Statistical properties of endomorphisms and compact group extensions

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

25 March 2003

Abstract

We consider the statistical properties of endomorphisms under the assumption that the associated Perron-Frobenius operator is quasicompact. In particular we consider the central limit theorem, weak invariance principle and law of the iterated logarithm for sufficiently regular observations. Our approach clarifies the role of the usual assumptions of ergodicity, weak-mixing and exactness.

We also give sufficient conditions for quasicompactness of the Perron-Frobenius operator to lift to the corresponding equivariant operator on a compact group extension of the base. This leads to statistical limit theorems for equivariant observations on compact group extensions.

Examples considered include compact group extensions of piecewise uniformly expanding maps (for example Lasota-Yorke maps), and subshifts of finite type, as well as systems that are nonuniformly expanding or nonuniformly hyperbolic.

1 Introduction

It is well known since the work of Sinai, Ruelle and Bowen (see for example [8, 33, 31]) that mixing uniformly expanding and uniformly hyperbolic dynamical systems enjoy strong statistical properties, such as exponential decay of correlations and the central limit theorem, for sufficiently regular observations. Since then, there has a been a great deal of effort to extend these results to more general classes of dynamical systems and observations (see for example the recent surveys in [1, 2, 35]).

One approach is to study the rate of decay of certain transfer operators or Perron-Frobenius operators when restricted to suitable function spaces. In many important cases, it is possible to prove that these operators are quasicompact leading to exponential decay rates. Exponential decay of correlations follows immediately, while the central limit theorem follows from an idea of Gordin [17]. A functional version of the central limit theorem is also known to hold in such situations and it was recently noticed [16] that the upper half of the law of the iterated logarithm is valid (provided the relevant function space can be chosen to lie in L^∞).

In a different direction, [16] considered compact group extensions of uniformly hyperbolic diffeomorphisms and, motivated by [28], restricted to a class of equivariant observations. Using a combination of existing techniques and new ideas, it was shown that (improved versions) of the statistical properties described above are inherited by the group extension. In particular, we note the improved results on nondegeneracy of the central limit theorem in [16].

In this paper, we describe new results in the theory of statistical properties of (a) dynamical systems, and (b) their compact group extensions. In direction (a), we give an account of the implications of the quasicompactness of the Perron-Frobenius operator for the statistical behaviour of endomorphisms. Our approach clarifies the usual assumptions of ergodicity, weak mixing, and exactness. In particular, in contrast to most of the literature, our results do not require weak mixing or exactness.

In direction (b), we show that certain axioms for quasicompactness described in Keller & Liverani [24] lift to the compact group extension setting, at the level of equivariant observations. We are claiming not that quasicompactness automatically lifts, but that certain sufficient axioms for quasicompactness lift. Since these axioms hold very generally, we are able to consider a large collection of examples. Thus, we greatly generalise the applicability of the ideas in [16], at the same time relaxing certain assumptions such as weak mixing.

The remainder of this paper is organised as follows. In §2, we discuss growth rates of ergodic sums at the level of isometries on a Hilbert space H. We show how a quasicompactness assumption leads directly to a martingale approximation, existence of the variance, and square root growth in L^2. (In this section, there is no dynamical system. The setting is analogous to the von Neumann mean ergodic theorem.)

In §3, we specialize to the case $H = L^2(X)$ where $f : X \to X$ is an ergodic map satisfying a quasicompactness property. The central limit theorem (and much more) follows directly from the martingale approximation in §2.

In §4, we discuss the axiomatic framework of [24] guaranteeing quasicompactness of the Perron-Frobenius operator. Compact group extensions are introduced in §5 and it is shown that the axioms for quasicompactness in §4 lift to spaces of equivariant observations.

Finally, a number of applications are considered in §6.

2 Quasicompact operators on Hilbert space

Let H be a Hilbert space and $U : H \to H$ an isometry. We are particularly interested in the case when U is not invertible. Let $U^* : H \to H$ be the adjoint of U and note that $U^*U = I$. We recall the following basic properties of the spectra of U and U^*:

- The spectrum of U lies on the unit circle and the spectrum of U^* lies in the closed unit disk.
- If $\alpha \in \mathbb{C}$ and $|\alpha| = 1$, then $Uv = \alpha v$ if and only if $U^*v = \overline{\alpha}v$. (One direction follows immediately from the fact that $U^*U = I$. In the other direction, compute directly that $\langle Uv - \alpha v, Uv - \alpha v\rangle = 0$.)

(We note that in general U^* may have eigenvalues α with $|\alpha| < 1$.)

Given $v \in H$, we define $v_N = \sum_{j=1}^N U^j v$. By the mean ergodic theorem, $v_N = N\pi v + o(N)$ as $N \to \infty$ where $\pi : H \to \ker(U - I)$ is the orthogonal projection. (That is, $\lim_{N\to\infty} \frac{1}{N}\|v_N - N\pi v\|_H = 0$.)

In this section, we are interested in obtaining more precise information on the growth of $v_N - N\pi v$ under a certain "quasicompactness" hypothesis:

Definition 2.1 Let $F \subset H$ be a Banach space such that U and U^* restrict to bounded operators on F. The operator $U^* : F \to F$ is *quasicompact* if $U^* : F \to F$ has essential spectral radius $\rho < 1$.

It follows from quasicompactness that U^* has at most finitely many eigenvalues on the unit circle, that these eigenvalues have finite multiplicity, and that the rest of the spectrum is contained in a disk around the origin of radius less than 1. (We note that our definition is slightly nonstandard, since we do not require that there exist eigenvalues α with $\rho < |\alpha| \le 1$.)

By quasicompactness, there is a closed U^*-invariant splitting $F = F_1 \oplus F_2$ where $F_1 = \ker(U - I) = \ker(U^* - I)$. Again it is the case that $v_N = Nv$ for all $v \in F_1$ and $\|v_N\|_H = o(N)$ for all $v \in F_2$.

Theorem 2.2 *Regarding U as an operator on F and U^* as an operator on F_2, we have $F_2 = \ker U^* \oplus \mathrm{Im}(U - I)$.*

Proof First, we show that $\ker U^* \cap \mathrm{Im}(U - I) = \{0\}$. Suppose that $U^*v = 0$ and $v = Uy - y$ where $y \in F$. Then $0 = U^*v = y - U^*y$. It follows that $Uy = y$ (since $U^*y = y$ implies $Uy = y$) and hence $v = 0$ as required.

Second, we show that $F_2 = \ker U^* + \mathrm{Im}(U - I)$. Let $F_\alpha \subset F_2$ be the sum of the eigenspaces corresponding to eigenvalues on the unit circle (other than 1). Note that F_α is finite dimensional (possibly trivial). The quasicompactness hypothesis guarantees that we have the further closed U^*-invariant splitting $F_2 = F_\alpha \oplus F_3$ where $U^* : F_3 \to F_3$ has spectral radius $\rho < 1$.

If $v \in F_\alpha$ is an eigenfunction of U, so $Uv = \alpha v$, then $v = (\alpha - 1)^{-1}(U - I)v$. Hence $F_\alpha \subset \mathrm{Im}(U - I)$.

It remains to show that $F_3 = \ker U^* + \mathrm{Im}(U - I)$. Given $v \in F_3$, define $y = \sum_{j \geq 1} (U^*)^j v$. It follows from the spectral radius assumption that the series converges to $y \in F_3$. Now compute that $U^*\{v - Uy + y\} = 0$. ∎

For reasons that will become clear in Section 3, we say that $w \in F$ is a *martingale* if $U^* w = 0$. By Theorem 2.2, if $v \in F_2$, then there is a *martingale approximation*

$$v = w + Uy - y, \quad \text{where } w, y \in F_2 \text{ and } U^* w = 0.$$

Since $v_N = w_N + U^n y - y$, many statistical properties for v follow from the corresponding property for the martingale w. One result in this direction is the following:

Corollary 2.3 *Let $v \in F_2$ and write $v = w + Uy - y$ as above. Define $\sigma = \|w\|_H$. Then $\|w_N\|_H = \sqrt{N}\sigma$ and*

$$\|v_N\|_H = \sqrt{N}\sigma + O(1).$$

In particular, $\sigma = \lim_{N \to \infty} \frac{1}{\sqrt{N}} \|v_N\|_H$.

Proof Since $U^* w = 0$, we compute that if $j > k$ then

$$\langle U^j w, U^k w \rangle = \langle U^{j-k} w, w \rangle = \langle w, (U^*)^{j-k} w \rangle = 0$$

and similarly for $j < k$. On the other hand, $\langle U^j w, U^j w \rangle = \langle w, w \rangle$. It follows that $\|w_N\|_H = \sqrt{N} \|w\|_H$. Next, consider the remainder term $r = Uy - y$. Then $r_N = U^N y - y$ and so $\|r_N\|_H \leq 2\|y\|_H$. ∎

3 Statistical properties of dynamical systems

In this section, we apply the results in §2 to the case where U is the Koopman operator associated to a measure-preserving transformation. Suppose that X is a probability space with measure m and σ-algebra $\mathcal{B}$. Let $f : X \to X$ be a (noninvertible) measure-preserving transformation. We take $H = L^2(X)$ and define the isometry $U : H \to H$ by $Uv = v \circ f$. The adjoint U^* satisfies $U^* U = I$ as before. In addition $UU^* v = E(v | f^{-1}\mathcal{B})$, where $E(\cdot | f^{-1}\mathcal{B})$ is the conditional expectation operator.

