le |1_b e^{p_k}|_\infty |v(bx) - v(by)| + |1_b e^{p_k}|_\infty |h_k^{-1}(bx) - h_k^{-1}(by)|\,|1_b v|_\infty \\ &\quad + |1_b e^{p_k}|_\infty \left| \frac{e^{p_k(bx)}}{e^{p_k(by)}} - 1 \right| |1_b v|_\infty\end{aligned}$$

and so by (3.1) and Proposition 5.1,

$$|(M_b v)(x) - (M_b v)(y)| \le D\mu(b)\{\theta^k |1_b v|_\theta + \theta(1-\theta)^{-1}|h|_\theta |1_b v|_\infty + D|1_b v|_\infty\} d_\theta(x, y).$$

Hence $|M_b v|_\theta \le D\mu(b)\{\theta^k |1_b v|_\theta + \theta(1-\theta)^{-1}|h|_\theta |1_b v|_\infty + D|1_b v|_\infty\}$. Combining this with the estimate for $|M_b|_\infty$ yields the required result. ∎

Remark 5.3 It can be seen from the proof of Lemma 5.2 that if $k = 1$, then $|h^Y|_\theta$ can be replaced by $|1_{\overline{b}} h^Y|_\theta$.

Corollary 5.4 *(a) Let $E > 0$ be the constant in Lemma 5.2. For all $\omega \in S^1$, the linear operator $R(\omega) : \mathcal{B} \to \mathcal{B}$ is bounded, and satisfies*

$$\|R(\omega)^k v\| \le E((1 + |h^Y|_\theta)|v|_\infty + \theta^k |v|_\theta),$$

for all $k \ge 1$.
(b) $\|R_n\| = O\Big(\sum_{\overline{b} \in Z_n} \mu(\overline{b})(1 + |1_{\overline{b}} h_n|_\theta)\Big)$, *where the sum is over all 1-cylinders $\overline{b} \in Z_n$.*

Proof Recall from Proposition 4.2(b) that $R(\omega)v = R(e^{ir\omega}(h^Y)^{-1}v)$ where R is the transfer operator for $f^Y : Y \to Y$. Write $R(\omega)^k = \sum_{\overline{b}} M_{\overline{b}} e^{ir_k\omega}$ where the sum is over all k-cylinders $\overline{b}$. Since r_k is constant on $\overline{b}$, the term $e^{ir_k\omega}$ does not contribute to the Hölder estimates. Hence the required estimate for $R(\omega)^k$ follows from Lemma 5.2.

Next, recall that $R_n v = R(1_{Z_n}(h^Y)^{-1}v)$. Hence, summing up the estimates in Lemma 5.2 over 1-cylinders $\overline{b} \subset Z_n$ yields $\|R_n\| \le E\sum_{\overline{b} \subset Z_n} \mu(\overline{b})(1 + |h^Y|_\theta)$. By Remark 5.3, the improved estimate is obtained. ∎

Theorem 5.5 *Let $f_h : X \times G \to X \times G$ be a mixing compact group extension of a Markov map $f : X \to X$. Suppose that there exists $Y \subset X$ such that the induced transformation $f^Y : Y \to Y$ is Gibbs-Markov. Suppose that $h : X \to G$ is a cocycle and that*

$$\sum_{j \ge n} \sum_{\overline{b} \in Z_j} \mu(\overline{b})(1 + |1_{\overline{b}} h_j|_\theta) = O(n^{-\beta}),$$

for some $\beta > 1$, where the second summation is over 1-cylinders $\bar{b} \in Z_j$.

Let G act orthogonally on $\mathbb{R}^d$ with $\operatorname{Fix} G = \{0\}$. Then hypotheses (H1) and (H2) of Theorem 2.1 (and Theorem 4.3) are satisfied.

