Statistical properties of endomorphisms and compact group extensions

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

24 October, 2002

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, law of the iterated logarithm and exponential decay of correlations 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 and exponential decay of correlations 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, 34, 32]) 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, 36]).

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 [29], 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, almost all of our results do not require weak mixing or exactness.

In direction (b), we show that certain axioms for quasicompactness described in Keller and Liverani [25] lift to the compact group extension setting, at the level of equivariant observations. This extension is not immediate, and we are not claiming 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.

Two aspects that we do not focus on in this paper are (i) subexponential decay rates and (ii) invertible dynamical systems. The results in Section 3 do not require exponential decay rates of the Perron-Frobenius operator (with the obvious caveat that exponential decay of correlations is weakened accordingly). However, quasicompactness is required in the passage from endomorphisms to compact group extensions. A natural question is to determine whether the methods used to derive subexponential decay rates (such as Hilbert cones [28, 36] or coupling [41]) generalize to equivariant extensions.

The study of invertible dynamical systems often proceeds by constructing a Markov extension (the symbolic dynamics of Bowen [8] or the tower constructions of Hofbauer [22] and Young [40]). Quotienting out the contracting directions leads in these situations to an endomorphism to which the methods described above can be described. The results for the underlying dynamical system can then be recovered by standard techniques. In this paper, we focus attention on the results for endomorphisms.

The remainder of this paper is organised as follows. In Section 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 Section 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 Section 2. Under the additional assumption of weak mixing, we obtain exponential decay of correlations.

In Section 4, we discuss the axiomatic framework of [25] guaranteeing quasicompactness of the Perron-Frobenius operator. Compact group extensions are introduced in Section 5 and it is shown that the axioms for quasicompactness in Section 4 lift to spaces of equivariant observations.

In Section 6, we show that statistical limit laws such as the central limit theorem on the compact group extension $X \times G$ restrict to $X \times \{g_0\}$ for each $g_0 \in G$. (We note that the limit law on $X \times G$ follows directly from the limit law on $X \times \{g_0\}$ but the converse is nontrivial.) This section can be read independently from the rest of the paper.

Finally, a number of applications are considered in Section 7.

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

(i) U^* has at most finitely many eigenvalues on the unit circle,

(ii) these eigenvalues have finite multiplicity,

(iii) 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 $\rho < |\alpha| \le 1$.)

An immediate consequence of quasicompactness is that 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 $v \in F_2$, but now a stronger result is possible, as stated in Corollary 2.3 below. Without loss, we now restrict attention to the Banach space F_2.

Theorem 2.2 *Regarding U as an operator on F and U^* as on 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 $F_\alpha \in \mathrm{Im}(U - I)$.

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

It follows that if $v \in F_2$, then there is a decomposition

$$v = w + Uy - y,$$

where $w, y \in F_2$ and $U^*w = 0$.

Corollary 2.3 *Let $v = w + Uy - y$ be as in Theorem 2.2. Define $\sigma = \|w\|_H$. Then $\|w_N\|_H = \sqrt{N}\sigma$ and*

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

In particular,

$$\sigma = \lim_{N\to\infty} \tfrac{1}{\sqrt{N}}\|v_N\|_H.$$

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 $\|r_N\|_H \le 2\|y\|_H$. ∎

Remark 2.4 Let $v \in F_2$. The sequence v_N either exhibits $\sqrt{N}$ growth in H (if $\sigma > 0$) or is bounded in H (if $\sigma = 0$).

3 Statistical properties of dynamical systems

In this section, we apply the results in Section 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 (non-invertible) 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

$$UU^* = E(\cdot|f^{-1}\mathcal{B}),$$

where $E(v|f^{-1}\mathcal{B})$ is the conditional expectation operator.

Again, we define for $v \in H$,

$$v_N = \sum_{j=1}^{N} U^j v = \sum_{j=1}^{N} v \circ f^j.$$

By the mean ergodic theorem, $v_N = N\pi v + o(N)$ in H where $\pi : H \to H$ is the orthogonal projection onto $\ker(U - I)$.

Remark 3.1 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*

$$\sigma^2 = \lim_{N\to\infty} \frac{1}{N} \|v_N - N\pi v\|_H^2.$$

(a) Statistical limit laws

In this subsection, we show how the CLT, WIP and upper LIL follow from quasicompactness and ergodicity.

Theorem 3.2 (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

$$v_N = w_N + U^N y - y,$$

almost everywhere, where $w, y \in F$ and $U^* w = 0$. Moreover, since $y \in L^2$ it follows from the pointwise ergodic theorem that $U^N y = y \circ f^N = o(N^{1/2})$ almost everywhere. 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 [33], we obtain a bi-infinite ergodic stationary martingale $\{Y_j : \ j \in \mathbb{Z}\}$ where $Y_{-j} = w \circ f^j$ for $j \geq 0$ (cf. [16, Remark 3.12]). It follows from Billingsley [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 $\sigma^2 = \|w\|_H^2$. ∎

The CLT is said to be degenerate if $\sigma^2 = 0$. We have the following criteria for degeneracy.

Proposition 3.3 *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 almost everywhere. Indeed, $|v_N|_\infty \le 2|y|_\infty$.

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

(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 $v_N = o(N^{1/p})$ almost everywhere.

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 $w = 0$ almost everywhere, 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 3.2, 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 almost everywhere, and hence in distribution. But since f is ergodic, Theorem 3.2 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 3.4 In the literature, it is standard to state the nondegeneracy condition in the CLT in the form $\sigma^2 = 0$ if and only if v is an L^2 coboundary (that is, the cobounding function y lies in L^2). Our conditions in Proposition 3.3 follow [16] and are a substantial improvement on the standard L^2 condition.

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 $N \ge 1$.

If $z \in X$ is a periodic point, $f^p(z) = z$, then $v_{jp}(z) = j v_p(z)$ for all $j \ge 1$. Thus the condition $|v_N(x)| \le C$ implies that the periodic data $v_p(z)$ vanishes. We deduce that if z is a periodic point of period p and $v_p(z) \ne 0$, then $\sigma^2 > 0$.

The idea that nonvanishing periodic data implies nondegeneracy can be applied even when the continuity assumptions on f and v are relaxed. Suppose that $F \subset L^\infty$ and there exists an open subset $U \subset X$ on which f and v are continuous. Let $z \in U$ be a periodic point of period p such that $\{z, fz, \ldots, f^{p-1}z\} \subset U$. If $v_p(z) \neq 0$, then $\sigma^2 > 0$.

Set $W_N(0) = 0$, and

$$W_N(t) = \tfrac{1}{\sqrt{N}} v_{Nt} = \frac{1}{\sqrt{N}} \sum_{j=0}^{Nt-1} v \circ f^j, \quad t = 1/N, 2/N, \ldots$$

Linearly interpolating on each interval $[(r-1)/N, r/N]$, $r \geq 1$, we obtain a sequence of random elements $W_N \in C([0,\infty), \mathbb{R}^d)$.

The weak invariance principle (which is a refinement of the CLT) states that the sequence $\{W_N\}$ converges weakly to Brownian motion. Recall that Brownian motion is a stochastic process $W : [0,\infty) \times \Omega \to \mathbb{R}$ such that (i) $W(0) = 0$ almost surely, (ii) there exists $\sigma > 0$ such that $W(t)$ has distribution $N(0, t\sigma)$ for each $t \geq 0$, and (iii) for each $0 \leq t_1 < t_2 < \cdots < t_k$, the increments $W(t_1)$, $W(t_2) - W(t_1), \ldots, W(t_k) - W(t_{k-1})$ are independent random vectors. It is a basic property of Brownian motion that the random element W lies almost surely in $C([0,\infty), \mathbb{R})$.

Theorem 3.5 (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. (In other words, the measures induced by W_N on $C([0,\infty), \mathbb{R})$ converge weakly to Wiener measure.)*

Proof Billingsley [6] proves the WIP for stationary ergodic L^2 martingales, so the result follows along the lines of Theorem 3.2. ∎

Theorem 3.6 (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*

$$\limsup_{N\to\infty} \frac{v_N}{\sqrt{2N \log\log N}} \leq \sigma,$$

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, Section 3(c)], the condition $U^* w = 0$ implies that $\{w_N\}$ is a "weakly multiplicative sequence". Since

$w \in F \in L^\infty$, the result follows by a result of Takahashi [35] for uniformly bounded weakly multiplicative sequences. ∎

Remark 3.7 (a) For future reference, we note that 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) If we pass to the natural extension [33], then it follows from the methods in [16] that the LIL can be proved in backwards time. Indeed, [16] obtain the almost sure invariance principle (ASIP) in backwards times, and the ASIP implies the CLT, WIP, LIL, and much more besides. Moreover, the L^∞ assumption is not required.
(c) In [16], the required statistical properties are obtained in the correct time direction for ergodic compact Lie group extensions of Axiom A base dynamics. (It is assumed in [16] that the group extension is weak mixing and the group is connected, but these hypotheses are relaxed in the present paper.) In addition to the structure assumed in this paper, the approach in [16] requires two further 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. Note that condition (ii) means that it suffices to derive statistical properties in backwards time.
(d) The difficulty with time directions described above is a possibly serious limitation of the martingale approximation approach to proving the ASIP utilised in [16]. The same issue arises in Conze and le Borgne [14]. An alternative approach to proving the ASIP is presented in Hofbauer and Keller [23].