Again, if $v \in H$ we define $v_N = \sum_{j=1}^N U^j v = \sum_{j=1}^N v \circ f^j$. By the mean ergodic theorem, $v_N = N\pi v + o(N)$ in H where $\pi : H \to H$ is the orthogonal projection onto $\ker(U - I)$. (If $f : X \to X$ is ergodic, then $\pi v = \int_X v \, dm$.)

As in Definition 2.1, we assume quasicompactness, so there is a Banach space $F \subset H$ such that U and U^* restrict to bounded operators on F and $U^* : F \to F$ has essential spectral radius less than one. By Corollary 2.3, for all $v \in F$, we can define the *variance* $\sigma^2 = \lim_{N\to\infty} \frac{1}{N}\|v_N - N\pi v\|_H^2$.

Theorem 3.1 (Central Limit Theorem (CLT)) *Assume quasicompactness and that $f : X \to X$ is ergodic. Let $v \in F$ with $\int_X v\,dm = 0$. Then $\frac{1}{\sqrt{N}}v_N$ converges in distribution to a normal distribution with mean zero and variance σ^2. That is*

$$m\big\{x \in X : \tfrac{1}{\sqrt{N}}v_N < b\big\} \to \frac{1}{\sqrt{2\pi}\,\sigma}\int_{-\infty}^{b} e^{-y^2/2\sigma^2}\,dy$$

as $N \to \infty$ for all $b \in \mathbb{R}$.

Proof By Theorem 2.2, we can write $v_N = w_N + U^N y - y$ a.e., where $w, y \in F$ and $U^* w = 0$. Moreover, since $y \in L^2$ it follows from the pointwise ergodic theorem that $U^N y = y \circ f^N = o(N^{1/2})$ a.e. Hence it suffices to prove that $\frac{1}{\sqrt{N}}w_N$ converges in distribution to a normal distribution with mean zero and variance σ^2.

But $U^* w = 0$ implies that $E(w|f^{-1}\mathcal{B}) = 0$. Passing to the natural extension [32], we obtain a bi-infinite ergodic stationary martingale $\{Y_j : j \in \mathbb{Z}\}$ where $Y_{-i} = w \circ f^i$ for $i \ge 0$ (cf. [16, Remark 3.12]). It follows from [5] that $\frac{1}{\sqrt{|N|}}\sum_{j=0}^{N-1} Y_j$ converges to a normal distribution with mean zero and variance $\int Y_1^2$ as $N \to \pm\infty$. In particular, $\frac{1}{\sqrt{N}}\sum_{j=1}^{N} w \circ f^j$ satisfies the CLT with mean zero and variance $\sigma^2 = \|w\|_H^2$. ∎

We have the following criteria for degeneracy in the CLT ($\sigma^2 = 0$).

Proposition 3.2 *Suppose that $v \in F$ and $\pi v = 0$. Then*

(a) $\sigma^2 = 0$ if and only if $v = y \circ f - y$ for some $y \in F$.

(b) If $F \subset L^\infty$, then $\sigma^2 = 0$ if and only if $v_N = O(1)$ uniformly a.e. Indeed, $|v_N|_\infty \le 2|y|_\infty$.

(c) If $f : X \to X$ is ergodic, then $\sigma^2 = 0$ if and only if $v_N = o(\sqrt{N})$ a.e.

(d) If $f : X \to X$ is ergodic, and $F \subset L^p$ for some $p > 2$, then $\sigma^2 = 0$ if and only if $v_N = o(N^{1/p})$ a.e.

Proof Recall the decomposition $v = w + y \circ f - y$ where $U^* w = 0$. By Corollary 2.3, $\sigma^2 = 0$ if and only if $w = 0$ a.e. proving part (a).

By part (a), if $\sigma^2 = 0$, then $v_N = y \circ f^N - y$. Part (b) follows immediately. As in the proof of Theorem 3.1, the pointwise ergodic theorem guarantees that $y \circ f^N =$

$o(N^{1/2})$ proving one direction of part (c). To prove the reverse direction, note that if $v_N = o(\sqrt{N})$, then $\frac{1}{\sqrt{N}} v_N$ converges to zero a.e., and hence in distribution. But since f is ergodic, Theorem 3.1 guarantees convergence in distribution to a normal distribution with variance σ^2. Hence, this is the degenerate normal distribution with $\sigma^2 = 0$. Part (d) is proved in the same way as part (c). ∎

Remark 3.3 It is well-known that the CLT is degenerate if and only if $v = y \circ f - y$ for some $y \in L^2$. Our conditions in Proposition 3.2, which follow [16], are a substantial improvement. For example, suppose that X is a topological space with open sets having positive measure, and that $f : X \to X$, $v \to \mathbb{R}$ are continuous. Under the assumptions of quasicompactness, ergodicity of $f : X \to X$, and $F \subset L^\infty$, we have that $\sigma^2 = 0$ if and only if there is a constant C such that $|v_N(x)| \le C$ for all $x \in X$ and $N \ge 1$. In particular, if x is a periodic point of period p, then $v_p(x) = 0$.

Set $W_N(0) = 0$, and $W_N(t) = \frac{1}{\sqrt{N}} v_{Nt} = \frac{1}{\sqrt{N}} \sum_{j=0}^{Nt-1} v \circ f^j, \quad t = 1/N, 2/N, \ldots$ Linearly interpolating on each interval $[(r-1)/N, r/N]$, $r \ge 1$, we obtain a sequence of random elements $W_N \in C([0,\infty), \mathbb{R})$. We have the weak invariance principle (which is a refinement of the CLT):

Theorem 3.4 (Weak invariance principle (WIP)) *The sequence $\{W_N\}$ converges weakly in $C([0,\infty), \mathbb{R})$ to an n-dimensional Brownian motion with variance σ^2.*

Proof Billingsley [6] proves the WIP for stationary ergodic L^2 martingales, so the result follows along the lines of Theorem 3.1. ∎

Theorem 3.5 (Upper law of the iterated logarithm (Upper LIL)) *Assume quasicompactness and that $f : X \to X$ is ergodic. Suppose further that $F \subset L^\infty$. Let $v \in F$ with $\int_X v \, dm = 0$. Then*

$$\limsup_{N \to \infty} \frac{v_N}{\sqrt{2N \log \log N}} \le \sigma \quad \textit{almost surely.}$$

Proof Again, we write $v_N = w_N + U^n y - y$ where $U^* w = 0$ and it suffices to prove the upper LIL for the sequence $\{w_N\}$. As pointed out in [16, §3(c)], the condition $U^* w = 0$ implies that $\{w_N\}$ is a "weakly multiplicative sequence". Since $w \in F \subset L^\infty$, the result follows by [34]. ∎

Remark 3.6 (a) The full law of the iterated logarithm (LIL) is the similar conclusion with $\le$ replaced by $=$. We do not know whether the LIL holds under our hypotheses, or whether it is possible to remove the L^∞ assumption.

(b) Passing to the natural extension [32], it follows from the methods in [16] that the LIL (and much more, including the almost sure invariance principle (ASIP)) can be proved in backwards time. Moreover, the L^∞ assumption is not required.

In certain situations, such as for ergodic compact Lie group extensions of Axiom A base dynamics, these statistical properties can be deduced *a fortiori* in the correct time direction [16]. The approach in [16] requires two ingredients: (i) that there is a method for passing from invertible transformations to noninvertible transformations without losing too much regularity in the observations, and (ii) that the system of dynamical systems is closed under time reversal.

(c) The difficulty with time directions described above is a possibly serious limitation of the martingale approximation approach to proving the ASIP [16]. The same issue arises in Conze & le Borgne [14]. An alternative approach to proving the ASIP is presented in Hofbauer & Keller [22].

Remark 3.7 Assume quasicompactness and in addition that $f : X \to X$ is weak mixing. Then by standard arguments, decay of correlations holds for observations in F. Indeed, there are constants $C > 0$ and $\rho \in (0, 1)$ such that $\|(U^*)^n v - \int v\| \le C\rho^n \|v\|$, for all $v \in F$ and $n \ge 1$. Hence

$$\left| \int_X v \cdot w \circ f^n \, dm - \int_X v \, dm \int_X w \, dm \right| \le C\rho^n \|v\| |w|_2,$$

for all $v \in F$, $w \in L^2$, $n \ge 1$. It follows easily that if $v \in F$ and $\int_X v \, dm = 0$, then the variance is given by $\sigma^2 = \int_X v^2 \, dm + 2 \sum_{j=1}^\infty \int_X v \cdot v \circ f^j \, dm$.