Proof Hypothesis (H1) is immediate from Corollary 5.4(b). By a standard argument, the unit ball in $\mathcal{B}$ is compact in L^∞. This combined with Corollary 5.4(a) implies, by Hennion [11], that the essential spectral radius of $R(\omega) : \mathcal{B} \to \mathcal{B}$ is bounded above by $\theta < 1$ for all $\omega \in S^1$. By Proposition 4.4, 1 is not an eigenvalue of $R(\omega) : L^2(Y) \to L^2(Y)$, and $\mathcal{B} \subset L^2(Y)$, so we conclude that 1 does not lie in the spectrum of $R(\omega)$, establishing (H2). ∎

As an immediate consequence of Theorems 5.5 and 4.3, we obtain decay of correlations for observations (one Lipschitz, one L^∞) supported in Y.

Remark 5.6 If h is Hölder, then $\sum_{j\ge n}\sum_{\bar{b}\in Z_j}\mu(\bar{b})(1+|1_{\bar{b}}h_j|_\theta) \le \sum_{j\ge n}\mu(Z_j)(1+j|h|_\theta)$. Hence it suffices that $\mu(r > n) = O(n^{-(\beta+1)})$, thus proving Theorem 1.1 in the case when observations are supported in $Y \times G$. If in addition h is supported in Y or h is locally constant, then we obtain the improved $O(n^{-(\beta+1)})$ estimate mentioned in Remark 1.2(b).

6 Observations defined on the whole of $X \times G$

In this section we prove Theorem 1.1 by proving the corresponding result for a Young tower $F : \Delta \to \Delta$. We refer to Section 3(a) for background material and notation. However, we will let μ denote the invariant measure on Δ, reserving $m = \mu \times \nu$ for the product measure on the G-extension $\Delta \times G$.

Let L denote the transfer operator corresponding to $F : \Delta \to \Delta$. Given the cocycle $h : \Delta \to G$, we define $L_h v = L(h^{-1}v)$. Then $(L_h^n v)(x) = \sum_{F^n z = x} g_n(z) h_n(z)^{-1} v(z)$ where $g_n(z)$ is the inverse of the Jacobian of F^n at z. It follows from the definition of the tower that $g_1(z) = 1$ if $z \in \Delta_{j,\ell}$ for $0 \le j < r_j$. Moreover, if $z \in \Delta_{j,0}$, then g_{r_j} is Lipschitz and coincides with the inverse of the Jacobian of $F_0 : \Delta_{j,0} \to \Delta_0$.

Define $A_n : L^\infty(\Delta_0, \mathbb{C}^d) \to L^1(\Delta, \mathbb{C}^d)$ by

$$(A_n v)(x) = \sum_{\substack{F^n z = x \\ z \in \Delta_0 \\ Fz \notin \Delta_0, \dots, F^n z \notin \Delta_0}} g_n(z) h_n(z)^{-1} v(z).$$

Note that $A_n v$ is supported on level n of the tower, and that for x in level n we have $(A_n v)(x) = h_n(z)^{-1} v(z)$ where z is the unique point in Δ_0 with $F^n z = x_0$.

For brevity, we let $|A_n|_{\infty,1}$ denote the operator norm of $A_n : L^\infty(\Delta_0, \mathbb{C}^d) \to L^1(\Delta, \mathbb{C}^d)$.

Lemma 6.1 *For all $n \ge 1$, $|A_n|_{\infty,1} \le \sum_{r_j>n} \mu(\Delta_{j,0})$*

Proof Since $|h_n^{-1}| \equiv 1$, it is immediate that $|A_n v|_\infty \le |v|_\infty$. Also, the support of $A_n v$ (contained in level n of the tower) has measure at most $\sum_{r_j>n} \mu(\Delta_{j,0}) = \mu(r > n)$. The result follows. ∎

