Ergod. Th. & Dynam. Sys. (2005), **XX**, 1–21
Printed in the United Kingdom

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

HENK BRUIN†, MARK HOLLAND† and IAN MELBOURNE†

† *Department of Mathematics and Statistics, University of Surrey, Guildford GU2 7XH, UK*
(e-mail: h.bruin@surrey.ac.uk, mark.holland@surrey.ac.uk, i.melbourne@surrey.ac.uk)

(*Received 23 July 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 invariant measure μ. If $\phi, \psi : X \to \mathbb{R}$ lie in $L^2(X)$, we define the correlation function

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

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 [**24**] 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 [**5, 14, 23**] and probabilistic coupling [**25**]. In particular, Young towers with subexponential return asymptotics have subexponential (stretched exponential or polynomial) decay of correlations [**25**]. Sarig [**21**] introduced operator renewal sequences to obtain also lower bounds for decay of correlations and this method was sharpened by Gouëzel [**8**]. In particular, the results of [**8, 21**] show that the subexponential decay rates of Young [**25**] 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 [**7**] for the case when X is uniformly hyperbolic). We take mixing as given in this paper, and direct our 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 now defined to be

$$\rho_{\phi,\psi}(n)=\big|\textstyle\int_{X\times G}\phi(\psi^T\circ f_h^n)dm-\int_{X\times G}\phi\,dm\int_{X\times G}\psi^T\,dm\big|.$$

We view elements of $\mathbb{R}^d$ as column vectors, and so $\rho_{\phi,\psi}(n)$ takes values in the space of $d\times d$ matrices.

Results on exponential decay of correlations for equivariant observations on compact group extensions were obtained by [**6**] in the case when X is uniformly hyperbolic. This was extended in [**16**] to include more general situations where the transfer operator is quasicompact for the X dynamics.

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. 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 operator renewal sequence method of Sarig [**21**] and Gouëzel [**8**]. In the nonequivariant context, the renewal method gives optimal decay rates for Hölder observations supported on a certain subset of X, see [**21, 8**]. This can be used to obtain decay rates for general Hölder observations as was pointed out in Gouëzel [**10**], see also [**9**]. 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 [**8, 21**] vanishes in the equivariant case.)

1.1. *Statement of the main result* Let (X, d) be a 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 corresponding return map $f^Y : Y \to Y$ is given by $f^Y(y) = f^{r(y)}(y)$. By condition (2), f^Y is uniformly expanding. It can be shown (see Section 5.1) that there is a unique invariant probability measure μ on X that is equivalent to η.

Remark 1.1. We note that the return times $r(y)$ need not be first returns and so f^Y need not be the first return map to Y. In order to apply renewal theory, a preliminary step is to model $f : X \to X$ by a *Young tower* $F : \Delta \to \Delta$ built over a base $\Delta_0 \subset \Delta$ that is a copy of Y. The tower map F is Markov and the first return map for Δ_0 is precisely the uniformly expanding map $f^Y : Y \to Y$. Hence, renewal theory can be used to study $F : \Delta \to \Delta$ and thereby to study $f : X \to X$. Throughout Sections 3 and 4, the return times are first return times. The Young tower model appears in Section 5.

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 return map on $Y \times G$ given by $(y, g) \mapsto (f^Y y, g h^Y(y))$ 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). Let $|\,|$ denote a choice of norm on $\mathbb{R}^d$. 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 1.2. *Let $f_h : X \times G \to X \times G$ be a mixing compact group extension of a nonuniformly expanding map as above, where $h : X \to G$ is a Hölder cocycle. Assume that*

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

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 2.1).

REMARK 1.3. (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 [**8, 21**] so we focus on the case $\mathrm{Fix}\, G = \{0\}$.
(b) Suppose that $\mathrm{Fix}\, G = \{0\}$ and that ϕ and ψ are 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 (more precisely, h is constant on $f^\ell(Y_j)$ for each $j \geq 1$ and $0 \leq \ell \leq r_j - 1$), or (ii) supported in Y. See Remark 5.4.

It is clear that the results in this paper hold for more general classes of piecewise Hölder observations/cocycles (possibly with weaker decay results). For ease of exposition, we consider primarily the uniformly Hölder case.

The remainder of the paper is organised as follows: In Section 2, we prove a simplified special case of the Renewal Theorem of [**8, 21**] which suffices for our purposes. In Section 3, we introduce operator renewal sequences in the context of group extensions. In Section 4, we prove a version of Theorem 1.2 in the context of Markov maps with a Gibbs-Markov first return map $f^Y : Y \to Y$, for Hölder observations supported in $Y \times G$. In Section 5, we prove Theorem 1.2. This is done by reducing to the case of a Young tower [**25**] which is itself a special case of a Markov map with a Gibbs-Markov first return map. We also consider in Section 5 compact group extensions of systems that are nonuniformly hyperbolic in the sense of Young [**24**].

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 $\|T_n\| = O(|n|^{-\beta})$.*

By definition, the negative Fourier coefficients of $R(\omega)$ vanish, but this property need not hold for T. If it does, then a sharper result is possible.

(H3) $T_n = 0$ for $n \le -1$.