Vector-valued observations We now generalise to the case of vector-valued observations $v : X \to \mathbb{R}^d$. We continue to consider a measure-preserving transformation $f : X \to X$ with $H = L^2(X) = L^2(X, \mathbb{R})$, and assume that $F \subset H$ is a Banach space such that $U^* : F \to F$ is quasicompact. Define $H^d = L^2(X, \mathbb{R}^d)$ and $F^d = \{v = (v_1, ..., v_d) : X \to \mathbb{R}^d \mid v_j \in F, \ j = 1, \dots, d\}$.

The operator $Uv = v \circ f$ acts on vector-valued observations and defines an isometry on H^d and a linear operator on F^d. Similarly, U^* acts component-wise on H^d and F^d. It is immediate that $U^* : F^d \to F^d$ has essential spectral radius $\rho < 1$ so that quasicompactness holds with $F^d \subset H^d$. Hence, the results of §2 apply to functions $v \in F^d$.

In particular, the scalar variance $\sigma^2 = \lim_{N\to\infty} \frac{1}{N} \|v_N\|_{H^d}^2$ is defined for all $v \in F_2^d$. However, it is natural (following Field *et al* [16]) to define the $d \times d$ covariance matrix

$$\Sigma = \lim_{N\to\infty} \frac{1}{N} \int_X v_N \cdot v_N^T \, dm \quad \text{(outer product)}.$$

Note that this limit is well-defined when $U^*v = 0$, in which case $\Sigma = \int_X v \cdot v^T = \frac{1}{N}\int_X v_N \cdot v_N^T$ for all N. Hence, Σ is well-defined for all $v \in F_2^d$. Moreover, $\Sigma_{jk} = E(Y_j Y_k)$, hence Σ is symmetric and $\langle \Sigma x, x\rangle \geq 0$ for all $x \in \mathbb{R}^d$.

Next, we suppose further that $f : X \to X$ is ergodic. Let $v \in F_2^d$ so that $\int_X v\, dm = 0$. Given $c \in \mathbb{R}^d$, we have the decomposition $c \cdot v = c \cdot w + o(N^{1/2})$ a.e., where $U^*(c \cdot w) = 0$. As in Theorem 3.1, $\frac{1}{\sqrt{N}} c \cdot v_N$ converges in distribution to a normal distribution with mean zero and variance $\sigma_c^2 = |c \cdot w|_2^2 = c^T \Sigma c$. By the Cramer-Wold technique (see for example [7, Theorem 29.4]) this implies that $\frac{1}{\sqrt{N}} v_N$ converges in distribution to a d-dimensional normal distribution with mean zero and covariance matrix Σ. The distribution is *nondegenerate* if Σ is nonsingular.

Similarly, the d-dimensional version of the WIP is valid for all $v \in F^d$ with $\int_X v\, dm = 0$ when $f : X \to X$ is ergodic. In addition, if $F \subset L^\infty$, then the upper LIL holds for $c \cdot v$ for all $c \in \mathbb{R}^d$. If $f : X \to X$ is weak mixing, then we obtain a d-dimensional analogue of exponential decay of correlations [16].

4 Quasicompactness and Perron-Frobenius

In this section, we interpret the operator $U^* : L^2 \to L^2$ as the Perron-Frobenius operator $P : L^1 \to L^1$ associated to the measure-preserving transformation $f : X \to X$. Following Keller & Liverani [24], we give an axiomatic approach to quasicompactness.

As before, we let X be a probability space with measure m and σ-algebra $\mathcal{B}$ and $f : X \to X$ is a (noninvertible) measure-preserving transformation. Given $1 \leq p \leq \infty$, define the Koopman operator $U : L^p \to L^p$ by $Uv = v \circ f$. For each p, $U : L^p \to L^p$ is an isometry.

Given $v \in L^1$, we define Pv by demanding that

$$\int_X Pv \cdot w\, dm = \int_X v \cdot Uw\, dm = \int_X v \cdot w \circ f\, dm$$

for all $w \in L^\infty$. The operator $P : L^1 \to L^1$ is called the Perron-Frobenius operator. Clearly, this also defines $P : L^p \to L^p$ for all $1 \leq p \leq \infty$ (restricting to $w \in L^q$ where $1/p + 1/q = 1$). When $p \geq 2$, the operator P coincides with the operator U^* in §2.

We assume the following:

(F1) F and F' are Banach spaces, $F \subset F' \subset L^1$ with norms $\| \, \| \geq |\, |$ respectively, and F is densely embedded in F'.

(F2) $F \subset L^2$ and $\| \, \| \geq |\, |_2$.

(F3) $\{v \in F : \|v\| = 1\}$ is compact in F'.

(F4) The Perron-Frobenius operator P restricts to a bounded operator on F and F'. Moreover, there is a constant $C_0 > 0$ such that $|P^n| \leq C_0$ for all $n \geq 1$.

(F5) (Lasota-Yorke inequality) For some $n_0 \geq 1$, there are constants $D_0 > 0$ and $\theta_0 \in (0,1)$ such that $\|P^{n_0}v\| \leq D_0|v| + \theta_0\|v\|$ for all $v \in F$.

Remark 4.1 If $F' = L^p$ for some $1 \leq p \leq \infty$ or $F' = C(X)$, then hypothesis (F4) is automatically satisfied with $C_0 = 1$.

Example 4.2 In §6, we consider a number of examples where (F1)–(F5) are satisfied. One such example is Hölder observations on a one-sided subshift of finite type. For details, see Ruelle [33], Bowen [8], or Parry & Pollicott [31]. Let $\sigma : X \to X$ denote an irreducible (not necessarily aperiodic) subshift of finite type. Here $X \subset \{1, \dots, k\}^{\mathbb{N}}$ for some k. Fix $\theta \in (0,1)$ and define $d_\theta(x,y) = \theta^N$ where $N \geq 1$ is least such that $x_i = y_i$ for $i < N$. Let F_θ be the space of functions that are Lipschitz with respect to this metric. Let $|g|_\theta$ denote the Lipschitz constant for $g \in F_\theta$ and define the norm $\|g\|_\theta = |g|_\infty + |g|_\theta$. Then F_θ is a Banach space. Moreover, taking $F = F_\theta$ and $F' = C(X)$, it is immediate that (F1) and (F2) are valid, while (F3) follows from Arzela-Ascoli. By Remark 4.1, (F4) is automatic with $C_0 = 1$. Finally, the "basic inequality" [31, Proposition 2.1] guarantees that (F5) holds for a large class of measures. (Technically speaking, (F5) holds whenever m is an equilibrium measure corresponding to a potential $g \in F$. The Perron-Frobenius operator is the Ruelle transfer operator corresponding to a normalized version of g.)

We note that F_θ is a Banach algebra. Indeed, if $f, g \in F_\theta$, then $|fg|_\theta \leq |f|_\infty \|g\|_\theta + |g|_\infty \|f\|_\theta$. Since $|fg|_\infty \leq |f|_\infty |g|_\infty$, it follows that $\|fg\|_\theta \leq \|f\|_\theta \|g\|_\theta$.

Example 4.3 Another example is piecewise expanding maps of an interval [10, 25]. Recall that a function $g : [0,1] \to \mathbb{R}$ is of bounded variation, $g \in BV$, if

$$\operatorname{var}(g) = \sup_{0 \leq t_0 < t_1 < \cdots < t_k \leq 1} \sum_{j=1,\dots,k} |g(t_j) - g(t_{j-1})| < \infty.$$

The norm $\|g\|_{BV} = |g|_1 + \operatorname{var}(g)$ is equivalent to $|g|_\infty + \operatorname{var}(g)$ (since $|g|_1 \leq |g|_\infty \leq |g|_1 + \operatorname{var}(g)$) and BV is a Banach space. If $f, g \in F_\theta$, then $\operatorname{var}(fg) \leq |f|_\infty \operatorname{var}(g) + |g|_\infty \operatorname{var}(f)$. A straightforward calculation shows that $\|fg\|_{BV} \leq 2\|f\|_{BV}\|g\|_{BV}$. Hence, BV is a Banach algebra.

Again, (F1), (F2) and (F4) are immediate, while (F3) is standard (see for example [22, Lemma 5]). Condition (F5) is discussed in detail in §6.

Proposition 4.4 *Suppose that (F4) is valid. Then (F5) is equivalent to the condition that there are constants $E > 0$, $\theta \in (0,1)$ such that $\|P^n v\| \leq E(|v| + \theta^n\|v\|)$, for all $v \in F$, $n \geq 1$.*

Proof If the conclusion holds, then choose n_0 large so that $\theta_0 = \theta^{n_0} E < 1$.