Define $\mathcal{B}(\Delta)$ to consist of globally Lipschitz functions $v : \Delta \to \mathbb{C}^d$ with $\|v\|_\theta = |v|_\infty + |v|_\theta$ where $|v|_\theta$ is the Lipschitz constant with respect to d_θ. Define $\mathcal{B}(\Delta_0)$ in the same way for functions $v : \Delta_0 \to \mathbb{C}^d$, so $\mathcal{B}(\Delta_0)$ coincides with the space $\mathcal{B}$ from Section 5. Define $B_n : \mathcal{B}(\Delta) \to \mathcal{B}(\Delta_0)$ by

$$(B_n v)(x) = \sum_{\substack{F^n z = x \\ z \notin \Delta_0, \dots, F^{n-1} z \notin \Delta_0 \\ F^n z \in \Delta_0}} g_n(z) h_n(z)^{-1} v(z).$$

Lemma 6.2 *There is a constant $C > 0$ such that*

$$\|B_n\| \le C \sum_{r_j>n} \mu(\Delta_{j,0})(1 + |1_{\Delta_{j,r_j-n}} h_n|_\theta),$$

for all $n \ge 1$.

Proof It follows from the definition of B_n that each preimage z lies in Δ_{j,r_j-n} for some j with $r_j > n$. If z and z' are compatible preimages of x and x', then $|v(z) - v(z')| \le |v|_\theta d_\theta(z, z') = \theta |v|_\theta d_\theta(x, x')$. Moreover, $g_n(z) = e^{p(y)}$ where $y \in \Delta_{j,0}$ with $F_0 y = x$ and p is the Lipschitz potential for F_0. Hence the estimates are obtained in the same way as was done for $\|R\|_n$ in Corollary 5.4(b). ∎

We can now estimate $L_h^n : \mathcal{B}(\Delta) \to \mathcal{B}(\Delta)$.

Corollary 6.3 *Suppose that $\mu(r > n) = O(n^{-(\beta+1)})$ and that in addition*

$$\sum_{r_j>n} \mu(\Delta_{j,0}) |1_{\Delta_{j,r_j-n}} h_n|_\theta = O(n^{-\beta}), \tag{6.1}$$

and

$$\sum_{r_j \ge n} \mu(\Delta_{j,0}) |1_{\Delta_{j,0}} h_{r_j}|_\theta = O(n^{-\beta}), \tag{6.2}$$

for some $\beta > 1$. Then $\|L_h^n\| = O(n^{-\beta})$.

Proof Recall that $T_n v = \sum_{\substack{F^n z = x \\ x,z\in\Delta_0}} g_n(z)h_n(z)^{-1}v(z)$ and so

$$L_h^n = \sum_{i+j+k=n} A_i T_j B_k + C_n,$$

where A_i, B_k are as defined above and

$$(C_n v)(x) = \sum_{\substack{F^n z = x \\ z\notin\Delta_0,\ldots,F^n z\notin\Delta_0}} g_n(z)h_n(z)^{-1}v(z).$$

Hence

$$\begin{aligned}|L_h^n v|_1 &\le \sum_{i+j+k=n} |A_i|_{\infty,1}|T_j B_k v|_\infty + |C_n|_{\infty,1}|v|_\infty \\ &\le \sum_{i+j+k=n} |A_i|_{\infty,1}\|T_j\|\,\|B_k\|\,\|v\| + |C_n|_{\infty,1}\|v\|.\end{aligned}$$

By condition (6.2) and the results of the previous sections, $\|T_n\| = O(n^{-\beta})$. By condition (6.1) and Lemmas 6.1 and 6.2, $|A_n|_{\infty,1} = O(n^{-(\beta+1)})$ and $\|B_n\| = O(n^{-\beta})$. Hence the convolution of the sequences A_n, T_n, B_n is $O(n^{-\beta})$. Also, arguing as in Lemma 6.1,

$$|C_n|_{\infty,1} \le \sum_{\substack{r_j>n \\ n<\ell<r_j}} \mu(\Delta_{j,\ell}) = \sum_{r_j>n}(r_j-n)\mu(\Delta_{j,\ell}) = O(n^{-\beta}),$$

giving the required estimate for the C_n term. ∎

Typically, one would expect that condition (6.1) is weaker than condition (6.2), but we have to allow for situations where cancellations mean that condition (6.2) is satisfied even though condition (6.1) fails.

