Functional Correlation Bounds and Deterministic Homogenisation of Fast-slow Systems

by

Nicholas Fleming

Submitted to the University of Warwick for the degree of

Doctor of Philosophy in Mathematics (Research)

Mathematics Institute

May, 2023

Contents

Acknowledgements

First of all, I would like to thank my supervisor, Ian Melbourne, for all the support and guidance that he has given me during my time as a student. Thanks Ian for your patience and kindness, I couldn't have asked for a better supervisor.

Thanks to Mark Pollicott and Mike Todd for agreeing to examine this thesis. I am also grateful to Alexey Korepanov and Nicolò Paviato for many interesting mathematical discussions and anecdotes.

I would also like to thank all the members of the Warwick maths department that have made my time here so enjoyable. I am especially grateful to the seminar organisers for arranging so many stimulating talks. Of course, I would also like to acknowledge the University of Warwick for funding my PhD, as well as all of the aforementioned seminars.

Lastly, I would like to thank my parents for checking up on me so often and for always being there when I needed their support.

Declarations

The main results of Chapter 3 are presented in the article [FV22], which has been published in Communications in Mathematical Physics. The results in Chapter 4 are in the process of being written as a paper, as are the results in Chapter 5.

I declare that the work in this thesis is my own. This thesis has not been submitted for a degree at any other university.

Abstract

The main topic of this thesis is the problem of proving homogenisation (convergence to a stochastic differential equation) for fast-slow systems with deterministic fast direction. Kelly & Melbourne used rough path theory to show that this problem reduces to verifying certain statistical properties for the fast dynamics.

It is natural to consider the case where the fast dynamics is given by a nonuniformly hyperbolic diffeomorphism in the sense of L.-S. Young. Indeed, this class covers a wide variety of examples such as dispersing billiards, intermittent maps and Hénon attractors. Moreover, our understanding of the statistical properties of this class is generally very complete. However, until now it was not possible to verify one of the statistical properties required to apply rough path theory for certain nonuniformly maps with slow decay of correlations.

Our first result is that this property is satisfied by nonuniformly hyperbolic maps in the sense of Young under optimal assumptions on the rate of decay of correlations. The proof splits in two steps. First we prove that the map satisfies a condition introduced by Leppänen which we call the *Functional Correlation Bound*. Then we use a weak dependence argument to show that the required property follows from the Functional Correlation Bound.

Our second main result is that the Functional Correlation Bound is in fact a sufficient condition for homogenisation. Since the Functional Correlation Bound is an elementary condition that is easy to write down, this could be useful for non-dynamicists interested in applying homogenisation results. More generally, we give elementary and explicit sufficient conditions for homogenisation in the case where the fast dynamics is given by a family of dynamical systems.

Finally, we consider the problem of proving rates of convergence in the multidimensional weak invariance principle.

Chapter 1

Introduction

This thesis concerns statistical properties of chaotic dynamical systems, with a particular emphasis on the properties that arise when considering the deterministic homogenisation problem for fast-slow systems. Before introducing this problem, let us briefly discuss some well-known statistical properties of dynamical systems.

Let μ be a Borel probability measure on a metric space M and let $T : M \to M$ be an ergodic μ-preserving transformation. Let $v : M \to \mathbb{R}^d$, $d \geq 1$ be integrable. Then by Birkhoff's ergodic theorem, $n^{-1} \sum_{i=0}^{n-1} v \circ T^i \to \int v \, d\mu$ almost surely. This is analogous to the strong law of large numbers, which says that if $(X_n)_{n \geq 1}$ is a sequence of integrable iid (independent and identically distributed) random vectors, then $n^{-1} \sum_{i=1}^{n} X_i \to \mathbb{E}[X_1]$ almost surely.