Conversely, suppose that (F4) and (F5) hold. By induction,

$$\|P^{jn_0}v\| \le C_0D_0(1+\theta_0+\cdots+\theta_0^j)|v|+\theta_0^j\|v\| \le C'|v|+\theta^{jn_0}\|v\|,$$

where $C' = C_0D_0/(1-\theta_0)$ and $\theta = \theta_0^{1/n_0}$. Write $n = jn_0 + k$ where $k < n_0$. Then $\|P^nv\| \le C'|P^kv| + \theta^{-1}\theta^n\|P^kv\|$ so the result follows with $E \ge C'C_0$ and $E \ge \theta_0^{-1}\max\{\|P\|, \|P^2\|, \ldots, \|P^{n_0}\|\}$. ∎

It follows easily that P has spectral radius at most 1 in F and F'.

Theorem 4.5 *Assume (F1)–(F5). Then the essential spectral radius $\rho_{\rm ess}$ of $P : F \to F$ is strictly less than 1. In fact $\rho_{\rm ess} \le \theta_0^{1/n_0}$ where θ_0, n_0 are as in (F5) and $\rho_{\rm ess} \le \theta$ where θ is as in Proposition 4.4.*

Proof See Hennion [18]. (See also [24].) ∎

Thanks to Theorem 4.5, we can apply the results of §2 (with $H = L^2$, F as given, and F' disregarded from now on) to investigate the sequence of partial sums $v_N = \sum_{j=0}^{N-1} v \circ f^j$ where $v \in F$. Consider the P-invariant splitting $F = F_1 \oplus F_2$ where $F_1 = \ker(U-I) = \ker(P-I)$ (recall $P = U^*$). Since $v_N = Nv$ when $v \in F_1$, we restrict attention to $v \in F_2$. Then Theorem 2.2 and Corollary 2.3 are valid. In particular, the variance $\sigma^2 = \lim_{N\to\infty} \frac{1}{N}|v_N|_2^2$ is defined for all $v \in F_2$, and $|v_N|_2 = \sqrt{N}\sigma + O(1)$.

If we assume further that $f : X \to X$ is ergodic, then the conclusions of §3 hold:

Theorem 4.6 *Assume (F1)–(F5) and that $f : X \to X$ is ergodic. Suppose that $v \in F$ has mean zero. Define $\sigma^2 = \lim_{N\to\infty} \frac{1}{N}|v_N|_2^2$. Then*

- *(a) $\{v_N\}$ satisfies the CLT and WIP with variance σ^2.*
- *(b) $\sigma^2 = 0$ if and only if v is a coboundary in F, and if and only if $v_N = o(N^{1/2})$ a.e. If $F \subset L^\infty$, then $\sigma^2 = 0$ if and only if $v_N = O(1)$ uniformly a.e.*
- *(c) If $F \subset L^\infty$, then the upper LIL holds for $\{v_N\}$.*
- *(d) If T is weak mixing, then we obtain exponential decay of correlations.*

Vector-valued observations In this section, we have deduced statistical properties of real-valued observations from quasicompactness of the operator $P : F \to F$ where F is a Banach space of real-valued observations. Statistical properties of vector valued observations in F^d can be proved just as in §3(c).

It is also immediate that hypotheses (F1)–(F5), which imply quasicompactness for $P : F \to F$, also imply quasicompactness for $P : F^d \to F^d$. (Clearly, the essential spectral radius on F is less than one if and only if the essential spectral radius on F^d is less than one.)

Another way to see quasicompactness on F^d is to note that hypotheses (F1)–(F5) immediately extend to the Banach spaces $F^d \subset (F')^d \subset L^1(X, \mathbb{R}^d)$ so that Theorem 4.5 can be applied directly on F^d. The observation that (F1)–(F5) hold for $F^d \subset (F')^d$ is crucial in §5.

Transfer operators So far, we have considered the Perron-Frobenius operator P corresponding to an f-invariant measure m. In applications, often quasicompactness is proved first for a preliminary Perron-Frobenius operator A corresponding to a "reference measure" ℓ that is not f-invariant. See §6 for examples. We shall refer to A as a *transfer operator* to distinguish it from the Perron-Frobenius operator P. The two measures are related by $dm = \varphi d\ell$ where the density function $\varphi > 0$ is a fixed point for the the transfer operator ($A\varphi = \varphi$). Hence $Pv = \varphi^{-1}A(\varphi v)$.

For many (but not all) applications, quasicompactness for A immediately implies quasicompactness for P. If multiplication by φ induces linear isomorphisms on F' and F, then we have the following useful result:

Lemma 4.7 *Suppose that A, B and M are linear operators on both F and F' such that $M : F' \to F'$ and $M : F \to F$ are linear isomorphisms and $B = M^{-1}AM$. Then $B : F' \to F'$ satisfies (F4) and/or (F5) if and only if A does.*

Proof First, suppose that A satisfies (F4). Since $|A^n| \le C_0$ for all n, $|B^n| = |M^{-1}A^nM| \le C_0|M^{-1}||M|$, so that B satisfies (F4) with constant $C_0|M^{-1}||M|$. Next, we verify (F5). By Proposition 4.4, $\|A^n v\| \le E(|v| + \theta^n\|v\|)$ and so

$$\begin{aligned}\|B^n v\| = \|M^{-1}A^nMv\| &\le \|M^{-1}\|\|A^nMv\| \le E\|M^{-1}\|(|Mv| + \theta^n\|Mv\|)\\ &\le E\|M^{-1}\|(|M||v| + \theta^n\|M\|\|v\|) \le E\|M^{-1}\|(|M| + \|M\|)(|v| + \theta^n\|v\|),\end{aligned}$$

so that B satisfies the condition in Proposition 4.4 with constant $E\|M^{-1}\|(|M| + \|M\|)$. Hence, B satisfies (F5). ∎

5 Compact group extensions

In this section, we consider compact group extensions. The aim is to establish statistical properties of *equivariant* vector-valued observations which were introduced in [28] and occur naturally in applications to dynamical systems with symmetry.

Our treatment closely follows [16] who considered group extensions of Axiom A diffeomorphisms. In [16], use was made of the equivariant Ruelle operator which was studied by Parry & Pollicott in [30, 31]. More generally, we consider an equivariant Perron-Frobenius operator which is a twisted version of the usual Perron-Frobenius operator. Our main results generalise results of [16] in the Axiom A setting, with the improvements that we do not require weak mixing nor that G is connected.

Suppose that f is a measure-preserving transformation on (X,μ). We assume that the Banach spaces $F \subset F' \subset L^1(X,\mathbb{R})$ satisfy axioms (F1)–(F5) in §4. These properties are inherited by $F^d \subset (F')^d \subset L^1(X,\mathbb{R}^d)$. From now on, we write F instead of F^d and F' instead of $(F')^d$, so that $F \subset F' \subset L^1(X,\mathbb{R}^d)$.

Let G be a compact Lie group with Haar measure ν. Given $h : X \to G$ measurable, consider the skew product $T : X \times G \to X \times G$ given by $T(x,g) = (fx, gh(x))$. Then T is a measure preserving transformation on $(X \times G, m)$ where $m = \mu \times \nu$.

Suppose that G acts orthogonally on $\mathbb{R}^d$. For each $g \in G$, write $M_g v = gv$ and given $h : X \to G$, write $(M_h v)(x) = h(x)v(x)$. Note that M_h is an isometry on $L^p(X,\mathbb{R}^d)$ for all $p \geq 1$. Let $h_n(x) = h(x)h(fx)\cdots h(f^{n-1}x)$. To obtain control over the norms $|\;|$ and $\|\;\|$ on F' and F under the action of M_{h_n}, we restrict to measurable cocycles $h : X \to G$ that satisfy:

(G1) For all $n \geq 1$, $M_{h_n}^{-1}$ is a bounded operator on F'. Moreover, there is a constant $C_1 > 0$ such that $|M_{h_n}^{-1}| \leq C_1$ for all $n \geq 1$.

(G2) Let n_0 and θ_0 be as in (F5). Then $M_* = M_{h_{n_0}}^{-1}$ is a bounded operator on F, and moreover there exist constants $D_1 > 0$ and $c \in (0, \frac{1}{\theta_0})$ such that $\|M_* v\| \leq D_1|v| + c\|v\|$ for all $v \in F$.

Remark 5.1 Suppose that $f : X \to X$ and $h : X \to G$ are measurable. Since G is compact, automatically $h \in L^\infty(X,G)$ with $|h|_\infty = 1$. If $F' = L^p(X,\mathbb{R}^d)$ for some $1 \leq p \leq \infty$, then (G1) is satisfied with $C_1 = 1$. The same is true if f and h are continuous and $F' = C(X,\mathbb{R}^d)$. Hence (G1) is satisfied in Examples 4.2 and 4.3.