Proof of Theorem 1.1 It follows from the properties of the projection $\pi : \Delta \to X$ in Section 3(a) that the equivariant Hölder observation $\phi : X \times G \to \mathbb{R}^d$ lifts to an equivariant Lipschitz observation $\phi\circ\pi : \Delta\times G \to \mathbb{R}^d$ with $\|\phi\circ\pi\| \le C\|\phi\|_\gamma$ where C is a universal constant. Similarly ψ lifts to $\psi\circ\pi$ with $|\psi\circ\pi|_\infty = |\psi|_\infty$.

Similarly, the Hölder cocycle $h : X \to G$ lifts to a Lipschitz cocycle on Δ (which we denote also by h). Conditions (6.1) and (6.2) are then automatic, just as in Remark 5.6. By Corollary 6.3, we obtain $\|L_h^n\| = O(n^{-\beta})$ on the tower. Hence by the same argument as in the proof of Theorem 4.3, it follows that on the tower,

$$\rho_{\phi\circ\pi,\psi\circ\pi}(n) \le K\|\phi\circ\pi\|\,|\psi\circ\pi|_\infty n^{-\beta} \le CK\|\phi\|_\gamma|\psi|_\infty n^{-\beta}.$$

Since π is measure preserving, it is immediate that $\rho_{\phi,\psi}(n) = \rho_{\phi\circ\pi,\psi\circ\pi}(n)$ giving the required result. ∎

A The renewal equation

In this appendix, we recall standard results about induced maps that were used in Section 4(a). Since the result has nothing to do with group extensions, we write Ω instead of $X \times G$ and $Z \subset \Omega$ instead of $Y \times G$. So the set up is that (Ω, μ) is a probability space, $f : \Omega \to \Omega$ is a measure preserving transformation, and $Z \subset \Omega$ is a measurable subset with $\mu(Z) > 0$. Let $f^Z : Z \to Z$ denote the induced transformation and let $Z = \cup_{n \geq 1} Z_n$ denote the corresponding partition. Let $\widehat{L}$, $\widehat{R}$ denote the transfer operators corresponding to f and f^Z and define

$$\widehat{T}_n \phi = \widehat{L}^n(1_Z \phi) 1_Z, \qquad \widehat{R}_n \phi = \widehat{L}^n(1_{Z_n} \phi) 1_Z.$$

We also define the return time function $r : Z \to \mathbb{N}$ by $r|_{Z_n} \equiv n$ and the twisted transfer operator

$$\widehat{R}_\omega \, \phi = \widehat{R}(e^{ir\omega} \phi),$$

for $\omega \in S^1$.

Proposition A.1 *(a)* $\widehat{T}_n = \sum_{i_1 + \cdots + i_k = n} \widehat{R}_{i_k} \cdots \widehat{R}_{i_2} \widehat{R}_{i_1}$.

(b) Restricting to $L^1(Z)$*, we have* $\widehat{R}_\omega = \sum_{n=1}^\infty \widehat{R}_n e^{in\omega}$.

Proof Compute that $\int \widehat{R}_n \phi \, \psi = \int \widehat{L}^n(1_{Z_n} \phi) \, 1_Z \psi = \int 1_{Z_n} \phi \, (1_Z \psi) \circ f^n$. But $f^n = f^Z$ when restricted to Z_n, so we have

$$\int \widehat{R}_n \phi \, \psi = \int 1_{Z_n} \phi \, (1_Z \psi) \circ f^n = \int 1_{Z_n} \phi \, (1_Z \psi) \circ f^Z. \tag{A.1}$$

Applying (A.1) inductively yields

$$\int \widehat{R}_{i_k} \cdots \widehat{R}_{i_2} \widehat{R}_{i_1} \phi \, \psi = \int_A \phi \, (1_B \psi) \circ f^{i_1 + \cdots + i_k},$$

where $A = Z_{i_1} \cap (f^Z)^{-1} Z_{i_2} \cap \cdots (f^Z)^{-(k-1)} Z_{i_k}$. Hence

$$\int \Big(\sum_{i_1 + \cdots + i_k = n} \widehat{R}_{i_k} \cdots \widehat{R}_{i_2} \widehat{R}_{i_1} \phi \Big) \, \psi = \int \phi 1_Z \, (1_Z \psi) \circ f^n = \int \widehat{T}_n \phi \, \psi,$$

proving part (a).