More generally, if $T : M \to M$ is 'chaotic' enough and v is sufficiently regular, then the sequence $(v \circ T^n)_{n \geq 1}$ obeys many of the same statistical limit laws as a sequence of iid random vectors. Let v satisfy $\int v \, d\mu = 0$. We say that v satisfies the *central limit theorem* if $n^{-1/2} \sum_{i=0}^{n-1} v \circ T^i$ weakly converges to a normal distribution. Define a piecewise constant random process W_n by

$$W_n(t) = n^{-1/2} \sum_{i=0}^{[nt]-1} v \circ T^i \quad \text{for } t \in [0, 1]. \tag{1.0.1}$$

We view W_n as a random element of the space $D([0, 1], \mathbb{R}^d)$, that is the space of functions $h : [0, 1] \to \mathbb{R}^d$ that are right continuous and have left limits. We say that v satisfies the *weak invariance principle (WIP)* if W_n weakly converges to a Brownian motion. Note that if v satisfies the WIP, then it also satisfies the central limit theorem.

Since the 1960s, both of these limit theorems have been studied for many classes of chaotic dynamical systems [Sin60, Rat73, HK82, Dol04]. For example, if

T is an Anosov diffeomorphism and μ is a Gibbs measure, then the WIP is satisfied provided that v is Hölder.

1.1 Fast-slow systems

A wide variety of systems in applied mathematics can be split into variables that evolve at radically different time-scales. Often one wishes to derive a simplified equation for the slow variable that is valid in the limit as the time-scale separation goes to infinity. There are many rigorous, widely applicable results in this direction in the case where the fast variable is given by a stochastic differential equation (see [PS08, Section 18.4]). In contrast, the case where the fast variable is given by a chaotic dynamical system is less well-understood.

We consider discrete-time fast-slow systems on $\mathbb{R}^d \times M$ of the form

$$\begin{cases} x_{k+1}^{(n)} = x_k^{(n)} + n^{-1} a(x_k^{(n)}, y_k) + n^{-1/2} b(x_k^{(n)}, y_k), \\ y_{k+1} = T y_k, \end{cases} \tag{1.1.1}$$

where $x_0^{(n)} \equiv \xi \in \mathbb{R}^d$ is fixed and y_0 is drawn randomly from (M, μ). Assume that $a, b : \mathbb{R}^d \times M \to \mathbb{R}^d$ are regular and that $\int_M b(x,y) d\mu(y) = 0$ for all $x \in \mathbb{R}^d$. We call this system *deterministic* because the only source of randomness is the initial condition y_0. We are interested in finding an approximate description for the slow dynamics $x^{(n)}$ that is valid in the limit as $n \to \infty$. More precisely, we wish to characterise the limiting behaviour of the random process X_n defined by $X_n(t) = x_{[nt]}^{(n)}$, $t \in [0,1]$. In order to see what kind of limiting behaviour we should expect, let us consider a very special case for the slow dynamics.

Example 1.1.1. *Let $\xi = 0$, $a \equiv 0$ and $b(x,y) = v(y)$. Then*

$$X_n(t) = n^{-1/2} \sum_{i=0}^{[nt]-1} v \circ T^i.$$

Hence if v satisfies the WIP, then X_n weakly converges to Brownian motion.

More generally, we are interested in proving that X_n weakly converges to a stochastic differential equation driven by Brownian motion for all sufficiently regular a, b. We refer to this as *(deterministic) homogenisation.*

In [Dol04] Dolgopyat proved homogenisation for a class of partially hyperbolic diffeomorphisms. In [MS11, GM13] Melbourne, Stuart & Gottwald initiated a programme where homogenisation is derived from statistical properties of the fast dynamics. In particular, [GM13] proved that for certain special choices of the coefficients a and b homogenisation follows from the WIP. However, it is not difficult to find examples of slow dynamics where this approach breaks down:

Example 1.1.2. *Take $d = 2$ and let $\xi = 0$, $a(x, y) \equiv 0$ and $b(x, y) = \begin{pmatrix} 1 & 0 \\ 0 & x^1 \end{pmatrix} \begin{pmatrix} v^1(y) \\ v^2(y) \end{pmatrix}$. Write $X_n = (X_n^1, X_n^2)$ and let $W_n = (W_n^1, W_n^2)$ be as defined in (1.0.1). Denote $W_n^1(s^-) = \lim_{r \uparrow s} W_n^1(r)$. Then $X_n^1(t) = W_n^1(t)$ and*

$$X_n^2(t) = \sum_{i=0}^{[nt]-1} W_n^1(\tfrac{i}{n})(W_n^2(\tfrac{i+1}{n}) - W_n^2(\tfrac{i}{n})) = \int_0^t W_n^1(s^-) dW_n^2(s),$$

where the above integral is interpreted in the Riemann-Stieltjes sense. Unfortunately, knowing that W_n weakly converges to Brownian motion does not determine the weak limit of $\int_0^t (W_n^1)(s^-) dW_n^2(s)$.