Example 5.2 Continuing Example 4.2, if $F = F_\theta(X,\mathbb{R}^d)$ and $F' = L^\infty(X,\mathbb{R}^d)$, then it is natural to restrict to cocycles $h : X \to G$ that are Lipschitz with respect to the metric d_θ on X. We denote the space of such cocycles by $F_\theta(X,G)$. Then $|M_* v|_\theta \leq |M_*|_\theta |v|_\infty + |M_*|_\infty |v|_\theta = |M_*|_\theta |v|_\infty + |v|_\theta$, and so $\|M_* v\|_\theta = |M_* v|_\infty + |M_* v|_\theta \leq \|h_{n_0}^{-1}\|_\theta |v|_\infty + \|v\|_\theta$. Hence (G2) is satisfied for $h \in F_\theta(X,G)$ with $D_1 = \|h_{n_0}^{-1}\|_\theta$ and $c = 1$.

Example 5.3 Continuing Example 4.3, if $F = BV$ (with $X = [0,1]$), then it is natural to restrict to BV cocycles $h : X \to G$. In particular, for certain classes of piecewise monotone functions $f : X \to X$, including functions whose domains of monotonicity form a finite partition of $[0,1]$, it is easily seen that $M_{h_n}^{-1} : BV \to BV$ is a bounded operator for each n, and $\|M_{h_n}^{-1}\| = \|h_n^{-1}\|_{BV} = 1 + \mathrm{var}(h_n^{-1})$.

In contrast to the previous example, it is necessary to add a further restriction on h. Let n_0 and θ_0 be the constants in (F5). We assume that $\mathrm{var}(h_{n_0}^{-1}) < \frac{1}{\theta_0} - 1$. (Equivalently $\|h_{n_0}^{-1}\|_{BV} = c < \frac{1}{\theta_0}$.) Then $\|M_*\| \le c < \frac{1}{\theta_0}$ and (G2) is satisfied with $D_1 = c = \|h_{n_0}^{-1}\|_{BV}$.

Alternatively, if $f : X \to X$ is piecewise Lipschitz, then we can take $h : X \to G$ to be piecewise Lipschitz with no restriction on the Lipschitz constant $L(h)$. The crucial estimate is $\mathrm{var}(M_* v) \le L(h_{n_0}^{-1})|v|_1 + \mathrm{var}(v)$. (Note that the L^1 norm appears in this estimate, whereas before we had only the L^∞ norm. The function spaces $F' = L^1$, $F = BV$ are as before.)

For $h : X \to G$ measurable, we define the *equivariant Perron-Frobenius operator* $P_h : L^1(X, \mathbb{R}^d) \to L^1(X, \mathbb{R}^d)$ by $P_h v = P M_h^{-1} v$. Observe that $P_h^n = P^n M_{h_n}^{-1}$.

Proposition 5.4 *Assume (F1), (F4), (F5), (G1) and (G2). Then*

(a) There is a constant $C_1' > 0$ such that $|P_h^n| \le C_1'$ for all $h \in H$, $n \ge 1$.

(b) There are constants $D_1' > 0$ and $\theta_0' \in (0,1)$ such that $\|P_h^{n_0} v\| \le D_1'|v| + \theta_0'\|v\|$ for all $v \in F$.

Proof For $v \in F'$, $|P_h^n v| = |P^n M_{h_n}^{-1} v| \le C_0 |M_{h_n}^{-1} v| \le C_0 C_1 |v|$, proving (a). To prove (b), compute that

$$\|P_h^{n_0} v\| = \|P^{n_0} M_\star v\| \le D_0 |M_\star v| + \theta_0 \|M_\star v\| \le D_0 C_1 |v| + \theta_0 (D_1 |v| + c\|v\|)$$

so the result follows with $D_1' = D_0 C_1 + \theta_0 D_1$ and $\theta_0' = c\theta_0$. ∎

Let L_G^p consist of *equivariant observations* $\phi : X \times G \to \mathbb{R}^d$ of the form $\phi(x, g) = M_g v(x)$ where $v \in L^p(X, \mathbb{R}^d)$ and *define* $|\phi|_p = |v|_p$. Symbolically, we can write $L_G^p = g \cdot L^p(X, \mathbb{R}^d)$.

Proposition 5.5 *Let $\widehat{P} : L^1(X \times G, \mathbb{R}^d) \to L^1(X \times G, \mathbb{R}^d)$ denote the Perron-Frobenius operator corresponding to the G-extension $T : X \times G \to X \times G$. Then $\widehat{P}|_{L_G^1} = M_g P_h M_g^{-1}$.*

Proof Let $\phi = g \cdot v \in L_G^1$ and $\psi = g \cdot w \in L_G^\infty$. Then

$$\begin{aligned}\int_{X\times G} M_g P_h M_g^{-1} \phi \cdot \psi^\tau \, dm &= \int_X P_h v \cdot w^\tau d\mu = \int_X P M_h^{-1} v \cdot w^\tau d\mu \\ &= \int_X M_h^{-1} v \cdot w^\tau \circ f \, d\mu = \int_X v \cdot (M_h w \circ f)^\tau d\mu \\ &= \int_{X\times G} M_g v \cdot (M_h w \circ f)^\tau M_g^\tau \, dm = \int_{X \times G} \phi \cdot \psi^\tau \circ T \, dm.\end{aligned}$$

∎

Similarly, define the spaces $F_G = g \cdot F$ with norm $\|g \cdot v\| = \|v\|$ and so on.

Proposition 5.6 *The operator $\widehat{P} = M_g P_h M_g^{-1}$ restricts to F'_G and F_G and satisfies properties (F1)–(F5).*

Proof Properties (F1)–(F3) are immediate since F'_G and F_G are isomorphic to F' and F. Properties (F4) and (F5) are Proposition 5.4(a) and (b) respectively. ∎

We can now apply Theorem 4.5 to deduce that $\widehat{P} : F_G \to F_G$ is quasicompact. We obtain the following conclusions:

Theorem 5.7 *Assume (F1)–(F5) for $f : X \to X$ and (G1), (G2) for $h : X \to G$. Suppose that $T : X \times G \to X \times G$ is ergodic and that $\phi \in F_G$ has mean zero. Then*

(a) The $d \times d$ covariance matrix $\Sigma = \lim_{N\to\infty} \frac{1}{N} \int_{X\times G} \phi_N \cdot \phi_N^T \, dm$ is well-defined and $\Sigma : \mathbb{R}^d \to \mathbb{R}^d$ commutes with the action of G on $\mathbb{R}^d$.

(b) $\{\phi_N\}$ satisfies d-dimensional versions of the CLT and WIP on $X \times G$ with covariance matrix Σ.

(c) $\det \Sigma = 0$ if and only if there is a G-invariant subspace $V \subset \mathbb{R}^d$ such that $\pi_V \phi$ is a coboundary in F_G, and if and only if $\pi_V \phi_N = o(N^{1/2})$. If $F \subset L^\infty$, then $\sigma^2 = 0$ if and only if there is a G-invariant subspace V such that $\pi_V \phi_N = O(1)$.

(d) If $F \subset L^\infty$, then the upper LIL holds for each component of $\{\phi_N\}$.

(e) If T is weak mixing, then we obtain exponential decay of correlations.

Proof The definition of Σ and parts (b), (d) and (e) are immediate from quasicompactness and §3. The statement about Σ commuting with G and part (c) are proved as in [28] (see also [16]). ∎

Remark 5.8 By [27], we obtain the stronger results that the CLT, WIP and upper LIL hold also on $X \times \{g_0\}$ for each fixed $g_0 \in G$. In the remainder of this paper, we will not mention this explicitly.

6 Applications

In this section we give examples of dynamical systems to which our results apply.

Let $f : X \to X$ be a mapping with certain regularity properties. Consider a transfer operator A acting on sufficiently regular functions v as $(Av)(x) = \sum_{fy=x} g(y)v(y)$ where g is a positive bounded function (again with certain regularity properties).

Suppose that ℓ is a Borel probability measure satisfying $A^*\ell = \ell$ (that is $\int A^*v\, d\ell = \int v\, d\ell$ for all v). Then $\int Av\, w\, d\ell = \int v\, w \circ f\, d\ell$. Thus A is the Perron-Frobenius operator corresponding to the measure ℓ. We do not assume that ℓ is f-invariant.

Now suppose that $\varphi > 0$ is a fixed point ($A\varphi = \varphi$) for A acting on a suitable function space. Then we define the f-invariant measure $dm = \varphi\, d\ell$ and the corresponding Perron-Frobenius operator $Pv = \varphi^{-1}A(\varphi v)$.

In the rest of this section, we refer to ℓ as a reference measure and m as an equilibrium measure. Note that in earlier sections, we chose to work directly with P and m, bypassing A and ℓ.