(b) Decay of correlations

Recall that $f : X \to X$ is weak mixing if 1 is a simple eigenvalue for $U : L^2 \to L^2$ and there are no further eigenvalues on the unit circle.

Proposition 3.8 *Assume quasicompactness and that $f : X \to X$ is weak mixing. Then either*

- *(a) F contains constant functions: P has the simple eigenvalue one, and the remainder of the spectrum lies within a circle of radius less than 1; or*
- *(b) F does not contain constant functions: P has spectral radius strictly less than 1 and $\int_X v\,dm = 0$ for all $v \in F$.*

In particular, there are constants $C > 0$ and $\rho \in (0,1)$ such that

$$\|P^n v - \textstyle\int v\| = \|P^n(v - \int v)\| \le C\rho^n \|v\|,$$

for all $v \in F$ and $n \geq 1$.

Proof If $Pv = \alpha v$ where $v \in F \subset L^2$ and $|\alpha| = 1$, then $Uv = \overline{\alpha} v$. By weak mixing, $\alpha = 1$ and v is constant. Hence $P : F \to F$ has at most one eigenvalue on the unit circle. By Theorem 4.6, the remainder of the spectrum lies strictly inside the unit circle.

When F does not contain constant functions, there exist constants $C > 0$ and $\rho \in (0,1)$ such that $\|P^n v\| \leq C\rho^n \|v\|$ for all $n \geq 1$. In particular $|\int v| = |\int P^n v| \leq |P^n v|_1 \leq \|P^n v\| \to 0$.

When F contains constant functions, we have the P-invariant splitting $F = F_2 \oplus \mathbb{R}$ ($v = (v - \int v) + \int v$) where F_2 consists of functions of mean zero. We have shown that $P|_{F_2}$ has spectral radius strictly less than 1, and the result follows. ∎

Corollary 3.9 (Exponential decay of correlations) *Assume quasicompactness and that f is weak mixing. Then*

$$\Big| \int_X v \cdot w \circ f^n \, dm - \int_X v \, dm \int_X w \, dm \Big| \leq C\rho^n \|v\| |w|_2,$$

for all $v \in F$, $w \in L^2$, $n \geq 1$.

Proof

$$\begin{aligned} \big|\textstyle\int v \cdot w \circ f^n - \int v \int w\big| &= \big|\textstyle\int P^n v \cdot w - \int v \int w\big| = \big|\textstyle\int P^n (v - \int v) w\big| \\ &\leq |P^n(v - \textstyle\int v)|_2 |w|_2 \leq \|P^n(v - \textstyle\int v)\| |w|_2 \leq C\rho^n \|v\| |w|_2 \end{aligned}$$

∎

Remark 3.10 It follows from Corollary 3.9 that the variance σ^2 is given by

$$\sigma^2 = \textstyle\int_X v^2 \, dm + 2 \displaystyle\sum_{j=1}^{\infty} \textstyle\int_X v \cdot v \circ f^j \, dm.$$

The proof is standard (see for example [16, Proposition 3.3]).

(c) 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 $F^d = \{v = (v_1, ..., v_d) : X \to \mathbb{R}^d \mid v_j \in F,\ j = 1, \ldots, d\}$.

The operator $Uv = v \circ f$ acts equally well 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 Section 2 apply to functions $v \in F^d$.

In particular, the scalar variance $\sigma^2 = \lim_{N\to\infty} \frac{1}{N} \|v_N\|^2_{H^d}$ is defined for all $v \in F^d_2$. 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, it follows as before that Σ is well-defined for all $v \in F^d_2$. It is easily verified that $\Sigma_{jk} = E(Y_j Y_k)$, 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 $v \in F^d_2$ so that $\int_X v \, dm = 0$. Given $c \in \mathbb{R}^d$, we have the decomposition $c \cdot v = c \cdot w + o(N^{1/2})$ almost everywhere, where $U^*(c \cdot w) = 0$. As in Theorem 3.2, $\frac{1}{\sqrt{N}} c \cdot v_N$ 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. That is

$$\mu\{x : \tfrac{1}{\sqrt{N}} v_N(x) \in I\} \to \frac{1}{(2\pi)^{d/2} (\det \Sigma)^{1/2}} \int_I \exp\{-\tfrac{1}{2} \langle \Sigma^{-1} x, x \rangle\} dx_1 \cdots dx_d \,.$$

The distribution is said to be *nondegenerate* if Σ is nonsingular, (equivalently, $\langle \Sigma x, x \rangle > 0$ for all nonzero $x \in \mathbb{R}^d$).

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 $c \in \mathbb{R}^d$. Actually, we have the following stronger result:

Proposition 3.11 *Assume quasicompactness and that $f : X \to X$ is ergodic. Suppose further that $F \subset L^\infty$. Let $v \in F^d$ with $\int_X v \, dm = 0$. Then there exists a full measure set $\widetilde{X} \subset X$ such that*

$$\limsup_{N\to\infty} \; c \cdot v_N(x) / \sqrt{2N \log\log N} \leq \sigma_c \tag{3.1}$$

for all $x \in \widetilde{X}$, $c \in \mathbb{R}^d$.

Proof For each $c \in \mathbb{R}^d$, we can apply Theorem 3.6. So there exists a subset $E_c \subset X$ of full measure such that equation (3.1) holds for all $x \in E_c$. Let C be a countable dense subset of $\mathbb{R}^d$. Then we can choose a single full measure set $\widetilde{X} \subset X$ such that (3.1) holds for all $x \in \widetilde{X}$ and $c \in C$.

Fix an $x \in \widetilde{X}$ and define $Y_N = v_N(x)/\sqrt{2N \log\log N}$. Then $\limsup_{N\to\infty} c \cdot Y_N \le \sigma_c$ for all $c \in C$. In particular, $c \cdot Y_N$ is bounded and since C spans $\mathbb{R}^d$, Y_N is bounded. Hence there exists a constant M (depending only on x) and an integer N_0 such that $|Y_N| \le M$ for all $N \ge N_0$.

Suppose that $c \in \mathbb{R}^d$ and $\epsilon > 0$. Choose $c' \in C$ such that $|c - c'| < \epsilon/3M$ and $|\sigma_c - \sigma_{c'}| < \epsilon/3$. Increasing N_0 if necessary, we have $c' \cdot Y_N < \sigma_{c'} + \epsilon/3 < \sigma_c + 2\epsilon/3$ for all $N \ge N_0$ and hence

$$c \cdot Y_N = (c - c') \cdot Y_N + c' \cdot Y_N \le \epsilon/3 + \sigma_c + 2\epsilon/3 = \sigma_c + \epsilon.$$

It follows that $\limsup_{N\to\infty} c \cdot Y_N \le \sigma_c$. ∎

Remark 3.12 As discussed in Remark 3.7, it is sometimes possible to prove the full LIL for real-valued observations in F. In such cases, the d-dimensional version in Proposition 3.11 also holds with $\le$ replaced by $=$. To see this, define $\widetilde{X}$, C and Y_N as in the proof of Proposition 3.11. We have $\limsup_{N\to\infty}\ c \cdot Y_N = \sigma_c$ for all $c \in C$, and we obtain $\limsup_{N\to\infty}\ c \cdot Y_N \le \sigma_c$ for all $c \in \mathbb{R}^d$ as before. To obtain the reverse inequality, approximate $c \in \mathbb{R}^d$ by $c' \in C$ as before and observe that since $c' \cdot Y_N > \sigma_{c'} - \epsilon/3 > \sigma_c - 2\epsilon/3$ infinitely often, it follows that $c \cdot Y_N > \sigma_c - \epsilon$ infinitely often.

Finally, if the ergodicity of $f : X \to X$ is strengthened to weak mixing, then we obtain a d-dimensional analogue of exponential decay of correlations [16].

