Martingale-coboundary decomposition for families of dynamical systems

A. Korepanov Z. Kosloff I. Melbourne

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

4 August 2016; updated 23 August 2017.

Abstract

We prove statistical limit laws for sequences of Birkhoff sums of the type $\sum_{j=0}^{n-1} v_n \circ T_n^j$ where T_n is a family of nonuniformly hyperbolic transformations.

The key ingredient is a new martingale-coboundary decomposition for nonuniformly hyperbolic transformations which is useful already in the case when the family T_n is replaced by a fixed transformation T, and which is particularly effective in the case when T_n varies with n.

In addition to uniformly expanding/hyperbolic dynamical systems, our results include cases where the family T_n consists of intermittent maps, unimodal maps (along the Collet-Eckmann parameters), Viana maps, and externally forced dispersing billiards.

As an application, we prove a homogenization result for discrete fast-slow systems where the fast dynamics is generated by a family of nonuniformly hyperbolic transformations.

1 Introduction

The emergence of statistical and stochastic phenomena in deterministic dynamical systems is currently a very active area. Topics of sustained interest include central limit theorems, invariance principles (weak and almost sure convergence to Brownian motion), moment estimates, and homogenization (whereby deterministic systems with multiple timescales converge to a stochastic differential equation).

One of the standard techniques for investigating such phenomena is the martingale-coboundary decomposition method of Gordin [26] which has seen extensive development in both the probability theory literature (for example [31, 35, 41, 51]) and in the dynamical systems literature (for example [39, 54, 55]).

In this paper, we introduce a new version of the Gordin method and show that it has significant advantages over previous versions when applied to a wide range of questions in nonuniformly hyperbolic dynamics. Even in the case of a single nonuniformly hyperbolic transformation, there are are advantages to the new approach which seems more elementary and more powerful than the existing ones in the literature. In addition, our method is well suited for studying sequences of Birkhoff sums of the form $S_n = \sum_{j=0}^{n-1} v_n \circ T_n^j$ where $T_n^{j+1} = T_n^j \circ T_n$ which arise naturally in averaging and homogenization problems. Here,

$T_n : \Lambda_n \to \Lambda_n$, $n \ge 0$ is a sequence of measure-preserving transformations defined on probability spaces (Λ_n, μ_n). It is assumed that the transformations T_n are nonuniformly expanding/hyperbolic with uniform constants, but no restrictions are imposed on how T_n varies with n.

In the case of a single nonuniformly hyperbolic map T, our method applies directly to T bypassing any induced limit theorems for the associated induced uniformly hyperbolic map. Unlike other approaches [35, 39, 41, 54], no approximation arguments are required for the central limit theorem (CLT) and weak invariance principle (WIP) when the inducing time is not L^3. For moment estimates, the method does not require special arguments when the inducing time is not L^2 (cf. [22, 30]). In addition, we obtain a simple proof of an unexpected CLT for systems with nonsummable decay of correlations due to [28], whereas the previous proof relied on operator renewal theory and the Wiener lemma in noncommutative Banach algebras.

Still in the case of a single map T, our method is very well-adapted for obtaining a secondary martingale-coboundary decomposition for the square of the martingale in the decomposition mentioned above. This enables control on sums of squares as is often required in more sophisticated limit laws. To illustrate this, we consider an almost sure invariance principle with excellent error rates due to [21], and show that our method of applying their results leads to stronger conclusions in certain examples.

The main advantage of the approach, however, is that it allows explicit control on various constants associated with each transformation T, making the method especially useful for sums of the form $\sum_{j=0}^{n-1} v_n \circ T_n^j$. This in turn has applications to fast-slow systems of the type considered in [37]. Whereas [37] obtained rates of averaging, we prove results here on homogenization.

The remainder of this paper is organised as follows. In Section 2, we establish the new martingale-coboundary decomposition for nonuniformly expanding maps and show how this implies moment estimates and the WIP. In Section 3, we obtain a secondary martingale-coboundary decomposition and apply this to the almost sure invariance principle. In Section 4, we derive limit laws for families of nonuniformly expanding maps. This is extended to families of nonuniformly hyperbolic transformations in Section 5. In Section 6, we state and prove an abstract theorem on homogenization for discrete time fast-slow systems, generalising [27]. In Section 7, we verify the hypotheses in Section 6 when the fast dynamics is given by a family of nonuniformly hyperbolic transformations.

Notation We write $\to_{\mu_n}$ to denote weak convergence with respect to a specific family of probability measures μ_n on the left-hand-side. So $A_n \to_{\mu_n} A$ means that A_n is a family of random variables on (Λ_n, μ_n) and $A_n \to_w A$.

For $J \in \mathbb{R}^{m\times n}$, we write $|J| = \big(\sum_{i=1}^m \sum_{j=1}^n J_{ij}^2\big)^{1/2}$.

We use "big O" and $\ll$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \ll b_n$ if there is a constant $C > 0$ such that $a_n \le C b_n$ for all $n \ge 1$. As usual, $a_n = o(b_n)$ means that $\lim_{n\to\infty} a_n/b_n = 0$.

Recall that $v : \Lambda \to \mathbb{R}$ is a Hölder observable on a metric space (Λ, d) if $\|v\|_\eta = |v|_\infty + |v|_\eta < \infty$ where $|v|_\infty = \sup_\Lambda |v|$, $|v|_\eta = \sup_{x\ne y} \frac{|v(x)-v(y)|}{d(x,y)^\eta}$.

2 Martingale-coboundary decomposition for nonuniformly expanding maps

In this section, we prove our main theoretical result on martingale-coboundary decomposition for nonuniformly expanding maps. The notion of nonuniformly expanding map is recalled in Subsection 2.1. The martingale-coboundary decomposition is carried out in Subsection 2.2. Subsection 2.3 shows how certain limit laws follow from this decomposition.

2.1 Nonuniformly expanding maps

Let (Λ, d_Λ) be a bounded metric space with finite Borel measure ρ and let $T : \Lambda \to \Lambda$ be a nonsingular transformation. Let $Y \subset \Lambda$ be a subset of positive measure, and let α be an at most countable measurable partition of Y with $\rho(a) > 0$ for all $a \in \alpha$. We suppose that there is an integrable *return time* function $\tau : Y \to \mathbb{Z}^+$, constant on each a with value $\tau(a) \geq 1$, and constants $\lambda > 1$, $\eta \in (0, 1]$, $C_0, C_1 \geq 1$ such that for each $a \in \alpha$,

(1) $F = T^\tau$ restricts to a (measure-theoretic) bijection from a onto Y.

(2) $d_\Lambda(Fx, Fy) \geq \lambda d_\Lambda(x, y)$ for all $x, y \in a$.

(3) $d_\Lambda(T^\ell x, T^\ell y) \leq C_0 d_\Lambda(Fx, Fy)$ for all $x, y \in a$, $0 \leq \ell < \tau(a)$.

(4) $\zeta_0 = \frac{d\rho|_Y}{d\rho|_Y \circ F}$ satisfies $|\log \zeta_0(x) - \log \zeta_0(y)| \leq C_1 d_\Lambda(Fx, Fy)^\eta$ for all $x, y \in a$.

Such a dynamical system $T : \Lambda \to \Lambda$ is called *nonuniformly expanding*. We refer to $F = T^\tau : Y \to Y$ as the *induced map*. (It is not required that τ is the first return time to Y.) There is a unique absolutely continuous F-invariant probability measure μ_Y on Y and $d\mu_Y/d\rho \in L^\infty$.

Define the Young tower [57], $\Delta = \{(y, \ell) \in Y \times \mathbb{Z} : 0 \leq \ell \leq \tau(y) - 1\}$, and the tower map $f : \Delta \to \Delta$ where $f(y, \ell) = \begin{cases} (y, \ell + 1), & \ell \leq \tau(y) - 2 \\ (Fy, 0), & \ell = \tau(y) - 1 \end{cases}$. The projection $\pi_\Delta : \Delta \to \Lambda$, $\pi_\Delta(y, \ell) = T^\ell y$, defines a semiconjugacy from f to T. Define the ergodic f-invariant probability measure $\mu_\Delta = \mu_Y \times \{\text{counting}\}/\int_Y \tau \, d\mu_Y$ on Δ. Then $\mu = (\pi_\Delta)_* \mu_\Delta$ is an absolutely continuous ergodic T-invariant probability measure.

Remark 2.1 The above definition of nonuniformly expanding map covers many important classes of examples such as those mentioned in this paper. Indeed, it is generally true that nonuniform expansivity plus the existence of good statistical properties actually implies the existence of an inducing scheme satisfying the conditions above, see [4, 5]. See [7] for related results in the invertible setting (Section 5).

In this section, we work with a fixed nonuniformly expanding map $T : \Lambda \to \Lambda$, induced map $F = T^\tau : Y \to Y$, where $\tau \in L^p(Y)$ for some $p \geq 1$, and Young tower map $f : \Delta \to \Delta$. The corresponding ergodic invariant probability measures are denoted μ, μ_Y and μ_Δ. Throughout, $|\ |_p$ denotes the norm in $L^p(\mu)$ for functions on Λ, in $L^p(\mu_Y)$ for functions on Y, and in $L^p(\mu_\Delta)$ for functions on Δ. Also, $\|\ \|_\eta$ denotes the Hölder norm on Λ and Y.

Although the map T is fixed, the dependence of various constants on T is important in later sections. To simplify the statement of results in this section, we denote by C various constants depending continuously on $\operatorname{diam}\Lambda$, C_0, C_1, λ, η, p and $|\tau|_p$.

Let $L: L^1(\Delta) \to L^1(\Delta)$ and $P: L^1(Y) \to L^1(Y)$ denote the transfer operators corresponding to $f: \Delta \to \Delta$ and $F: Y \to Y$. (So $\int_\Delta Lv\, w\, d\mu_\Delta = \int_\Delta v\, w \circ f\, d\mu_\Delta$ for $v \in L^1(\Delta)$, $w \in L^\infty(\Delta)$, and $\int_Y Pv\, w\, d\mu_Y = \int_Y v\, w \circ F\, d\mu_Y$ for $v \in L^1(Y)$, $w \in L^\infty(Y)$.)

Let $\zeta = \frac{d\mu_Y}{d\mu_Y \circ F}$. Given $y \in Y$, let y_a denote the unique $y_a \in a$ with $Fy_a = y$. Then we have the pointwise expressions for P and L,

$$(P\psi)(y) = \sum_{a\in\alpha} \zeta(y_a)\psi(y_a), \quad (L\psi)(y,\ell) = \begin{cases} \sum_{a\in\alpha} \zeta(y_a)\psi(y_a, \tau(y_a)-1), & \ell = 0 \\ \psi(y, \ell-1), & \ell \geq 1 \end{cases}. \tag{2.1}$$

Proposition 2.2 *$\zeta(x) \leq C\mu_Y(a)$ and $|\zeta(x) - \zeta(y)| \leq C\mu_Y(a) d_\Lambda(Fx, Fy)^\eta$ for all $x, y \in a$, $a \in \alpha$.*

Proof By [38, Propositions 2.3 and 2.5], $\big|\log\zeta(x) - \log\zeta(y)\big| \ll d_\Lambda(Fx, Fy)^\eta$. In particular, $\zeta(x)/\zeta(y) \ll 1$. Hence

$$\mu_Y(a) = \textstyle\int_Y 1_a\, d\mu_Y = \int_Y P1_a\, d\mu_Y \geq \inf P1_a = \inf_a \zeta \gg \sup_a \zeta,$$

and so $\zeta(y) \ll \mu_Y(a)$.

Next, we note the inequality $|s-t| \leq \max\{s,t\}|\log s - \log t|$ which is valid for all $s, t > 0$. Hence $|\zeta(x) - \zeta(y)| \ll \sup_a \zeta d_\Lambda(Fx, Fy)^\eta \ll \mu_Y(a) d_\Lambda(Fx, Fy)^\eta$. ∎

2.2 The primary martingale-coboundary decomposition

Let $v: \Lambda \to \mathbb{R}^d$ be Hölder with $\int_\Lambda v\, d\mu = 0$, and define $\phi = v \circ \pi_\Delta: \Delta \to \mathbb{R}^d$. Define the *induced observable* $\phi': Y \to \mathbb{R}^d$ by $\phi'(y) = \sum_{\ell=0}^{\tau(y)-1} \phi(y,\ell)$.

Proposition 2.3 *$\|P\phi'\|_\eta \leq C\|v\|_\eta$.*

Proof Let $x, y \in Y$, $a \in \alpha$, with corresponding preimages $x_a, y_a \in a$. Then

$$|\phi'(x_a) - \phi'(y_a)| \leq \sum_{\ell=0}^{\tau(a)-1} \big|v(T^\ell x_a) - v(T^\ell y_a)\big| \ll |v|_\eta \tau(a) d_\Lambda(x,y)^\eta. \tag{2.2}$$

Also $|\phi'| \leq |v|_\infty \tau$. By (2.1) and Proposition 2.2,

$$\begin{aligned} |(P\phi')(x) - (P\phi')(y)| &\leq \sum_{a\in\alpha} |\zeta(x_a) - \zeta(y_a)|\, |\phi'(x_a)| + \sum_{a\in\alpha} \zeta(y_a)|\phi'(x_a) - \phi'(y_a)| \\ &\ll \|v\|_\eta \sum_{a\in\alpha} \mu_Y(a)\tau(a) d_\Lambda(x,y)^\eta \ll \|v\|_\eta d_\Lambda(x,y)^\eta. \end{aligned}$$

Hence $|P\phi'|_\eta \ll \|v\|_\eta$. Similarly, $|P\phi'|_\infty \ll |v|_\infty$. ∎

Define $\chi', m' : Y \to \mathbb{R}^d$ as follows:

$$\chi' = \textstyle\sum_{k=1}^{\infty} P^k \phi', \qquad \phi' = m' + \chi' \circ F - \chi'.$$

By Proposition 2.3 and [38, Corollary 2.4 and Proposition 2.5],

$$\|\chi'\|_\eta \le \sum_{k=0}^{\infty} \|P^k P\phi'\|_\eta \ll \|P\phi'\|_\eta \ll \|v\|_\eta,$$
$$|m'|_p \le |\phi'|_p + 2|\chi'|_\infty \le |v|_\infty |\tau|_p + 2|\chi'|_\infty \ll \|v\|_\eta.$$

Define $\chi, m : \Delta \to \mathbb{R}^d$ by

$$\chi(y,\ell) = \chi'(y) + \sum_{k=0}^{\ell-1} \phi(y,k), \qquad m(y,\ell) = \begin{cases} 0, & \ell \le \tau(y) - 2 \\ m'(y), & \ell = \tau(y) - 1 \end{cases}.$$

Proposition 2.4 *$|m|_p \le C\|v\|_\eta$ and $|\chi|_{p-1} \le C\|v\|_\eta$.*

Proof Compute that $\int_\Delta |m|^p \, d\mu_\Delta = |\tau|_1^{-1} \int_Y \sum_{\ell=0}^{\tau(y)-1} |m(y,\ell)|^p \, d\mu_Y = |\tau|_1^{-1} \int_Y |m'(y)|^p \, d\mu_Y \le |m'|_p^p \ll \|v\|_\eta^p$. Similarly, $|\chi(y,\ell)| \le |\chi'|_\infty + \ell|v|_\infty \ll \tau(y)\|v\|_\eta$ yielding the estimate for χ. ∎

Proposition 2.5 *$\phi = m + \chi \circ f - \chi$ and $m \in \ker L$.*

Proof If $\ell \le \tau(y) - 2$, then

$$\chi \circ f(y,\ell) - \chi(y,\ell) = \chi(y,\ell+1) - \chi(y,\ell) = \phi(y,\ell) = \phi(y,\ell) - m(y,\ell).$$

For $p = (y, \tau(y) - 1)$,

$$\begin{aligned} \chi \circ f(p) - \chi(p) &= \chi(Fy, 0) - \chi(y, \tau(y) - 1) = \chi'(Fy) - \chi'(y) - \textstyle\sum_{k=0}^{\tau(y)-2} \phi(y,k) \\ &= \phi'(y) - m'(y) - \textstyle\sum_{k=0}^{\tau(y)-2} \phi(y,k) = \phi(p) - m(p). \end{aligned}$$

Hence $\phi = m + \chi \circ f - \chi$.

By definition, $Pm' = P\phi' - \chi' + P\chi' \equiv 0$. Using (2.1), observe that $(Lm)(y,\ell) = m(y,\ell-1) = 0$ if $\ell \ge 1$, and

$$(Lm)(y,0) = \textstyle\sum_{a\in\alpha} \zeta(y_a) m(y_a, \tau(y_a) - 1) = \sum_{a\in\alpha} \zeta(y_a) m'(y_a) = (Pm')(y) = 0.$$

Hence $m \in \ker L$. ∎

Proposition 2.6 *$\max_{0\le k\le n} |\chi \circ f^k| = o(n^{1/p})$ a.e.*

Proof Since $\tau \in L^p$, it follows from the ergodic theorem that $\tau \circ F^n = o(n^{1/p})$ a.e., and hence that $\max_{0\le k\le n} \tau \circ F^k = o(n^{1/p})$ a.e.