COROLLARY 2.2. *Assume that (H1)–(H3) are valid with $\beta > 0$ and $\beta \ne 1$. Then the Fourier coefficients T_n of T satisfy $\|T_n\| = O(n^{-\beta})$.*

Proof. By Theorem 2.1, it remains to consider the cases $\beta \in \{2, 3, \ldots\}$. In particular, the Fourier coefficients $\|T_n\|$ are summable. Now apply [**8**, Lemma 4.5]. □

REMARK 2.3. (1) Much of the difficulty, and hence depth, of the work in [**8, 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, when (H3) holds it suffices to show that Theorem 2.1 holds for $\beta \in (0, 1)$ and that $\|T_n\|$ is summable for $\beta > 1$. For completeness, an independent proof of Theorem 2.1 is given in Subsection 2.2. 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).

2.1. *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) and (H3) are 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).

2.2. *Proof of Theorem 2.1* In the remainder of this section, we prove Theorem 2.1. Throughout, if $\beta \geq 0$ 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 [**8, 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. *Renewal sequences for 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, see Subsection 3.1. In Subsection 3.2, we introduce the operator renewal sequences T_n and R_n. In our context, these operators are twisted versions of the transfer operators introduced by Sarig [**21**]. In Subsection 3.3 we prove a partial result towards the verification of hypothesis (H2).

3.1. *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 first return map (or induced map) $f^Y : Y \to Y$ by setting $f^Y y = f^n y$ for $y \in Z_n$.

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 $d \geq 1$ (so $L(v_1, \ldots, v_d)^T = (Lv_1, \ldots, Lv_d)^T$). Similarly, we obtain a transfer operator $\hat{L} : L^1(X \times G, \mathbb{C}^d) \to L^1(X \times G, \mathbb{C}^d)$ 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 first return map $f^Y : Y \to Y$. We have the transfer operators $R : L^1(Y, \mathbb{C}^d) \to L^1(Y, \mathbb{C}^d)$ and $\hat{R} : L^1(Y \times G, \mathbb{C}^d) \to L^1(Y \times G, \mathbb{C}^d)$, 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 3.1. *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 $(\hat{L}\phi)(x, g) = g(L_h v)(x)$ and $(\hat{R}\phi)(y, g) = g(R_h v)(y)$.*

Proof. This is standard, see for example [**6**]. We give the details for completeness in the case of $\hat{L}$. (The argument for $\hat{R}$ is identical.) By the Peter-Weyl theorem and orthogonality relations for representations of compact Lie groups, it suffices to show that $\int_{X\times G} \hat{L}\phi\, \psi^T\, dm = \int_{X\times G} g(L_h v)\, \psi^T\, dm$ for all *equivariant* $\psi \in L^2(X \times G, \mathbb{C}^d)$. In other words, we may suppose that $\psi(x, g) = gw(x)$ where $w \in L^2(X, \mathbb{C}^d)$. Then we compute that

$$\begin{aligned}
&\textstyle\int_{X\times G} (\hat{L}\phi)(x, g)\psi(x, g)^T dm = \int_{X\times G} \phi(x, g)\psi(fx, gh(x))^T dm \\
&\quad = \textstyle\int_{X\times G} gv(x)w(fx)^T h(x)^T g^T dm = \int_{X\times G} gh(x)^{-1}v(x)w(fx)^T g^T dm \\
&\quad = \textstyle\int_{X\times G} g[L(h^{-1}v)](x)w(x)^T g^T dm = \int_{X\times G} g[L(h^{-1}v)](x)\psi(x)^T dm
\end{aligned}$$

as required. □

3.2. *Operator renewal sequences* Following [**8, 21**], we define the following bounded linear operators on $L^1(X, \mathbb{C}^d)$ for $n \geq 1$:

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

Proposition 3.2. *(a) $T_n = \sum_{i_1+\cdots+i_k=n} R_{i_k} \cdots R_{i_2} R_{i_1}$, for all $n \geq 1$.*

(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 $\hat{T}_n$, $\hat{R}_n$ on $L^1(X \times G)$ by

$$\hat{T}_n\phi = \hat{L}^n(1_{Y\times G}\phi)1_{Y\times G}, \qquad \hat{R}_n\phi = \hat{L}^n(1_{Z_n\times G}\phi)1_{Y\times G}. \tag{3.1}$$

It follows from Proposition A.1(a) that

$$\hat{T}_n = \sum_{i_1+\cdots+i_k=n} \hat{R}_{i_k} \cdots \hat{R}_{i_2} \hat{R}_{i_1}. \tag{3.2}$$

Let $\phi(x, g) = gv(x)$. Using Proposition 3.1 and the definitions in (3.1), we compute that

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

Similarly, $(\hat{R}_n\phi)(x,g) = g(R_n v)(x)$. Substituting into (3.2) yields part (a).

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

$$\hat{R}_\omega = \sum_{n\geq 1} \hat{R}_n e^{in\omega} \tag{3.3}$$

By Proposition 3.1, $(\hat{R}_\omega\phi)(x,g) = g(R_h(e^{ir\omega}v))(x)$ and again $(\hat{R}_n\phi)(x,g) = g(R_n v)(x)$. Substituting these into (3.3) yields part (b). □

3.3. *Ruling out eigenvalues for $R(\omega)$* The next result is a step towards verifying hypothesis (H2) from Section 2. Recall that $R(\omega) = \sum_{n\geq 1} R_n e^{in\omega}$.