Perhaps somewhat surprisingly, understanding the limiting behaviour of the iterated integrals $\int_0^t W_n^1(s^-) dW_n^2(s)$ is almost enough to prove homogenisation for much more general slow dynamics.

Fix $\eta \in (0, 1]$. Let $C_0^\eta(M) = \{v : M \to \mathbb{R} \,|\, v \text{ is } \eta\text{-Hölder}, \int_M v \, d\mu = 0\}$. Given $v = (v^1, \ldots, v^k) \in C_0^\eta(M)^k$ let W_n be as defined in (1.0.1) and define $\mathbb{W}_n \in D([0, 1], \mathbb{R}^{k \times k})$ by

$$\mathbb{W}_n^{\alpha,\beta}(t) = \int_0^t W_n^\alpha(s^-) dW_n^\beta(s) = n^{-1} \sum_{0 \le i < j < [nt]} v^\alpha \circ T^i \, v^\beta \circ T^j$$

for $1 \le \alpha, \beta \le k$, $t \in [0, 1]$.

Theorem 1.1.3 ([KM16, Che$^+$22])**.** *Assume that*

(i) *(Iterated WIP) For all $v \in C_0^\eta(M)^k$, $k \ge 1$ we have that $(W_n, \mathbb{W}_n)$ weakly converges to $(W, \mathbb{W})$ in $D([0, 1], \mathbb{R}^k \times \mathbb{R}^{k \times k})$, where W is a Brownian motion and $\mathbb{W}$ is a suitable matrix-valued process.*

(ii) *(Iterated moment bounds) There exists $p > 1$ such that*

$$\left| \sum_{0 \le i < j \le n} v \circ T^i \, w \circ T^j \right|_{L^p} = O(n)$$

for all $v, w \in C_0^\eta(M)$.

Let $a, b : \mathbb{R}^d \times M \to \mathbb{R}^d$ be regular. Then X_n weakly converges to the solution of a stochastic differential equation driven by Brownian motion.

The above theorem is a special case of [Che$^+$22, Theorem 2.10] and is proved by rough path theory (see [Che$^+$19] for a good survey on how rough path theory is applied in this context). In [KM16] a similar theorem was proved under the stronger assumption that moment bounds hold with $p > 3$.

Now that we have a sufficient condition for homogenisation in terms of statistical properties for $T : M \to M$, the goal is to check these properties for a wide class of fast dynamics. It is particularly natural to consider the class of nonuniformly expanding/hyperbolic maps modelled by Young towers. This framework was introduced by L.-S. Young [You98, You99] in order to study the statistical properties of a broad generalisation of uniformly expanding/hyperbolic (Axiom A) dynamics. In particular, it covers many physical examples such as Hénon attractors, dispersing billiards, intermittent interval maps and the Poincaré map of the Lorenz attractor.

The statistical properties of a system modelled by a Young tower are determined by the tails of returns to the base of the tower. In particular, if we assume that the tails decay at rate $O(n^{-\beta})$ with $\beta > 2$, then the WIP is satisfied by any Hölder mean zero observable so it is natural to consider homogenisation. (This assumption is optimal; for $\beta \leq 2$ there are examples where even the central limit theorem fails for generic Hölder observables [Gou04].) By [KM16, MV16] the iterated WIP is also satisfied by mean zero Hölder observables. However, obtaining iterated moment bounds in the full range $\beta > 2$ proved more problematic and even in the nonuniformly *expanding* case this was only achieved recently [KKM22].

In Chapter 3, we extend iterated moment bounds to the case where $T : M \to M$ is a nonuniformly hyperbolic map with $O(n^{-\beta})$ tails for the optimal range $\beta > 2$. By Theorem 1.1.3, this extends homogenisation results to examples such as Bunimovich flowers [Bun73, CZ05a], a class of billiards with flat points considered by Chernov & Zhang [CZ05a] and certain almost Anosov maps [EL21].