Next, $|\chi(y,\ell)| \le |\chi'|_\infty + \ell|v|_\infty \ll \tau(y)\|v\|_\eta$. For any $(y,\ell) \in \Delta$ and $n \ge 0$, there exists $k \in \{0, \dots, n\}$ and $\ell' \in \{0, \dots, \tau(F^k y) - 1\}$ such that $f^n(y,\ell) = (F^k y, \ell')$. Hence $|\chi(f^n(y,\ell))| \ll$

$\|v\|_\eta \max_{0\le k\le n}\tau(F^ky)$ and so $\max_{0\le k\le n}|\chi(f^k(y,\ell))| \ll \|v\|_\eta \max_{0\le k\le n}\tau(F^ky) = o(n^{1/p})$ a.e. ∎

No uniformity of constants is claimed in Proposition 2.6. It is straightforward to show that $\big|\max_{0\le k\le n}|\chi\circ f^k|\big|_{p-1} \le C\|v\|_\eta n^{1/(p-1)}$ where C is a uniform constant. However, for various purposes (such as optimal moment estimates in Corollary 2.10) we require the following more delicate estimate.

Proposition 2.7 *$\big|\max_{1\le k\le n}|\chi\circ f^k-\chi|\big|_p \le C\|v\|_\eta n^{1/p}$. Moreover,*

$$\big|\max_{1\le k\le n}|\chi\circ f^k-\chi|\big|_p \le C\|v\|_\eta(n^{1/q}+n^{1/p}|1_{\{\tau\ge n^{1/q}\}}\tau|_p) \quad \textit{for all } n\ge 0,\ q\ge p.$$

Proof Define $t_a=|1_{\{\tau\ge a\}}\tau|_p$, $a\ge 0$. Then

$$\begin{aligned}\textstyle\sum_{k\ge n}k^{p-1}\mu_Y(\tau\ge k) &= \textstyle\sum_{k\ge n}\sum_{j\ge k}k^{p-1}\mu_Y(\tau=j)\\ &= \textstyle\sum_{j\ge n}\mu_Y(\tau=j)\sum_{k=n}^{j}k^{p-1}\le \sum_{j\ge n}j^p\mu_Y(\tau=j)=t_n^p. \end{aligned}\tag{2.3}$$

Let $\Delta_n=\{(y,\ell)\in\Delta:\ell=n\}$ and $A_n=\{(y,\ell)\in\Delta:\ell<\tau(y)-n\}$. Then $\mu_\Delta(\Delta_n)=|\tau|_1^{-1}\mu_Y(\tau\ge n)$ and $\mu_\Delta(A_n)=\mu_\Delta(\cup_{k\ge n}\Delta_k)=|\tau|_1^{-1}\sum_{k\ge n}\mu_Y(\tau\ge k)$. By (2.3),

$$n^{p-1}\mu_\Delta(A_n)=n^{p-1}|\tau|_1^{-1}\textstyle\sum_{k\ge n}\mu_Y(\tau\ge k)\le\sum_{k\ge n}k^{p-1}\mu_Y(\tau\ge k)\le t_n^p.$$

If $(y,\ell)\in A_n$, then $\max_{1\le k\le n}\big|(\chi\circ f^k-\chi)(y,\ell)\big|\le n|v|_\infty$. Therefore

$$\begin{aligned}\big|1_{A_n}\max_{1\le k\le n}|\chi\circ f^k-\chi|\big|_p &\le n|v|_\infty[\mu_\Delta(A_n)]^{1/p}\\ &= n^{1/p}|v|_\infty[n^{p-1}\mu_\Delta(A_n)]^{1/p}\le |v|_\infty n^{1/p}t_n. \end{aligned}\tag{2.4}$$

For all $(y,\ell)\in\Delta$, we have $|\chi(y,\ell)|\ll\tau(y)\|v\|_\eta$ and so $|\chi\circ f^k|\ll\|v\|_\eta\max_{0\le j\le k}\tau\circ F^j$. Let $a>0$ and denote $\tau_a=1_{\{\tau>a\}}\tau$. Since $\tau^p\le a^p+\tau_a^p$,

$$\begin{aligned}\|v\|_\eta^{-p}\max_{1\le k\le n}|\chi(f^k(y,\ell))-\chi(y,\ell)|^p &\le 2^p\|v\|_\eta^{-p}\max_{0\le k\le n}|\chi(f^k(y,\ell))|^p\\ &\ll \max_{0\le k\le n}\tau^p(F^ky)\le a^p+\textstyle\sum_{0\le k\le n}\tau_a^p(F^ky). \end{aligned}\tag{2.5}$$

Suppose that $\psi:\Delta\to\mathbb{R}$ has the form $\psi(y,\ell)=\psi_0(y)$ where $\psi_0:Y\to\mathbb{R}$. Then

$$\textstyle\int_{\Delta\setminus A_n}|\psi|\,d\mu_\Delta=|\tau|_1^{-1}\int_Y\min\{\tau,n\}|\psi_0|\,d\mu_Y\le\int_Y\min\{\tau,n\}|\psi_0|\,d\mu_Y. \tag{2.6}$$

Taking $v=1$ in Proposition 2.3 (resulting in $\phi=1$, $\phi'=\tau$) and using that P is a contraction yields the estimate $|P^k\tau|_\infty\le|P\tau|_\infty\ll\|1\|_\eta=1$ for all $k\ge1$. Then by equations (2.5) and (2.6),

$$\begin{aligned}\|v\|_\eta^{-p}\textstyle\int_{\Delta\setminus A_n}\max_{1\le k\le n}|\chi\circ f^k-\chi|^p\,d\mu_\Delta &\ll a^p+\textstyle\sum_{0\le k\le n}\int_{\Delta\setminus A_n}\tau_a^p(F^ky)\,d\mu_\Delta(y,\ell)\\ &\le a^p+\textstyle\sum_{0\le k\le n}\int_Y\min\{\tau,n\}\,\tau_a^p\circ F^k\,d\mu_Y\le a^p+n|\tau_a^p|_1+\sum_{k=1}^n|\tau\,\tau_a^p\circ F^k|_1\\ &= a^p+n|\tau_a^p|_1+\textstyle\sum_{k=1}^n|P^k\tau\,\tau_a^p|_1\ll a^p+n|\tau_a^p|_1=a^p+nt_a^p.\end{aligned}$$

Hence

$$\big|1_{\Delta\setminus A_n}\max_{1\le k\le n}|\chi\circ f^k-\chi|\big|_p\ll\|v\|_\eta(a^p+nt_a^p)^{1/p}\le\|v\|_\eta(a+n^{1/p}t_a).$$

We take $a=n^{1/q}$. Combining with (2.4) and using $t_n\le t_{n^{1/q}}$ completes the proof. ∎

Corollary 2.8 $\left|\max_{1\le k\le n}|\chi\circ f^k-\chi|\right|_p = o(n^{1/p})$. ∎

The next result justifies calling $\phi = m+\chi\circ f-\chi$ a *martingale-coboundary decomposition.* Let U denote the Koopman operator corresponding to f, i.e. $Uv = v\circ f$.

Proposition 2.9 *Fix $n \ge 1$. Let $\mathcal{M}$ denote the underlying σ-algebra on (Δ,μ_Δ) and define $\mathcal{G}_j = f^{-(n-j)}\mathcal{M}$, $1\le j\le n$. Then $\{m\circ f^{n-j}, \mathcal{G}_j;\ 1\le j\le n\}$ is a sequence of* martingale differences. *That is, $\mathcal{G}_1\subset\cdots\subset\mathcal{G}_n$, $m\circ f^{n-j}$ is $\mathcal{G}_j$-measurable for each j, and $\mathbb{E}(m\circ f^{n-j}|\mathcal{G}_{j-1})=0$ for each j.*

Proof Since $f^{-1}\mathcal{M}\subset\mathcal{M}$, it follows that $\mathcal{G}_j\subset\mathcal{G}_{j+1}$. Measurability of $m\circ f^{n-j}$ with respect to $\mathcal{G}_j$ is clear. It is standard, and easy to check, that $UL=\mathbb{E}(\,\cdot\,|f^{-1}\mathcal{M})$. Hence

$$\mathbb{E}(m\circ f^{n-j}|\mathcal{G}_{j-1}) = \mathbb{E}(m|f^{-1}\mathcal{M})\circ f^{n-j} = (ULm)\circ f^{n-j} = 0,$$

since $m\in\ker L$. ∎

2.3 Some limit theorems

Suppose that $v:\Lambda\to\mathbb{R}^d$ is Hölder and $\int_\Lambda v\,d\mu = 0$. By the results from Subsection 2.2 we have the decomposition $v\circ\pi_\Delta = \phi = m+\chi\circ f-\chi$, where m, χ satisfy the estimates in Propositions 2.4 and 2.7.

Corollary 2.10 (Moments) *If $p\le 2$, then $\left|\max_{j\le n}|\sum_{k=0}^{j-1} v\circ T^k|\right|_p \le C\|v\|_\eta n^{1/p}$ and $\left|\max_{j\le n}|\sum_{k=0}^{j-1} m\circ f^k|\right|_p \le C\|v\|_\eta n^{1/p}$ for all $n\ge 1$.*

If $p\ge 2$, then $\left|\max_{j\le n}|\sum_{k=0}^{j-1} v\circ T^k|\right|_{2(p-1)} \le C\|v\|_\eta n^{1/2}$ for all $n\ge 1$.

Proof Since $\{m\circ f^{n-j};\ 1\le j\le n\}$ is a sequence of martingale differences with respect to the filtration $\mathcal{G}_j = f^{n-j}\mathcal{M}$ for each $n\ge 1$ (Proposition 2.9), it follows from Burkholder's inequality [15] and Proposition 2.4 that for $p\le 2$,

$$\left|\max_{j\le n}|\textstyle\sum_{k=1}^{j} m\circ f^{n-k}|\right|_p \ll n^{1/p}|m|_p \ll n^{1/p}\|v\|_\eta.$$

Writing $\sum_{k=0}^{j-1} m\circ f^k = \sum_{k=1}^{n} m\circ f^{n-k} - \sum_{j=1}^{n-j} m\circ f^{n-k}$, we obtain that $\left|\max_{j\le n}|\sum_{k=0}^{j-1} m\circ f^k|\right|_p \ll n^{1/p}\|v\|_\eta$. Combining this with Proposition 2.7 yields $\left|\max_{j\le n}|\sum_{k=0}^{j-1}\phi\circ f^k|\right|_p \ll n^{1/p}\|v\|_\eta$ and the result for $p\le 2$ follows.

When $p\ge 2$, we use Rio's inequality [53] following [44]. See [50, Proposition 7] for a statement of Rio's inequality. Let $X_j = \phi\circ f^{n-j}$. For $1\le j\le\ell\le n$, by Proposition 2.9, $\sum_{k=j}^{\ell}\mathbb{E}(X_k|\mathcal{G}_j) = m\circ f^{n-j} + \mathbb{E}(\chi\circ f^{n+1-\ell}|\mathcal{G}_j) - \chi\circ f^{n-j}$. By Proposition 2.4, $\max_{1\le j\le\ell\le n}|\sum_{k=j}^{\ell}\mathbb{E}(X_k|\mathcal{G}_j)|_{p-1} \ll \|v\|_\eta$. Hence $\max_{1\le j\le\ell\le n}|X_j\sum_{k=j}^{\ell}\mathbb{E}(X_k|\mathcal{G}_j)|_{p-1} \le |\phi|_\infty \max_{1\le j\le\ell\le n}|\sum_{k=j}^{\ell}\mathbb{E}(X_k|\mathcal{G}_j)|_{p-1} \ll \|v\|_\eta^2$. The result follows by Rio's inequality. ∎

Remark 2.11 The moment estimates for $p\ge 2$ were first obtained in [44] and the results for $p<2$ are due to [22, 30].

Corollary 2.10 is easily seen to be optimal given the formulation of our results in this paper in terms of the integrability of the return time p. Often a tail estimate of the form $\mu_Y(\tau > n) = O(n^{-p})$ is available, and this gives rise to some interesting subtleties; such issues are also resolved in [22, 30]. On the other hand, these references do not explicitly address the uniformity of the constant C which is required in later sections.

Corollary 2.12 (Covariance) *Suppose that $p \geq 2$. Then $\lim_{n\to\infty} n^{-1}\int_\Lambda(\sum_{j=0}^{n-1} v \circ T^j)(\sum_{j=0}^{n-1} v \circ T^j)^T d\mu = \int_\Delta m\, m^T d\mu_\Delta$.*

Proof Write $S_n v = \sum_{j=0}^{n-1} v \circ T^j$ and similarly define $S_n\phi$ and $S_n m$. Since $m \in \ker L$, $\int_\Delta S_n m\, S_n m^T\, d\mu_\Delta = n\int_\Delta m\, m^T\, d\mu_\Delta$.

By Corollary 2.10, $|S_n\phi|_2 \ll n^{1/2}\|v\|_\eta$ and $|S_n m|_2 \ll n^{1/2}\|v\|_\eta$. Hence,

$$\begin{aligned}\Big|n^{-1}\int_\Lambda S_n v\, S_n v^T\, d\mu - \int_\Delta m\, m^T\, d\mu_\Delta\Big| &= n^{-1}\Big|\int_\Delta S_n\phi\, S_n\phi^T\, d\mu_\Delta - \int_\Delta S_n m\, S_n m^T\, d\mu_\Delta\Big| \\ &\le n^{-1}|S_n\phi\, S_n\phi^T - S_n m\, S_n m^T|_1 \le n^{-1}(|S_n\phi|_2 + |S_n m|_2)|S_n\phi - S_n m|_2 \\ &\ll n^{-1/2}|\chi \circ f^n - \chi|_2\|v\|_\eta \to 0\end{aligned}$$

by Corollary 2.8. ∎

For $n \geq 1$, define the process $W_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} v \circ T^j$ on (Λ, μ).

Corollary 2.13 (WIP) *Suppose that $p \geq 2$. Then $W_n \to_\mu W$ where W is Brownian motion with covariance $\Sigma = \lim_{n\to\infty} n^{-1}\int_\Lambda(\sum_{j=0}^{n-1} v \circ T^j)(\sum_{j=0}^{n-1} v \circ T^j)^T\, d\mu$.*

Proof By Corollary 2.12, $\Sigma = \int_\Delta m\, m^T\, d\mu_\Delta = \int_\Delta UL(m\, m^T)\, d\mu_\Delta$. By the ergodic theorem, $n^{-1}\sum_{j=0}^{[nt]-1}\{UL(mm^T)\} \circ f^j \to t\Sigma$ a.e. as $n \to \infty$ for all $t > 0$. Hence we can apply Theorem A.1 to the process $M_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} m \circ f^j$ to deduce that $M_n \to_{\mu_\Delta} W$.

Next define $\Phi_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} \phi \circ f^j$. For any $T > 0$,

$$\big|\sup_{t\in[0,T]}|\Phi_n(t) - M_n(t)|\big|_2 \le n^{-1/2}\big|\max_{1\le j\le nT}|\chi \circ f^j - \chi|\big|_2 \to 0,$$

by Corollary 2.8. Hence $\Phi_n \to_{\mu_\Delta} W$. Finally, π_Δ is a measure-preserving semiconjugacy, so the result follows. ∎

Remark 2.14 As mentioned in the introduction, previous methods [39, 41, 54] require special techniques for $p \leq 3$. In particular, a sequence of martingale approximations is needed, whereas we work with a single martingale-coboundary decomposition.

Alternatively, [29, 43] obtained a single martingale-coboundary decomposition at the level of the induced map. The resulting CLT/WIP can then be lifted back to the original system by [46, 49].

Finally, we show how to recover a result of [28] where the WIP holds somewhat unexpectedly. Our method of proof is significantly simpler than in [28]. On the other hand, [28] also obtained unexpectedly fast decay of correlations in this situation.

Corollary 2.15 (WIP with $p = 1$) *Suppose that $\tau : Y \to \mathbb{Z}^+$ is the first return to Y and that* $\operatorname{supp} v \subset Y$*. (We continue to suppose that v is Hölder with mean zero.) Then the WIP above holds for all $p \ge 1$.*

Proof The assumptions on τ and v ensure that $\phi(y, \ell) = 0$ for $\ell \ge 1$. But then the definition of χ reduces to $\chi(y,\ell) = \begin{cases} \chi'(y), & \ell = 0 \\ \chi'(y) + \phi(y,0), & \ell \ge 1 \end{cases}$, and it follows that $|\chi|_\infty \ll \|v\|_\eta$ and hence that $|m|_\infty \ll \|v\|_\eta$. The arguments above go through (with numerous simplifications). ∎

Remark 2.16 The results in Subsections 2.2 and 2.3 were proved for observables $\phi = v \circ \pi_\Delta$ where $v : \Lambda \to \mathbb{R}^d$ is Hölder and mean zero. It is easy to check that the only properties of ϕ that were used are (i) $\int_\Delta \phi \, d\mu_\Delta = 0$, (ii) $\phi \in L^\infty$, (iii) $\|P\phi'\|_\eta < \infty$. For such observables $\phi : \Delta \to \mathbb{R}^d$, all the results go through with $\|v\|_\eta$ replaced by $|\phi|_\infty + \|P\phi'\|_\eta$.

3 Secondary martingale-coboundary decomposition and the ASIP

In this section, we derive a secondary martingale-coboundary decomposition for nonuniformly expanding maps. As an illustration of its utility, we obtain an ASIP for nonuniformly expanding maps with improved error rates over those in the literature.

We continue to suppose that $v : \Lambda \to \mathbb{R}^d$ is Hölder with $\int_\Lambda v \, d\mu = 0$ and that $\phi = v \circ \pi_\Delta : \Delta \to \mathbb{R}^d$. In addition, we suppose that $p \ge 2$.

Let

$$\phi = m + \chi \circ f - \chi, \quad \phi' = m' + \chi' \circ f - \chi'$$

be the decompositions from Subsection 2.2. Define $\breve{\phi} : \Delta \to \mathbb{R}^{d \times d}$,

$$\breve{\phi} = UL(m\, m^T) - \textstyle\int_\Delta m\, m^T \, d\mu_\Delta,$$

where U and L are the Koopman and transfer operators for f.