PROPOSITION 3.3. *Suppose that* $\operatorname{Fix} G = \{0\}$ *and that* $f_h : X \times G \to X \times G$ *is 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 3.2(b),

$$R(\omega)v = \sum_{n=1}^{\infty} R_n v e^{in\omega} = R_h(e^{ir\omega}v) = R(e^{ir\omega}(h^Y)^{-1}v).$$

Hence, 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 [**16**, 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

$$\phi \circ f^Y_{h^Y} = e^{ir\omega}\phi. \tag{3.4}$$

Denoting $r|_{Z_n} = n$, 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 (3.4) 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 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 $w(x) \in \operatorname{Fix} G = \{0\}$ for all $x \in X$ contradicting the assumption that v is nonzero. □

4. *The Gibbs-Markov setting*

In this section, we obtain a version of Theorem 1.2 under the additional assumptions that (i) the underlying dynamical system $f : X \to X$ is Markov with a first return map $f^Y : Y \to Y$ that is Gibbs-Markov, and (ii) the observations are supported in $Y \times G$.

The notions of Markov and Gibbs-Markov map are recalled in Subsection 4.1. In Subsection 4.2 we obtain some basic estimates for the twisted operators R_n that arise for compact group extensions in the (Gibbs-)Markov setting. In Subsection 4.3 we state and prove the version of Theorem 1.2 mentioned above.

4.1. *Markov maps and Gibbs-Markov maps* A background reference for Markov maps is [**1**, Section 4.7]. We follow the presentation in [**8**]. 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 $\bigvee_{i\geq 0} f^{-i}\alpha$ generates the σ-algebra of measurable sets. 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 $Y = \cup_{a\in\tilde\alpha} a$ is a union of elements of α 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 first return (induced) map $f^Y : Y \to Y$ by setting $f^Y y = f^n y$ for $y \in Z_n$. 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 $a \in \tilde\alpha, \xi_1, \dots, \xi_{n-1} \notin \tilde\alpha$.

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 $[b_0, \dots, b_{n-1}]_Y$.

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

(i) *Big images:* 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 $b \in \beta$.

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

$$\left| \frac{e^{p_k(x)}}{e^{p_k(y)}} - 1 \right| \leq D\theta^{-k} d_\theta(x, y) \quad \text{and} \quad D^{-1} \leq \frac{\mu[b_0, \dots, b_{k-1}]_Y}{e^{p_k(x)}} \leq D, \tag{4.1}$$

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

Let $\mathcal{B}$ denote the Banach space of functions $v : Y \to \mathbb{C}^d$ with norm $\|v\| = |v|_\infty + |v|_\theta < \infty$, where $|v|_\theta$ denotes the Lipschitz constant (with respect to the metric d_θ). 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

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

4.2. *Estimates for group extensions* 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. Viewing $h^Y : Y \to G$ as a map $h^Y : Y \to \mathbb{R}^{d^2}$ we can speak of the Hölder constant $|\ |_\theta$ with respect to the metric d_θ on Y and any choice of norm on $\mathbb{R}^{d^2}$. In fact, we do not assume that $|h^Y|_\theta$ is finite, but we do require that $|1_{Z_n} h_n|_\theta < \infty$ for all $n \geq 1$.

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

PROPOSITION 4.1. *Let $x, y \in Y$. Then $|h^Y_k(\bar{b}x) - h^Y_k(\bar{b}y)| \leq \sum_{j=0}^{k-1} \theta^{k-j} |1_{b_j} h^Y|_\theta d_\theta(x, y)$.*

Proof. To simplify notation, we write $H = h^Y$ and $F = f^Y$. Compute that

$$\begin{aligned}|H_k(\bar{b}x) - H_k(\bar{b}y)| &\leq \sum_{j=0}^{k-1} |H|_\infty^{k-1} |H \circ F^j(\bar{b}x)) - H \circ F^j(\bar{b}y))| \\ &\leq \sum_{j=0}^{k-1} |1_{b_j} H|_\theta d_\theta(F^j(\bar{b}x), F^j(\bar{b}y)).\end{aligned}$$

The result follows since $d_\theta(F^j(\bar{b}x), F^j(\bar{b}y)) \leq \theta^{k-j} d_\theta(x, y)$. □

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 (4.1) that $|M_{\bar{b}}|_\infty \leq D\mu(\bar{b})$.