In order to prove moment bounds, we first prove that $T : M \to M$ satisfies an elementary condition introduced by Leppänen [Lep17], which we call the *Functional Correlation Bound*. Let us now motivate this condition, which is the main technical tool used in this thesis.

By [MT14], $T : M \to M$ enjoys decay of correlations with rate $O(n^{-(\beta-1)})$, that is the correlation function

$$C_n(v, w) = \int_M v\, w \circ T^n \, d\mu - \int_M v \, d\mu \int_M w \, d\mu$$

satisfies $|C_n(v,w)| = O(n^{-(\beta-1)})$ for all Hölder $v, w : M \to \mathbb{R}$. Note that

$$C_n(v,w) = \int_M G(x, T^n x) d\mu(x) - \int_{M^2} G(x_1, T^n x_2) d\mu(x_1) d\mu(x_2)$$

where $G(x,y) = v(x)w(y)$. Hence, *very* loosely speaking, decay of correlations can be thought of as quantifying how close the states of our system at times 0 and n are to being independent. The Functional Correlation Bound generalises decay of correlations by giving analogous bounds on expressions of the form

$$\int_M G(x, Tx, \dots, T^n x, T^m x, T^{m+1} x, \dots, T^p x) d\mu(x) \\ - \int_{M^2} G(x_1, Tx_1, \dots, T^n x_2, T^m x_2, T^{m+1} x_2, \dots, T^p x_2) d\mu(x_1) d\mu(x_2)$$

for $n \ll m \leq p$ and general regular multivariable $\mathbb{R}$-valued functions G.

In Section 3.4, we prove that the Functional Correlation Bound implies a weak dependence lemma that is used throughout this thesis. We then finish Chapter 3 by proving that this lemma implies iterated moment bounds.

In Chapter 4, we prove that the Functional Correlation Bound is a sufficient condition for homogenisation by showing that it also implies the iterated WIP. Since the Functional Correlation Bound is an elementary condition, this could be useful for non-dynamicists interested in applying homogenisation results. We also provide elementary sufficient conditions for homogenisation for a generalisation of (1.1.1) where the fast dynamics is given by a family of dynamical systems $T_n : M \to M$. We then check these conditions for examples of families including intermittent Baker's maps and externally forced dispersing billiards.

1.2 Rates in the weak invariance principle

A different problem, which is the subject of Chapter 5, is to quantify the rate of convergence in the WIP. Sharp convergence rates have been established in the central limit theorem for many classes of dynamical systems [CP90, Gou05]. In contrast, the first results on rates in the WIP for dynamical systems were obtained only recently [AM19, LW22, Pav23] and give significantly worse rates than those available for iid random vectors.

Again, we consider systems that satisfy the Functional Correlation Bound with a sufficiently fast polynomial rate. We prove rates in the WIP in the Wasserstein-1 metric for $\mathbb{R}^d$-valued Hölder observables. Our argument is based on Bernstein's classical 'big block-small block' method, whereas existing results for dynamical systems use the martingale approximation method.

The rates that we obtain are independent of d and improve on those available in the literature for $d > 1$ [Pav23]. In particular, for systems that satisfy the Functional Correlation Bound with a superpolynomial rate (including systems modelled by Young towers with superpolynomial tails) we obtain a rate of the form $O(n^{-(1/4-\delta)})$ for any $\delta > 0$, whereas the rate obtained in [Pav23] is at best $O(n^{-1/6})$.

1.3 Thesis outline

The chapters of this thesis are structured as follows:

- In Chapter 2, we recall some basic preliminary material on weak convergence.
- Chapter 3 is the first containing new results. First, we introduce the Functional Correlation Bound and prove that is satisfied by nonuniformly hyperbolic maps with polynomial tails. Then we show that the Functional Correlation Bound implies iterated moment bounds.
- In Chapter 4, we provide elementary and explicit conditions (including the Functional Correlation Bound) for homogenisation in the case where the fast dynamics is given by a family of maps. We then verify these conditions for certain families of nonuniformly hyperbolic maps.
- In Chapter 5, we prove rates of convergence in the multidimensional weak invariance principle for systems that satisfy the Functional Correlation Bound.
- In Appendix A, we prove an iterated weak invariance principle for arrays of random vectors.