There are a number of limit laws in the literature that require control of Birkhoff sums corresponding to $\breve{\phi}$. As examples, we mention the martingale CLT/WIP [13] (see Appendix A) and an ASIP for reverse martingale differences [21] discussed at the end of this section. A third application [9] is to the estimate of convergence rates in the WIP. To control the Birkhoff sums of $\breve{\phi}$, a martingale-coboundary decomposition for $\breve{\phi}$ is of great utility; this is the topic of the current section.

Proposition 3.1 $|\breve{\phi}|_\infty \le \|v\|_\eta^2$ *and* $\|P\breve{\phi}'\|_\eta \le C\|v\|_\eta^2$.

Proof Let U_F and P denote the Koopman and transfer operators on $L^1(Y)$ for F (so P is as in Subsection 2.1 and $U_F v = v \circ F$). A calculation shows that

$$(L(mm^T))(y, \ell) = \begin{cases} (P(m' m'^T))(y) & \ell = 0 \\ 0 & \ell \ge 1 \end{cases},$$

and hence that

$$(UL(mm^T))(y,\ell) = \begin{cases} 0 & \ell \le \tau(y)-2 \\ (U_F P(m'm'^T))(y) & \ell = \tau(y)-1 \end{cases}.$$

By Proposition 2.4, $|m|_2 \ll \|v\|_\eta$. Also, $\|\chi'\|_\eta \ll \|v\|_\eta$ so $|m'| \le |\phi'| + 2|\chi'|_\infty \ll \tau\|v\|_\eta$. It follows that $|P(m'm'^T)(y)| \le \sum_{a\in\alpha} \zeta(y_a)|m'm'^T(y_a)| \ll \sum_{a\in\alpha} \mu_Y(a)\tau(a)^2\|v\|_\eta^2 \ll \|v\|_\eta^2$. Hence

$$|\breve\phi|_\infty \le |UL(mm^T)|_\infty + |\textstyle\int_\Delta mm^T\, d\mu_\Delta| \le |P(m'm'^T)|_\infty + |m|_2^2 \ll \|v\|_\eta^2.$$

It remains to estimate $\|P\breve\phi'\|_\eta$. Now $\breve\phi'(y) = \sum_{\ell=0}^{\tau(y)-1} \breve\phi(y,\ell) = (U_F P(m'm'^T))(y) - \tau(y)\int_\Delta m\, m^T\, d\mu_\Delta$, so

$$P\breve\phi' = P(m'm'^T) - (P\tau)|m|_2^2.$$

In particular, $|P\breve\phi'|_\infty \le |P(m'm'^T)|_\infty + |P\tau|_\infty |m|_2^2 \ll \|v\|_\eta^2$.

Let $x,y \in Y$ and $a\in\alpha$. Using equation (2.2),

$$\begin{aligned}|m'(x_a)-m'(y_a)| &\le |\phi'(x_a)-\phi'(y_a)| + |\chi'(x)-\chi'(y)| + |\chi'(x_a)-\chi'(y_a)| \\ &\ll \tau(a)\|v\|_\eta d_\Lambda(x,y)^\eta,\end{aligned}$$

and so

$$\begin{aligned}\big|m'(x_a)m'(x_a)^T - m'(y_a)m'(y_a)^T\big| &\le \big(|m'(x_a)|+|m'(y_a)|\big)|m'(x_a)-m'(y_a)| \\ &\ll \tau(a)^2\|v\|_\eta^2 d_\Lambda(x,y)^\eta.\end{aligned}$$

As in the proof of Proposition 2.3,

$$\begin{aligned}&|(P(m'm'^T))(x) - (P(m'm'^T))(y)| \\ &\quad \le \sum_{a\in\alpha} |\zeta(x_a)-\zeta(y_a)||(m'm'^T)(x_a)| + \sum_{a\in\alpha}\zeta(y_a)|(m'm'^T)(x_a)-(m'm'^T)(y_a)| \\ &\quad \ll \|v\|_\eta^2 \sum_{a\in\alpha}\mu_Y(a)\tau(a)^2 d_\Lambda(x,y)^\eta \ll \|v\|_\eta^2 d_\Lambda(x,y)^\eta.\end{aligned}$$

Hence $|P(m'm'^T)|_\eta \ll \|v\|_\eta^2$. A simpler computation shows that $|P\tau|_\eta \ll 1$. It follows that $|P\breve\phi'|_\eta \ll \|v\|_\eta^2$, and so $\|P\breve\phi'\|_\eta \ll \|v\|_\eta^2$. ∎

Proposition 3.1 together with Remark 2.16 allows us to write $\breve\phi = \breve m + \breve\chi\circ f - \breve\chi$, as in Subsection 2.2. In particular,

$$\begin{aligned}&|\breve m|_p \le C\|v\|_\eta^2, \quad \max_{0\le k\le n}|\breve\chi\circ f^k| = o(n^{1/p}) \text{ a.e.}, \\ &\big|\max_{1\le k\le n}|\breve\chi\circ f^k - \breve\chi|\big|_p \le C\|v\|_\eta^2\, n^{1/p} \text{ for all } n\ge 1.\end{aligned} \tag{3.1}$$

We refer to $\breve\phi = \breve m + \breve\chi\circ f - \breve\chi$ as a *secondary martingale-coboundary decomposition.*

Corollary 3.2 $\big|\max_{1\le k\le n}|\sum_{j=0}^{k-1}\breve\phi\circ f^j|\big|_p \le C\|v\|_\eta^2\, n^{1/2}$.

Proof This follows from the argument for Corollary 2.10. ∎

Remark 3.3 In a previous version of this paper, we obtained a similar decomposition for $\breve\phi_1 = m\,m^T - \int_\Delta m\,m^T\,d\mu_\Delta$. The main difference is that $\breve\phi_1' = m'm'^T - \tau\int_\Delta m\,m^T\,d\mu_\Delta$. The estimate for $\|P\breve\phi_1'\|_\eta$ is unchanged. However, we obtain the inferior estimate $|\breve m|_{p/2} \ll \|v\|_\eta^2$, resulting in a weaker estimate in Corollary 3.2.

Since probabilistic results in the literature are typically stated in terms of the conditional variances $ULmm^T = \mathbb{E}(mm^T|f^{-1}\mathcal{M})$ (where $\mathbb{E}$ and $\mathcal{M}$ are as in Proposition 2.9), we have chosen to omit the decomposition for $\breve\phi_1$ in this paper.

Corollary 3.4 (ASIP) *Suppose that $d=1$. Define $\sigma^2 = \lim_{n\to\infty} n^{-1}\int_\Lambda(\sum_{j=0}^{n-1} v\circ T^j)^2\,d\mu$ and suppose that $\sigma^2>0$. Then there exists a probability space Ω supporting a sequence of random variables $\{S_n\}$ with the same joint distributions as $\{\sum_{j=0}^{n-1} v\circ f^j\}$ and a sequence $\{Z_n\}$ of i.i.d. random variables with distribution $N(0,\sigma^2)$ such that almost everywhere as $n\to\infty$*

$$\sup_{1\le k\le n}\big|S_k - \textstyle\sum_{j=1}^k Z_j\big| = \begin{cases} o\big((n\log\log n)^{1/2}\big) & p=2,\\ o\big(n^{1/p}(\log n)^{1/2}\big) & p\in(2,4),\\ O\big(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}\big) & p\ge 4\end{cases}$$

Proof Since $m\in\ker L$, it follows as in Proposition 2.9 that $\mathbb{E}(m\circ f^n|f^{-(n+1)}\mathcal{M}) = \mathbb{E}(m|f^{-1}\mathcal{M})\circ f^n = 0$ for all $n\ge 0$. That is, $\{m\circ f^n\}$ is a sequence of *reverse martingale differences* with respect to the nonincreasing sequence $\{f^{-n}\mathcal{M}\}$ of σ-algebras.

We apply results of [21] to deduce the conclusion of the corollary with the sequence $\{\sum_{j=0}^{n-1} v\circ T^j\}$ replaced by the sequence $\{\sum_{j=0}^{n-1} m\circ f^j\}$. Suppose that this is the case. By Proposition 2.6,

$$\max_{1\le k\le n}\big|\textstyle\sum_{j=0}^{k-1}(\phi\circ f^j - m\circ f^j)\big| \le \max_{1\le k\le n}|\chi\circ f^k-\chi| = o(n^{1/p})\ \textit{a.e.}$$

Enlarging the probability space Ω (cf. [51, p. 23]), there exists a sequence $\{S_n'\}$ with the same joint distributions as $\{\sum_{j=0}^{n-1}\phi\circ f^j\}$ so that $\sup_{1\le k\le n}|S_k' - \sum_{j=1}^k Z_j|$ satisfies the desired estimates. Finally, π_Δ is a measure-preserving semiconjugacy, so the joint distributions of $\{\sum_{j=0}^{n-1} v\circ T^j\}$ also coincide with those of $\{S_n'\}$.

It remains to prove the ASIP (with the appropriate error rates) for the sequence $\{\sum_{j=0}^{n-1} m\circ f^j\}$. The case $p=2$ is immediate from [21, Corollary 2.5].

When $p>2$, we require almost sure estimates for the sequence $A_n = \sum_{j=0}^{n-1}(\mathbb{E}(m^2\circ f^j|\mathcal{G}_{j+1}) - \sigma^2)$. For this we use the secondary martingale-coboundary decomposition $ULm^2 - \sigma^2 = \breve m + \breve\chi\circ f - \breve\chi$. Then

$$A_n = \textstyle\sum_{j=0}^{n-1}((ULm^2)\circ f^j - \sigma^2) = \sum_{j=0}^{n-1}\breve m\circ f^j + \breve\chi\circ f^n - \breve\chi.$$

Since $\{\breve m\circ f^n\}$ is a sequence of L^2 reverse martingale differences, the result for $p=2$ implies an ASIP for $\{\sum_{j=0}^{n-1}\breve m\circ f^j\}$ with error rate $o((n\log\log n)^{1/2})$. This is sufficient to deduce the law of the iterated logarithm $\sum_{j=0}^{n-1}\breve m\circ f^j = O((n\log\log n)^{1/2})$ a.e. Also $\breve\chi\circ f^n - \breve\chi = o(n^{1/p})$ a.e., so $A_n = O((n\log\log n)^{1/2})$ a.e.

For $p \in (2,4)$, the desired ASIP for m now follows from [21, Corollary 2.7] (taking $b(n) \equiv 1$). For $p \geq 4$, the desired ASIP for m follows from [21, Corollary 2.8]. ∎

For the class of (Markovian) nonuniformly expanding maps as defined in Section 2.1, with Hölder observables v, our results improve existing results in the literature. The best previous result that we are aware of is [21, Theorem 3.5] who obtained the error rate $O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})$ for $p > 6$ (this constraint is required to ensure that $|L^n v|_4$ decays faster than $n^{-5/4}$ for mean zero Hölder v so that condition (3.2) in [21] is satisfied), whereas we require only that $p \geq 4$. (For one-dimensional dynamical systems and certain classes of (unbounded) observables, [21] obtain much better results.)

Example 3.5 Consider the map $T : [0,1] \to [0,1]$ of intermittent type [52] studied by [40], namely $T(x) = \begin{cases} x(1+2^\gamma x^\gamma) & x \in [0, \frac{1}{2}) \\ 2x-1 & x \in [\frac{1}{2}, 1] \end{cases}$. It is standard that T is a nonuniformly expanding map with absolutely continuous invariant probability measure μ for each $\gamma \in (0,1)$. The inducing time τ lies in L^p if and only if $p < 1/\gamma$. Hence, we obtain the error rate $O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})$ for all mean zero Hölder observables v provided $\gamma < \frac{1}{4}$; previously this was known only for $\gamma < \frac{1}{6}$.

Example 3.6 We consider a family of planar periodic dispersing billiards introduced by [17]. The scatterers have smooth strictly convex boundaries with nonvanishing curvature, except that the curvature vanishes at two points. Moreover, it is assumed that there is a periodic orbit that runs between the two flat points, and that the boundary near these flat points has the form $\pm(1+|x|^b)$ for some $b > 2$.

By [17], quotienting out stable manifolds leads to a nonuniformly expanding map T with inducing time τ lying in L^p for all $p < (b+2)/(b-2)$. Hence at the level of the quotient map we obtain the error rate $O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4})$ for all mean zero Hölder observables v provided $b < \frac{10}{3}$; previous results require $b < \frac{14}{5}$. We conjecture that these results go over to the full (unquotiented map); this is the topic of future work.

Example 3.7 Bunimovich flowers [14] are billiards where the boundary components of the billiard table are either dispersing, or focusing arcs of circles, subject to some technical constraints. By [19], the quotient map T (obtained by quotienting out stable manifolds) is nonuniformly expanding with inducing time $\tau \in L^p$ for all $p < 3$. Hence, at least at the level of the quotient map we obtain the error rate $o(n^{1/q})$ for all $q < 3$.

4 Limit laws for families of nonuniformly expanding maps

In this section, we show how the martingale-boundary decompositions from the previous sections apply to Birkhoff sums of the type $\sum_{j=0}^{n-1} v_n \circ T_n^j$ where the dynamical systems T_n vary with n.

Suppose that $T_n : \Lambda_n \to \Lambda_n$, $n \geq 0$, is a family of nonuniformly expanding maps as defined in Section 2.1, with absolutely continuous ergodic T_n-invariant probability measures μ_n. Let $\tau_n : Y_n \to \mathbb{Z}^+$ and $F_n : Y_n \to Y_n$ be the corresponding inducing times and induced maps with ergodic F_n-invariant probability measures μ_{Y_n}. We say that $T_n : \Lambda_n \to \Lambda_n$ is a *uniform family of order* $p \geq 1$ if

(i) $\sup_{n\ge0}\operatorname{diam}\Lambda_n<\infty$ and the constants $C_0, C_1 \ge 1$, $\lambda > 1$, $\eta \in (0,1]$ can be chosen independent of $n \ge 0$.

(ii) The family $\{\tau_n^p,\, n \ge 0\}$ is uniformly integrable. (For this it suffices that $\sup_{n\ge0}\int_{Y_n}\tau_n^q\,d\rho<\infty$ for some $q>p$.)

Let $v_n : \Lambda_n \to \mathbb{R}^d$ be a family of Hölder observables with $\int_{\Lambda_n} v_n\,d\mu_n = 0$. We suppose that $\sup_{n\ge0}\|v_n\|_\eta<\infty$.

Let Δ_n be the corresponding family of Young towers defined as in Section 2.1, with maps $f_n : \Delta_n \to \Delta_n$, invariant probability measures μ_{Δ_n} and semiconjugacies $\pi_{\Delta_n} : \Delta_n \to \Lambda_n$.

Let $\phi_n = v_n \circ \pi_{\Delta_n} : \Delta_n \to \mathbb{R}^d$. By the results from Section 2.2, we have the primary martingale-coboundary decomposition

$$v_n \circ \pi_{\Delta_n} = \phi_n = m_n + \chi_n \circ f_n - \chi_n, \tag{4.1}$$

where m_n, χ_n satisfy the estimates in Propositions 2.4 and 2.7 uniformly in n.

Lemma 4.1 *If $p \le 2$, then $\big|\max_{j\le n}|\sum_{k=0}^{j-1} v_n \circ T_n^k|\big|_p \le C\|v_n\|_\eta n^{1/p}$ and $\big|\max_{j\le n}|\sum_{k=0}^{j-1} m_n \circ f_n^k|\big|_p \le C\|v_n\|_\eta n^{1/p}$ for all $n \ge 0$.*

If $p \ge 2$, then $\big|\max_{j\le n}|\sum_{k=0}^{j-1} v_n \circ T_n^k|\big|_{2(p-1)} \le C\|v_n\|_\eta n^{1/2}$ for all $n \ge 0$.

Proof This is immediate from Corollary 2.10 since the constant C there is independent of n. ∎

From now on, we suppose that $p \ge 2$. By Corollary 2.12, we can define the family of covariance matrices

$$\Sigma_n = \lim_{k\to\infty} k^{-1}\int_{\Lambda_n}\Big(\sum_{j=0}^{k-1} v_n\circ T_n^j\Big)\Big(\sum_{j=0}^{k-1} v_n\circ T_n^j\Big)^T d\mu_n = \int_{\Delta_n} m_n\, m_n^T\, d\mu_{\Delta_n}. \tag{4.2}$$

Remark 4.2 It follows from the proof of Corollary 2.12 that the convergence in (4.2) is uniform in n.

By the results from Section 3, we have the secondary martingale coboundary decomposition

$$U_n L_n(m_n m_n^T) - \Sigma_n = \breve{\phi}_n = \breve{m}_n + \breve{\chi}_n \circ f_n - \breve{\chi}_n, \tag{4.3}$$

where U_n and L_n are the Koopman and transfer operators for f_n.

Proposition 4.3 *The family $\{|m_n|^2,\, n \ge 0\}$ is uniformly integrable.*

Proof We start from the primary decomposition (4.1), with $\phi_n' = m_n' + \chi_n' \circ F_n - \chi_n'$ (as in Section 2.2).

Since $|v_n|_\infty$ is bounded, it is immediate from condition (ii) and the definition $\phi_n' = \sum_{j=0}^{\tau-1} v_n \circ \pi_{\Delta_n} \circ f_n^j$ that $\{|\phi_n'|^2\}$ is uniformly integrable. Next, $|\chi_n'|_\infty$ is bounded and hence $\{|m_n'|^2\}$ is uniformly integrable. It follows from the proof of Proposition 2.4 that the uniform integrability of $\{|m_n'|^2\}$ is inherited by $\{|m_n|^2\}$. ∎

Let $W_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} v_n \circ T_n^j$.