LEMMA 4.2. *There is a constant $E \geq 1$ such that*

$$\|M_{\bar{b}}v\| \leq E\mu(\bar{b}) \Big\{ \big(1 + \sum_{j=0}^{k-1} \theta^{k-j} |1_{b_j} h^Y|_\theta \big) |v|_\infty + \theta^k |v|_\theta \Big\},$$

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. We again write $H = h^Y$ and also, we write b instead of $\bar{b}$. Let $x, y \in Y$ and compute that

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

and so by (4.1) and Proposition 4.1,

$$|(M_b v)(x) - (M_b v)(y)| \le D\mu(b)\{\theta^k|1_b v|_\theta + \sum_{j=0}^{k-1}\theta^{k-j}|1_{b_j}h^Y|_\theta|1_b v|_\infty + D|1_b v|_\infty\}d_\theta(x,y).$$

Hence $|M_b v|_\theta \le D\mu(b)\{\theta^k|1_b v|_\theta + \sum_{j=0}^{k-1}\theta^{k-j}|1_{b_j}h^Y|_\theta|1_b v|_\infty + D|1_b v|_\infty\}$. Combining this with the estimate for $|M_b|_\infty$ yields the required result with $E = D^2 + D$. □

COROLLARY 4.3. *(a) There is a constant $E' \ge 1$ such that for all $\omega \in S^1$, the linear operator $R(\omega) : \mathcal{B} \to \mathcal{B}$ satisfies*

$$\|R(\omega)^k v\| \le E'\{(1 + \sum_{n\ge1}\mu(Z_n)|1_{Z_n}h_n|_\theta)|v|_\infty + \theta^k|v|_\theta\},$$

for all $k \ge 1$.
(b) $\|R_n\| = O\Big(\mu(Z_n)(1 + |1_{Z_n}h_n|_\theta)\Big)$.

Proof. Recall from Proposition 3.2(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 $R(\omega)^k = \sum_{\bar b} M_{\bar b}e^{ir_k\omega}$ where $r_k = \sum_{j=0}^{k-1} r \circ (f^Y)^j$ and the sum is over all k-cylinders $\bar b$. Since r_k is constant on $\bar b$, the term $e^{ir_k\omega}$ does not contribute to the Hölder estimates. Hence by Lemma 4.2, to prove part (a) it remains to estimate $\sum_{\bar b=[b_0,\dots,b_{k-1}]}\mu(\bar b)\sum_{j=0}^{k-1}\theta^{k-j}|1_{b_j}h^Y|_\theta$.

Let n denote the symbol corresponding to Z_n (that is $b_0 = n$ precisely when $[b_0] = Z_n$). We compute that

$$\begin{aligned}
\sum_{\bar b=[b_0,\dots,b_{k-1}]}\mu(\bar b)\sum_{j=0}^{k-1}\theta^{k-j}|1_{b_j}h^Y|_\theta &= \sum_{j=0}^{k-1}\sum_{\bar b=[b_0,\dots,b_{k-1}]}\mu(\bar b)\theta^{k-j}|1_{b_j}h^Y|_\theta \\
&= \sum_{j=0}^{k-1}\theta^{k-j}\sum_{n\ge1}\sum_{\bar b=[b_0,\dots,b_{k-1}],\,b_j=n}\mu(\bar b)|1_{b_j}h^Y|_\theta \\
&= \sum_{n\ge1}|1_{Z_n}h_n|_\theta\sum_{j=0}^{k-1}\theta^{k-j}\mu((f^Y)^{-j}Z_n) \\
&\le \theta(1-\theta)^{-1}\sum_{n\ge1}\mu(Z_n)|1_{Z_n}h_n|_\theta,
\end{aligned}$$

as required (with $E' = E\theta(1-\theta)^{-1}$).

Next, recall that $R_n v = R(1_{Z_n}(h^Y)^{-1}v)$. Hence, summing up the estimates in Lemma 4.2 over 1-cylinders $\bar b \subset Z_n$ yields the estimate for $\|R_n\|$. □

4.3. *Decay rates for observations supported in $Y \times G$*

THEOREM 4.4. *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 first* *return map $f^Y : Y \to Y$ is Gibbs-Markov. Suppose that $h : X \to G$ is a cocycle satisfying*

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

for some $\beta > 1$.

Let G act orthogonally on $\mathbb{R}^d$ with $\mathrm{Fix}\, G = \{0\}$. Then there is a constant $C > 0$ such that for all equivariant observations $\phi, \psi : Y \times G \to \mathbb{R}^d$ of the form $\phi(x, g) = gv(x)$, $\psi(x, g) = gw(x)$ 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

$$R(\omega) = \sum_{n\ge1} R_n e^{in\omega}, \qquad T(\omega) = (I - R(\omega))^{-1}.$$

We begin by verifying hypotheses (H1)–(H3) stated in Section 2.

Hypothesis (H1) is immediate from Corollary 4.3(b). By a standard argument, the unit ball in $\mathcal{B}$ is compact in L^∞. This combined with Corollary 4.3(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 3.3, 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).

To prove (H3), we extend R to an analytic map $z \mapsto R(z) = \sum_{n\ge1} R_n z^n$ for $z \in \mathbb{C}$, $|z| \le 1$. (This is a slight abuse of notation since $R(\omega)$ is now written $R(e^{i\omega})$.) The map $I - R(z)$ is invertible for $|z| = 1$ by (H2), and invertibility for $|z| < 1$ is simpler: the inequality in Corollary 4.3(a) generalises to $\|R(z)^n\| \le E''|z|^n$ for $|z| \le 1$, so the spectral radius of $R(z)$ is at most $|z|$. Hence $T(z) = (I - R(z))^{-1}$ is analytic and we can write $T(z) = \sum_{n\ge0} T_n z^n$ establishing (H3).