1.4 Notation

- We endow $\mathbb{R}^d$ with the norm $|y| = \sum_{i=1}^d |y_i|$. For $a, b \in \mathbb{R}^d$ we denote $a \otimes b = ab^T$.
- We write $a_n = O(b_n)$ or $a_n \ll b_n$ if there exists a constant $C > 0$ such that $a_n \le Cb_n$ for all $n \ge 1$. We write $a_n = o(b_n)$ if $\lim_n a_n/b_n = 0$ and $a_n \sim b_n$ if $\lim_n a_n/b_n = 1$.
- Let $\eta \in (0, 1]$. We say that a function $v\colon M \to \mathbb{R}$ on a metric space (M, d) is η-Hölder, and write $v \in \mathcal{C}^\eta(M)$, if $\|v\|_\eta = |v|_\infty + [v]_\eta < \infty$, where $|v|_\infty =$

$\sup_M |v|$ and $[v]_\eta = \sup_{x\neq y} |v(x) - v(y)|/d(x,y)^\eta$. If $\eta = 1$ we call v Lipschitz and write $\mathrm{Lip}(v) = [v]_1$.

- For $1 \leq p \leq \infty$ we use $|\cdot|_p$ to denote the L^p norm.

Chapter 2

Preliminaries

2.1 Weak convergence

Let $\mathcal{X}$ be a metric space and let μ, μ_n be Borel probability measures on $\mathcal{X}$ for $n \geq 1$. We say that μ_n *converges weakly to* μ and write $\mu_n \xrightarrow{w} \mu$ if

$$\lim_{n\to\infty} \int_{\mathcal{X}} f \, d\mu_n = \int_{\mathcal{X}} f \, d\mu_n$$

for all bounded, continuous functions $f : \mathcal{X} \to \mathbb{R}$.

The following lemma provides a useful characterisation of weak convergence:

Lemma 2.1.1 (Portmanteau theorem)**.** *Let μ_n, μ be Borel probability measures on $(\mathcal{X}, \rho)$ for $n \geq 1$. The following are equivalent:*

(i) $\mu_n \xrightarrow{w} \mu$*;*

(ii) $\int_{\mathcal{X}} f \, d\mu_n \to \int_{\mathcal{X}} f \, d\mu$ *for all bounded, Lipschitz* $f : \mathcal{X} \to \mathbb{R}$*;*

(iii) $\limsup_n \mu_n(F) \leq \mu(F)$ *for all closed* $F \subseteq \mathcal{X}$*;*

(iv) $\limsup_n \mu_n(U) \geq \mu(U)$ *for all open* $U \subseteq \mathcal{X}$*;*

(v) $\lim_{n\to\infty} \mu_n(B) = \mu(B)$ *for all Borel sets* $B \subseteq \mathcal{X}$ *with* $\mu(\partial B) = 0$.

The equivalence of conditions (i)-(iv) is proved in [Bog18, Theorem 2.2.5]. By [Bog18, Theorem 2.4.1], (v) is equivalent to (i).

A *random element* of $\mathcal{X}$ is a measurable mapping $X : \Omega \to \mathcal{X}$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is some probability space. The distribution of X is the probability measure $\mathbb{P}_X = X_*\mathbb{P}$. In this thesis, we consider the cases where $\mathcal{X} = \mathbb{R}^d$ or $\mathcal{X} = D([0,1], \mathbb{R}^d)$ with the metric induced by the supremum norm. If $\mathcal{X} = D([0,1], \mathbb{R}^d)$ then we call $\mathcal{X}$ a *random process.*

Let X, X_n be random elements of $\mathcal{X}$ for $n \geq 1$. We say that $X_n \xrightarrow{w} X$ if $\mathbb{P}_{X_n} \xrightarrow{w} \mathbb{P}_X$.