4 Quasicompactness and the Perron-Frobenius operator

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 [25], 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 $1 \le p \le \infty$, 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 $P : L^1 \to L^1$ defined in this way 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 Section 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 Section 7, we consider a number of examples where (F1)–(F5) are satisfied. One such example is provided by Hölder observations on a one-sided subshift of finite type. For details, we refer to the books by Ruelle [34], Bowen [8], or Parry and Pollicott [32]. Let $\sigma : X \to X$ denote an irreducible (not necessarily aperiodic) subshift of finite type. Note that $X \subset \{1, \ldots, k\}^{\mathbb{N}}$ for some k. Fix $\theta \in (0, 1)$ and define the metric $d_\theta(x, y) = \theta^N$ where $N \geq 1$ is least such that $x_i = y_i$ for $i < N$. We define F_θ to be be the space of functions that are Lipschitz continuous with respect to this metric. Let $|g|_\theta$ denote the Lipschitz constant for $g \in F_\theta$ and define the norm $\|g\|_\theta = |f|_\infty + |f|_\theta$. 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 the Arzela-Ascoli theorem. By Remark 4.1, (F4) is automatic with $C_0 = 1$. Finally, the "basic inequality" [32, 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.)

For later reference, we note that F_θ is a Banach algebra. Indeed, if $f, g \in F_\theta$, then $|fg|_\theta \le |f|_\infty \|g\|_\theta + |g|_\infty \|f\|_\theta$. Since $|fg|_\infty \le |f|_\infty |g|_\infty$, it follows that $\|fg\|_\theta \le \|f\|_\theta \|g\|_\theta$.

Example 4.3 Another example is given by piecewise expanding maps on an interval [10, 26]. We defer details of the maps until Section 7 and just discuss (F1)–(F4) here. Recall that a function $g : [0,1] \to \mathbb{R}$ is of bounded variation, $g \in BV$, if

$$\mathrm{var}(g) = \sup_{0 \le t_0 < t_1 < \cdots < t_k \le 1} \sum_{j=1,\ldots,k} |g(t_j) - g(t_{j-1})|$$

is finite. We note that BV is a Banach space with the norm $\|g\|_{BV} = |g|_1 + \mathrm{var}(g)$ and that this norm is equivalent to $|g|_\infty + \mathrm{var}(g)$ (since $|g|_1 \le |g|_\infty \le |g|_1 + \mathrm{var}(g)$).

Again, (F1), (F2) and (F4) are immediate, while (F3) is standard (see for example [23, Lemma 5]).

As in the previous example, we note that BV is a Banach algebra. Indeed, if $f, g \in F_\theta$, then $\mathrm{var}(fg) \le |f|_\infty \mathrm{var}(g) + |g|_\infty \mathrm{var}(f)$. A straightforward calculation shows that $\|fg\|_{BV} \le 2\|f\|_{BV}\|g\|_{BV}$.

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*

$$\|P^n v\| \le E(|v| + \theta^n \|v\|), \tag{4.1}$$

for all $v \in F$, $n \ge 1$.

Proof If (4.1) holds, then choose n_0 large so that $\theta_0 = \theta^{n_0} E < 1$.

Conversely, suppose that (F4) and (F5) hold. An induction argument shows that

$$\|P^{jn_0} v\| \le C_0 D_0 (1 + \theta_0 + \cdots + \theta_0^j)|v| + \theta_0^j \|v\| \le C'|v| + \theta^{jn_0}\|v\|,$$

where $C' = C_0 D_0/(1 - \theta_0)$ and $\theta = \theta_0^{1/n_0}$. Write $n = jn_0 + k$ where $k < n_0$. Then $\|P^n v\| \le C'|P^k v| + \theta^{-1}\theta^n \|P^k v\|$ so the result follows with $E \ge C' C_0$ and $E \ge \theta_0^{-1} \max\{\|P\|, \|P^2\|, \ldots, \|P^{n_0}\|\}$. ∎

Proposition 4.5 *Assume (F4) and (F5). Then P has spectral radius at most 1 in F and F'.*

Proof The result for F' follows from (F4). It follows from (4.1) that $\|P^n\| \le E + 1$ for all sufficiently large n yielding the result for F. ∎

Theorem 4.6 *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 [25].) ∎

Thanks to Theorem 4.6, we can apply the results of Section 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, Corollary 2.3 and Remark 2.4 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 1 is a simple eigenvalue for the operator $P : L^2 \to L^2$ (with constant eigenfunctions) and so F_1 is of dimension 0 or 1. Moreover F_2 consists precisely of those $v \in F$ with $\int_X v\, dm = 0$. Hence the conclusions of Section 3 hold for all observations $v \in F$ with mean zero.

Theorem 4.7 *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 $v_N = o(N^{1/2})$ almost everywhere. If $F \subset L^\infty$, then $\sigma^2 = 0$ if and only if $v_N = O(1)$ uniformly almost everywhere.

(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, as in Corollary 3.9.

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 Section 3(c).

Remark 4.8 It is 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 4.6 can be applied directly on F^d.

The observation that (F1)–(F5) hold for $F^d \subset (F')^d$ is crucial in Section 5.

Transfer operators So far, we have considered the Perron-Frobenius operator P corresponding to an f-invariant measure m. In applications, it is often the case that quasicompactness is proved first for a preliminary Perron-Frobenius operator A corresponding to a "reference measure" ℓ that is not f-invariant. See Section 7 for examples. The operator A is often called a *transfer operator* and we shall use this terminology 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 ($A\varphi = \varphi$). It follows easily that P is related to A by the formula $Pv = \varphi^{-1}A(\varphi v)$.

For many (but not all) applications, quasicompactness for A immediately implies quasicompactness for P. Indeed, it is often the case that multiplication by φ induces linear isomorphisms on F' and F, in which case we have the following useful result:

Lemma 4.9 *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*

(i) $B : F' \to F'$ satisfies (F4) if and only if A does.

(ii) $B : F \to F$ satisfies (F5) if and only if A does.

Proof First, suppose that A satisfies (F4). Since $|A^n| \le C_0$ for all n, we deduce that $|B^n| = |M^{-1}A^nM| \le C_0|M^{-1}||M|$, 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 group extensions $X \times G$ where G is a compact Lie group with Haar measure ν. The aim is to establish statistical properties of a class of *equivariant* vector-valued observations. This class of observations was introduced in [29] and occurs 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 is a twisted version of the usual Ruelle operator and which was studied by [31, 32]. More generally, we consider an equivariant Perron-Frobenius operator which is a twisted version of the usual Perron-Frobenius operator.

Again, our main results generalise results of [16] in the Axiom A setting, with the improvements that we do not require weak mixing (except when proving exponential decay of correlations) 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 Section 4. As pointed out in Remark 4.8, 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. Note that we do not assume that G is connected. Given a measurable map $h : X \to G$, we 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 $M_g v = gv$. Similarly, given $h : X \to G$, we have $(M_h v)(x) = h(x)v(x)$. If $v \in L^p(X, \mathbb{R}^d)$, then $M_h v \in L^p(X, \mathbb{R}^d)$ and $|M_h v|_p = |v|_p$. We require control over the norms $|\;|$ and $\|\;\|$ on F' and F under the action of M_{h_n} where

$$h_n(x) = h(x)h(fx) \cdots h(f^{n-1}x).$$

In particular, 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 $v \in F$.

Since G is compact, automatically $h \in L^\infty(X, G)$ with $|h|_\infty = 1$.

Remark 5.1 Suppose that $f : X \to X$ and $h : X \to G$ are measurable. If $F' = L^p(X, \mathbb{R}^d)$ for some $1 \le p \le \infty$, 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 continuous with respect to the metric d_θ on X. We denote the space of such cocycles by $F_\theta(X, G)$. Then $|M_* v|_\theta \le |M_*|_\theta |v|_\infty + |M_*|_\infty |v|_\theta = |M_*|_\theta |v|_\infty + |v|_\theta$, and so $\|M_* v\|_\theta = |M_* v|_\infty + |M_* v|_\theta \le \|h_{n_0}^{-1}\|_\theta |v|_\infty + \|v\|_\theta$. 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 $F = BV$ consists of functions of bounded variation (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 $\|M_{h_n}^{-1}\| = \|h_n^{-1}\|_{BV} = 1 + \mathrm{var}(h_n^{-1})$.

Nevertheless, 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.)

Equivariant Perron-Frobenius operator

For each $h : X \to G$, 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 $P_h^j = P^j M_{h_j}^{-1}$ where $h_j = h \cdot h \circ f \cdots h \circ f^{j-1}$.

Proposition 5.4 *Assume (F1), (F4), (F5) and suppose that $h : X \to G$ satisfies (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 $\theta_0' = c\theta_0$. ∎

Equivariant observations

Let L^p_G consist of all functions $\phi : X \times G \to \mathbb{R}^d$ 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 $L^p_G = g \cdot L^p(X, \mathbb{R}^d)$.

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 $\widehat{P}|_{L^1_G} = M_g P_h M_g^{-1}$.*

Proof Let $\phi = g \cdot v \in L^1_G$ and $\psi = g \cdot w \in L^\infty_G$. 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 $F_G = g \cdot F$, $F_G' = g \cdot F'$, etc with norm $\|g \cdot v\| = \|v\|$, $|g \cdot v| = |v|$, etc.