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 $\beta > 1$, $T(\omega) = (I - R(\omega))^{-1}$ has an absolutely convergent Fourier series $T(\omega) = I + \sum_{|n|\ge1} \tilde{T}_n e^{in\omega}$. By (H3) and Corollary 2.2, there is a constant $C > 0$ such that $\|\tilde{T}_n\| \le Cn^{-\beta}$. Equating coefficients in $T(I - R) = I$ yields $\tilde{T}_n = \sum_{i_1+\cdots+i_k=n} R_{i_k} \cdots R_{i_1}$ so it follows from Proposition 3.2(a) that $T_n = \tilde{T}_n$. Hence $\|T_n\| \le Cn^{-\beta}$.

The remainder of the proof is a straightforward computation using Proposition 3.1:

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

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 $|T_n v|_1 \le \|T_n v\| \le Cn^{-\beta}\|v\|$ as required. □

5. *Proof of the main theorem*

In this section, we complete the proof of Theorem 1.2. In Subsection 5.1, we reduce to the case of a Young tower satisfying the hypotheses of Theorem 4.4. Decay of correlations follows for Hölder observations supported in $Y \times G$. In Subsection 5.2, we obtain decay rates for Hölder observations supported on the whole of $X \times G$. In Subsection 5.3, we consider group extensions of systems that are nonuniformly hyperbolic in the sense of [**24**].

5.1. *Reduction to a Young tower* Suppose that $f : X \to X$ is a nonuniformly expanding map with ergodic measure η and uniformly expanding return map $f^Y : Y \to Y$ as described in the introduction. The map f can be modelled by a *Young tower* $F : \Delta \to \Delta$ where $\pi : \Delta \to X$ is a Markov extension with certain properties [**25**]. We recall the construction now.

Let Δ_0 be a copy of the subset $Y \subset X$ and let $\Delta_{j,0} = Y_j$. Recall that the return time function $r : Y \to \mathbb{N}$ is constant on partition elements $\Delta_{j,0}$ with value $r|_{\Delta_{j,0}} = r_j \geq 1$. Let $\Delta = \{(y, \ell) : y \in \Delta_0,\ \ell = 0, \dots, r(y) - 1\}$, so Δ is the disjoint union of r_j copies of each $\Delta_{j,0}$. Define the *tower map* $F : \Delta \to \Delta$ by setting $F(y, \ell) = (y, \ell + 1)$ for $0 \leq \ell < r(y) - 1$ and $F(y, r(y) - 1) = (f^Y y, 0)$. Note that the return map $f^Y : Y \to Y$ is identified with the *first* return map $f^Y : \Delta_0 \to \Delta_0$. In particular, the return times r_j for the map f are first return times for the map F.

We can write Δ as the disjoint union $\Delta = \bigcup_{j\geq 1} \bigcup_{\ell=0}^{r_j-1} \Delta_{j,\ell}$, where $\Delta_{j,\ell} = \Delta_{j,0} \times \{\ell\}$. Then F is Markov with respect to the partition $\{\Delta_{j,\ell}\}$. Also f^Y is Markov with respect to the partition $\{\Delta_{j,0}\}$. (In the notation of Section 4.1, $\{\Delta_{j,\ell}\}$ plays the role of α and $\{\Delta_{j,0}\}$ plays the role of β.)

As in Section 4.1, we define a metric d_θ on Δ_0 in terms of the separation times under the first return map f^Y. So if $x, y \in \Delta_{j,0}$ for some j, then $s(x, y)$ is the greatest integer $n \geq 0$ such that $(f^Y)^n x$ and $(f^Y)^n y$ lie in the same partition element of Δ_0 and $d_\theta(x, y) = \theta^{s(x,y)}$.

The separation time function and metric extend to Δ as follows: If x, y lie in distinct partition elements, then $s(x, y) = 0$. 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 $s(x, y) = s(x_0, y_0)$. Hence $d_\theta(x, y) = \theta^{s(x,y)}$ defines a metric on Δ for any choice of $\theta \in (0, 1)$.

Define the projection $\pi : \Delta \to X$ by $\pi(y, \ell) = f^\ell(y)$. Clearly, π is a semi-conjugacy between F and f.

PROPOSITION 5.1. *Assume that f is nonuniformly expanding, and hence satisfies conditions (1)–(5) in Section 1.1. Let $\theta = 1/\lambda^\gamma$. Then*

(i) There is a constant $C' \geq 1$ such that $d(\pi x, \pi y) \leq C'[d_\theta(x, y)]^{1/\gamma}$ for all $x, y \in \Delta$.

(ii) There exist (unique) ergodic F-invariant and f-invariant probability measures μ' on Δ and μ on X, with μ equivalent to η, such that the semi-conjugacy $\pi : \Delta \to X$ is measure-preserving.

(iii) The first return map $f^Y : \Delta_0 \to \Delta_0$ *is Gibbs-Markov with respect to the partition* $\{\Delta_{j,0}\}$.

Proof. First suppose that $x, y \in \Delta_0$ and let $k = s(x,y)$. Then it follows from condition (2) that $d((f^Y)^k x, (f^Y)^k y) \geq \lambda^k d(x,y)$. Hence $d(x,y) \leq \operatorname{diam}(Y)/\lambda^{s(x,y)}$.

If x, y lie in $\Delta_{j,\ell}$, then write $x = f^\ell x_0$, $y = f^\ell y_0$. By condition (3),

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

This proves part (i).