Proposition 4.4 *Suppose that* $\lim_{n\to\infty}\Sigma_n = \Sigma$ *where* $\Sigma \in \mathbb{R}^{d\times d}$*. Then* $W_n \to_{\mu_n} W$ *in* $D([0,\infty),\mathbb{R}^d)$ *where* W *is Brownian motion with covariance* Σ*.*

Proof Define processes $\Phi_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1}\phi_n\circ f_n^j$, $M_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} m_n\circ f_n^j$.

By Proposition 4.3, the family $\{|m_n|^2,\, n\ge 0\}$ is uniformly integrable. Next,

$$n^{-1}\sum_{j=0}^{[nt]-1}\{U_nL_n(m_nm_n^T)\}\circ f_n^j \,-\, t\Sigma = n^{-1}\sum_{j=0}^{[nt]-1}\breve{\phi}_n\circ f_n^j + n^{-1}[nt]\Sigma_n - t\Sigma \to_{\mu_{\Delta_n}} 0,$$

by Corollary 3.2. By Theorem A.1, $M_n \to_{\mu_{\Delta_n}} W$.

Let $T>0$. Since T_n is a uniform family, $\big|\max_{1\le k\le nT}|\chi_n\circ f_n^k - \chi_n|\big|_2 = o(n^{1/2})$ by Proposition 2.7. Also,

$$\begin{aligned}\sup_{t\in[0,T]}|\Phi_n(t)-M_n(t)| &\le n^{-1/2}\max_{1\le k\le nT}|\textstyle\sum_{j=0}^{k-1}(\phi_n\circ f_n^j - m_n\circ f_n^j)|\\ &\le n^{-1/2}\max_{1\le k\le nT}|\chi_n\circ f_n^k-\chi_n|.\end{aligned}$$

Hence $\lim_{n\to\infty}\big|\sup_{t\in[0,T]}|\Phi_n(t)-M_n(t)|\big|_2 = 0$ for each $T>0$. It follows that $\Phi_n \to_{\mu_{\Delta_n}} W$. Also, π_{Δ_n} is a measure-preserving semiconjugacy for each n, so $W_n\to_{\mu_n} W$. ∎

Define $\mathcal{W}\subset D([0,\infty),\mathbb{R}^d)$ to be the set of weak limits of $\{W_n,\, n\ge 0\}$ and let $\mathcal{S}\subset\mathbb{R}^{d\times d}$ be the set of limit points of $\{\Sigma_n,\, n\ge 0\}$. By Proposition 4.3, $\{\Sigma_n,\, n\ge 0\}$ is bounded and hence $\mathcal{S}\ne\emptyset$.

Theorem 4.5 *(i)* $\{W_n,\, n\ge 0\}$ *is tight, (ii)* $W\in\mathcal{W}$ *if and only if* W *is a Brownian motion with covariance matrix in* $\mathcal{S}$*.*

Proof Given a covariance matrix $\Sigma\in\mathbb{R}^{d\times d}$, let $W(\Sigma)$ denote Brownian motion with covariance Σ.

Since $\{\Sigma_n\}$ is bounded, for any subsequence W_{n_k} we can pass to a subsubsequence along which $\Sigma_{n_k}\to\Sigma$ for some $\Sigma\in\mathcal{S}$. By Proposition 4.4, we then have that $W_{n_k}\to_{\mu_{n_k}} W(\Sigma)$. This shows that $\{W_n\}$ is tight and that all weak limits have the form $W(\Sigma)$, $\Sigma\in\mathcal{S}$.

Conversely, if $\lim_{k\to\infty}\Sigma_{n_k}=\Sigma$ for some subsequence n_k, then $W_{n_k}\to_{\mu_n} W(\Sigma)$ by Proposition 4.4. ∎

Corollary 4.6 *Suppose in Theorem 4.5 that* $W_{n_k} \to_{\mu_{n_k}} W$ *as* $k\to\infty$*. Then* $\lim_{k\to\infty} n_k^{-q/2}\int_{\Lambda_{n_k}}|\sum_{j=0}^{n_k-1} v_{n_k}\circ T_{n_k}^j|^q\, d\mu_{n_k} = \mathbb{E}|W(1)|^q$ *for all* $q<2(p-1)$*.*

Proof This follows immediately from Lemma 4.1 (cf. [47, Lemma 2.1(e)]). ∎

Remark 4.7 It is not difficult to formulate conditions under which the weak limits of $\{W_n\}$ are nondegenerate. One possibility is to suppose that there is a limiting nonuniformly expanding map T_∞ with corresponding observable v_∞ and covariance matrix Σ_∞. Typically $\det\Sigma_\infty>0$. Under certain conditions (see for example Section 7), it can be shown that $W_n\to_{\mu_n} W$ where W is Brownian motion with covariance Σ_∞.

An alternative mechanism for nondegenerate limits is the following. Suppose that $d = 1$. Recall that $|\chi'_n|_2 \le |\chi'_n|_\infty \le C\|v_n\|_\eta \le C\|v_n\|_\eta|\tau_n|_1$, where $C > 0$ is a constant depending only on the induced maps F_n. Hence

$$\sigma_n = |m_n|_2 = |m'_n|_2/|\tau_n|_1 \ge (|\phi'_n|_2 - 2|\chi'_n|_2)/|\tau_n|_1 \ge |\phi'_n|_2/|\tau_n|_1 \; - \; C\|v_n\|_\eta.$$

If we arrange that $|\tau_n|_2 \ge 2C|\tau_n|_1$ for all n, then it follows that

$$\sigma_n \ge C(2|\phi'_n|_2/|\tau_n|_2 \; - \; \|v_n\|_\eta).$$

Suppose for simplicity that $F_n = T_n^{\tau_n}$ is the first return map for each n. Let k_n be largest such that $\sum_{j=1}^{k_n} j\mu_{Y_n}(\tau_n = j) > \sum_{j=k_n+1}^{\infty} j\mu_{Y_n}(\tau_n = j)$, There is a unique observable $v_n : \Lambda_n \to \mathbb{R}$ taking values ± 1 such that $\phi'_n(y) = \tau_n(y)$ if $\tau_n(y) \le k_n$ and $\phi'_n(y) = -\tau_n(y)$ if $\tau_n(y) > k_n$. By construction $\int_{Y_n} \phi'_n \, d\mu_{Y_n} \approx 0$ and $|\phi'_n|_2 = |\tau_n|_2$. A slight modification produces $\int_{Y_n} \phi'_n \, d\mu_{Y_n} = 0$ and $|\phi'_n|_2 \approx |\tau_n|_2$ so that $\sigma_n \gtrsim C$. Hence this is a robust mechanism for producing nondegenerate limits in Theorem 4.5.

We end this section with some examples of nonuniformly expanding maps where uniformity of the constants can be verified, and hence to which the results in this section apply.

Example 4.8 Fix a sequence $\lambda_n \in \mathbb{Z}$ such that $\lambda_n \ge 2$ for all $n \ge 0$. Define the family of uniformly expanding maps $T_n : [0,1] \to [0,1]$ given by $T_n x = \lambda_n x \bmod 1$. Clearly T_n is a uniform family of order p for any p, with $Y_n = [0,1]$, $\tau_n = 1$ and $\mu_n =$ Lebesgue.

This example emphasizes that tightness in Theorem 4.5 is unrelated to any accumulation properties of the dynamical systems T_n and is governed purely by accumulation of the bounded set of covariance matrices.

Example 4.9 Fix a sequence $\gamma_n \in (0,1)$ and let $T_n : [0,1] \to [0,1]$ be the corresponding intermittent map defined in Example 3.5. As verified in [37, Example 5.1], T_n is a uniform family of order p for any $p < \sup \gamma_n^{-1}$.

Example 4.10 Consider the family of quadratic maps $T_n : [-1,1] \to [-1,1]$ given by $T_n(x) = 1 - a_n x^2$, $a_n \in [0,2]$. We assume that there exists $b, c > 0$ such that the Collet-Eckmann condition [20] $|(T_n^k)'(1)| \ge ce^{bn}$ holds for all k, $n \ge 0$. (By [32, 12], the set of parameters a_n for which this condition holds has positive Lebesgue measure for b, c sufficiently small.) As verified in [37, Example 5.2] (based on arguments of [25]), T_n is a uniform family of order p for any p. This example generalises to multimodal maps (see [37, Example 5.3].

Example 4.11 Viana [55] introduced a C^3 open class of multi-dimensional nonuniformly expanding maps $T_\epsilon : M \to M$. For definiteness, we restrict attention to the case $M = S^1 \times \mathbb{R}$. Fix $\lambda_n \in \mathbb{Z}$, $\lambda_n \ge 16$, and let $S_n : M \to M$ be the map $S_n(\theta, y) = (\lambda_n \theta \bmod 1, a_0 + a \sin 2\pi\theta - y^2)$. Here a_0 is chosen so that 0 is a preperiodic point for the quadratic map $y \mapsto a_0 - y^2$ and a is fixed sufficiently small. Let T_n be a family of C^3 maps each of which is sufficiently close to S_n. It follows from [1, 8] that there is an interval $I \subset (-2, 2)$ such that, for each $n \ge 0$, there is a unique absolutely continuous T_n-invariant ergodic probability measure μ_n

supported in the interior of $S^1 \times I$. Moreover the invariant set $\Lambda_\epsilon = \operatorname{supp} \mu_\epsilon$ attracts almost every initial condition in $S^1 \times I$.

As verified in [37, Example 5.4] (based on arguments of [2, 5]), $T_n : \Lambda_n \to \Lambda_n$ is a uniform family of nonuniformly expanding maps of order p for any p.

5 Limit laws for families of nonuniformly hyperbolic transformations

In this section, we show the results from Section 4 pass over to the invertible setting. The notion of nonuniformly hyperbolic transformation is recalled in Subsection 5.1. In Subsection 5.2, we recall how to quotient to a nonuniformly expanding map. In Subsection 5.3, we prove limit laws for families of nonuniformly hyperbolic transformations.

5.1 Nonuniformly hyperbolic transformations

Let $T : \Lambda \to \Lambda$ be a diffeomorphism (possibly with singularities) defined on a Riemannian manifold (Λ, d_Λ). We assume that T is nonuniformly hyperbolic in the sense of Young [56, 57]. The precise definitions are somewhat technical; here we are content to focus on the parts necessary for understanding this paper, referring to [56, 57] for further details.

As part of this set up, there is a measurable (with respect to the Riemannian measure) set $Y \subset M$, a measurable partition $\{Y_j\}$ of Y, and an inducing time $\tau : Y \to \mathbb{Z}^+$ constant on partition elements such that $T^{\tau(y)}(y) \in Y$ for all $y \in Y$. We refer to $F = T^\tau : Y \to Y$ as the induced map. The separation time $s(y, y')$ of points $y, y' \in Y$ is the least integer $n \geq 0$ such that $F^n y, F^n y'$ lie in distinct partition elements of Y.

In addition, there exist integers d_s, $d_u \geq 1$ with $d_s + d_u = \dim M$, a measurable partition $\mathcal{W}^s$ of Y consisting of embedded d_s-dimensional disks (called "stable leaves") and an embedded d_u-dimensional disk W^u (called an "unstable leaf") such that W^u intersects each element of $\mathcal{W}^s$ in a single point. If $y \in Y$, the leaf in $\mathcal{W}^s$ that contains y is labelled W^s_y. Let ρ denote the measure on W^u induced by the Riemannian measure.

We assume that there are constants D_0, $D_1 \geq 1$, $\gamma \in (0, 1)$, $p \geq 1$, such that

(A1) Each Y_j is a union of elements of $\mathcal{W}^s$ (in particular, τ is constant on stable leaves), and $F(W^s_y) \subset W^s_{Fy}$ for all $y \in Y$.

(A2) (i) $d_\Lambda(T^j y, T^j y') \leq D_0 \gamma^j$ for all $y \in Y$, $y' \in W^s_y$,

(ii) $d_\Lambda(T^j y, T^j y') \leq D_0 \gamma^{s(y,y') - \psi_j(y)}$ for all $y, y' \in W^u$,

for all $j \geq 0$, where $\psi_j(y) = \#\{k = 0, \ldots, j-1 : T^k y \in Y\}$ is the number of visits of y to Y by time j.

(A3) $\int_Y \tau^p \, d\rho < \infty$.

Let $\bar{Y} = Y/\sim$ where $y \sim y'$ if $y' \in W^s_y$, and let $\bar{\pi} : Y \to \bar{Y}$ denote the natural projection. By (A1), we obtain well-defined functions $\tau : \bar{Y} \to \mathbb{Z}^+$ and $\bar{F} : \bar{Y} \to \bar{Y}$. Let $\bar{\rho} = \bar{\pi}_* \rho$. Let α be the countable partition of $\bar{Y}$ consisting of the partition elements Y_j quotiented by W^s. We assume:

(A4) $\bar{F}$ restricts to a bijection from a onto $\bar{Y}$ for all $a \in \alpha$ and $\zeta_0 = \frac{d\bar{\rho}}{d\bar{\rho}\circ\bar{F}}$ satisfies $|\log\zeta_0(y) - \log\zeta_0(y')| \le D_1\gamma^{s(y,y')}$ for all $y, y' \in a$.

There is a unique absolutely continuous $\bar{F}$-invariant probability measure $\bar{\mu}_Y$ on $\bar{Y}$ and $d\bar{\mu}_Y/d\bar{\rho} \in L^\infty$. By for instance [10, Section 6.1], there is a unique ergodic F-invariant probability measure μ_Y on Y such that $\bar{\pi}_*\mu_Y = \bar{\mu}_Y$.

As in Section 4, we define a tower map $f : \Delta \to \Delta$ with semiconjugacy $\pi_\Delta : \Delta \to \Lambda$ from f to T, and ergodic f-invariant probability measure $\mu_\Delta = \mu_Y \times \text{counting}/\int_Y \tau\, d\mu_Y$. Then $\mu = (\pi_\Delta)_*\mu_\Delta$ is an ergodic T-invariant probability measure on M.

Remark 5.1 For simplicity, we restrict to the case where T contracts exponentially along stable manifolds. It is also possible to consider polynomial (but summable) contraction as in [3], as well as the general situation [48] where contraction and expansion is assumed only on returns to Y. (The arguments to treat this general situation are correspondingly longer.)

Next, we introduce the quotient tower map $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ defined in the same way as $f : \Delta \to \Delta$ but starting from $\bar{F} : \bar{Y} \to \bar{Y}$ instead of $F : Y \to Y$. The projection $\bar{\pi} : Y \to \bar{Y}$ extends to a projection $\bar{\pi} : \Delta \to \bar{\Delta}$, $\bar{\pi}(y, \ell) = (\bar{\pi}y, \ell)$, and we have the ergodic $\bar{f}$-invariant probability measure $\bar{\mu}_\Delta = \bar{\pi}_*\mu_\Delta = \bar{\mu}_Y \times \text{counting}/\int_{\bar{Y}} \tau\, d\bar{\mu}_Y$.

The separation time s on Y projects to a separation time on $\bar{Y}$. For $\theta \in (0, 1)$ we define the symbolic metric d_θ on $\bar{Y}$, setting $d_\theta(y, y') = \theta^{s(y,y')}$. This extends to a metric on $\bar{\Delta}$, where $d_\theta((y, \ell), (y', \ell')) = \begin{cases} d_\theta(y, y') & \ell = \ell' \\ 1 & \ell \ne \ell' \end{cases}$.

Proposition 5.2 *Choose $\theta \in [\gamma, 1)$. Then $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ is a nonuniformly expanding map on the metric space $(\bar{\Delta}, d_\theta)$ with induced map $\bar{F} : \bar{Y} \to \bar{Y}$, partition α, and constants $\lambda > 1$, $\eta \in (0, 1]$, $C_0, C_1 \ge 1$ given by $\lambda = \theta^{-1}$, $\eta = C_0 = 1$, $C_1 = D_1$.*

Proof By (A4), $\bar{F}$ maps partition elements bijectively onto $\bar{Y}$. By definition of d_θ, if $y, y' \in a$, $a \in \alpha$, then $d_\theta(\bar{F}y, \bar{F}y') = \theta^{-1}d_\theta(y, y')$ and $d_\theta(\bar{f}^\ell y, \bar{f}^\ell y') = d_\theta(y, y') \le d_\theta(\bar{F}y, \bar{F}y')$ for all $0 \le \ell < \tau(a)$. Finally, by (A4), $|\log\zeta_0(x) - \log\zeta_0(y)| \le D_1\gamma^{s(y,y')} \le D_1 d_\theta(\bar{F}y, \bar{F}y')$. ∎

5.2 Quotienting step

In this subsection, we recall a standard procedure for reducing limit laws for nonuniformly hyperbolic transformations to the noninvertible (nonuniformly expanding) setting. We work throughout with Hölder observables $v \in C^\eta$, where $\eta \in (0, 1]$ is fixed.

First, define a projection $Y \to W^u$ by setting $\hat{y} = W^s_y \cap W^u$ for $y \in Y$. This extends to a projection on Δ by setting $\hat{p} = (\hat{y}, \ell)$ for $p = (y, \ell) \in \Delta$.