Proposition 5.6 *The operator $\widehat{P}$ restricts to an operator on F_G' and F_G and on these spaces $\widehat{P} = M_g P_h M_g^{-1}$. Moreover, properties (F1)–(F5) are inherited by $\widehat{P}$ on these spaces.*

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

We can now apply Theorem 4.6 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 assume (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 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 $\{\phi_N\}$ as in Proposition 3.11.

(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 Section 3. The statement about Σ commuting with G and part (c) are proved as in [29] (see also [16]). We note that $\Sigma : \mathbb{R}^d \to \mathbb{R}^d$ commutes with the action of G on $\mathbb{R}^d$ ∎

6 A refinement of the CLT, WIP and LIL

In Section 5, we proved equivariant statistical limit theorems for equivariant observations $\phi : X \times G$. These results are phrased in terms of the measure $m = \mu \times \nu$ on $X \times G$. It is natural to ask whether such limit theorems remain valid when restricted to the measure μ on $X \times \{g_0\}$ for some fixed $g_0 \in G$. That is, does a statistical law for the sequence $\{\phi_N(x, g)\}$ defined on $X \times G$. imply the corresponding statement for $\{\phi_N(x, g_0)\}$ defined on $X \times \{g_0\}$. It turns out that the answer is affirmative. This is easily seen for the LIL, but is less easily proved for the CLT. (We note that the reverse implication is immediate by equivariance. The implication we seek is not immediate and clearly need not be true if $T : X \times G \to X \times G$ is not ergodic.)

This section, which is independent from the rest of the paper, contains the precise statement and proof of these statements.

(a) Central limit theorem on $X \times \{g_0\}$

Theorem 6.1 *Suppose that the G-extension $T : X \times G \to X \times G$ is ergodic and that $\phi \in L^2(X \times G, \mathbb{R}^d)$ is a G-equivariant observation.*

If the sequence $\{\frac{1}{\sqrt{N}}\phi_N(x,g)\}$ defined on the probability space $(X \times G, m)$ converges in distribution to a d-dimensional normal distribution with mean zero and covariance matrix Σ, then for any $g_0 \in G$, the sequence $\{\frac{1}{\sqrt{N}}\phi_N(x,g_0)\}$ defined on the probability space (X,μ) converges in distribution to the same d-dimensional normal distribution with mean zero and covariance matrix Σ.

Proof It simplifies our argument considerably to assume from the start that f and hence T are invertible measure-preserving transformations. This assumption is with no loss of generality as we could always form the natural extension of f. Obviously, it suffices to prove the result for $g_0 = e$.

Recall that $T(x,g) = (fx, gh(x))$ where $h : X \to G$ is measurable. Throughout, we write $Z_N = \frac{1}{\sqrt{N}}\phi_N$. Also, we write $\mu(Z_N \in E)$ instead of $\mu\{x \in X : Z_N(x,e) \in E\}$ and we write $m(Z_N \in E)$ instead of $m\{(x,g) \in X \times G : Z_N(x,g) \in E\}$.
Idea of proof

Let $C_\rho \subset \mathbb{R}^d$ be a ball of radius ρ.

Our proof is based upon the observation that for any set $S \subset G$, for μ a.e $y \in X$ the sequence $\{h_j(f^{-j}y)\}$ is 'well-distributed' in G in the sense that

$$\lim_{J\to\infty} \frac{1}{J}\sum_{j=0}^{J_1} I_S(h_j(f^{-j}y) = \nu(S).$$

This is a consequence of the ergodicity of T. If we take a cover $\{S_i\}_{i=1}^M$ of G by sets of small diameter and equal Haar measure we may choose J large enough that for most points in X, for $1 \le u, v \le M$

$$\frac{1}{J}\sum_{j=0}^{J-1} I_{S_u}(h_j(f^{-j}y) \simeq \frac{1}{J}\sum_{j=0}^{J-1} I_{S_v}(h_j(f^{-j}y)$$

where $h_j(x) = h(x)h(fx)\cdots h(f^{j-1}x)$. Define $X_j = T^j(X \times \{e\})$. Then the set $\cup_{j=0}^{J-1} T^j(X \times \{e\}) = \cup_{j=0}^{J-1} X_j$ is 'well-distributed' in $X \times G$.

Note that for large N, $Z_N \simeq Z_N \circ T^j$ for $0 \le j \le J-1$ (uniformly in (x,e)). Hence, if we define

$$\sigma(Z_N \in E) = \frac{1}{J}\sum_{j=0}^{J-1} \mu(Z_N \circ T^j \in E)$$

then $\sigma(Z_N \in E) \simeq \mu(Z_N \in E)$.

Since points in $\cup_{j=0}^{J-1} X_j$ are well-distributed, and the measure σ averages over this set we have $\sigma(Z_N \in S_i C_\rho) \simeq \sigma(Z_N \in S_j C_\rho)$. But since $\sigma(Z_N \in E) \simeq \mu(Z_N \in E)$ we also have $\mu(Z_N \in S_i C_\rho) \simeq \mu(Z_N \in S_j C_\rho)$.

Step 1: Main steps of proof.

Let C_ρ be a closed ball of radius $\rho > 0$. Let $\epsilon > 0$.

Choose ρ sufficiently small and $\delta < \rho/4$, so that

$$0 \leq \int_{C_{\rho+4\delta}} \mathcal{N} dx - \int_{C_{\rho-4\delta}} \mathcal{N} dx < \epsilon$$

where $\mathcal{N}$ is the Gaussian kernel of the standard d-dimensional normal distribution. Since Z_N converges in distribution to $\mathcal{N}$ (on $X \times G$), we may choose $N(\delta)$ sufficiently large that for all $N \geq N(\delta)$

$$0 \leq m(Z_N \in C_{\rho+4\delta}) - m(Z_N \in C_{\rho-4\delta}) < 2\epsilon. \tag{6.1}$$

If $B \subset G$ define $B(\alpha) = \{g \in G : d(g, B) \leq \alpha\}$. We claim that for all sufficiently small $r > 0$ there exists α independent of r and a collection of compact sets $\{B_i\}_{i=1}^M$ in G with $\mathrm{diam}(B_i) < r$, such that for all $1 \leq i, j, k, l \leq M$,

(a) $\nu(B_i) = \frac{1}{M}$,

(b) $e \in B_1$ and $\nu(B_i \cap B_j) = 0$ unless $i = j$.

(c1) If $\nu(B_i \cap B_j B_k B_l) \neq 0$ then $B_i \subset B_j(\alpha r) B_k B_l$.

(c2) If $\nu(B_i \cap B_j B_k^{-1} B_l) \neq 0$ then $B_j B_k^{-1} B_l \subset B_i(\alpha r)$.

(d1) $B_i(\alpha r) C_{\rho - \frac{\delta}{2}} \subset B_i C_\rho$.

(d2) $B_i(\alpha r) C_\rho \subset B_i(r) C_{\rho + \frac{3\delta}{2}}$.

In other words $\{B_i\}_{i=1}^M$ form a measurable partition of G satisfying the additional properties $(c1), (c2), (d1), (d2)$.

Furthermore, we claim that this collection of sets $\{B_i\}$ has the property that

$$m(Z_N \in C_{\rho+4\delta}) \geq \sum_{i=1}^{M} \nu(B_i) \mu(Z_N \in B_i C_{\rho+3\delta}) \tag{6.2}$$

$$\frac{1}{M} \sum_{i=1}^{M} \mu(Z_N \in B_i C_{\rho-3\delta}) \geq m(Z_N \in C_{\rho-4\delta}) \tag{6.3}$$

and that there exists $N(\epsilon)$ such that

$$\mu(Z_N \in B_i C_{\rho-3\delta}) \leq \mu(Z_N \in B_1 C_\rho) + 3\epsilon \tag{6.4}$$
$$\mu(Z_N \in B_1 C_\rho) \leq \mu(Z_N \in B_i C_{\rho+3\delta}) + 3\epsilon \tag{6.5}$$

for all $N \geq N(\epsilon)$ and $1 \leq i \leq M$.

Suppose that the claim is valid. Then by (6.4) and (6.5),

$$\frac{1}{M}\sum_i \mu(Z_N \in B_i C_{\rho-3\delta}) \leq \mu(Z_N \in B_1 C_\rho) + 3\epsilon \tag{6.6}$$
$$\mu(Z_N \in B_1 C_\rho) \leq \frac{1}{M}\sum_i \mu(Z_N \in B_i C_{\rho+3\delta}) + 3\epsilon. \tag{6.7}$$

Hence,

$$\begin{aligned}
m(Z_N \in C_{\rho+4\delta}) &\geq \frac{1}{M}\sum_{i=1}^{M} \mu(Z_N \in B_i C_{\rho+3\delta}) \text{ from } (6.2)\\
&\geq \mu(Z_N \in B_1 C_\rho) - 3\epsilon \text{ from } (6.7)\\
&\geq \frac{1}{M}\sum_{i=1}^{M} \mu(Z_N \in B_i C_{\rho-3\delta}) - 6\epsilon \text{ from } (6.6)\\
&\geq m(Z_N \in C_{\rho-4\delta}) - 6\epsilon \text{ from } (6.3).
\end{aligned}$$

In particular,

$$m(Z_N \in C_{\rho+4\delta}) + 3\epsilon \geq \mu(Z_N \in B_1 C_\rho) \geq m(Z_N \in C_{\rho-4\delta}) - 3\epsilon.$$

Combining this with (6.1), since ϵ and δ are arbitrary,

$$\lim_{N\to\infty} \mu(Z_N \in C_\rho) = \lim_{N\to\infty} m(Z_N \in C_\rho)$$

proving the lemma.