Define an ergodic measure η' on Δ by setting $\eta'|_{\Delta_{j,\ell}} = \eta|_{Y_j}$. We note that Δ satisfies the technical assumptions to be a Young tower [**25**]: by condition (1), each partition element $\Delta_{j,0}$ is mapped by f^Y onto the whole of Δ_0; the required distortion condition is immediate by condition (4) and part (i) of the proposition. Since the return time function r is integrable (condition(5)), it is standard (see for example [**25**, Theorem 1]) that there is an F-invariant probability measure μ' on Δ, with μ' equivalent to η. Now define $\mu = \pi_* \mu'$ to obtain the required measure on X. This completes the proof of part (ii).

Part (iii) is immediate again by part (i) and conditions (1) and (4). (Indeed $\beta' = \{\Delta_0\}$ and p is globally Lipschitz.) □

Define the projection $\pi : \Delta \times G \to X \times G$ by setting $\pi(x,g) = (\pi x, g)$. This is a measure-preserving semi-conjugacy between the group extensions $F_{h\circ\pi} : \Delta \times G \to \Delta \times G$ and $f_h : X \times G \to X \times G$ (with respect to the product measures $m' = \mu' \times \nu$ and $m = \mu \times \nu$ where ν is Haar measure on G). It is immediate that mixing properties for $F_{h\circ\pi}$ are inherited by f_h. We have the following partial converse:

PROPOSITION 5.2. *Suppose that the group extension* $f_h : X \times G \to X \times G$ *is mixing and that* $\gcd\{r_j, j \geq 1\} = 1$. *Then* $F_{h\circ\pi} : \Delta \times G \to \Delta \times G$ *is mixing.*

Sketch proof. First consider the case when G is absent. Then it is clear that $\gcd\{r_j, j \geq 1\} = 1$ is a necessary condition for $F : \Delta \to \Delta$ to be mixing and Young [**24**, Lemma 5] shows that this condition is also sufficient. The idea is that the required mixing takes place at the base Δ_0 of the tower (where the return map f^Y is identical for f_h and $F_{h\circ\pi}$) and the greatest common divisor condition ensures that simultaneous returns to the base occur. The arguments are identical when G is present so we omit the details. □

We obtain a special case of Theorem 1.2 as a consequence of Theorem 4.4.

COROLLARY 5.3. *Theorem 1.2 holds under the additional hypothesis that the observations* ϕ *and* ψ *are supported on* $Y \times G$.

Proof. As above, we have the measure-preserving projection $\pi : \Delta \times G \to X \times G$ where Δ is a Young tower. By Proposition 5.1(i), Hölder cocycles $h_0 : X \to G$ lift to Lipschitz cocycles $h \circ \pi : \Delta \to G$ and similarly Hölder observations on $Y \times G$ lift to Lipschitz observations on $\Delta_0 \times G$. A simple argument [**24**, p. 607] shows that it is no loss of generality to assume that $\gcd\{r_j, j \geq 1\} = 1$. Then Proposition 5.2 guarantees that $F_{h\circ\pi} : \Delta \times G \to \Delta \times G$ is mixing.

Hence it suffices to prove decay of correlations under the assumptions that $F_h : \Delta \times G \to \Delta \times G$ is a mixing group extension defined by a Lipschitz cocycle $h : X \to G$ and $\phi(x,g) = gv(x)$, $\psi(x,g) = gw(x)$ are equivariant observations supported on $\Delta_0 \times G$ where $v : \Delta_0 \to \mathbb{R}^d$ is Lipschitz (and hence lies in the Banach space $\mathcal{B}$) and $w \in L^\infty(\Delta_0, \mathbb{R}^d)$.

By Theorem 4.4, it remains to show that $\sum_{j\geq n} \mu(Z_j)(1 + |1_{Z_j} h_j|_\theta) = O(n^{-\beta})$. Here Z_n is the union of partition elements of Δ_0 on which r takes the value of n. If $x, y \in Z_n$, then using the definition of d_θ on Δ we compute that

$$|h_n(x) - h_n(y)| \leq \sum_{\ell=0}^{n-1} |h(f^\ell x) - h(f^\ell y)| \leq |h|_\theta \sum_{\ell=0}^{n-1} d_\theta(f^\ell x, f^\ell y) = n|h|_\theta d_\theta(x,y).$$

Hence $|1_{Z_n} h_n|_\theta \leq n|h|_\theta$. By assumption $\sum_{j\geq n} \mu(Z_j) = O(n^{-(\beta+1)})$ and so we obtain the required $O(n^{-\beta})$ estimate. □

Remark 5.4. 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 1.3(b).

5.2. *Decay of correlations for observations on $X \times G$* In this subsection, we complete the proof of Theorem 1.2 by extending Corollary 5.3 to the case where observations are defined on the whole of $X \times G$. Again, it suffices to work at the level of a tower group extension $F_h : \Delta \times G \to \Delta \times G$. As before, $h : \Delta \to G$ is a Lipschitz cocycle. We consider observations $\phi, \psi : \Delta \times G \to \mathbb{R}^d$ of the form $\phi(x,g) = gv(x)$, $\psi(x,g) = gw(x)$, where $v : \Delta \to \mathbb{R}^d$ is Lipschitz and $w \in L^\infty(\Delta, \mathbb{R}^d)$.