(a) One dimensional monotone maps

Hofbauer & Keller [22] have analysed a class of endomorphisms of a totally ordered, order complete set X (usually X is taken to be the unit interval $[0,1]$). Here $f : X \to X$ is piecewise monotonic and order-continuous and the transfer operator A acts on the space of bounded measurable functions by $Av(x) = \sum_{fy=x} g(y)v(y)$ where g is a function of bounded variation on $[0,1]$ with $0 < g(x) \le d < 1$. They show the existence of a reference Borel probability measure ℓ on X satisfying $A^*\ell = \ell$ in the sense that $\int Av\, d\ell = \int v\, d\ell$ for all bounded measurable $v : X \to \mathbb{R}$. Under these assumptions, there exists a maximal absolutely continuous invariant measure $dm = \varphi\, d\ell$ where φ is a density of bounded variation.

Examples of the systems considered in [22] include:

(i) Lasota-Yorke maps: piecewise monotonic C^2 transformations f of the unit interval $[0,1]$ which satisfy $|f'| > 1$. In this setting, ℓ is Lebesgue measure and $g(x) = \frac{1}{|f'(x)|}$. For earlier results on these maps see [25, 26, 38, 36, 9].

(ii) Piecewise monotonic transformations f on $[0,1]$ with $h_{\mathrm{top}}(f) > 0$ [20, 21]. Setting $g(x) = \exp(-h_{\mathrm{top}}(f)) < 1$, it can be shown that there exists ℓ such that $A^*\ell = \ell$, leading to an f-invariant measure $dm = \varphi\, d\ell$ of maximal entropy.

(iii) The β-transformation $fx = \beta x \pmod 1$ on $[0,1]$. Walters [37] constructed equilibrium measures corresponding to Lipschitz potentials $\phi : [0,1] \to \mathbb{R}$. The class of allowable potentials is extended in [22], and defining $g(x) = \exp(\sum_{i=1}^{n-1} \phi(f^i x))/\lambda^n$ for suitable choices of $\lambda > 0$ and $n \ge 1$, it is shown that there exists a Borel probability measure ℓ such that $A^*\ell = \ell$. Again, this leads to an f-invariant equilibrium measure $dm = \varphi\, d\ell$ with potential function ϕ.

For the class of transformations (i) and (iii) above the density φ, which is of bounded variation, is bounded above and below on the support of m. That is, there exists $C \ge 1$ such that $0 < \frac{1}{C} < \varphi < C$. The same is true for class (ii) under the assumption that $\frac{l(fI)}{l(I)}$ is bounded over intervals $I \subset [0,1]$ (here l is Lebesgue measure). This is proved in an unpublished preprint of Keller [23] in the context of Lasota-Yorke maps but it is easily seen that the proof generalizes to class (iii) and to class (ii) if we also require $\sup_{I\subset[0,1]}\frac{l(fI)}{l(I)}$ bounded. We let class (ii') denote the subset of class (ii) maps for which $\sup_{I\subset[0,1]}\frac{l(fI)}{l(I)}$ is bounded.

Lemma 6.1 *Let $f:[0,1]\to[0,1]$ be a one-dimensional piecewise monotone map and let $g\in BV$ with transfer operator $A:BV\to BV$ given by $Av(x)=\sum_{fy=x}g(y)v(y)$. Suppose that f and g fall into one of the three classes (i), (ii) or (iii). Let ℓ and m be the corresponding reference and equilibrium measures with density $\varphi>0$ in BV. Then A satisfies hypotheses (F1)–(F5) with $F'=L^1(m)$ and $F=BV$. In particular, $A:BV\to BV$ is quasicompact.*

Suppose further that $\varphi^{-1}\in BV$ (certainly the case for the classes (i), (ii'), (iii)). Define the Perron-Frobenius operator $Pv=\varphi^{-1}A(\varphi v)$ corresponding to the invariant measure m. Then P satisfies hypotheses (F1)–(F5) with $F'=L^1(m)$ and $F=BV$. In particular, $P:BV\to BV$ is quasicompact.

Proof We largely follow Hofbauer & Keller [22]. Conditions (F1)–(F3) are already discussed in Example 4.3. Since $\varphi,\varphi^{-1}\in F$, it follows that $v\mapsto\varphi v$ is a linear isomorphism on F' and F. (Indeed, it is clear that $|\varphi v|_1\le|\varphi|_\infty|v|_1$, and $\|\varphi v\|_{BV}\le 2\|\varphi\|_{BV}\|v\|_{BV}$ was established in Example 4.3.) Hence by Lemma 4.7 it suffices to verify (F4) and (F5) either for A or P.

Condition (F4) is immediate for P by Remark 4.1. (Alternatively, see [22, Lemma 6].) The crucial condition (F5) is proved for A in [22, Lemma 7]. ∎

We are now in a position to apply the results in §§2, 3 and 4. Assume that the equilibrium measure m in Lemma 6.1 is ergodic. Let $v\in BV$ with $\int_X v\,dm=0$.

- The variance $\sigma^2=\lim_{N\to\infty}\frac{1}{N}\int_X v_N^2\,dm$ exists, and $\sigma^2=0$ if and only if $v=\chi\circ f-\chi$ for some $\chi\in BV$ (or equivalently, v_N is uniformly bounded). In particular, if $x\in X$ is a periodic point of period k, f and v are continuous at f^jx for $j=1,\dots,k$, and $v_k(x)\ne 0$, then $\sigma^2>0$.
- The sequence of partial sums v_N satisfies the CLT, WIP, and upper LIL.
- If in addition m is weak mixing, then we obtain exponential decay of correlations.

Remark 6.2 Under the assumption of weak mixing, various statistical results were obtained in [22]. We have extended the CLT, WIP and upper LIL to the case where m is ergodic but not necessarily weak mixing. Also, we have obtained strong conditions for nondegeneracy that are not present in [22]. [1]

[1]Whilst writing this paper, we learned of independent work of [19] who obtain the CLT also without assuming weak mixing. The methods in [19] are different from ours and the results are also somewhat different. For example, they prove the CLT with error estimates and they prove the local CLT. The WIP is not explicitly stated, but should follow from their methods. On the other

G-extensions of one dimensional monotone maps We continue to suppose that $X = [0,1]$ and that $f : X \to X$ is piecewise monotone. Suppose that m is an ergodic measure belonging to classes (i), (ii') or (iii) above.

Now let G be a compact Lie group acting on $\mathbb{R}^d$. Let $h : X \to G$ be a BV cocycle and form the compact group extension $T(x,g) = (fx, gh(x))$. We suppose further that $\mathrm{var}(h)$ is sufficiently small, in the sense of Example 5.3, guaranteeing that h satisfies (G1) and (G2). In addition, we assume that $T : X \times G \to X \times G$ is ergodic with respect to $m \times \nu$ where ν is Haar measure on G.

As in §5, we define the space BV_G of G-equivariant observations $\phi(x,g) = gv(x)$ where $v : X \to \mathbb{R}^d$ is BV. Suppose that $\phi \in BV_G$ and $\int_{X\times G} \phi \, d(m \times \nu) = 0$.

By Theorem 5.7 we have the following results:

- The covariance matrix $\Sigma = \lim_{N\to\infty} \frac{1}{N} \int_{X\times G} \phi_N \phi_N^T \, d(m \times \nu)$ exists. Moreover, $\det \Sigma = 0$ if and only if there is a component of ϕ_N that is uniformly bounded a.e.
- The sequence of partial sums ϕ_N satisfies the d-dimensional CLT and WIP.
- The components of ϕ satisfy the upper LIL.
- If in addition $m \times \nu$ is weak mixing, then we obtain exponential decay of correlations.

Remark 6.3 Results of [29] show that compact group extensions of Lasota-Yorke maps are weak mixing for a residual, prevalent subset of Hölder compact group extensions of Lasota-Yorke maps. If G is semisimple then compact group extensions of Lasota-Yorke maps are weak mixing for an open, dense and prevalent subset of extensions. Hence the hypotheses required for our probabilistic properties to hold are 'typically' valid for extensions of Lasota-Yorke maps.

(b) Nonuniformly hyperbolic diffeomorphisms

In this section we follow very closely the original exposition of Young [39] referring also to Baladi [1, §4.3]. The tower approach is applicable to the following $C^{1+\epsilon}$ diffeomorphisms $\mathcal{F}$ defined on a Riemannian manifold M:
(i) Lozi maps and certain piecewise hyperbolic maps [39, 12],
(ii) a class of Hénon maps [3, 4],
(iii) Poincaré maps of billiards with convex scatterers [39] and certain other dispersing billiards [13],
(iv) some partially hyperbolic diffeomorphisms with a mostly contracting direction [11, 15].

---

hand, they do not prove the upper LIL and their results on nondegeneracy seem weaker than ours in general. The axiomatic frameworks are somewhat different and it is not clear that the hypotheses in [19] apply for all of the examples considered in this section.

The approach also applies to

(v) C^2 unimodal maps satisfying conditions (H1) and (H2) of Young [39, §9.1].