Given $v : \Lambda \to \mathbb{R}^d$ Hölder, define $\psi : \Delta \to \mathbb{R}^d$,

$$\psi(p) = \textstyle\sum_{j=0}^\infty \{v \circ \pi_\Delta(f^j p) - v \circ \pi_\Delta(f^j \hat{p})\}.$$

Proposition 5.3 *Let $\theta \in [\gamma^\eta, 1)$. Then $|\psi|_\infty \le D_0^\eta(1-\theta)^{-1}|v|_\eta$, for all Hölder $v : \Lambda \to \mathbb{R}^d$.*

Proof Let $p = (y, \ell) \in \Delta$. Then $\pi_\Delta(f^j p) = T^{j+\ell} y \in T^{j+\ell} W^s_y$ and $\pi_\Delta(f^j \hat{p}) = T^{j+\ell}\hat{y} \in T^{j+\ell} W^s_{\hat{y}} = T^{j+\ell} W^s_y$. In particular, $d_\Lambda(\pi_\Delta(f^j p), \pi_\Delta(f^j \hat{p})) \le D_0 \gamma^{j+\ell} \le D_0\gamma^j$ by (A2)(i). Hence $|\psi(p)| \le |v|_\eta \sum_{j=0}^\infty d_\Lambda(\pi_\Delta(f^j p), \pi_\Delta(f^j\hat{p}))^\eta \le D_0^\eta (1-\theta)^{-1}|v|_\eta$. ∎

A calculation shows that $v \circ \pi_\Delta = \hat{v} + \psi - \psi \circ f$, where $\hat{v} \in L^\infty(\Delta)$ is given by

$$\hat{v}(p) = v \circ \pi_\Delta(\hat{p}) + \textstyle\sum_{j=0}^\infty \{ v \circ \pi_\Delta(f^{j+1}\hat{p}) - v \circ \pi_\Delta(f^j \widehat{fp})\}.$$

Note that $\hat{v}$ is constant along fibres $\bar{\pi}^{-1}(y)$, $y \in \bar{Y}$. Hence we can write $\hat{v} = \bar{v} \circ \bar{\pi}$ where $\bar{v} : \bar{\Delta} \to \mathbb{R}^d$. The observables $v : \Lambda \to \mathbb{R}^d$ and $\bar{v} : \bar{\Delta} \to \mathbb{R}^d$ are related by the equation

$$v \circ \pi_\Delta = \bar{v} \circ \bar{\pi} + \psi - \psi \circ f. \tag{5.1}$$

Let $\|\bar{v}\|_\theta = |\bar{v}|_\infty + |\bar{v}|_\theta$ where $|\bar{v}|_\theta$ denotes the d_θ-Lipschitz constant of $\bar{v}$.

Proposition 5.4 *Let $\theta \in [\gamma^{\eta/2}, 1)$. Then $\|\bar{v}\|_\theta \le 6D_0^\eta \theta^{-2}(1-\theta^2)^{-1}\|v\|_\eta$, for all Hölder $v : \Lambda \to \mathbb{R}^d$.*

Proof By Proposition 5.3, $|\bar{v}|_\infty \le 2D_0^\eta(1-\theta)^{-1}\|v\|_\eta$.

Next, let $p = (y, \ell)$, $q = (z, \ell) \in \Delta$. Write $|\hat{v}(p) - \hat{v}(q)| \le A_1 + A_2 + A_3 + A_4$ where

$$A_1 = \sum_{j=N}^\infty |v \circ \pi_\Delta(f^{j+1}\hat{p}) - v \circ \pi_\Delta(f^j \widehat{fp})|, \quad A_2 = \sum_{j=N}^\infty |v \circ \pi_\Delta(f^{j+1}\hat{q}) - v \circ \pi_\Delta(f^j \widehat{fq})|,$$

$$A_3 = \sum_{j=0}^N |v \circ \pi_\Delta(f^j\hat{p}) - v \circ \pi_\Delta(f^j \hat{q})|, \qquad A_4 = \sum_{j=0}^{N-1} |v \circ \pi_\Delta(f^j\widehat{fp}) - v \circ \pi_\Delta(f^j \widehat{fq})|.$$

We take $N = [\frac12 s(y,z)]$ and show that $A_j \le D_0^\eta \theta^{-2}(1-\theta^2)^{-1}|v|_\eta\, \theta^{s(y,z)}$ for $j = 1, \dots, 4$. Hence $|\hat{v}|_\theta \le 4D_0^\eta\theta^{-2}(1-\theta^2)^{-1}|v|_\eta d_\theta(p,q)$ and the proof is complete.

First, we estimate A_1. If $\ell < \tau(y) - 2$ then $f\hat{p} = \widehat{fp} = (\hat{y}, \ell+1)$ and $A_1 = 0$. Otherwise $f\hat{p} = (F\hat{y}, 0)$ and $\widehat{fp} = (\widehat{Fy}, 0)$. In particular, $\pi_\Delta(f\hat{p})$, $\pi_\Delta(\widehat{fp}) \in W^s_{Fy}$, so it follows from (A2)(i) that

$$d_\Lambda(\pi_\Delta(f^{j+1}\hat{p}), \pi_\Delta(f^j\widehat{fp})) = d_\Lambda(T^j(\pi_\Delta(f\hat{p})), T^j(\pi_\Delta(\widehat{fp}))) \le D_0\gamma^j.$$

Hence $A_1 \le |v|_\eta D_0^\eta \sum_{j=N}^\infty \theta^{2j} \le D_0^\eta(1-\theta^2)^{-1}\theta^{2N}|v|_\eta \le D_0^\eta\theta^{-2}(1-\theta^2)^{-1}|v|_\eta\,\theta^{s(y,z)}$. Similarly $A_2 \le D_0^\eta\theta^{-2}(1-\theta^2)^{-1}|v|_\eta\,\theta^{s(y,z)}$.

Next, we estimate A_3. For each j, we have $\pi_\Delta(f^j\hat{p}) = T^{j+\ell}\hat{y} = T^L F^J \hat{y}$ where $0 \le J \le j$ and $0 \le L < \tau(F^J\hat{y})$. Note that $F^j y$ and $F^j z$ lie in the same partition element of Y for all $j \le N$, so $\pi_\Delta(f^j\hat{q}) = T^{j+\ell}\hat{z} = T^L F^J\hat{z}$. By (A2)(ii),

$$d_\Lambda(\pi_\Delta(f^j\hat{p}), \pi_\Delta(f^j\hat{q})) = d_\Lambda(T^LF^J\hat{y}, T^LF^J\hat{z}) \le D_0\gamma^{s(y,z)-J} \le D_0\gamma^{s(y,z)-j}.$$

Hence $A_3 \le D_0^\eta(1-\theta^2)^{-1}|v|_\eta\theta^{2(s(y,z)-N)} \le D_0^\eta(1-\theta^2)^{-1}|v|_\eta\theta^{s(y,z)}$. Similarly, $A_4 \le D_0^\eta(1-\theta^2)^{-1}|v|_\eta\theta^{s(y,z)}$. ∎

Corollary 5.5 *Suppose that $v : \Lambda \to \mathbb{R}^d$ is Hölder with $\int_\Lambda v\, d\mu = 0$. Let $\bar{v} : \bar{\Delta} \to \mathbb{R}^d$ be the corresponding d_θ-Lipschitz observable. Then*

(a) $\left|\max_{j\le n}|\sum_{k=0}^{j-1} v\circ T^k|\right|_{L^{p^*}(\mu)} \le C\|v\|_\eta\, n^{\max\{\frac12,\frac1p\}}$ *for all* $n\ge 1$*, where* $C\ge 1$ *is a constant that depends continuously on* D_0, D_1, γ*, and* $p^*=\max\{p,2(p-1)\}$*.*

(b) *Suppose that* $p\ge 2$*. Then the limits*

$$\Sigma=\lim_{n\to\infty} n^{-1}\int_\Lambda S_nv\, S_nv^T\,d\mu=\lim_{n\to\infty} n^{-1}\int_{\bar\Delta} S_n\bar v\, S_n\bar v^T\,d\bar\mu_\Delta,$$

exist and coincide, where $S_nv=\sum_{j=0}^{n-1} v\circ T^j$*,* $S_n\bar v=\sum_{j=0}^{n-1}\bar v\circ \bar f^j$*.*

Proof Define $S_n(v\circ\pi_\Delta)$ and $S_n(\bar v\circ\bar\pi)$ similarly.

By Propositions 5.2 and 5.4, $\bar v$ is a mean zero Lipschitz observable for the nonuniformly expanding map $\bar f:\bar\Delta\to\bar\Delta$ on the metric space $(\bar\Delta,d_\theta)$. Moreover, $\|\bar v\|_\theta\ll\|v\|_\eta$. Hence, by Corollary 2.10, $\left|\max_{j\le n}|S_j\bar v|\right|_{L^{p^*}(\bar\mu_\Delta)}\ll\|v\|_\eta\, n^{\max\{\frac12,\frac1p\}}$.

By (5.1) and Proposition 5.3, and using that $\pi_\Delta:\Delta\to\Lambda$ and $\bar\pi:\Delta\to\bar\Delta$ are measure-preserving,

$$\begin{aligned}\Big|\max_{j\le n}|S_jv|\Big|_{L^{p^*}(\mu)}&=\Big|\max_{j\le n}|S_j(v\circ\pi_\Delta)|\Big|_{L^{p^*}(\mu_\Delta)}\le\Big|\max_{j\le n}|S_j(\bar v\circ\bar\pi)|\Big|_{L^{p^*}(\mu_\Delta)}+2|\psi|_{L^{p^*}(\mu_\Delta)}\\&=\Big|\max_{j\le n}|S_j\bar v|\Big|_{L^{p^*}(\bar\mu_\Delta)}+2|\psi|_{L^{p^*}(\mu_\Delta)}\ll\|v\|_\eta\, n^{\max\{\frac12,\frac1p\}},\end{aligned}$$

proving part (a).

Next, by Corollary 2.12, $n^{-1}\int_{\bar\Delta}S_n\bar v\,S_n\bar v^T d\bar\mu_\Delta$ converges. Also,

$$\begin{aligned}&\Big|\int_\Lambda S_nv\,S_nv^T d\mu-\int_{\bar\Delta}S_n\bar v\,S_n\bar v^T d\bar\mu_\Delta\Big|\\&\quad=\Big|\int_\Delta\big(S_n(v\circ\pi_\Delta)\,S_n(v\circ\pi_\Delta)^T-S_n(\bar v\circ\bar\pi)\,S_n(\bar v\circ\bar\pi)^T\big)\,d\mu_\Delta\Big|\\&\quad\le|S_n(v\circ\pi_\Delta)-S_n(\bar v\circ\bar\pi)|_{L^2(\mu_\Delta)}\big(|S_n(v\circ\pi_\Delta)|_{L^2(\mu_\Delta)}+|S_n(\bar v\circ\bar\pi)|_{L^2(\mu_\Delta)}\big)\\&\quad\le 2|\psi|_{L^2(\mu_\Delta)}\big(|S_n(v\circ\pi_\Delta)|_{L^2(\mu_\Delta)}+|S_n(\bar v\circ\bar\pi)|_{L^2(\mu_\Delta)}\big)\ll|v|_\eta\|v\|_\eta n^{1/2},\end{aligned}$$

by Proposition 5.3 and the estimates in the proof of part (a). Part (b) follows. ∎

5.3 Families of nonuniformly hyperbolic transformations

Suppose that $T_n:\Lambda_n\to\Lambda_n$, $n\ge 0$, is a family of nonuniformly hyperbolic transformations with induced maps $F_n=T_n^{\tau_n}:Y_n\to Y_n$. Let $p\ge 1$. We say that this is a *uniform family of order* p if

(i) The constants $D_0, D_1\ge 1$, $\gamma\in(0,1)$ can be chosen independent of $n\ge 0$.
(ii) The family $\{\tau_n^p,\ n\ge 0\}$ is uniformly integrable.

Let $v_n:\Lambda\to\mathbb{R}^d$ be a family of Hölder observables with $\int_{\Lambda_n} v_n\,d\mu_n=0$. We suppose that $\sup_{n\ge0}\|v_n\|_\eta<\infty$.

Proposition 5.6 $\left|\max_{j\le n}|\sum_{k=0}^{j-1}v_n\circ T_n^k|\right|_{L^{p^*}(\mu_n)}\le C\|v_n\|_\eta\, n^{\max\{\frac12,\frac1p\}}$ *for all* $n\ge 0$*, where* $p^*=\max\{p,2(p-1)\}$*.*

Proof This is immediate from Corollary 5.5(a). ∎

Proceeding as in Subsection 5.1, we construct metric spaces $(\bar\Delta_n, d_{\theta,n})$ and families $\psi_n : \Delta_n \to \mathbb{R}^d$, $\bar v_n : \bar\Delta_n \to \mathbb{R}^d$ where $|\psi_n|_\infty \ll |v_n|_\eta$, $\|\bar v_n\|_{\theta,n} \ll \|v_n\|_\eta$ such that

$$v_n \circ \pi_{\Delta_n} = \bar v_n \circ \bar\pi_n + \psi_n - \psi_n \circ f_n.$$

By Corollary 5.5(b), for $p \geq 2$ we can define the family of covariance matrices

$$\begin{aligned}\Sigma_n &= \lim_{k\to\infty} k^{-1} \int_{\Lambda_n} \big(\textstyle\sum_{j=0}^{k-1} v_n \circ T_n^j\big)\big(\sum_{j=0}^{k-1} v_n \circ T_n^j\big)^T \, d\mu_n,\\ &= \lim_{k\to\infty} k^{-1} \int_{\bar\Delta_n} \big(\textstyle\sum_{j=0}^{k-1} \bar v_n \circ \bar f_n^j\big)\big(\sum_{j=0}^{k-1} \bar v_n \circ \bar f_n^j\big)^T \, d\bar\mu_{\Delta_n}.\end{aligned}$$

Let $W_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} v_n \circ T_n^j$. Define $\mathcal{W} \subset D([0,\infty),\mathbb{R}^d)$ to be the set of weak limits of $\{W_n,\, n \geq 0\}$ and let $\mathcal{S} \subset \mathbb{R}^{d\times d}$ be the set of limit points of $\{\Sigma_n,\, n \geq 0\}$.

Theorem 5.7 *Suppose that $p \geq 2$. Then (i) $\{W_n,\, n \geq 0\}$ is tight, (ii) $W \in \mathcal{W}$ if and only if W is a Brownian motion with covariance matrix in $\mathcal{S}$.*

In particular, if $\lim_{n\to\infty}\Sigma_n = \Sigma \in \mathbb{R}^{d\times d}$, then $W_n \to_w W$ where W is Brownian motion with covariance Σ.

Proof Define the process $\overline{W}_n(t) = n^{-1/2}\sum_{j=0}^{[nt]-1} \bar v_n \circ \bar f_n^j$ on $(\bar\Delta_n, \bar\mu_{\Delta_n})$.

By Proposition 5.2, $\bar f_n : \bar\Delta_n \to \bar\Delta_n$ is a uniform family of nonuniformly expanding maps. By Proposition 5.4, $\bar v_n$ is a family of mean zero Lipschitz observables satisfying $\|\bar v_n\|_{\theta,n} \ll \|v_n\|_\eta$. Hence Theorem 4.5 characterises the weak limits of $\{\overline{W}_n,\, n \geq 0\}$.

It remains to show that the weak limits of $\{W_n,\, n \geq 0\}$ coincide with those of $\{\overline{W}_n,\, n \geq 0\}$. Since π_{Δ_n} and $\bar\pi_n$ are measure-preserving semiconjugacies, the weak limits of $\{W_n\}$ coincide with those of $\{W_n \circ \pi_{\Delta_n}\}$, and the weak limits of $\{\overline{W}_n\}$ coincide with those of $\{\overline{W}_n \circ \bar\pi_n\}$. Also

$$\sup_{[0,T]} |W_n \circ \pi_{\Delta_n} - \overline{W}_n \circ \bar\pi_n|_\infty \leq 2n^{-1/2}|\psi_n|_\infty \ll n^{-1/2}|v_n|_\eta,$$

so $W_n \circ \pi_{\Delta_n} - \overline{W}_n \circ \bar\pi_n \to_{\mu_{\Delta_n}} 0$ completing the proof. ∎

Example 5.8 The classical solenoid construction of Smale & Williams can be used as in [6] to construct nonuniformly hyperbolic families from each of the nonuniformly expanding families in the examples in Section 4. It is immediate that our results apply to such families.

Example 5.9 Collision maps for dispersing billiards under small external forces are nonuniformly hyperbolic with uniform constants for all p by [18]. (See the proof of [18, Proposition 6.4] where uniformity of constants is mentioned explicitly.) Hence the results in this section apply to such examples.

6 An abstract homogenization theorem

In [27], a homogenization theorem was proved for fast-slow systems of the form

$$x_\epsilon(n+1) = x_\epsilon(n) + \epsilon^2 a(x_\epsilon(n), y(n)) + \epsilon b(x_\epsilon(n))\, v(y(n)),$$

where the fast dynamics $y(n+1) = Ty(n)$ is generated by an ergodic transformation $T : \Lambda \to \Lambda$ and the slow variables $x(n)$ lie in $\mathbb{R}^d$. The main assumptions are that either $d = 1$ or $b \equiv I_d$ (or more generally that b is exact, see below), and that $v : \Lambda \to \mathbb{R}^d$ is mean zero and satisfies the WIP for T. The corresponding result for flows was obtained in [45].

In this section, we show how to generalise to the case where the single map T generating the fast dynamics is replaced by a family of maps. As in [27, 45], the setting is completely abstract, with no hyperbolicity assumptions on the fast dynamics.