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 4.

Recall that $L_h v = L(h^{-1}v)$ where L denote the transfer operator corresponding to the tower map $F : \Delta \to \Delta$. Following Gouëzel [**10, 9**], we estimate the norm of $L_h^n : \mathcal{B}(\Delta) \to L^1(\Delta)$.

Write $(L_h^n v)(x) = \sum_{F^n z = x} g_n(z) h_n(z)^{-1} v(z)$ 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 $0 \leq \ell \leq r_j - 2$. 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

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

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 $F^n z = x$.

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 5.5. $|A_n|_{\infty,1} = O(n^{-(\beta+1)})$.

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 $B_n : \mathcal{B}(\Delta) \to \mathcal{B}(\Delta_0)$ 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 5.6. $\|B_n\| = O(n^{-\beta})$.

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 we obtain estimates of the form

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

in the same way as was done for $\|R_n\|$ in Corollary 4.3(b). The calculation in the proof of Corollary 5.3 shows that $|1_{\Delta_{j,r_j-n}} h_n|_\theta \le n|h|_\theta$ and the result follows. □

We can now estimate $L_h^n : \mathcal{B}(\Delta) \to L^1(\Delta)$.

COROLLARY 5.7. *There exists $C \ge 1$ such that $|L_h^n v|_1 \le C n^{-\beta} \|v\|_\theta$ for all $v \in \mathcal{B}(\Delta)$.*

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

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

Hence

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

As shown in the proof of Theorem 4.4, $\|T_n\| = O(n^{-\beta})$. By Lemmas 5.5 and 5.6, $|A_n|_{\infty,1} = O(n^{-(\beta+1)})$ 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 5.5,

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

giving the required estimate for the C_n term. □

Proof of Theorem 1.2. As in Subsection 5.1, we first reduce to the case of a mixing group extension $F_h : \Delta \times G \to \Delta \times G$ where $F : \Delta \to \Delta$ is a tower map and $h : \Delta \to G$ is a Lipschitz cocycle. This time we consider equivariant observations $\phi(x,g) = gv(x)$, $\psi(x,g) = gw(x)$ defined on the whole of $\Delta \times G$, so $v \in \mathcal{B}(\Delta)$ and $w \in L^\infty(\Delta, \mathbb{R}^d)$.

By Corollary 5.7, we have $\|L_h^n\| = O(n^{-\beta})$. Decay of correlations on $\Delta \times G$ follows immediately by the same argument as in the last three lines of the proof of Theorem 4.4. □

5.3. *Group extensions of nonuniformly hyperbolic systems* In this subsection, we consider the case of a compact group extension $M \times G$ where $f : M \to M$ is nonuniformly hyperbolic in the sense of Young [**24**]. Part of the set up in [**24**] is that there is a "physical" f-invariant ergodic probability measure μ and a "uniformly hyperbolic" subset $\Lambda \subset M$ with an integrable return time function $r : \Lambda \to \mathbb{N}$. If $h : M \to G$ is a Hölder cocycle, we define the group extension $f_h : M \times G \to M \times G$. Consider equivariant observations of the form $\phi(x,g) = gv(x)$ where $v : M \to \mathbb{R}^d$. As before, ϕ is said to be γ-Hölder if v is γ-Hölder, and we define $\|\phi\|_\gamma = |v|_\infty + |v|_\gamma$.

THEOREM 5.8. *Let $f_h : M \times G \to M \times G$ be a mixing compact group extension of a nonuniformly hyperbolic system as above, where $h : M \to G$ is a Hölder cocycle. Assume that*

$$\mu(y \in \Lambda : r(y) \ge n) = O(n^{-(\beta+1)}),$$

for some $\beta > 1$. Let G act orthogonally on $\mathbb{R}^d$. Then for any $\beta' < \beta$, there exists a constant $C > 0$ such that for all Hölder equivariant observations $\phi, \psi : M \times G \to \mathbb{R}^d$,

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

As in the nonuniformly expanding case, $f : M \to M$ can be modelled by a tower $F : \Delta \to \Delta$. This tower has an F-invariant foliation by stable disks, leading to a quotient tower map $\bar F : \bar\Delta \to \bar\Delta$. For each $\theta \in (0,1)$, define a metric d_θ on $\bar\Delta$ as in Subsection 5.1. (We note that the quotient tower map $\bar F : \bar\Delta \to \bar\Delta$ corresponds to the tower map $F : \Delta \to \Delta$ in Subsection 5.1.) There are invariant measures $m = \mu \times \nu$ on $M \times G$, $m' = \mu' \times \nu$ on $\Delta \times G$, and $\bar m = \bar\mu \times \nu$ on $\bar\Delta \times G$, and measure-preserving semiconjugacies $\pi : \Delta \times G \to M \times G$ and $\bar\pi : \Delta \times G \to \bar\Delta \times G$.

The reduction to the nonuniformly expanding case breaks into three main steps (cf. [**6**] in the uniformly hyperbolic case).