The remainder of the proof is structured as follows. In Step 2, we prove (6.2) and (6.3). The sets $\{B_i\}$ are constructed in Step 3. Steps 4–6 contain the proof of (6.4) and (6.5).

Step 2: Proof of (6.2) and (6.3).

For sufficiently small $r > 0$ if $B \subset G$ is a set of diameter less than r then

$$gC_{\rho+4\delta} \supset BC_{\rho+3\delta} \supset BC_{\rho-3\delta} \supset gC_{\rho-4\delta} \tag{6.8}$$

for all $g \in B$.

If $A \subset \mathbb{R}^d$, then $m(Z_N \in A) = \int_G \mu(Z_N \in gA) d\nu(g) = \sum_{i=1}^M \int_{B_i} \mu(Z_N \in gA) d\nu(g)$. It follows from the inequalities (6.8) that

$$\begin{aligned} m(Z_N \in C_{\rho+4\delta}) &= \sum_{i=1}^M \int_{B_i} \mu(Z_N \in gC_{\rho+4\delta}) d\nu(g) \geq \sum_{i=1}^M \nu(B_i)\mu(Z_N \in B_i C_{\rho+3\delta}) \\ &\geq \sum_{i=1}^M \nu(B_i)\mu(Z_N \in B_i C_{\rho-3\delta}) \geq \sum_{i=1}^M \int_{B_i} \mu(Z_N \in gC_{\rho-4\delta}) d\nu(g) \\ &= m(Z_N \in C_{\rho-4\delta}). \end{aligned}$$

This proves (6.2) and (6.3).
Step 3: Construction of the sets B_i.

Lemma 6.2 *For all sufficiently small $r > 0$ there exists α, independent of r and a collection of compact sets $\{B_i\}_{i=1}^M$, $B_i \subset G$ of equal measure, $diam(B_i) < r < \epsilon$ and satisfying $(a), (b), (c1)(c2), (d1), (d2)$ of Step 1.*

Since G is a compact Riemannian manifold G has a triangulation into sets $\{A_i\}_{i=1}^m$ with $\mathrm{diam}(A_i) < r$ for all i.

By perturbing the boundaries of the triangulation we may assume that the numbers $\{\nu(A_i)\}$ are rational, and hence have a greatest common denominator q. Express the numbers $\nu(A_i)$ with respect to the common denominator so that $\nu(A_i) = \frac{p_i}{q}$. Then subdivide each set A_i into p_i compact subsets $A_{i,j}, j = 1, ..., p(i)$ of equal measure. Finally relabel the sets $\{A_{i,j}\}$ as sets $B_1, \ldots, B_M$, ensuring (by relabelling again if necessary) that $e \in B_1$. Given a set B_i we let $A_{r(i)}$ denote the unique triangle which contains B_i. Since G is compact and acts smoothly on itself there exists a α' such that for all $r > 0$ if $\mathrm{diam}(A_i) < r$, $i = 1, \ldots, M$ then $\mathrm{diam}(A_j A_k^{-1} A_l) < \alpha' r$. Suppose that $\nu(B_i \cap B_j B_k^{-1} B_l) \neq 0$. Then $B_i \cap B_j B_k^{-1} B_l \neq \emptyset$. Since $B_j B_k^{-1} B_l \subset A_{r(j)} A_{r(k)}^{-1} A_{r(l)}$ and $\mathrm{diam}(A_{r(j)} A_{r(k)}^{-1} A_{r(l)}) \leq \alpha' r$ we have $A_{r(j)} A_{r(k)}^{-1} A_{r(l)} \subset B_i(2\alpha' r)$. Thus $B_j B_k^{-1} B_l \subset B_i(2\alpha' r)$. We now define $\alpha = 2\alpha'$. This proves (c1) and (c2) is proved similarly.

For fixed ρ, δ we may choose $0 < r < \epsilon$ sufficiently small such that if $B \subset G$ is a set of diameter less than r then $B_i(\alpha r) C_{\rho - \frac{\delta}{2}} \subset B_i C_\rho$ and $B_i(\alpha r) C_\rho \subset B_i C_{\rho + \frac{3\delta}{2}}$. This proves (d1) and (d2).
Step 4: Ergodicity.
Since T^{-1} is ergodic there exists an integer J and group element $g_0 \in G$ such that for all $i = 1, \ldots, M$

$$\mu\Big\{x \in X : \Big|\frac{1}{J}\sum_{j=1}^J I_{X \times g_0 B_i^{-1}} \circ T^{-j}(x, e) - \frac{1}{M}\Big| \leq \epsilon\Big\} \geq 1 - \epsilon.$$

To see this, note that by Fubini's theorem there exists a 'good' g_0, such that for all $i = 1, \ldots, M$, for μ a.e. $x \in X$

$$\lim_{J\to\infty} \frac{1}{J} \sum_{j=1}^{J} I_{X\times B_i^{-1}} \circ T^{-j}(x, g_0^{-1}) = \frac{1}{M}.$$

(In fact, this holds for ν a.e. $g_0 \in G$.)

It is no loss of generality and also notationally simpler to redefine B_i as $B_i g_0^{-1}$ and after relabelling assume $e \in B_1$. Thus we assume that there exists an integer J such that for all $i = 1, \ldots, M$,

$$\mu\Big\{x \in X : \Big|\frac{1}{J} \sum_{j=1}^{J} I_{B_i^{-1}} \circ T^{-j}(x, e) - \frac{1}{M}\Big| \le \epsilon\Big\} \ge 1 - \epsilon. \tag{6.9}$$

Recall that $T^j(x, e) = (f^j x, h_j(x))$ where $h_j(x) = h(x)h(fx)\cdots h(f^{j-1}x)$. Similarly, $T^{-j}(x, e) = (f^{-j}x, [h_j(f^{-j}x)]^{-1})$. By (6.9), it follows that

$$\mu\Big\{x \in X : \Big|\frac{1}{J} \sum_{j=1}^{J} I_{B_i}(h_j(f^{-j}x)) - \frac{1}{M}\Big| \le \epsilon\Big\} \ge 1 - \epsilon.$$

Hence, there exists a subset $\widetilde{X} \subset X$ with $\mu(\widetilde{X}) > 1 - \epsilon$ such that

$$\Big|\frac{1}{J} \sum_{j=1}^{J} I_{B_i} \circ h_j(f^{-j}x) - \frac{1}{M}\Big| \le \epsilon$$

for all $x \in \widetilde{X}$ and $i = 1, \ldots, M$.

Step 5: A new measure on $X \times G$.

Let J be the integer fixed in Step 4. Define the measure σ on $X \times G$ by

$$\sigma(E) = \frac{1}{J} \sum_{i=0}^{J-1} \mu\{x : T^j(x, e) \in E\}.$$

Again, if $C \subset \mathbb{R}^d$, we write $\sigma(Z_N \in C)$ instead of $\sigma\{(x, g) \in X \times G : Z_N(x, g) \in C\}$.

For each $0 \le j \le J$, $\big|Z_N \circ T^j(x, e) - Z_N(x, e)\big| \le 2J|\phi|_\infty/\sqrt{N}$. If necessary, increase $N(\delta)$ so that for all $N \ge N(\delta)$, for all $0 \le j \le J$,$1 \le i \le M$,

$$\big|Z_N \circ T^j(x, e) - Z_N(x, e)\big| < \delta.$$

Then

$$\{Z_N \in B_i C_{\rho-3\delta}\} \subset \{Z_N \circ T^j \in B_i C_{\rho-2\delta}\} \subset \{Z_N \circ T^j \in B_i C_{\rho-\delta}\} \subset \{Z_N \in B_i C_\rho\},$$

so by definition of σ,

$$\mu(Z_N \in B_i C_{\rho-3\delta}) \leq \sigma(Z_N \in B_i C_{\rho-2\delta}) \leq \sigma(Z_N \in B_i C_{\rho-\delta}) \leq \mu(Z_N \in B_i C_\rho). \quad (6.10)$$

We claim that for all $N > N(\delta)$ and $1 \leq i \leq M$,

$$\sigma(Z_N \in B_i C_{\rho-2\delta}) \leq \sigma(Z_N \in B_1 C_{\rho-\delta}) + 3\epsilon. \quad (6.11)$$

By (6.10) and (6.11),

$$\begin{aligned}\mu(Z_N \in B_i C_{\rho-3\delta}) &\leq \sigma(Z_N \in B_i C_{\rho-2\delta}) \leq \sigma(Z_N \in B_1 C_{\rho-\delta}) + 3\epsilon \\ &\leq \mu(Z_N \in B_1 C_\rho) + 3\epsilon\end{aligned}$$

for all i proving (6.4). The proof of (6.5) is similar, and we omit the details. It now remains to prove (6.11).