Let $T_\epsilon : M \to M$, $\epsilon \in [0, \epsilon_0)$, be a family of maps defined on a topological space M, with T_ϵ-invariant Borel probability measures μ_ϵ. Consider the family of fast-slow equations

$$x_\epsilon(n+1) = x_\epsilon(n) + \epsilon^2 a_\epsilon(x_\epsilon(n), y_\epsilon(n)) + \epsilon b_\epsilon(x_\epsilon(n))\, v_\epsilon(y_\epsilon(n)), \quad x_\epsilon(0) = \xi_\epsilon, \tag{6.1}$$

where the fast dynamics is given by $y_\epsilon(n+1) = T_\epsilon\, y_\epsilon(n)$, and $\xi_\epsilon \in \mathbb{R}^d$ is the initial condition for the slow dynamics. The maps $v_\epsilon : M \to \mathbb{R}^d$, $a_\epsilon : \mathbb{R}^d \times M \to \mathbb{R}^d$ and $b_\epsilon : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ are defined and continuous for each $\epsilon \in [0, \epsilon_0)$. Moreover, $\int_M v_\epsilon \, d\mu_\epsilon = 0$.

Regularity assumptions We suppose that there is a constant $L \geq 1$ such that

$$|v_\epsilon|_\infty \leq L, \quad |a_\epsilon|_\infty \leq L, \quad \operatorname{Lip} a_\epsilon = \sup_{x \neq x'} \sup_y \frac{|a_\epsilon(x,y) - a_\epsilon(x',y)|}{|x - x'|} \leq L,$$

for all $\epsilon \in [0, \epsilon_0)$. Also, we assume that

$$\lim_{\epsilon \to 0} \sup_{x,y} |a_\epsilon(x,y) - a_0(x,y)| = 0, \qquad \lim_{\epsilon \to 0} \sup_y |v_\epsilon(y) - v_0(y)| = 0, \qquad \lim_{\epsilon \to 0} \xi_\epsilon = \xi_0.$$

Exactness We suppose that the multiplicative noise $b_\epsilon : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is *exact*. That is, $b_\epsilon(x) = [(dh_\epsilon)_x]^{-1}$ where $h_\epsilon : \mathbb{R}^d \to \mathbb{R}^d$ is a continuous family of C^3 diffeomorphisms $h_\epsilon : \mathbb{R}^d \to \mathbb{R}^d$ (with C^3 norm uniform in ϵ and $x \in \mathbb{R}^d$).

Remark 6.1 (a) The exactness assumption on b_ϵ can be removed, but then additional assumptions are required on the fast dynamics. When the fast dynamics T_ϵ is independent of ϵ, the corresponding result without exactness is proved for partially hyperbolic dynamics using standard pairs/martingale problems [24], and for nonuniformly hyperbolic dynamics using rough path theory [33, 34].
(b) The global nature of the regularity assumptions in $x \in \mathbb{R}^d$ is easily relaxed, see for example [27, Section 3.1].

Define $\hat{x}_\epsilon(t) = x_\epsilon([t\epsilon^{-2}])$, $t \geq 0$. We are interested in weak convergence of $\hat{x}_\epsilon$ to a solution X of an SDE. Convergence is in the space $D([0,\infty), \mathbb{R}^d)$ of càdlàg functions (right-continuous functions with left-hand limits, see for example [13, Chapter 3]) with the supremum norm.

Remark 6.2 One technical issue is that $D([0,\infty), \mathbb{R}^d)$ is not separable in the supremum norm. Since our limit processes have continuous sample paths, convergence in the supremum norm is equivalent to convergence in the standard Skorokhod topology which is metrizable and separable. Hence whenever we apply results where separability is required, we can momentarily work in this topology.

Dynamical assumptions So far, the assumptions on the fast-slow system (6.1) have been standard subject to the comments in Remark 6.1. Now we introduce mild assumptions on the fast dynamics that suffice for homogenization.

Define $\alpha_x : M \to \mathbb{R}^d$ for $x \in \mathbb{R}^d$,

$$\alpha_x(y) = a_0(x,y) - \tfrac{1}{2}\{(db_0)_u b_0(x)v_0(y)\}v_0(y). \tag{6.2}$$

Also, define the family of random elements $W_\epsilon : (M,\mu_\epsilon) \to D([0,\infty),\mathbb{R}^d)$,

$$W_\epsilon(t) = \epsilon \sum_{j=0}^{[t/\epsilon^2]-1} v_\epsilon \circ T_\epsilon^j.$$

Uniform mean ergodicity (UME) There exists $P : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\lim_{\epsilon\to 0}\int_M \Big|\epsilon^{1/2}\sum_{j=0}^{[\epsilon^{-1/2}]-1} \alpha_x \circ T_\epsilon^j \;-P(x)\Big|\, d\mu_\epsilon = 0 \quad \text{for all } x \in \mathbb{R}^d.$$

Weak invariance principle (WIP) $W_\epsilon \to_{\mu_\epsilon} W$ in $D([0,\infty),\mathbb{R}^d)$ as $\epsilon \to 0$, where W is Brownian motion with some covariance matrix $\Sigma \in \mathbb{R}^{d\times d}$.

Theorem 6.3 *Assume (UME) and (WIP). Then P is Lipschitz and $\hat{x}_\epsilon \;\to_{\mu_\epsilon}\; X$ in $D([0,\infty),\mathbb{R}^d)$ as $\epsilon \to 0$ where X is the solution to the Stratonovich SDE*

$$dX = P(X)\,dt + b_0(X) \circ dW, \quad X(0) = \xi. \tag{6.3}$$

Remark 6.4 Suppose that

(a) $\mu_\epsilon \to_w \mu_0$ as $\epsilon \to 0$ (statistical stability).

(b) $\displaystyle\lim_{\epsilon\to 0}\int_M \Big|\epsilon^{1/2}\sum_{j=0}^{[\epsilon^{-1/2}]-1} \alpha_x \circ T_\epsilon^j \;-\; \int_M \alpha_x\, d\mu_\epsilon\Big|\, d\mu_\epsilon = 0$ for all $x \in \mathbb{R}^d$.

Then (UME) holds, and

$$\begin{aligned} P(x) &= \textstyle\int_M a_0(x,y)\,d\mu_0(y) - \tfrac{1}{2}\int_M \{(db_0)_x b_0(x)v_0(y)\}v_0(y)\,d\mu_0(y) \\ &= \textstyle\int_M a_0(x,y)\,d\mu_0(y) - \tfrac{1}{2}\sum b_0^{\alpha\gamma}(x)(\partial_{x_\alpha} b_0^\beta)(x)\int_M v_0^\beta v_0^\gamma\, d\mu_0. \end{aligned}$$

Here, $b_0^{\alpha\gamma}$ denotes the (α,γ)'th entry of b_0 and b_0^β denotes the β'th column of b_0, while the summation is over indices $\alpha,\beta,\gamma = 1,\dots,d$.

Remark 6.5 (a) We focus attention on weak convergence with respect to the family of invariant measures μ_ϵ. If we assume *strong statistical stability* (so μ_ϵ is absolutely continuous with respect to a reference measure ρ for all ϵ and $d\mu_\epsilon/d\rho \to d\mu_0/d\rho$ in L^1), then it is immediate from Theorem 6.3 that $x_\epsilon \to_{\mu_0} X$. We will return to the issue of weak convergence with respect to a wider range of measures in subsequent work.

(b) The WIP is a necessary condition for Theorem 6.3 since it is equivalent to the case $a_\epsilon \equiv 0$, $b_\epsilon \equiv I_d$.

The remainder of this section is concerned with the proof of Theorem 6.3. We deal first with the special case $b_\epsilon \equiv I_d$, and then with the general case.

The special case $b_\epsilon \equiv I_d$ We extend the arguments in [27, 45] developed for the situation where T_ϵ is independent of ϵ.

Lemma 6.6 *Theorem 6.3 holds in the case $b_\epsilon \equiv I_d$.*

Proof Note that $\alpha_x(y) = a_0(x,y)$ is uniformly Lipschitz in x and hence that P is Lipschitz with constant L. Define $\tilde{a}(x,y) = a_0(x,y) - P(x)$.

To prove weak convergence in $D([0,\infty),\mathbb{R}^d)$, it suffices to prove convergence in $D([0,T],\mathbb{R}^d)$ for each fixed $T \ge 1$. Let $\delta(\epsilon) = \sup_{x,y}|a_\epsilon(x,y) - a_0(x,y)| + |\xi_\epsilon - \xi|$, so $\lim_{\epsilon\to 0}\delta(\epsilon) = 0$. Write

$$x_\epsilon(n) = \xi_\epsilon + \epsilon^2 \sum_{j=0}^{n-1} a_\epsilon(x_\epsilon(j), y_\epsilon(j)) + \epsilon \sum_{j=0}^{n-1} v_\epsilon(y_\epsilon(j)).$$

Hence

$$\begin{aligned}
\hat{x}_\epsilon(t) &= \xi + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} a_0(x_\epsilon(j), y_\epsilon(j)) + W_\epsilon(t) + Z_{\epsilon,1}(t) \\
&= \xi + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} P(\hat{x}_\epsilon(\epsilon^2 j)) + W_\epsilon(t) + Z_{\epsilon,1}(t) + Z_{\epsilon,2}(t),
\end{aligned}$$

where $W_\epsilon(t) = \epsilon \sum_{j=0}^{[t\epsilon^{-2}]-1} v_\epsilon(y_\epsilon(j))$ satisfies the WIP and

$$|Z_{\epsilon,1}(t)| \le T\delta(\epsilon), \quad Z_{\epsilon,2}(t) = \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} \tilde{a}(x_\epsilon(j), y_\epsilon(j)).$$

For t an integer multiple of ϵ^2, the term $\epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} P(\hat{x}_\epsilon(\epsilon^2 j))$ is the Riemann sum of a piecewise constant function and is precisely $\int_0^t P(\hat{x}_\epsilon(s))\,ds$. For general t,

$$\epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} P(\hat{x}_\epsilon(\epsilon^2 j)) = \int_0^t P(\hat{x}_\epsilon(s))\,ds + Z_{\epsilon,3}(t),$$

where $|Z_{\epsilon,3}(t)| \le \epsilon^2 |P|_\infty \le \epsilon^2 L$. Altogether,

$$\hat{x}_\epsilon(t) = \xi + \textstyle\int_0^t P(\hat{x}_\epsilon(s))\,ds + W_\epsilon(t) + Z_\epsilon(t),$$

where $Z_\epsilon = Z_{\epsilon,1} + Z_{\epsilon,2} + Z_{\epsilon,3}$.

We show below that $Z_{\epsilon,2} \to_{\mu_\epsilon} 0$ in $D([0,T],\mathbb{R}^d)$. It follows that $W_\epsilon + Z_\epsilon \to_{\mu_\epsilon} W$ in $D([0,T],\mathbb{R}^d)$.

Now consider the continuous map $\mathcal{G} : D([0,T],\mathbb{R}^d) \to D([0,T],\mathbb{R}^d)$ given by $\mathcal{G}(u) = z$ where z is the unique solution to the integral equation

$$z(t) = \xi + u(t) + \textstyle\int_0^t P(z(s))\,ds.$$

Define $z_\epsilon = \mathcal{G}(W_\epsilon + Z_\epsilon)$. Since continuous maps preserve weak convergence, it follows that $z_\epsilon \to_{\mu_\epsilon} \mathcal{G}(W)$. But $z_\epsilon = x_\epsilon$ by uniqueness of solutions, so $x_\epsilon \to_{\mu_\epsilon} \mathcal{G}(W)$. The result follows since $X = \mathcal{G}(W)$ satisfies the SDE $dX = P(X)\,dt + dW$, $X(0) = \xi$.

It remains to show that $Z_{\epsilon,2} \to_{\mu_\epsilon} 0$ in $D([0,T],\mathbb{R}^d)$. Note that $|\tilde{a}|_\infty \le 2L$ and $\operatorname{Lip}\tilde{a} \le 2L$. Let $N = [t\epsilon^{-3/2}]$ and write $Z_{\epsilon,2} = Y_\epsilon + I_0$, where

$$Y_\epsilon(t) = \epsilon^2 \sum_{0\le j<N\epsilon^{-1/2}} \tilde{a}(x_\epsilon(j), y_\epsilon(j)), \qquad I_0(t) = \epsilon^2 \sum_{N\epsilon^{-1/2}\le j\le [t\epsilon^{-2}]-1} \tilde{a}(x_\epsilon(j), y_\epsilon(j)).$$

We have

$$|I_0(t)| \le \epsilon^{3/2}|\tilde{a}|_\infty \le 2L\epsilon^{3/2}. \tag{6.4}$$

We now estimate Y_ϵ as follows:

$$\begin{aligned}
Y_\epsilon(t) &= \epsilon^2 \sum_{n=0}^{N-1} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \tilde{a}(x_\epsilon(j), y_\epsilon(j)) = I_1 + I_2 \\
I_1 &= \epsilon^2 \sum_{n=0}^{N-1} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \big(\tilde{a}(x_\epsilon(j), y_\epsilon(j)) - \tilde{a}(x_\epsilon(n\epsilon^{-1/2}), y_\epsilon(j))\big) \\
I_2 &= \epsilon^2 \sum_{n=0}^{N-1} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \tilde{a}(x_\epsilon(n\epsilon^{-1/2}), y_\epsilon(j)).
\end{aligned}$$

For $n\epsilon^{-1/2} \le j < (n+1)\epsilon^{-1/2}$, we have $|x_\epsilon(j) - x_\epsilon(n\epsilon^{-1/2})| \le (|a_\epsilon|_\infty + |v_\epsilon|_\infty)\epsilon^{1/2} \le 2L\epsilon^{1/2}$. Hence

$$|I_1| \le N\epsilon^{3/2}\operatorname{Lip}\tilde{a}\, 2L\epsilon^{1/2} \le 4L^2T\epsilon^{1/2}. \tag{6.5}$$

Next, $I_2 = \epsilon^{3/2}\sum_{n=0}^{N-1} J_n$, where

$$J_n = \epsilon^{1/2} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \tilde{a}(x_\epsilon(n\epsilon^{-1/2}), y_\epsilon(j)).$$

Hence

$$|I_2| \le \epsilon^{3/2} \sum_{n=0}^{[T\epsilon^{-3/2}]-1} |J_n|. \tag{6.6}$$

For $u \in \mathbb{R}^d$ fixed, define

$$\tilde{J}_n(u) = \epsilon^{1/2} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \tilde{a}(u, y_\epsilon(j)) = \epsilon^{1/2} \sum_{n\epsilon^{-1/2}\le j<(n+1)\epsilon^{-1/2}} \alpha_u \circ T_\epsilon^j \; - \; P(u).$$

Note that $\tilde{J}_0$ has $[\epsilon^{-1/2}]$ terms, and $\tilde{J}_n(u)$ has at most one term more or one term less than $\tilde{J}_0(u)$. Hence

$$\textstyle\int_M |\tilde{J}_n(u)|\,d\mu_\epsilon = \int_M |\tilde{J}_0(u)|\,d\mu_\epsilon + E_n(u), \quad \text{where} \quad |E_n(u)| \le \epsilon^{1/2}|\tilde{a}|_\infty \le 2L\epsilon^{1/2}.$$

Let $Q > 0$ and write $I_2 = K_{Q,1} + K_{Q,2}$ where

$$K_{Q,1} = I_2 1_{B_\epsilon(Q)}, \quad K_{Q,2} = I_2 1_{B_\epsilon(Q)^c}, \quad B_\epsilon(Q) = \big\{\max_{[0,T]} |x_\epsilon| \le Q\big\}.$$

For any $\sigma > 0$, there exists a finite subset $S \subset \mathbb{R}^d$ such that $\mathrm{dist}(x, S) \le \sigma/(2L)$ for any x with $|x| \le Q$. Then $1_{B_\epsilon(Q)}|J_n| \le \sum_{u\in S} |\tilde{J}_n(u)| + \sigma$ for all $n \ge 0$, $\epsilon > 0$, Hence by (6.6),

$$\begin{aligned}
\int_M \max_{[0,T]} |K_{Q,1}|\, d\mu_\epsilon &\le \epsilon^{3/2} \sum_{n=0}^{[T\epsilon^{-3/2}]-1} \sum_{u\in S} \int_M |\tilde{J}_n(u)|\, d\mu_\epsilon + T\sigma \\
&= \epsilon^{3/2} \sum_{n=0}^{[T\epsilon^{-3/2}]-1} \sum_{u\in S} \Big(\int_M |\tilde{J}_0(u)|\, d\mu_\epsilon + E_n(u)\Big) + T\sigma \\
&\le T \sum_{u\in S} \int_M |\tilde{J}_0(u)|\, d\mu_\epsilon + 2\epsilon^{1/2} T|S|L + T\sigma.
\end{aligned}$$

By (UME), $\int_M |\tilde{J}_0(u)|\, d\mu_\epsilon \to 0$ as $\epsilon \to 0$ for each u. Since $\sigma > 0$ is arbitrary, we obtain for each fixed Q that $\max_{[0,T]} |K_{Q,1}| \to 0$ in $L^1(\mu_\epsilon)$, and hence in probability, as $\epsilon \to 0$.

Next, since $x_\epsilon - W_\epsilon$ is bounded on $[0, T]$, for Q sufficiently large

$$\mu_\epsilon\big\{\max_{[0,T]} |K_{Q,2}| > 0\big\} \le \mu_\epsilon\big\{\max_{[0,T]} |x_\epsilon| \ge Q\big\} \le \mu_\epsilon\big\{\max_{[0,T]} |W_\epsilon| \ge Q/2\big\}.$$

Fix $c > 0$. Increasing Q if necessary, we can arrange that $\mu_\epsilon\{\max_{[0,T]} |W| \ge Q/2\} < c/4$. By the continuous mapping theorem, $\max_{[0,T]} |W_\epsilon| \to_d \max_{[0,T]} |W|$. Hence there exists $\epsilon_1 > 0$ such that $\mu_\epsilon\{\max_{[0,T]} |W_\epsilon| \ge Q/2\} < c/2$ for all $\epsilon \in (0, \epsilon_1)$. For such ϵ,

$$\mu_\epsilon\big\{\max_{[0,T]} |K_{Q,2}| > 0\big\} < c/2.$$

Shrinking ϵ_1 if necessary, we also have that $\mu_\epsilon\{\max_{[0,T]} |K_{Q,1}| > c/2\} < c/2$. Hence $\mu_\epsilon\{\max_{[0,T]} |I_2| > c\} < c$, and so $\max_{[0,T]} |I_2| \to 0$ in probability. Combining this with estimates (6.4) and (6.5), we obtain that $\max_{[0,T]} |Z_{\epsilon,2}| \to 0$ in probability as required. ∎

The case of exact noise Now we consider the general case of exact multiplicative noise, following [27].