For these systems a Markov tower $(\overline{f}, \overline{\Delta})$ is constructed. The set $\overline{\Delta}$ is partitioned into countably many levels $\{\overline{\Delta}_j\}_{j=0}^{\infty}$. The base $\overline{\Delta}_0$ is further partitioned into countably many subsets $\{\overline{\Delta}_{0,j}\}$ by a return time function $R : \overline{\Delta}_0 \to \mathbb{N}$ such that $R|\overline{\Delta}_{0,j} = R_j$ is constant on each $\overline{\Delta}_{0,j}$. The map $\overline{f}$ moves each set $\overline{\Delta}_{0,j}$ up the tower until the level $\overline{\Delta}_{R_j-1}$ is reached, and $\overline{f}^{R_j}$ maps $\overline{\Delta}_{0,j}$ bijectively onto $\overline{\Delta}_0$. The levels Δ_l are further subdivided so that the partition $P = \{\overline{\Delta}_{l,j}\}$ has the Markov property. A separation time $s(\cdot,\cdot)$ is defined for all pairs x, y in the same $\overline{\Delta}_{l,j}$; $s(x,y)$ is the largest $n \geq 0$ such that $\overline{f}^n x$ lies in the same element of P as $\overline{f}^n y$. It is assumed [39, Condition P4] that there exists $0 < \alpha < 1$ such that $d(\mathcal{F}^n x, \mathcal{F}^n y) \leq C\alpha^{s(x,y)-n}$ for all $y \in \gamma^u(x)$ (here γ^u is an unstable disk or manifold [39, Definition 1]).

A non-invertible tower (f, Δ) is derived by quotienting $(\overline{f}, \overline{\Delta})$ along stable manifolds (the quotiented tower is not necessary for the unimodal maps described in (v)). Denote this projection $\overline{\pi} : \overline{\Delta} \to \Delta$ and write corresponding objects under this quotient map without bars. The map $f^{R_j} : \Delta_{0,j} \to \Delta_0$ is uniformly expanding. A reference measure ℓ, equivalent to Lebesgue, is constructed on Δ [39, §3].

By studying the transfer operator A with weight $g = \frac{1}{\mathrm{Jac}(f)}$ acting on a suitable space of functions, Young obtains an absolutely continuous invariant measure $dm = \varphi\, d\ell$, with density $C^{-1} \leq \varphi \leq C$ bounded above and below.

The measure m lifts to an invariant measure $\overline{m}$ on $\overline{\Delta}$ and thence to an invariant measure μ for $\mathcal{F} : M \to M$. We assume, following [39], that there exists $\epsilon > 0$ such that $\sum_{l=0}^{\infty} m(\Delta_l) e^{2\epsilon l} < \infty$ (equivalently $\sum_{l=0}^{\infty} \ell(\Delta_l) e^{2\epsilon l} < \infty$) . (The scaling factor $2\epsilon l$ rather than ϵl ensures that the Banach space F (to be defined) satisfies $F \subset L^2_m$.)

The underlying observations $v : M \to \mathbb{R}$ are assumed to be Hölder continuous with fixed exponent $\gamma \in (0,1)$. Take $0 < \beta < 1$ such that $\beta \geq \max\{\sqrt{\alpha}, \alpha^{\gamma}\}$. This implies that $d(x,y) \leq \beta^{s(x,y)}$. In fact $\beta^{s(x,y)}$ defines a metric on Δ, as $\beta^{s(x,y)} = 0$ implies $x = y$ and the triangle inequality is immediate from the definition of $s(x,y)$.

For $v : \Delta \to \mathbb{R}$ measurable, define

$$\|v\|_\infty = \sup_{l,j} \sup_{\Delta_{l,j}} |v| e^{-l\epsilon}, \qquad \|v\|_\beta = \sup_{l,j} |v|_{l,j,\beta}, \quad \text{where}$$

$$|v|_{l,j,\beta} = \Big(\sup\nolimits_{x,y \in \Delta_{l,j}} \frac{|v(x) - v(y)|}{\beta^{s(x,y)}}\Big) e^{-\epsilon l}.$$

Define $\|v\| = \|v\|_\infty + \|v\|_\beta$ and (see [1, §3.4, p. 203]) define F to be the Banach space of functions $v : \Delta \to \mathbb{R}$ with $\|v\| < \infty$. Let $F' = L^1_m(\Delta)$. The transfer operator A is well-defined on F and on F'. As sketched below, conditions (F1)–(F5) are valid. Since $\beta^{s(x,y)}$ defines a metric on Δ, elements of F are Lipschitz on each level Δ_l.

The density φ lies in F, as does φ^{-1}. Additionally, φ and φ^{-1} are uniformly bounded and uniformly Lipschitz [39, Lemma 2]. Hence $v \mapsto \varphi v$ is a linear isomorphism on F' and F so that Lemma 4.7 is applicable.

(F1) It suffices to work on each level Δ_l separately. Note that F restricted to Δ_l is densely embedded in F' (similarly restricted) since the space of Lipschitz functions on a compact measure metric space (X, μ) is densely embedded in $L^1_\mu(X)$.

(F2) If $v \in F$, then $|v|_{\Delta_l} \leq \|v\|_\infty e^{\epsilon l}$ and thus $\sum_l e^{2\epsilon l} m(\Delta_l) < \infty$ implies $F \subset L^2_m$. A suitable scaling of the norm $\| \, \|$ yields $\| \, \| \geq | \, |_2$.

(F3) Suppose that $\{v_n\}$ is a sequence in F with $\|v_n\| \leq 1$. Restricting to Δ_l, we have $\|v_n|\Delta_l\| \leq e^{\epsilon \ell}$. By Arzela-Ascoli, there exists $w : \Delta_l \to \mathbb{R}$ with $\|w\| \leq e^{\epsilon \ell}$ and a subsequence with $\|v_n|\Delta_l - w\|_\infty \to 0$. Altogether, we obtain a function $w : \Delta \to \mathbb{R}$ with $w \in F$ and $\|w\| \leq 1$. By a standard diagonal argument there is a single subsequence such that $\|(v_n - w)|\Delta_l\|_\infty \to 0$ for all l. Since v_n, w are bounded on Δ_l, $\|(v_n - w)|\Delta_l\|_1 \to 0$ for each l. Since $\sum_{l=0}^\infty m(\Delta_l) < \infty$, it follows that $\|v_n - w|_1 \to 0$.

(F4) It is immediate that P satisfies (F4) by Remark 4.1. (By Lemma 4.7, it is also the case that A satisfies (F4).)

(F5) This condition is verified for A in [39, Lemma 3] and [1, Lemma 3.7]. Applying Lemma 4.7, we have that P satisfies (F5).

We are now once again in a position to apply the results in §§2, 3 and 4. Assume that m is ergodic and that $v \in F$ with $\int_\Delta v \, dm = 0$.

- The variance $\sigma^2 = \lim_{N \to \infty} \frac{1}{N} \int_\Delta v_N^2 \, dm$ exists, and $\sigma^2 = 0$ if and only if $v = \chi \circ f - \chi$ for some $\chi \in F$ (in particular, v_N is uniformly bounded on each level Δ_ℓ). If $x \in \Delta$ is a periodic point of period k, and $v_k(x) \neq 0$, then $\sigma^2 > 0$.
- v_N satisfies the CLT and the WIP.
- If in addition m is weak mixing, then we obtain exponential decay of correlations.

Remark 6.4 The CLT and exponential decay of correlations is obtained in Young [39] when m is weak mixing. The CLT for m ergodic is new, as are the strong nondegeneracy results. The WIP is not stated in [39] but follows in a standard way from the set up there. As far as we know, the upper LIL remains open. Note that the condition $F \subset L^\infty$ in Theorem 3.5 is violated.

G-extensions Suppose that a quotiented tower (f, Δ, m) has been constructed as above. Let G be a compact Lie group acting on $\mathbb{R}^d$. Let $h : \Delta \to G$ be uniformly Lipschitz and let $T : \Delta \times G \to \Delta \times G$ denote the corresponding G-extension.

Hypothesis (G1) is immediate by Remark 5.1. It is easily seen that (G2) is satisfied with $D_1 = 1 + L(h_n^{-1})$ and $c = 1$. Define the space F_G of G-equivariant observations $\phi(x, g) = gv(x)$ where $v : \Delta \to \mathbb{R}^d$ lies in F. Suppose that $m \times \nu$ is ergodic where ν

is Haar measure on G. Let $\phi \in F_G$ with $\int_{\Delta\times G} \phi \, d(m \times \nu) = 0$. By Theorem 5.7 we have the following results:

- The covariance matrix $\Sigma = \lim_{N\to\infty} \frac{1}{N} \int_{\Delta\times G} \phi_N \phi_N^T \, d(m \times \nu)$ exists. Moreover, $\det \Sigma = 0$ if and only if there is a nonzero vector $c \in \mathbb{R}^d$ such that $c\cdot\phi = c\cdot\psi\circ f - c\cdot\psi$ where $\psi \in F_G$.
- ϕ_N satisfies the d-dimensional CLT and WIP on $\Delta \times G$.
- If in addition $m\times\nu$ is weak mixing, then we obtain exponential decay of correlations.