We will also use the follow results:

Lemma 2.1.2 (Continuous mapping theorem)**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be metric spaces and let X, $(X_n)_{n\geq 1}$ be random elements of $\mathcal{X}$. Assume that $h : \mathcal{X} \to \mathcal{Y}$ is continuous. If $X_n \xrightarrow{w} X$, then $h(X_n) \xrightarrow{w} h(X)$.*

Lemma 2.1.3 (Slutsky's Theorem [Bas11, Proposition A.42])**.** *Let $(X_n)_{n\geq 1}, (Y_n)_{n\geq 1}$ be sequences of random vectors defined on the same probability space. Suppose that $X_n \xrightarrow{w} X$ and $Y_n \xrightarrow{w} c$ where c is constant. Then $X_n + Y_n \xrightarrow{w} X + c$.*

Chapter 3

The functional correlation bound and iterated moment bounds for nonuniformly hyperbolic maps

3.1 Introduction

As discussed in the introduction of this thesis, we are interested in proving homogenisation for systems modelled by Young towers. Let $T : M \to M$ be a nonuniformly hyperbolic map modelled by a Young tower with invariant probability measure μ. As part of the setup, there is a positive measure set $Y \subset M$ and a return time $\phi : Y \to \mathbb{Z}^+$. Define $F : Y \to Y$ by $F(y) = T^{\phi(y)}(y)$. Roughly speaking, F is assumed to be uniformly hyperbolic. In the special case where F is uniformly expanding (Gibbs-Markov), T is called nonuniformly expanding.

The statistical properties of $T : M \to M$ are determined by the tails $\mu(\phi > n)$. Assume that $\mu(\phi > n) = O(n^{-\beta})$ with $\beta > 2$. Then Hölder mean zero observables satisfy the weak invariance principle (WIP) so it is natural to consider homogenisation. In Theorem 1.1.3 we saw that proving homogenisation reduces to verifying two statistical properties, namely the "iterated WIP" and "iterated moment bounds".

In [KM16, MV16] it was shown that the iterated WIP is satisfied by nonuniformly expanding/hyperbolic maps in the full range $\beta > 2$. However, obtaining iterated moment bounds in the full range $\beta > 2$ proved more problematic. Indeed, even in the nonuniformly *expanding* case this was only achieved recently [KKM22].

The case where T is nonuniformly *hyperbolic* is more challenging. In [You98] it is assumed that T is exponentially contracting along stable manifolds. Under this assumption, iterated moment bounds follow straightforwardly from the corresponding bounds in the nonuniformly expanding case (see Remark 3.3.3 for more details). However, this assumption fails for examples such as slowly-mixing billiards.

Without assuming uniform contraction along stable manifolds, previously it was only possible to show iterated moment bounds for $\beta > 5$ [DMN20]. In this chapter, we extend iterated moment bounds to the optimal range $\beta > 2$.

3.1.1 Illustrative examples

Many examples of invertible dynamical systems are modelled by Young towers. For example, Axiom A (uniformly hyperbolic) diffeomorphisms, Hénon attractors and the finite-horizon Sinai billiard are modelled by Young towers with exponential tails and exponential contraction along stable manifolds [You98]. Hence for such systems deterministic homogenisation results follow from [KM16]. We now give some examples of slowly-mixing nonuniformly hyperbolic dynamical systems for which it was not previously possible to show deterministic homogenisation, due to a lack of control of iterated moments. We start with an example which is easy to write down:

- **Intermittent Baker's maps.** Let $\alpha \in (0,1)$. Define $g\colon [0,1/2] \to [0,1]$ by $g(x) = x(1 + 2^\alpha x^\alpha)$. The Liverani-Saussol-Vaienti map $\bar{T}\colon [0,1] \to [0,1]$,

$$\bar{T}x = \begin{cases} g(x), & x \leq 1/2, \\ 2x - 1, & x > 1/2 \end{cases}$$

  is a prototypical example of a slowly-mixing nonuniformly expanding map [LSV99]. As in [MV16, Exa. 4.1], consider an intermittent Baker's map $T\colon M \to M$, $M = [0,1]^2$ defined by

$$T(x_1, x_2) = \begin{cases} (\bar{T}x_1, g^{-1}(x_2)), & x_1 \in [0, \frac{1}{2}],\ x_2 \in [0,1], \\ (\bar{T}x_1, (x_2+1)/2), & x_1 \in (\frac{1}{2}, 1],\ x_2 \in [0,1]. \end{cases}$$

  Let π denote the projection onto the first coordinate. There is a unique T-invariant probability measure μ such that $\pi_*\mu = \bar{\mu}$. The map T is nonuniformly hyperbolic and has a neutral fixed point at $(0,0)$ whose influence increases with α. In particular, T is modelled by a Young tower with tails of the form $\sim n^{-\beta}$ where $\beta = 1/\alpha$ (see Subsection 4.4.1 for more details).