Proof of Theorem 6.3 Define $z_\epsilon(n) = h_\epsilon(x_\epsilon(n))$. Using Taylor's theorem to expand the C^3 map h_ϵ, we obtain

$$\begin{aligned}
&z_\epsilon(n+1) - z_\epsilon(n) = h_\epsilon(x^{(\epsilon)}(n+1)) - h_\epsilon(x^{(\epsilon)}(n)) \\
&= (dh_\epsilon)_{x_\epsilon(n)}\big(x_\epsilon(n+1) - x_\epsilon(n)\big) \qquad (6.7) \\
&+ \tfrac{1}{2}\{(d^2h_\epsilon)_{x_\epsilon(n)}(x_\epsilon(n+1) - x_\epsilon(n))\}(x_\epsilon(n+1) - x_\epsilon(n)) + o(|x_\epsilon(n+1) - x_\epsilon(n)|^2).
\end{aligned}$$

Here, we are identifying $(d^2h)_x$ as an element of $L(\mathbb{R}^d, L(\mathbb{R}^d, \mathbb{R}^d))$ for each $x \in \mathbb{R}^d$. The last term is uniformly $o(\epsilon^2)$.

Substituting for $x_\epsilon(n+1) - x_\epsilon(n)$ using equation (6.1) and the fact that $b_\epsilon = [dh_\epsilon]^{-1}$, equation (6.7) becomes

$$\begin{aligned} z_\epsilon(n+1) - z_\epsilon(n) &= \epsilon^2 \Big\{ (dh_\epsilon)_{x_\epsilon(n)} a_\epsilon(x_\epsilon(n), y_\epsilon(n)) \\ &\quad + \tfrac{1}{2}\{(d^2 h_\epsilon)_{x_\epsilon(n)} b_\epsilon(x_\epsilon(n)) v_\epsilon(y_\epsilon(n))\} b_\epsilon(x_\epsilon(n)) v_\epsilon(y_\epsilon(n)) + o(1) \Big\} + \epsilon v_\epsilon(y_\epsilon(n)). \end{aligned}$$

In other words,

$$z_\epsilon(n+1) - z_\epsilon(n) = \epsilon^2 A_\epsilon(z_\epsilon(n), y_\epsilon(n)) + \epsilon v_\epsilon(y_\epsilon(n)),$$

where

$$A_\epsilon(z,y) = (dh_\epsilon)_{h_\epsilon^{-1} z} \Big\{ a_\epsilon(h_\epsilon^{-1} z, y) - \frac{1}{2}\{(db_\epsilon)_{h_\epsilon^{-1} z} b_\epsilon(h_\epsilon^{-1} z) v_\epsilon(y)\} v_\epsilon(y) + o(1) \Big\}$$

uniformly in z, y as $\epsilon \to 0$.

The regularity assumptions on v_ϵ, a_ϵ and h_ϵ ensure that A_ϵ is bounded and globally Lipschitz in z. Similarly, it is easily checked that $\lim_{\epsilon\to 0} \sup_{z,y} |A_\epsilon(z,y) - A_0(z,y)| = 0$. Notice also that $A_0(z,y) = (dh_0)_{h_0^{-1} z} \alpha_{h_0^{-1} z}(y)$.

Hence we are in the situation of Lemma 6.6, and it follows that $\tilde z_\epsilon(t) = z_\epsilon([t\epsilon^{-2}])$ converges weakly to solutions Z of the SDE

$$dZ = (dh_0)_{h_0^{-1} z} P(h_0^{-1} Z)\, dt + dW, \tag{6.8}$$

where $P(Z)$ is the limit function in (UME).

Next

$$\sup_t |h_\epsilon^{-1}(z_\epsilon(t)) - h_0^{-1}(z_\epsilon(t))| \le \sup_z |h_\epsilon^{-1}(z) - h_0^{-1}(z)| \to 0,$$

so by the continuous mapping theorem,

$$x_\epsilon = h_\epsilon^{-1}(z_\epsilon) = h_0^{-1}(z_\epsilon) + \{h_\epsilon^{-1}(z_\epsilon) - h_0^{-1}(z_\epsilon)\} \to_{\mu_\epsilon} h_0^{-1}(Z).$$

Hence it remains to determine $X = h_0^{-1}(Z)$. Clearly $X(0) = \xi$. Since the Stratonovich integral transforms according to the standard laws of calculus,

$$\begin{aligned} dX &= [(dh_0)_X]^{-1} \circ dZ = [(dh_0)_X]^{-1} \circ [(dh_0)_{h_0^{-1} Z} P(h_0^{-1} Z)\, dt + dW] \\ &= P(X)\, dt + b_0(X) \circ dW, \end{aligned}$$

as required. ∎

7 Homogenization for uniform families of fast-slow systems

In this section, we apply the abstract homogenization theorem from Section 6 to the case where the fast dynamics is generated by a uniform family of nonuniformly hyperbolic transformations. We consider first the noninvertible case (Subsection 7.1) and then the invertible case (Subsection 7.2).

7.1 Nonuniformly expanding fast dynamics

Let (M, d_M) be a bounded metric space with finite Borel measure ρ. For each $\epsilon \in [0, \epsilon_0)$ we suppose that $T_\epsilon : M \to M$ is a nonuniformly expanding map as in Section 2.1, with induced map $F_\epsilon = T_\epsilon^{\tau_\epsilon} : Y_\epsilon \to Y_\epsilon$, and absolutely continuous ergodic T_ϵ-invariant and F_ϵ-invariant Borel probability measures μ_ϵ and μ_{Y_ϵ}. (The metric space (M, d_M) and finite Borel measure ρ are fixed independent of ϵ. This is natural for the purposes of this current section, but is easily relaxed, see Remark 7.4.)

We assume that T_ϵ is a uniform family of order $p \geq 2$ (cf. Section 4), so the various constants in the definition of nonuniformly expanding can be chosen independent of $\epsilon \in [0, \epsilon_0)$, and $\{\tau_\epsilon^2,\, \epsilon \in [0, \epsilon_0)\}$ is uniformly integrable. Moreover, we suppose that μ_0 is *statistically stable*: $\mu_\epsilon \to_w \mu_0$ as $\epsilon \to 0$.

Let $v_\epsilon : M \to \mathbb{R}^d$, $\epsilon \in [0, \epsilon_0)$, be a family of Hölder observables with $\int_M v_\epsilon \, d\mu_\epsilon = 0$. We require that v_ϵ and T_ϵ satisfy

$$\sup_{\epsilon\in[0.\epsilon_0)} \|v_\epsilon\|_\eta < \infty, \qquad \lim_{\epsilon\to 0} |v_\epsilon - v_0|_\infty = 0, \tag{7.1}$$

and

$$\int_M v_0 \circ T_0^j \, (v_0 \circ T_0^k)^T \, (d\mu_\epsilon - d\mu_0) \to 0, \qquad T_\epsilon^j \to_{\mu_\epsilon} T_0^j, \tag{7.2}$$

for all $j, k \geq 0$, as $\epsilon \to 0$. (The last part of condition (7.2) means that $\mu_\epsilon\{y \in M : d_M(T_\epsilon^j y, T_0^j y) > a\} \to 0$ for all $a > 0$.)

Consider the family of fast-slow equations (6.1) where $y_\epsilon(n+1) = T_\epsilon y_\epsilon(n)$. We assume that a_ϵ and b_ϵ satisfy the regularity conditions in Section 6 and that b_ϵ is exact. Let $\hat{x}_\epsilon = x_\epsilon([t\epsilon^{-2}])$.

Theorem 7.1 *Let* $P(x) = \int_M a_0(x, y)\, d\mu_0(y) - \frac{1}{2}\int_M \{(db_0)_x b_0(x) v_0(y)\} v_0(y)\, d\mu_0(y)$. *Let* W *denote* d*-dimensional Brownian motion with covariance*

$$\Sigma = \lim_{n\to\infty} \tfrac{1}{n} \int_M \big(\textstyle\sum_{j=0}^{n-1} v_0 \circ T_0^j\big)\big(\textstyle\sum_{j=0}^{n-1} v_0 \circ T_0^j\big)^T \, d\mu_0.$$

Then $\hat{x}_\epsilon \to_{\mu_\epsilon} X$ *in* $D([0,\infty), \mathbb{R}^d)$ *as* $\epsilon \to 0$ *where* X *is the solution to the Stratonovich SDE*

$$dX = P(X)\, dt + b_0(X) \circ dW, \quad X(0) = \xi.$$

Remark 7.2 In very general situations, [2, 8] show that μ_0 is strongly statistically stable. The first part of condition (7.2) follows immediately. Moreover, in the conclusion of Theorem 7.1 we obtain in addition that $\hat{x}_\epsilon \to_{\mu_0} X$ by Remark 6.5(a).

Examples 7.3 It is straightforward to choose the examples in Section 4 to be strongly statistically stable. Theorem 7.1 and Remark 7.2 then apply.

For instance, in the case of the intermittent maps, Example 4.9, fix $\gamma_0 \in (0, \frac{1}{2})$ and choose $\gamma_\epsilon \to \gamma_0$. Let T_ϵ, $0 \leq \epsilon < \epsilon_0$, be the corresponding family of intermittent maps. Then μ_0 is strongly statistically stable by [11, 36], while (UME) and (WIP) follow from Section 4.

Similar comments apply to Examples 4.10 and 4.11 with statistical stability following from [25] and [8] respectively (see the corresponding examples in [37] for details).

Remark 7.4 Various conditions — namely independence of M and ρ on ϵ, $\lim_{\epsilon\to 0}|v_\epsilon - v_0|_\infty = 0$, and conditions (7.2) — are used only in the proof of continuity of certain covariance matrices Σ_ϵ, see Proposition 7.6. It is easy to check that the results in this section go through with these assumptions removed, provided $\operatorname{diam} M$ is bounded independent of ϵ and the conclusion of Proposition 7.6 holds. We note that [23] gives general conditions under which Σ_ϵ varies continuously.

In the remainder of this subsection, we prove Theorem 7.1. By Theorem 6.3, it suffices to verify (UME) and (WIP) and to identify P and Σ.

Proposition 7.5 *Condition (UME) is satisfied with $P(x)$ as stated in Theorem 7.1.*

Proof We apply Remark 6.4. Condition (a) is automatic, so it remains to verify condition (b).

Recall from (6.2) that $\alpha_x(y) = a_0(x,y) - \frac{1}{2}\{(db_0)_x b_0(x)v_0(y)\}v_0(y)$. Define $\beta_{x,\epsilon} = \alpha_x - \int_M \alpha_x\, d\mu_\epsilon$. Then $\beta_{x,\epsilon} : M \to \mathbb{R}^d$ is family of Hölder observables with $\int_M \beta_{x,\epsilon}\, d\mu_\epsilon = 0$ such that $\|\beta_{x,\epsilon}\|_\eta \ll \|v_0\|_\eta^2$ uniformly in ϵ. (The estimate is also uniform in x, but that is not needed.) By Lemma 4.1, $\lim_{\epsilon\to 0}\int_M |\sum_{j=0}^{[\epsilon^{-1/2}]-1} \beta_{x,\epsilon}\circ T_\epsilon^j|\, d\mu_\epsilon = 0$ for all $x\in\mathbb{R}^d$ as required. ∎

As in Section 4, we can define the family of covariance matrices

$$\Sigma_\epsilon = \lim_{n\to\infty} n^{-1}\int_M S_n v_\epsilon\, S_n v_\epsilon^T\, d\mu_\epsilon, \quad S_n v_\epsilon = \sum_{j=0}^{n-1} v_\epsilon\circ T_\epsilon^j, \quad \epsilon\in[0,\epsilon_0).$$

Proposition 7.6 $\lim_{\epsilon\to 0}\Sigma_\epsilon = \Sigma_0$.

Proof Write $I_{\epsilon,n} = \int_M S_n v_\epsilon\, S_n v_\epsilon^T\, d\mu_\epsilon$.

Let $\delta > 0$. By Remark 4.2, there exists $N \ge 1$ such that $|N^{-1}I_{\epsilon,N} - \Sigma_\epsilon| < \delta$ for all $\epsilon\in[0,\epsilon_0)$. Next

$$I_{\epsilon,N} - I_{0,N} = \int_M (S_N v_\epsilon\, S_N v_\epsilon^T - S_N v_0\, S_N v_0^T)\, d\mu_\epsilon + \int_M S_N v_0\, S_N v_0^T\,(d\mu_\epsilon - d\mu_0).$$

By condition (7.2), $\lim_{\epsilon\to 0}\int_M S_N v_0\, S_N v_0^T\,(d\mu_\epsilon - d\mu_0) = 0$. Also,

$$\begin{aligned}|S_N v_\epsilon\, S_N v_\epsilon^T - S_N v_0\, S_N v_0^T|_{L^1(\mu_\epsilon)} &\le \big(|S_N v_\epsilon|_{L^2(\mu_\epsilon)} + |S_N v_0|_{L^2(\mu_\epsilon)}\big)|S_N v_\epsilon - S_N v_0|_{L^2(\mu_\epsilon)} \\ &\le N(|v_\epsilon|_\infty + |v_0|_\infty)|S_N v_\epsilon - S_N v_0|_{L^2(\mu_\epsilon)}.\end{aligned}$$

But

$$|S_N v_\epsilon - S_N v_0| \le \sum_{j=0}^{N-1}|v_\epsilon - v_0|\circ T_\epsilon^j + \sum_{j=0}^{N-1}|v_0\circ T_\epsilon^j - v_0\circ T_0^j| \le N|v_\epsilon - v_0|_\infty + |v_0|_\eta g_{\epsilon,N},$$

where $g_{\epsilon,N}(y) = \sum_{j=0}^{N-1} d_M(T_\epsilon^j y, T_0^j y)^\eta$. By (7.1) and condition (7.2), we obtain that $\lim_{\epsilon\to 0}|S_N v_\epsilon\, S_N v_\epsilon^T - S_N v_0\, S_N v_0^T|_{L^1(\mu_\epsilon)} = 0$. Hence $\lim_{\epsilon\to 0} I_{\epsilon,N} = I_{0,N}$ and so $\limsup_{\epsilon\to 0}|\Sigma_\epsilon - \Sigma_0| < 2\delta$. Since δ is arbitrary, the result follows. ∎

Corollary 7.7 *Condition (WIP) holds with Σ as stated in Theorem 7.1.*

Proof This follows immediately from Propositions 4.4 and 7.6. ∎

7.2 Nonuniformly hyperbolic fast dynamics

Now we show how to extend Theorem 7.1 to the invertible setting. We assume the same set up as in Subsection 7.1 except that T_ϵ is now a uniform family of nonuniformly hyperbolic transformations as in Subsection 5.1

Theorem 7.8 *The conclusion $\hat{x}_\epsilon \to_{\mu_\epsilon} X$ of Theorem 6.3 remains valid. If in addition μ_0 is strongly statistically stable then $\hat{x}_\epsilon \to_{\mu_0} X$.*

Remark 7.9 The comments in Remark 7.4 apply equally in the current context.

To prove Theorem 7.8, it again suffices to verify condition (UME) and (WIP) in Theorem 6.3.

The proof of (UME) is identical to that of Proposition 7.5 with Corollary 5.5(a) replacing Lemma 4.1. By Corollary 5.5(b), we can define

$$\Sigma_\epsilon = \lim_{n\to\infty} \tfrac{1}{n} \int_M \big(\textstyle\sum_{j=0}^{n-1} v_\epsilon \circ T_\epsilon^j \big) \big(\textstyle\sum_{j=0}^{n-1} v_\epsilon \circ T_\epsilon^j \big)^T \, d\mu_\epsilon.$$

By Remark 4.2 and the proof of Corollary 5.5(b), the convergence is uniform in ϵ. Hence the argument in the proof of Proposition 7.6 shows that $\lim_{\epsilon\to 0} \Sigma_\epsilon = \Sigma_0$. Condition (WIP) with $\Sigma = \Sigma_0$ follows from Theorem 5.7.

Example 7.10 By [23], statistical stability holds for the families of externally forced dispersing billiards in Example 5.9, and hence Theorem 7.8 holds.

We note that the stronger linear response property can be established in certain situations [16], but that linear response is not required for the purposes of this paper.

A WIP for martingale difference arrays

In this appendix, we recast a classical martingale CLT/WIP of [13] into a form that is convenient for ergodic stationary martingale difference arrays of the type commonly encountered in the deterministic setting.

Let $\{(\Delta_n, \mathcal{M}_n, \mu_n)\}$ be a sequence of probability spaces. Suppose that $f_n : \Delta_n \to \Delta_n$ is a sequence of measure-preserving transformations with transfer operators L_n and Koopman operators U_n. Suppose that $m_n : \Delta_n \to \mathbb{R}^d$ lies in $L^2(\Delta_n)$ and that $\int_{\Delta_n} m_n \, d\mu_n = 0$ and $m_n \in \ker L_n$.