Observations on M We now relate the quasicompactness of the equivariant Perron-Frobenius operator on the quotiented tower $\Delta \times G$ to the statistical properties of equivariant observations on $M \times G$.

Define the space F_G of G-equivariant observations on $M \times G$, $\phi(x, g) = gv(x)$ where $v : M \to \mathbb{R}^d$ is Hölder of exponent η. Let $h : M \to G$ lie in the space of Lipschitz cocycles. We form the G-extension $\mathcal{T}(x, g) = (\mathcal{F}x, gh(x))$.

We let $\overline{h}$ denote the lift of h to $\overline{\Delta}$ and similarly define $\overline{v}, \overline{\phi}$. Let $\mathcal{B}$ denote the σ-algebra on Δ and define $\mathcal{B}_s = \{\overline{\pi}^{-1}A : A \in \mathcal{B}\}$ where $\overline{\pi} : \overline{\Delta} \to \Delta$ is the projection. Following Young [39, §5.2] we let $h_0 = E_{\overline{m}}(\overline{h}|\mathcal{B}_s)$ denote the conditional expectation of $\overline{h}$ with respect to $\mathcal{B}_s$. Similarly, define v_0 and $\phi_0(x, g) = gv_0(x)$. This defines v_0, h_0 on Δ as well and ϕ_0 on $\Delta \times G$. The assumption that $\beta \geq \max\{\sqrt{\alpha}, \alpha^\eta\}$ implies that $v_0 \in F$ and $\phi_0 \in F_G$.

Form the G-extension of $(\overline{f}, \overline{\Delta}, \overline{m})$ by defining $\overline{T}(\overline{x}, g) = (\overline{f}\overline{x}, g\overline{h}(\overline{x}))$. We construct the natural extension of $\overline{T} : \overline{\Delta} \times G \to \overline{\Delta} \times G$, with invariant measure $\widehat{m} \times \nu$, and denote it $\widehat{T} : Y \to Y$, where $Y = (\overline{\Delta} \times G)^{\mathbb{Z}^+}$. Let $\widehat{\pi} : Y \to \overline{\Delta} \times G$ denote the natural projection and let $\widehat{\mathcal{B}}$ denote the σ-algebra $(\widehat{\pi}^{-1}\mathcal{B}_s) \times \mathcal{B}_G$, where $\mathcal{B}_G$ is the usual Borel algebra on G. Lift $\overline{\phi}$ to Y and denote the lift $\widehat{\phi}(y) = \overline{\phi}(\widehat{\pi}y)$. To establish the CLT we apply Gordin [17] as done in Young [39].

Let $\widehat{\phi}_j = E_{\widehat{m}}(\widehat{\phi}|\widehat{T}^j\widehat{\mathcal{B}})$. Since we have uniform contraction on stable manifolds by [39, Condition (P3)] $|\widehat{\phi}_j - \widehat{\phi}| \leq C\alpha^{j\eta}$ and hence $\sum_{j\geq 0} |\widehat{\phi}_j - \widehat{\phi}|_2 < \infty$ and so the first condition of Gordin's theorem is satisfied.

We now consider $\widehat{\phi}_{-j}$. Since the order of conditioning commutes, $E_{\widehat{m}}(\widehat{\phi}|\widehat{T}^{-j}\widehat{\mathcal{B}}) = E_m(\phi_0|F^{-j}\mathcal{B})|$ so it suffices to prove $\sum_{j\geq 0} |E_m(\phi_0|f^{-j}\mathcal{B})|_2 < \infty$. This follows immediately from quasicompactness of the equivariant transfer operator, hence proving the CLT. The WIP follows by standard techniques. Exponential decay of correlations follows as in [39, §4.1] or [1, Proposition 4.2].

Hence we have the following theorem for Hölder G-equivariant observations $\phi : M \times G \to \mathbb{R}^d$ with mean zero and ergodic Lipschitz cocycles $h : M \to G$:

- The covariance matrix $\Sigma = \lim_{N\to\infty} \frac{1}{N} \int_{M\times G} \phi_N \phi_N^T \, d(\mu \times \nu)$ exists. Moreover, $\det \Sigma = 0$ if and only if there is a nonzero vector $c \in \mathbb{R}^d$ such that $c\cdot\phi = c\cdot\psi\circ\mathcal{F} - c\cdot\psi$

where $\psi : M \times G \to \mathbb{R}^d$ is G-equivariant and Hölder.
• ϕ_N satisfies the d-dimensional CLT and WIP on $M \times G$.
• If in addition $\mu \times \nu$ is weak mixing, then we obtain exponential decay of correlations.

Acknowledgements We thank Gerhard Keller for very helpful discussions.

References

[1] V. Baladi. *Positive Transfer Operators and Decay of Correlations.* Advanced Series in Nonlinear Dynamics **16**, World Scientific, 2000.

[2] V. Baladi. *Decay of Correlations.* In: Smooth Ergodic Theory and Its Applications. *Proc. Symp. Pure Math.* **69** (2001) 297–325.

[3] M. Benedicks and L.-S. Young. Absolutely continuous invariant measures and random perturbations for certain one-dimensional maps. *Ergod. Th. & Dynam. Sys.* **12** (1992) 13–37.

[4] M. Benedicks and L.-S. Young. Sinai-Bowen-Ruelle measures for certain Hénon maps. *Invent. Math.* **112** (1993) 541–576.

[5] P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12** (1961) 788–792.

[6] P. Billingsley. *Convergence of Probability Measures.* Wiley, New York, 1968.

[7] P. Billingsley. *Probability and Measure.* Wiley, New York, 1986.

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

[9] R. Bowen. Bernoulli maps of an interval to itself. *Israel J. Math.* **28** (1978) 298–314.

[10] A. Boyarsky and P. Gora. *Laws of Chaos: Invariant measures and dynamical systems in one dimension.* Probability and Its Applications, Birkhäuser, 1997.

[11] A. Castro. Backward inducing and exponential decay of correlations for partially hyperbolic attractors with mostly contracting direction Ph. D. Thesis, IMPA (1998).

[12] N. Chernov. Statistical properties of piecewise smooth hyperbolic systems in high dimensions. *Discrete Contin. Dynam. Systems* **5** (1999) 425–448.

[13] N. Chernov. Decay of correlations and dispersing billiards. *J. Statist. Phys.* **94** (1999) 513–556.

[14] J.-P. Conze and S. Le Borgne. Méthode de martingales et flow géodésique sur une surface de courbure constante négative. *Ergod. Th. & Dynam. Sys.* **21** (2001) 421–441.

[15] D. Dolgopyat. On dynamics of mostly contracting diffeomorphisms. *Commun. Math. Phys.* **213** (2000) 181–201.

[16] 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.

[17] M. I. Gordin. The central limit theorem for stationary processes. *Soviet Math. Dokl.* **10** (1969) 1174–1176.

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

[19] H. Hennion and L. Hervé. *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness.* Lecture Notes in Math. **1766**, Springer, Berlin, 2001.

[20] F. Hofbauer. On intrinsic ergodicity of piecewise monotonic transformations with positive entropy. *Israel J. Math* **34** (1979) 213–237.

[21] F. Hofbauer. On intrinsic ergodicity of piecewise monotonic transformations with positive entropy, II. *Israel J. Math* **38** (1981) 107–115.

[22] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Zeit.* **180** (1982) 119–140.

[23] G. Keller. Piecewise monotonic transformations and exactness. *Preprint.* (1978).

[24] G. Keller and C. Liverani. Stability of the spectrum for transfer operators. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze* **XXVIII** (1999) 141–152.

[25] A. Lasota and J. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186** (1973) 481–488.

[26] T. Li and J. Yorke. Ergodic transformations from an interval into itself. *Trans. Amer. Math. Soc.* **235** (1978) 183–192.

[27] I. Melbourne and M. Nicol. Statistical limit laws for equivariant observations. In preparation.

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

[29] M. Nicol and A. Scott. Livšic theorems for group extensions of hyperbolic systems with singularities. Preprint, University of Surrey, 2001.

[30] W. Parry and M. Pollicott. The Chebotarev theorem for Galois coverings of Axiom A flows. *Ergod. Th. & Dynam. Sys.* **6** (1986) 133–148.

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

[32] V. A. Rohlin. Exact endomorphisms of a Lebesgue space. *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961) 499–530.

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

[34] S. Takahashi. Notes on the law of the iterated logarithm. *Studia Sci. Math. Hung.* **7** (1972) 21–24.

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

[36] G. Wagner. The ergodic behaviour of piecewise monotonic transformations. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **46** (1979) 317–324.

[37] P. Walters. Equilibrium states for β-transformations and related transformations. *Math. Z.* **159** (1978) 65–88.

[38] S. Wong. Some metric properties of piecewise monotonic mappings of the unit interval. *Trans. Amer. Math. Soc.* **246** (1978) 493–500.

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