Restricting to $L^1(Z)$ and summing (A.1) over n yields

$$\int \Big(\sum_{n=1}^\infty \widehat{R}_n \phi \Big) \, \psi = \int \phi \, \psi \circ f^Z = \int \widehat{R} \phi \, \psi,$$

so that $\widehat{R} = \sum_{n=1}^{\infty} \widehat{R}_n$. Hence $\widehat{R}_\omega \phi = \widehat{R}(e^{ir\omega}\phi) = \sum_{n=1}^{\infty} \widehat{R}_n(e^{ir\omega}\phi)$. But

$$\widehat{R}_n(e^{ir\omega}\phi) = \widehat{L}^n(1_{Z_n} e^{ir\omega}\phi) = \widehat{L}^n(1_{Z_n} e^{in\omega}\phi) = \widehat{R}_n(\phi)e^{in\omega},$$

and part (b) follows. ∎

Acknowledgements We are grateful to Sébastien Gouëzel for very helpful discussions and for pointing out an error in a previous version of this paper. This research was supported in part by EPSRC Grant GR/S11862/01.

References

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

[2] V. Baladi. Decay of correlations. *Smooth Ergodicity Theory and its Applications* (A. Katok et al., ed.). Proc. Symp. Pure Math. **69**, Amer. Math. Soc., 2001, pp. 297–325.

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

[4] D. Dolgopyat. On mixing properties of compact group extensions of hyperbolic systems. *Israel J. Math.* **130** (2002) 157–205.

[5] M. J. Field, I. Melbourne, and A. Török. Stable ergodicity for smooth compact Lie group extensions of hyperbolic basic sets. *Ergod. Th. & Dynam. Sys.* To appear.

[6] M. J. Field, I. Melbourne, and A. Török. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergod. Th. & Dynam. Sys.* **23** (2003) 87–110.

[7] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.* **139** (2004) 29–65.

[8] S. Gouëzel. Statistical properties of a skew product with a curve of neutral points. Preprint, 2004.

[9] S. Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. Preprint, 2004.

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

[11] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. *Proc. Amer. Math. Soc.* **118** (1993) 627–634.

[12] H. Hu. Statistical properties of some almost hyperbolic systems. *Smooth Ergodicity Theory and its Applications* (A. Katok et al., ed.), Proc. Symp. Pure Math. **69**, Amer. Math. Soc., 2001, pp. 367–384.

[13] Y. Katznelson. *An Introduction to Harmonic Analysis.* Dover, New York, 1976.

[14] C. Liverani. Decay of correlations. *Ann. of Math.* **142** (1995) 239–301.

[15] I. Melbourne and M. Nicol. Statistical properties of endomorphisms and compact group extensions. *J. London Math. Soc.* **70** (2004) 427–446.

[16] I. Melbourne and M. Nicol. In preparation.

[17] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.* To appear.

[18] M. Nicol, I. Melbourne, and P. Ashwin. Euclidean extensions of dynamical systems. *Nonlinearity* **14** (2001) 275–300.

[19] W. Parry and M. Pollicott. *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics.* Astérique **187-188**, Société Mathématique de France, Montrouge, 1990.

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

[21] O. M. Sarig. Subexponential decay of correlations. *Invent. Math.* **150** (2002) 629–653.

[22] M. Viana. *Stochastic dynamics of deterministic systems.* Col. Bras. de Matemática, 1997.

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

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