Step 6: Proof of (6.11)**.**

Since f is invertible and measure preserving,

$$\mu\{x \in X : T^j(x,e) \in E\} = \mu\{y \in X : (y, h_j(f^{-j}y)) \in E\}.$$

Hence,

$$\sigma(E) = \frac{1}{J}\sum_{j=0}^{J-1} \mu\{y \in X : (y, h_j(f^{-j}y)) \in E\}.$$

The group elements $h_j(f^{-j}y)$ induce a coding of the point y in the symbol space $\{1,\ldots,M\}^J$ by $w(y) = (w_0(y), w_1(y), \ldots, w_J(y))$ where $w_j(y) = s \in \{1,\ldots,M\}$ if $h_j(f^{-j}y) \in B_s$. We define cylinder sets $\{S_i\}$, $i = 1,\ldots,r$, say, where each cylinder set consists of points in $\widetilde{X}$ with the same coding i.e the cylinder sets are elements of the partition of $\widetilde{X}$ induced by the equivalence relation $y \simeq z$ if and only if $w(y) = w(z)$. Each cylinder set is defined by a word in $\{1,\ldots,M\}^J$ and $\widetilde{X}$ is partitioned by $\{S_i\}_{i=1}^r$. Note that to each letter $w_j(y)$ of the word $w(y)$ there corresponds the point $(y, h_j(f^{-j}y)) \in X_j$ where $h_j(f^{-j}y) \in B_s$ if and only if $w_j(y) = s$. Furthermore by assumption, for all $p, q \in \{1,\ldots,M\}, y \in \widetilde{X}$

$$\frac{1}{J}|\#\{j : w_j(y) = p\} - \#\{j : w_j(y) = q\}| \leq \epsilon.$$

Our key observation is that by G-equivariance for any set U, $Z_N(y, h_j(f^{-j}y)) \in U$ implies $Z_N(y, h_k(f^{-k}y)) \in h_k(f^{-k}y)h_j^{-1}(f^{-j}y)U$.

Lemma 6.3 *Fix B_p. We claim that there is a permutation $\{t(1),\ldots,t(M)\}$ of $\{1,\ldots,M\}$ such that for each $s \in \{1,\ldots,M\}$, $B_t(s)B_s^{-1}B_p \subset B_1(\alpha r)$. In other words $t : \{1,\ldots,M\} \to \{1,\ldots,M\}$ is a bijection and $B_t(s)B_s^{-1}B_p \subset B_1(\alpha r)$.*

Proof To see we may construct such a map, let $g_p \in B_p$ and consider the disjoint partition of G given by the sets $\{g_p^{-1}B_1, g_p^{-1}B_2, \ldots, g_p^{-1}B_M\} = \{C_1, \ldots, C_M\}$. We say a collection of sets $\{B_{i_1}, \ldots, B_{i_r}\}$ is path connected if any sets B_{i_u}, B_{i_v} may be connected by a path contained in $\cup_{j=1}^r B_{i_j}$.

Our proof is by induction and starts at $k = 2$. Choose B_{i_1} and B_{i_2} which have a common boundary and hence are path connected. Now $\nu(B_{i_1} \cap C_j) \neq 0$ for some j. Choose such a j and define $t(i_1) = j$. Similarly $\nu(B_{i_2} \cap C_r) \neq 0$ for some r. We may choose $r \neq t(i_1)$ since otherwise $B_{i_2} \subset B_{t(i_1)}$ and because $\nu(B_{i_2}) = \nu(B_{t(i_1)})$ this implies $\nu(B_{t(i_1)}) \cap B_{i_1}) = \nu(B_{i_2} \cap B_{i_1}) = 0$, a contradiction.

Suppose we have chosen a path connected set $\{B_{i_1}, \ldots, B_{i_k}\}$ with corresponding permutation t defining a set $\{C_{t(i_1)}, \ldots, C_{t(i_k)}\}$, such that $\nu(B_{i_j} \cap C_{t(i_j)}) \neq 0$. We say B_{i_j} is labelled by $C_{t(i_j)}$. Now we choose $B_{i_{k+1}}$, path connected to $\{B_{i_1}, \ldots, B_{i_k}\}$. If $B_{i_{k+1}} \not\subset \cup_{j=1}^k C_{t(i_j)}$ then there exists C_s such that $\nu(C_s \cap B_{i_{k+1}}) \neq 0$ and we may choose $t(i_{k+1}) = s$ to extend the set to $k+1$ members. A problem arises only if $B_{i_{k+1}} \subset \cup_{j=1}^k C_{t(i_j)}$.

So suppose $B_{i_{k+1}} \subset \cup_{j=1}^k C_{t(i_j)}$. We define $S_k = \cup_{j=1}^k C_{t(i_j)}$. Although we will redefine the map t later we keep S_k fixed.

Choose $C_{t(i_{s_1})}$ such that $C_{t(i_{s_1})} \cap B_{i_{k+1}} \neq 0$ and label $B_{i_{k+1}}$ with this set by defining $t(i_{k+1}) = t(i_{s_1})$ and keeping $t(i_j)$ fixed for $j \notin \{k+1, s_1\}$. Now $B_{i_{s_1}}$ is unlabelled. If $B_{i_{s_1}} \not\subset S_k$ then there exists $C_t \subset S_k^c$ such that $\nu(C_t \cap B_{i_s}) \neq 0$ and we may define $t(i_s) = t$ to extend the number of labelled sets to $k+1$. If however $B_{i_{s_1}} \subset S_k$ we choose $C_{t(i_{s_2})}$ such that $C_{t(i_{s_2}} \cap B_{i_{s_1}} \neq 0$ and relabel $t(i_{s_1}) = t(i_{s_2})$. We continue this process of relabelling along all possible paths, hoping to find a set $B_{i_{s_r}} \not\subset S_k$, in which case we may extend the number of labelled sets to $k+1$. A possible problem is that the set of all such paths formed in this way starting at $B_{i_{k+1}}$ does not include a set $B_{i_{s_r}} \not\subset S_k$. To see that this is not the case suppose we encountered only sets $\{B_{i_{s_1}}, \ldots, B_{i_{s_t}}\}$ covered wholly by S_k. Then in fact

$$B_{i_{k+1}} \cup \cup_{j=1}^t B_{i_{s_j}} \subset \cup_{j=1}^k C_{t(i_{s_j})}$$

since otherwise we could form a path including another set $B_{i_j} \notin \{B_{i_{s_1}}, \ldots, B_{i_{s_t}}\}$. However $B_{i_{k+1}} \cup \cup_{j=1}^t B_{i_{s_j}} \subset \cup_{j=1}^k C_{t(i_{s_j})}$ leads to a contradiction as $B_{i_{k+1}} \cup \cup_{j=1}^t B_{i_{s_j}}$ has strictly greater measure than $\cup_{j=1}^k C_{t(i_{s_j})}$. Thus we may find a set $B_{i_{s_r}} \not\subset S_k$ and extend the number of labelled sets to $k+1$. This concludes the proof of the lemma. ∎

Thus we have inductively defined the permutation function t and defined a one-to-one function such that $\nu(g_p^{-1}B_s \cap B_{t(s)}) \neq 0$ and $e \in B_{t(s)}B_s^{-1}B_p$ for all s.

Recall that the partition $\{B_1, \ldots, B_M\}$ has the property that if $\nu(B_j \cap B_s B_t^{-1} B_u) \neq 0$ then $B_s B_t^{-1} B_u \subset B_j(\alpha r)$. Hence $B_{t(s)} B_s^{-1} B_p \subset B_1(\alpha r)$.