Define the sequence of processes $M_n : \Delta_n \to D([0,\infty), \mathbb{R}^d)$ by

$$M_n(t) = n^{-1/2} \sum_{j=0}^{[nt]-1} m_n \circ f_n^j, \quad t \geq 0.$$

Theorem A.1 *Suppose that the family $\{|m_n|^2, \, n \geq 0\}$ is uniformly integrable. Suppose also that there exists a constant matrix $\Sigma \in \mathbb{R}^{d\times d}$, such that $n^{-1} \sum_{j=0}^{[nt]-1} \{U_n L_n (m_n m_n^T)\} \circ f_n^j \to_{\mu_n} t\Sigma$ as $n \to \infty$ for each $t > 0$.*

Then $M_n \to_{\mu_n} W$ in $D([0,\infty), \mathbb{R}^d)$ where W is Brownian motion with covariance Σ.

Proof First we consider the case where Σ is not identically zero. By Prokhorov's Theorem, we must prove convergence of finite-dimensional distributions and tightness.

Finite-dimensional distributions Fix $0 = t_0 \le t_1 \le \cdots \le t_k$, $c_1, \dots, c_k \in \mathbb{R}^d$, $k \ge 1$. Define

$$Z_n = \textstyle\sum_{\ell=1}^k c_\ell^T (M_n(t_\ell) - M_n(t_{\ell-1})), \quad Z = \textstyle\sum_{\ell=1}^k c_\ell^T (W(t_\ell) - W(t_{\ell-1})).$$

We must show that $Z_n \to_{\mu_n} Z$.

Now $Z = N(0, V)$ where $V = \sum_{\ell=1}^k c_\ell^T \Sigma c_\ell (t_\ell - t_{\ell-1})$. Also,

$$Z_n = n^{-1/2} \sum_{\ell=1}^k c_\ell^T \sum_{j=[nt_{\ell-1}]}^{[nt_\ell]-1} m_n \circ f_n^j = \sum_{j=1}^{[nt_k]} X_{n,j},$$

where $X_{n,j} = n^{-1/2} d_{n,j}^T m_n \circ f_n^{[nt_k]-j}$ for appropriate choices of $d_{n,j} \in \{c_1, \dots, c_k\}$.

Define $\mathcal{G}_{n,j} = f_n^{-([nt_k]-j)} \mathcal{M}_n$ for $1 \le j \le [nt_k]$. The same calculation as in the proof of Proposition 2.9 shows that $\{X_{n,j}, \mathcal{G}_{n,j};\, 1 \le j \le [nt_k]\}$ is a martingale difference array. That is, $\mathcal{G}_{n,j} \subset \mathcal{G}_{n,j+1}$ for all $1 \le j \le [nt_k] - 1$, $X_{n,j}$ is $\mathcal{G}_{n,j}$-measurable for all $1 \le j \le [nt_k]$, and $\mathbb{E}(X_{n,j+1} | \mathcal{G}_{n,j}) = 0$ for all $1 \le j \le [nt_k] - 1$.

We now apply a CLT for martingale difference arrays [13, Theorem 18.1]. (See also [42, Theorem 2.3].) To show that $Z_n \to_d N(0, V)$ it suffices to show that

(B1) $\sum_{j=1}^{[nt_k]} \mathbb{E}(X_{n,j}^2 | \mathcal{G}_{n,j-1}) \to_{\mu_n} V$ as $n \to \infty$.

(B2) $\lim_{n\to\infty} \sum_{j=1}^{[nt_k]} \mathbb{E}(X_{n,j}^2 1_{\{|X_{n,j}| \ge \epsilon\}}) = 0$ for all $\epsilon > 0$.

Arguing as in the proof of Proposition 2.9,

$$\mathbb{E}(X_{n,j}^2 | \mathcal{G}_{n,j-1}) = n^{-1} \mathbb{E}((d_{n,j}^T m_n)^2 \circ f^{[nt_k]-j} | \mathcal{G}_{n,j-1}) = n^{-1} \{U_n L_n (d_{n,j}^T m_n)^2\} \circ f_n^{[nt_k]-j}.$$

Hence

$$\begin{aligned}
\sum_{j=1}^{[nt_k]} \mathbb{E}(X_{n,j}^2 | \mathcal{G}_{n,j-1}) &= n^{-1} \sum_{j=1}^{[nt_k]} \{U_n L_n (d_{n,j}^T m_n)^2\} \circ f_n^{[nt_k]-j} \\
&= n^{-1} \sum_{\ell=1}^k \sum_{j=[nt_{\ell-1}]}^{[nt_\ell]-1} \{U_n L_n c_\ell^T (m_n m_n^T) c_\ell\} \circ f_n^j \\
&= \sum_{\ell=1}^k c_\ell^T \Big(n^{-1} \sum_{j=0}^{[nt_\ell]-1} \{U_n L_n (m_n m_n^T)\} \circ f_n^j - n^{-1} \sum_{j=0}^{[nt_{\ell-1}]-1} \{U_n L_n (m_n m_n^T)\} \circ f_n^j \Big) c_\ell
\end{aligned}$$

which converges in probability to V. This proves (B1).

Next, $|X_{n,j}| \le K n^{-1/2} |m_n \circ f_n^{[nt_k]-j}|$ where $K = \max\{|c_1|, \dots, |c_k|\}$. Hence $X_{n,j}^2 1_{\{|X_{n,j}| \ge \epsilon\}} \le K^2 n^{-1} (|m_n|^2 1_{\{|m_n| \ge \epsilon' n^{1/2}\}}) \circ f_n^{[nt_k]-j}$ where $\epsilon' = \epsilon / K$, and

$$\begin{aligned}
\sum_{j=1}^{[nt_k]} \mathbb{E}(X_{n,j}^2 1_{\{|X_{n,j}| \ge \epsilon\}}) &\le K^2 n^{-1} \sum_{j=1}^{[nt_k]} \mathbb{E}(|m_n|^2 1_{\{|m_n| \ge \epsilon' n^{1/2}\}}) \\
&= K^2 n^{-1} [nt_k] \mathbb{E}(|m_n|^2 1_{\{|m_n| \ge \epsilon' n^{1/2}\}}),
\end{aligned}$$

which converges to zero by uniform integrability of $\{|m_n|^2\}$. This proves (B2) and completes the proof that $Z_n \to_{\mu_n} Z$, showing that finite-dimensional distributions converge.

Tightness Tightness of $\{M_n\}$ in $D([0,\infty),\mathbb{R}^d)$ is equivalent to tightness of each coordinate of $\{M_n\}$ in $D([0,T],\mathbb{R})$ for each $T>0$. Hence we fix $T>0$, and assume without loss that M_n is $\mathbb{R}$-valued and that $\frac{1}{n}\sum_{k=0}^{[nt]-1} m_n^2 \circ f_n^k \to_{\mu_n} t\sigma^2$ for some $\sigma^2 \geq 0$. Since Σ is nonzero, we can choose coordinates so that $\sigma^2>0$ in each coordinate. Without loss $\sigma^2=1$.

Since $M_n(0)\equiv 0$, proving tightness of $\{M_n\}$ is equivalent [13, Theorem 7.3] to showing that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mu_n\Bigg(\sup_{\substack{0\le s,t\le T\\ |s-t|\le\delta}}|M_n(t)-M_n(s)|>\epsilon\Bigg)=0 \quad \text{for every } \epsilon>0. \tag{A.1}$$

Define $M_n^-(t)=\sum_{j=1}^{[nt]}\xi_{n,j}$ where $\xi_{n,j}=n^{-1/2}m_n\circ f_n^{[nT]-j}$. We claim that the hypotheses of [13, Theorem 18.2] are satisfied and hence in particular that $\{M_n^-\}$ is tight. It follows that condition (A.1) is satisfied with M_n replaced by M_n^-. But $M_n(t)-M_n(s)=M_n^-(u_{n,s})-M_n^-(u_{n,t})$ where $u_{n,t}\in[0,T]$ is such that $[nu_{n,t}]=[nT]-[nt]$. Hence

$$\sup_{\substack{0\le s,t\le T\\ |s-t|\le\delta}}|M_n(t)-M_n(s)|\le \sup_{\substack{0\le s,t\le T\\ |s-t|\le\delta+\frac{2}{n}}}|M_n^-(t)-M_n^-(s)|,$$

and the result follows.

It remains to verify the claim. Consider the martingale difference array $\{X_{n,j},\mathcal{G}_{n,j};\,1\le j\le[nT]\}$ where $X_{n,j}=n^{-1/2}m_n\circ f_n^{[nT]-j}$, $\mathcal{G}_{n,j}=f_n^{-([nT]-j)}\mathcal{M}_n$. By [13, Theorem 18.2], it suffices to show that for each $t\in[0,T]$,

(B3) $\sum_{j=1}^{[nt]}\mathbb{E}(X_{n,j}^2|\mathcal{G}_{n,j-1})\to_{\mu_n} t$ as $n\to\infty$.

(B4) $\lim_{n\to\infty}\sum_{j=1}^{[nt]}\mathbb{E}(X_{n,j}^2 1_{\{|X_{n,j}|\ge\epsilon\}})=0$ for all $\epsilon>0$.

These are proved in exactly the same way as conditions (B1) and (B2) above.

The completely degenerate case When $\Sigma=0$, we consider the direct product of the underlying dynamics with a simple symmetric random walk on the integers. The product system leads to a WIP with one nondegenerate direction and the result reduces to the case $\Sigma\ne 0$.

More precisely, let $\Lambda'=\{\pm 1\}^{\mathbb{N}}$ with fair ($p=q=1/2$) Bernoulli measure μ', and consider the one-sided shift $f:\Lambda'\to\Lambda'$ and observable $m':\Lambda'\to\{\pm1\}$ where $m'(x)=x_0$. The process $M_n'(t)=n^{-1/2}\sum_{j=0}^{[nt]-1}m'\circ f'^j$ converges in $D([0,\infty),\mathbb{R})$ to Brownian motion W' with variance 1. Define the family of product systems $\Lambda_n''=\Lambda_n\times\Lambda'$, $\mu_n''=\mu_n\times\mu'$, $f_n''=f_n\times f':\Lambda_n''\to\Lambda_n''$, $m_n''=m_n\oplus m':\Lambda_n''\to\mathbb{R}^{d+1}$, $M_n''=M_n\oplus M_n'\in D([0,\infty),\mathbb{R}^{d+1})$. Let U_n'' and L_n'' denote the Koopman and transfer operators corresponding to f_n''. An easy calculation shows that $U''L''(m''m''^T)=UL(mm^T)\oplus 1$ and hence $\{m_n''\}$ satisfies the hypotheses of the theorem with $\Sigma''=0\oplus 1$. Consequently, $M_n''\to_{\mu_n''}W''=0\oplus W'$ in $D([0,\infty),\mathbb{R}^{d+1})$. In particular, $M_n\to_{\mu_n}0$. ∎

Acknowledgements This research was supported in part by a European Advanced Grant *StochExtHomog* (ERC AdG 320977).

References

[1] J. F. Alves. SRB measures for non-hyperbolic systems with multidimensional expansion. *Ann. Sci. École Norm. Sup.* **33** (2000) 1–32.

[2] J. F. Alves. Strong statistical stability of non-uniformly expanding maps. *Nonlinearity* **17** (2004) 1193–1215.

[3] J. F. Alves and D. Azevedo. Statistical properties of diffeomorphisms with weak invariant manifolds. *Discrete Contin. Dyn. Syst.* **36** (2016) 1–41.

[4] J. F. Alves, J. M. Freitas, S. Luzzatto and S. Vaienti. From rates of mixing to recurrence times via large deviations. *Adv. Math.* **228** (2011) 1203–1236.

[5] J. F. Alves, S. Luzzatto and V. Pinheiro. Markov structures and decay of correlations for non-uniformly expanding dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **22** (2005) 817–839.

[6] J. F Alves and V. Pinheiro. Slow rates of mixing for dynamical systems with hyperbolic structures. *J. Stat. Phys.* **131** (2008) 505–534.

[7] J. F. Alves and V. Pinheiro. Gibbs-Markov structures and limit laws for partially hyperbolic attractors with mostly expanding central direction. *Adv. Math.* **223** (2010) 1706–1730.

[8] J. F. Alves and M. Viana. Statistical stability for robust classes of maps with non-uniform expansion. *Ergodic Theory Dynam. Systems* **22** (2002) 1–32.

[9] M. Antoniou and I. Melbourne. Rate of convergence in the weak invariance principle for deterministic systems. In preparation.

[10] V. Araujo, M. J. Pacifico, E. R. Pujals and M. Viana. Singular-hyperbolic attractors are chaotic. *Trans. Amer. Math. Soc.* **361** (2009) 2431–2485.

[11] V. Baladi and M. Todd. Linear response for intermittent maps. *Comm. Math. Phys.* **347** (2016) 857–874.

[12] M. Benedicks and L. Carleson. On iterations of $1 - ax^2$ on $(-1, 1)$. *Ann. of Math.* **122** (1985) 1–25.

[13] P. Billingsley. *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999.

[14] L. A. Bunimovič. The ergodic properties of billiards that are nearly scattering. *Dokl. Akad. Nauk SSSR* **211** (1973) 1024–1026.

[15] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probability* **1** (1973) 19–42.

[16] N. Chernov and A. Korepanov. Spatial structure of Sinai-Ruelle-Bowen measures. *Phys. D* **285** (2014) 1–7.

[17] N. Chernov and H.-K. Zhang. A family of chaotic billiards with variable mixing rates. *Stoch. Dyn.* **5** (2005) 535–553.

[18] N. I. Chernov. Sinai billiards under small external forces. *Ann. Henri Poincaré* **2** (2001) 197–236.

[19] N. I. Chernov and H.-K. Zhang. Billiards with polynomial mixing rates. *Nonlinearity* **18** (2005) 1527–1553.

[20] P. Collet and J.-P. Eckmann. Positive Liapunov exponents and absolute continuity for maps of the interval. *Ergodic Theory Dynam. Systems* **3** (1983) 13–46.

[21] C. Cuny and F. Merlevède. Strong invariance principles with rate for "reverse" martingales and applications. *J. Theor. Probab.* (2015) 137–183.

[22] J. Dedecker and F. Merlevède. Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.* **125** (2015) 3401–3429.

[23] M. F. Demers and H.-K. Zhang. A functional analytic approach to perturbations of the Lorentz gas. *Comm. Math. Phys.* **324** (2013) 767–830.

[24] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689.

[25] J. M. Freitas and M. Todd. The statistical stability of equilibrium states for interval maps. *Nonlinearity* **22** (2009) 259–281.

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

[27] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. London A* (2013) 20130201.

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

[29] S. Gouëzel. Statistical properties of a skew product with a curve of neutral points. *Ergodic Theory Dynam. Systems* **27** (2007) 123–151.

[30] S. Gouëzel and I. Melbourne. Moment bounds and concentration inequalities for slowly mixing dynamical systems. *Electron. J. Probab.* **19** (2014) 1–30.

[31] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application.* Probability and Mathematical Statistics, Academic Press, New York, 1980.

[32] M. V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.* **81** (1981) 39–88.

[33] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520.

[34] D. Kelly and I. Melbourne. Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise. *J. Funct. Anal.* **272** (2017) 4063–4102.

[35] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** (1986) 1–19.

[36] A. Korepanov. Linear response for intermittent maps with summable and nonsummable decay of correlations. *Nonlinearity* **29** (2016) 1735–1754.

[37] A. Korepanov, Z. Kosloff and I. Melbourne. Averaging and rates of averaging for uniform families of deterministic fast-slow skew product systems. *Studia Math.* **238** (2017) 59–89.

[38] A. Korepanov, Z. Kosloff and I. Melbourne. Explicit coupling argument for nonuniformly expanding maps. *Proc. Roy. Soc. Edinburgh.* To appear.

[39] C. Liverani. Central limit theorem for deterministic systems. *International Conference on Dynamical Systems* (F. Ledrappier, J. Lewowicz and S. Newhouse, eds.), Pitman Research Notes in Math. **362**, Longman Group Ltd, Harlow, 1996, pp. 56–75.

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

[41] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** (2000) 713–724.

[42] D. L. McLeish. Dependent central limit theorems and invariance principles. *Ann. Probability* **2** (1974) 620–628.

[43] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005) 131–146.

[44] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **360** (2008) 6661–6676.

[45] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity* (2011) 1361–1367.

[46] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.* **144** (2004) 191–209.

[47] I. Melbourne and A. Török. Convergence of moments for Axiom A and nonuniformly hyperbolic flows. *Ergodic Theory Dynam. Systems* **32** (2012) 1091–1100.

[48] I. Melbourne and P. Varandas. A note on statistical properties for nonuniformly hyperbolic systems with slow contraction and expansion. *Stoch. Dyn.* **16** (2016) 1660012, 13 pages.

[49] I. Melbourne and R. Zweimüller. Weak convergence to stable Lévy processes for nonuniformly hyperbolic dynamical systems. *Ann Inst. H. Poincaré (B) Probab. Statist.* **51** (2015) 545–556.

[50] F. Merlevède, M. Peligrad and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** (2006) 1–36.

[51] W. Philipp and W. F. Stout. *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables.* Memoirs of the Amer. Math. Soc. **161**, Amer. Math. Soc., Providence, RI, 1975.

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

[53] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants.* Mathématiques & Applications (Berlin) [Mathematics & Applications] **31**, Springer-Verlag, Berlin, 2000.

[54] M. Tyran-Kamińska. An invariance principle for maps with polynomial decay of correlations. *Comm. Math. Phys.* **260** (2005) 1–15.

[55] M. Viana. Multidimensional nonhyperbolic attractors. *Inst. Hautes Études Sci. Publ. Math.* (1997) 63–96.

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

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