(1) Define $\hat{h} = h \circ \pi : \Delta \to G$ and $\hat{\phi} = \phi \circ \pi$, $\hat{\psi} = \psi \circ \pi : \Delta \times G \to \mathbb{R}^d$. It suffices to establish decay of correlations for the observations $\hat{\phi}, \hat{\psi}$ and group extension $F_{\hat{h}}$.
(2) A version of the "Sinai trick" [**22**] shows that $\hat{h}$ is cohomologous to a cocycle $\tilde{h} : \Delta \to G$ that "depends only on future coordinates". In other words, $\tilde{h}$ is constant on stable disks and projects down to a cocycle $\bar{h} : \bar{\Delta} \to G$. Moreover, $\bar{h}$ is Lipschitz with respect to the metric d_θ. Details in the case of $\mathbb{R}$-valued cocycles are given in [**17**], and the argument there generalises to compact group extensions using [**19**, Appendix II]. Group extensions by cohomologous cocycles are topologically conjugate, so there is almost no loss in carrying out this step. However, the observations $\hat{\phi}$ and $\hat{\psi}$ are slightly modified in the process, yielding new observations $\tilde{\phi}$ and $\tilde{\psi}$ (which need not be lifts of Hölder observations on M).
(3) Write $\tilde{\phi}(x,g) = g\tilde{v}(x)$, $\tilde{\psi}(x,g) = g\tilde{w}(x)$. To estimate $\int_{\Delta\times G} \tilde{\phi}\,(\tilde{\psi}^T \circ F_{\tilde{h}}^n)\, dm'$, it suffices to estimate $\int_\Delta \tilde{h}_n^{-1}\tilde{v}\,(\tilde{w}^T \circ F^n)\, d\mu'$. Let $k \geq 1$. By [**15**, Lemma 5.4], we can construct $v_k : \Delta \to \mathbb{R}^d$ such that

(a) v_k depends only on future coordinates and projects to $\overline{v_k} : \bar{\Delta} \to \mathbb{R}^d$.

(b) $|\overline{v_k}|_\infty \leq |v|_\infty$, $|\overline{v_k}|_\theta \leq 2\theta^{-2k}|v|_\infty$, $|\tilde{v} \circ F^k - v_k|_\infty \leq C_1 \|v\|_\gamma\, \alpha^k$.

Similarly, we define $w_k : \Delta \to \mathbb{R}^d$. Here, $\alpha \in (0,1)$ is a constant arising from the nonuniformly hyperbolic structure of f and the Hölder exponent γ.

Following [**15**, Section 5(b)], the expression $\int_\Delta \tilde{h}_n^{-1}\tilde{v}\,(\tilde{w}^T \circ F^n)\, d\mu'$ breaks up into three integrals, two of which are of order α^k by (b). The third integral is given by $I_3 = \int_\Delta \tilde{h}_n^{-1}(\tilde{h}_k v_k)\,((\tilde{h}_k w_k)^T \circ F^n)\, d\mu'$. The integrand projects down to $\bar{\Delta}$ and so $I_3 = \int_{\bar{\Delta}} L_{\bar{h}}^n(\bar{h}_k\overline{v_k})\,(\bar{h}_k\overline{w_k})^T\, d\bar{\mu}$. By our results in the nonuniformly expanding case, $|I_3| \leq C_2 n^{-\beta}\|\bar{h}_k\overline{v_k}\|_\theta |\bar{h}_k\overline{w_k}|_\infty$. By (b), this leads easily to $|I_3| \leq C_3 n^{-\beta}\theta^{-2k}|v|_\infty |w|_\infty$. Hence, decay of correlations is estimated by $C_4(\alpha^k + n^{-\beta}\theta^{-2k})$. Choosing $k = k(n)$ so that $\alpha^k \sim n^{-\beta}\theta^{-2k}$, and choosing θ close to 1, we obtain the required decay rate $n^{-\beta'}$ with β' close to β.

A. *The renewal equation*

In this appendix, we recall standard results about first return maps that were used in Section 3.2. 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 first return map and let $Z = \cup_{n\geq 1} Z_n$ denote the corresponding partition. Let $\hat{L}$, $\hat{R}$ denote the transfer operators corresponding to f and f^Z and define

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

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

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

for $\omega \in S^1$.

PROPOSITION A.1 ([**21**, PROPOSITION 1]) *(a)* $\hat{T}_n = \sum_{i_1+\cdots+i_k=n} \hat{R}_{i_k}\cdots\hat{R}_{i_2}\hat{R}_{i_1}$.

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

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

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

Applying (A.1) inductively yields

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

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

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

proving part (a).

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

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

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

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

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. We are also grateful to a referee for helpful comments and suggestions. 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**] P. Ferrero and B. Schmitt. Produits aléatoires d'opérateurs matrices de transfert. *Probab. Theory Related Fields* **79** (1988) 227–248.

[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] M. J. Field, I. Melbourne, and A. Török. Stable ergodicity for smooth compact Lie group extensions of hyperbolic basic sets. *Ergod. Th. & Dynam. Sys.* **25** (2005) 517–551.

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

[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. Rapid decay of correlations for nonuniformly hyperbolic flows. Preprint, 2005.

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

[17] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Commun. Math. Phys.* 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] Y. G. Sinai. Gibbs measures in ergodic theory. *Russ. Math. Surv.* **27** (1972) 21–70.

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

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

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