To sum up: for all j such that $w_j = t(s)$, for all $y \in S_i$ we have $h_j(f^{-j}y) \in B_{t(s)}$. Thus if $Z_N(y, h_j(f^{-j}y)) \in B_p C_{\rho - 2\delta}$ and $w_j = s$ then if $w_k = t(s)$ it follows that

$Z_N(y, h_k(f^{-k}y)) \in B_1(\alpha r)C_{\rho-2\delta} \subset B_1 C_{\rho-\delta}$. Hence

$$\begin{aligned}
&\frac{1}{J}\sum_{j=0}^{J-1} \mu\{y \in S_i : Z_N(y, h_j(f^{-j}y)) \in B_p C_{\rho-2\delta}\} \\
&\quad = \frac{1}{J}\sum_{s=1}^{M}\sum_{w_j=s} \mu\{y \in S_i : Z_N(y, h_j(f^{-j}y)) \in B_p C_{\rho-2\delta}\} \\
&\quad \le \frac{1}{J}\sum_{s=1}^{M}\sum_{w_j=t(s)} \mu\{y \in S_i : Z_N(y, h_j(f^{-j}y)) \in B_1(\alpha r) C_{\rho-2\delta}\} + \epsilon\mu(S_i) \\
&\quad = \frac{1}{J}\sum_{j=0}^{J-1} \mu\{y \in S_i : Z_N(y, h_j(f^{-j}y)) \in B_1(\alpha r) C_{\rho-2\delta}\} + \epsilon\mu(S_i) \\
&\quad \le \frac{1}{J}\sum_{j=1}^{J} \mu\{y \in S_i : Z_N(y, h_j(f^{-j}y)) \in B_1 C_{\rho-\delta}\} + \epsilon\mu(S_i).
\end{aligned}$$

This holds for all cylinder sets S_i so by summing over cylinder sets in $\widetilde{X}$ we have

$$\begin{aligned}
&\frac{1}{J}\sum_{j=0}^{J-1} \mu\{y \in \widetilde{X} : Z_N(y, h_j(f^{-j}y)) \in B_p C_{\rho-2\delta}\} \\
&\qquad \le \frac{1}{J}\sum_{j=0}^{J-1} \mu\{y \in \widetilde{X} : Z_N(y, h_j(f^{-j}y)) \in B_1 C_{\rho-\delta}\} + \epsilon.
\end{aligned}$$

Finally including the error from the set $\widetilde{X}$ we have

$$\sigma(Z_N \in B_p C_{\rho-2\delta}) \le \sigma(Z_N \in B_1 C_{\rho-\delta}) + 3\epsilon.$$

Since B_p is arbitrary, we have proved (6.11) as required. ∎

(b) Weak invariance principle on $X \times \{g_0\}$

Theorem 6.4 *Suppose that the G-extension $T : X \times G \to X \times G$ is ergodic and that $\phi \in L^2(X \times G, \mathbb{R}^d)$ is a G-equivariant observation.*

If the sequence $\{\frac{1}{\sqrt{N}}\phi_N(x, g)\}$ defined on the probability space $(X \times G, m)$ satisfies the WIP with covariance matrix Σ, then for all $g_0 \in G$ the sequence $\{\frac{1}{\sqrt{N}}\phi_N(x, g_0)\}$ defined on the probability space (X, μ) also satisfies the WIP with covariance matrix Σ.

Proof Weak convergence is equivalent [6] to weak convergence of the finite dimensional distributions coupled with a tightness condition. The proof of Theorem 6.1 combined with the Cramer-Wold technique (cf. [16]) establishes the first of these facts with no essential modification. By considering G-invariant subsets of $\mathbb{R}^d$, it is immediate that tightness for m is equivalent to tightness for μ. ∎

(c) Law of the iterated logarithm on $X \times \{g_0\}$

Theorem 6.5 *Suppose that the G-extension $T : X \times G \to X \times G$ is ergodic and that $\phi \in L^2(X \times G, \mathbb{R}^d)$ is a G-equivariant observation.*

If $c \cdot \phi : X \times G \to \mathbb{R}$ satisfies the LIL for all $c \in \mathbb{R}^d$, then there is a full measure set $\widetilde{X} \subset X$ such that for all $g_0 \in G$

$$\limsup_{N\to\infty} c \cdot \phi_N(x, g_0)/\sqrt{2N \log\log N} = \sigma_c \tag{6.12}$$

for all $x \in \widetilde{X}$ and $c \in \mathbb{R}^d$ (where $\sigma_c^2 = c^T \Sigma c$).

Proof By Proposition 3.11 and Remark 3.12, there is a full measure set $\widetilde{Y} \subset X \times G$ such that (6.12) holds for all $(x, g) \in \widetilde{Y}$ and $c \in \mathbb{R}^d$. By Fubini's theorem, there exists $\widetilde{X} \subset X$ of full measure and $g_0 \in G$ such that (6.12) holds for all $x \in \widetilde{X}$. Now, use the equivariance of ϕ_N to deduce that for each $g \in G$,

$$c \cdot \phi_N(x, g) = c \cdot g g_0^{-1} \phi_N(x, g_0) = c' \phi_N(x, g_0),$$

where $c' = g_0 g^{-1} c$. Hence $\limsup_{N\to\infty} c \cdot \phi_N(x, g)/\sqrt{2N \log\log N} = \sigma_{c'}$. But $\sigma_{c'} = (c')^T \Sigma c' = c^T \Sigma c = \sigma_c$ since Σ commutes with the action of G. ∎

Remark 6.6 Theorem 6.5 remains valid for the upper LIL (with $=$ replaced by $\leq$ in (6.12)).

7 Applications

In this section we give examples of dynamical systems to which our results apply. Some preliminary comments are required to make the connection with the approach in this paper. (Certain issues such as regularity are made precise in the specific examples.)

The starting point is that X is a topological space and $f : X \to X$ is 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 it is easily verified that $\int Av\, w\, d\ell = \int v\, w \circ f\, d\ell$. Thus A is the Perron-Frobenius operator corresponding to the measure ℓ. Note that we have not assumed that ℓ is f-invariant. (Indeed ℓ is f-invariant if and only if $A1 = 1$ which need not be the case.)

Now suppose that $\varphi > 0$ is a fixed point ($A\varphi = \varphi$) for A acting on a suitable function space. Then we define the measure $dm = \varphi\, d\ell$ and $Pv = \varphi^{-1}A(\varphi v)$. It is easily verified that $\int Pv\, w\, dm = \int v\, w \circ f\, dm$ for $v \in L^1(X, m)$, $w \in L^\infty(X, m)$ so P is the Perron-Frobenius operator corresponding to the measure m. Moreover, $P1 = 1$ and so m is an f-invariant measure with Perron-Frobenius operator P.

This construction leads to an f-invariant measure m (equivalent to the measure ℓ) which turns out to have desirable properties (often it is an equilibrium measure for a potential function related to the potential $\log g$).

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. In this paper, we are not concerned with the construction of m, but we want to refine existing methods based on quasicompactness to (i) obtain improved results on the statistical properties related to m and (ii) to generalise these results to compact group extensions.

(a) One dimensional monotone maps

Hofbauer and Keller [23] have analysed a class of endomorphisms of a totally ordered, order complete set X (usually X is taken to be the unit interval $[0, 1]$). They consider a transformation $f : X \to X$ which is piecewise monotonic and order-continuous and define a transfer operator A on the space of all bounded measurable functions v 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 $v : X \to \mathbb{R}$. It is not assumed that ℓ is f-invariant. 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 [23] 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 [26, 27, 39, 37, 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 [38] constructed equilibrium measures corresponding to Lipschitz continuous potentials $\phi : [0, 1] \to \mathbb{R}$. The class of allowable potentials is extended in [23], 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 $dm = \varphi\, d\ell$ corresponding to the 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 [24] 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 7.1 *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 Hofbauer and Keller [23] verify hypotheses (F1)–(F5) for the transfer operator A acting on $F = BV$ with $F' = L^1(X, m)$. Conditions (F1)–(F3) were discussed in Example 4.3. To verify (F4) and (F5) we use the assumptions that $\varphi, \varphi^{-1} \in F$ from which it easily 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 we can apply Lemma 4.9.

Condition (F4) is immediate for P by Remark 4.1 and hence holds also for A. (Alternatively, see [23, Lemma 6].) The crucial condition (F5) is proved for A in [23, Lemma 7], and holds also for P by Lemma 4.9. ∎

We are now in a position to apply the results in Sections 2, 3 and 4. Assume that the equilibrium measure m in Lemma 7.1 is ergodic. Let $v \in BV$ with $\int_X v\, dm = 0$, and define $v_N = \sum_{j=1}^N v \circ f^j$.

• 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 $f^j x$ for $j = 1, \dots, k$, and $v_k(x) \neq 0$, 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 7.2 Under the assumption of weak mixing, the statistical results above were obtained in [23]. Moreover, [23] proved exponential decay of correlations and the almost sure invariance principle – which implies the CLT and WIP, and the full LIL. In this paper, we have extended the last three results to the case where m is ergodic but not necessarily weak mixing. Also, following [16], we have obtained strong conditions for nondegeneracy that are not present in [23]. [1]

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. As discussed above, the Perron-Frobenius operator P satisfies hypotheses (F1)–(F5) with respect to the Banach spaces $F = BV$ and $F' = L^1$.

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 the product measure $m \times \nu$ where ν is Haar measure on G.

As in Section 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 almost everywhere uniformly bounded.
• The sequence of partial sums ϕ_N satisfies the d-dimensional CLT and WIP on $X \times G$ with covariance matrix Σ.

[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 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 components of ϕ satisfy the upper LIL.
• If in addition $m \times \nu$ is weak mixing, then we obtain exponential decay of correlations (against test functions in L^2_G).

Remark 7.3 By the results in Section 6, we obtain the stronger result that the CLT, WIP and upper LIL hold also on $X \times \{g_0\}$ for each fixed $g_0 \in G$.

(b) SRB measures and unimodal maps

In this section we follow very closely the original exposition of Young [40] referring also to Baladi [1]. As Baladi [1, Section 4.3] records, this 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 [40, 12]

(ii) a class of Hénon maps [3, 4]

(iii) Poincaré maps of billiards with convex scatterers [40] and certain other dispersing billiards [13]

(iv) some partially hyperbolic diffeomorphisms with a mostly contracting direction [11, 15].

The tower approach also applies to

(v) certain C^2 unimodal maps, namely those satisfying conditions (H1) and (H2) of Young [40, Section 9.1].

For these systems a Markov tower $(\overline{f}, \overline{\Delta})$ is constructed. We denote the projection from $\overline{\Delta}$ to M by $\pi : \overline{\Delta} \to M$. The set $\overline{\Delta}$ is partitioned into countably many subsets $\{\overline{\Delta}_j\}_{j=0}^{\infty}$ which are called levels. The base $\overline{\Delta}_0$ is partitioned into countably many subsets $\{\overline{\Delta}_{0,j}\}$ by a measurable 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 levels Δ_l are further subdivided to form a partition $P = \{\overline{\Delta}_{l,j}\}$ of $\overline{\Delta}$ which has the Markov property with respect to $\overline{f}$. 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 [40, 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 [40, Definition 1]).

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 $\overline{\Delta}_0$. From $(\overline{f}, \overline{\Delta})$ a non-invertible tower (f, Δ) is derived by quotienting along stable manifolds (note that the construction of the quotiented tower is not necessary for the one-dimensional non-invertible unimodal maps described in (v), so their analysis is somewhat simpler). We denote this projection $\overline{\pi} : \overline{\Delta} \to \Delta$. We 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, which we denote ℓ, is constructed on Δ [40, Section 3]. The measure ℓ is equivalent to Lebesgue.

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. Note that A is the Perron-Frobenius operator for the reference measure ℓ.

The measure m is lifted to an invariant measure $\overline{m}$ on $\overline{\Delta}$ and thence to an invariant measure $\mu = \pi^*\overline{m}$ for $\mathcal{F} : M \to M$. We assume, following [40], that there exists $\epsilon > 0$ such that $\sum_{l=0}^{\infty} m(\Delta_l)e^{2\epsilon l} < \infty$ (equivalently $\sum_{l=0}^{\infty} \ell(\Delta_l)e^{2\epsilon l} < \infty$) . (We take the scaling factor as $2\epsilon l$ rather than ϵl as we desire the Banach space F (to be defined) to satisfy $F \subset L^2_m$.)

We now discuss the relevant Banach spaces. 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 the separation time function $s(x, y)$.

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

$$\|v\|_\infty = \sup_{l,j} \sup_{\Delta_{l,j}} |v|e^{-l\epsilon},$$

$$\|v\|_\beta = \sup_{l,j} |v|_{l,j,\beta}, \quad \text{where}$$

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

We define $\|v\| = \|v\|_\infty + \|v\|_\beta$ and (see [1, Section 3.4, page 203]) define the Banach space F to be the space of functions $v : \Delta \to \mathbb{R}$ with $\|v\| < \infty$. Also, we define $F' = L^m_1$ to be the space of functions $v : \Delta \to \mathbb{R}$ that are integrable with respect to m. The transfer operator A is well-defined on F and on F'. As sketched below, conditions (F1)–(F5) may be deduced in a straightforward manner. We note that since $\beta^{s(x,y)}$ defines a metric on Δ, elements of F are Lipschitz continuous on each level Δ_l.

The density φ lies in F, as does φ^{-1}. It is important to note that additionally φ and φ^{-1} are uniformly bounded and uniformly Lipschitz [40, Lemma 2]. It follows easily that the operator $v \mapsto \varphi v$ is a linear isomorphism on F' and F so that Lemma 4.9 is applicable.

(F1) It suffices to verify condition (F1) on each level Δ_l separately. Note that F restricted to Δ_l is densely embedded in $L^1_m(\Delta_0)$ (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) We show $F \subset L^2_m$ and $\|\ \| \geq |\ |_2$. 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 $\|v_n\| \le 1$. Restricting to Δ_l, we have $\|v_n|\Delta_l\| \le e^{\epsilon\ell}$. The usual Arzela-Ascoli argument shows that there exists $w : \Delta_l \to \mathbb{R}$ with $\|w\| \le e^{\epsilon\ell}$ 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 $\|w\| \le 1$. 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 4.9, it is also the case that A satisfies (F4).)

(F5) This condition is verified for A in [40, Lemma 3] and [1, Lemma 3.7]. Applying Lemma 4.9, we have that P satisfies (F5).

We are now once again in a position to apply the results in Sections 2, 3 and 4. Assume that the equilibrium measure m is ergodic. Let $v \in F$ with $\int_\Delta v\, dm = 0$, and define $v_N = \sum_{j=1}^N v \circ f^j$.

• 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 $v_k(x) \ne 0$, then $\sigma^2 > 0$.
• The sequence of partial sums v_N satisfies the CLT and the WIP.
• If in addition m is weak mixing, then we obtain exponential decay of correlations.

Remark 7.4 The CLT and exponential decay of correlations is obtained in Young [40] when m is weak mixing. The CLT for m ergodic is new, as are the strong nondegeneracy results. The WIP is not stated in [40] but follows in a standard way from the set up there.

As far as we know, the upper LIL has not been proved in this context. Note that the upper LIL does not follow from the results in this paper since the condition $F \subset L^\infty$ 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 $c\cdot\phi_N = c\cdot\psi\circ f - c\cdot\psi$ where $\psi \in F_G$.

• The sequence of partial sums ϕ_N satisfies the d-dimensional CLT and WIP on $\Delta\times G$ with covariance matrix Σ. Moreover, the CLT and WIP hold also on $\Delta\times\{g_0\}$ for each fixed $g_0 \in G$.

• If in addition $m\times\nu$ is weak mixing, then we obtain exponential decay of correlations (against test functions in L^2_G).

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 $M\times G$. In this set-up we obtain the CLT and WIP and, assuming weak mixing, we obtain exponential decay of correlations.

Our line of argument to obtain the CLT and exponential decay of correlations follows Young [40]. To obtain the CLT she verifies the conditions of Gordin [17].

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 sigma-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 [40, Section 5.2] we let h_0 denote the conditional expectation of $\overline{h}$ with respect to $\mathcal{B}_s$, so that

$$h_0 = E_{\overline{m}}(\overline{h}|\mathcal{B}_s)$$

We similarly define v_0 and $\phi_0(x,g) = gv_0(x)$. Note that 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 $\phi_0 \in F_G$.

We 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 sigma-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 use Gordin's theorem [40, Theorem G] and follow Young's proof of the verification of its condition.

Let $\widehat{\phi}_j = E_{\widehat{m}}(\widehat{\phi}|\widehat{T}^j\widehat{\mathcal{B}})$. Since we have uniform contraction on stable manifolds by [40, 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 $\widehat{\phi}_{-j}$ and 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 the fact that the equivariant transfer operator is quasicompact and hence we have proved the CLT. Similarly the WIP follows by standard techniques. We omit the details.

The proof of exponential decay of correlations follows as in [40, Section 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 $c \cdot \phi_N = c \cdot \psi \circ \mathcal{F} - c \cdot \psi$ where $\psi : M \times G \to \mathbb{R}^d$ is G-equivariant and Hölder.
- The sequence of partial sums ϕ_N satisfies the d-dimensional CLT and WIP on $M \times G$ with covariance matrix Σ. Moreover, the CLT and WIP hold also on $M \times \{g_0\}$ for each fixed $g_0 \in G$.
- If in addition $\mu \times \nu$ is weak mixing, then we obtain exponential decay of correlations (against Hölder equivariant test functions).

Remark 7.5 Results of [30] 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.

Acknowledgements We thank Gerhard Keller and Andrew Török 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. *Ergodic Theory Dynam. Systems* **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. *Ergodic Theory Dynam. Systems* **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*, to appear.

[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. Piecewise invertible dynamical systems. *Probab. Theory Relat. Fields* **72** (1986) 359–386.

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

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

[25] 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.

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

[27] T. Li and J. Yorke. Ergodic transformations from an interval into itself. *Trans. AMS.* **235**, (1978) 183–192.

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

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

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

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

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

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

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

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

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

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

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

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

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

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