For $\beta > 2$ the WIP holds for all Hölder observables. For $\beta \leq 2$, the central limit theorem fails for generic Hölder observables even for the $\bar{T}_\alpha$ dynamics [Gou04], so it is natural to restrict to $\beta > 2$ when considering deterministic homogenisation. By [DMN20] it is possible to show iterated moment bounds for $\beta > 5$. Our results yield iterated moment bounds and hence deterministic homogenisation in the full range $\beta > 2$.

Chaotic billiards provide many examples of slowly-mixing nonuniformly hyperbolic maps. Markarian [Mar04], Chernov and Zhang [CZ05b] showed how to model many examples of chaotic billiards by Young towers with polynomial tails.

We give two classes of chaotic billiards for which it is now possible to show deterministic homogenisation:

- **Bunimovich flowers [Bun73].** By [CZ05b] the billiard map is modelled by a Young tower with tails of the form $O(n^{-3}(\log n)^3)$.
- **Dispersing billiards with vanishing curvature.** In [CZ05a] Chernov and Zhang introduced a class of billiards modelled by Young towers with tails of the form $O((\log n)^\beta n^{-\beta})$ for any prescribed value of $\beta \in (2, \infty)$.

Other examples of maps modelled by a Young tower with tails of the form $O(n^{-\beta})$ with $\beta \in (2, 5)$ include Wojtkowski's system of two falling balls [BBNV12] and certain almost Anosov maps [EL21].

3.2 Main results

Let (M, d) be a metric space. Fix $\eta \in (0, 1]$ and let $v : M \to \mathbb{R}$. Let $[v]_\eta = \sup_{x \neq y} |v(x) - v(y)|/d(x, y)^\eta$ denote the η-Hölder seminorm of v.

Definition 3.2.1. *Fix an integer $q \geq 1$. Given a function $G\colon M^q \to \mathbb{R}$ and $0 \leq i < q$ we denote*

$$[G]_{\eta,i} = \sup_{x_0,\ldots,x_{q-1} \in M} [G(x_0, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_{q-1})]_\eta.$$

We call G separately η-Hölder, and write $G \in \mathcal{H}^\eta_q(M)$, if $|G|_\infty + \sum_{i=0}^{q-1} [G]_{\eta,i} < \infty$.

Note that $\mathcal{H}^\eta_1(M) = \mathcal{C}^\eta(M)$ is the space of η-Hölder observables. Fix $\gamma > 0$. We consider dynamical systems which satisfy the following property:

Definition 3.2.2. *Let μ be a Borel probability measure and let $T : M \to M$ be an ergodic μ-preserving transformation. Suppose that there exists a constant $C > 0$*

such that for all integers $0 \le p < q,\ 0 \le n_0 \le \cdots \le n_{q-1}$*,*

$$\begin{aligned}\Big|\int_M & G(T^{n_0}x,\dots,T^{n_{q-1}}x)d\mu(x) \\ &- \int_{M^2} G(T^{n_0}x_0,\dots,T^{n_{p-1}}x_0,T^{n_p}x_1,\dots,T^{n_{q-1}}x_1)d\mu(x_0)d\mu(x_1)\Big| \\ &\le C(n_p - n_{p-1})^{-\gamma}\Big(|G|_\infty + \sum_{i=0}^{q-1}[G]_{\eta,i}\Big) \qquad (3.2.1)\end{aligned}$$

for all $G \in \mathcal{H}_q^\eta(M)$*. Then we say that* T *satisfies the Functional Correlation Bound with rate* $n^{-\gamma}$*.*

A similar condition was introduced by Leppänen in [Lep17] and further studied by Leppänen and Stenlund in [LS17, LS20]. In particular, [Lep17] showed that functional correlation decay implies a multi-dimensional CLT with bounds on the rate of convergence. We are now ready to state the main results which we prove in this chapter.

The rate of decay of correlations of a dynamical system modelled by a Young tower is determined by the tails of the return time to the base of the tower. Indeed, let T be a mixing transformation modelled by a two-sided Young tower with tails of the form $O(n^{-\beta})$ for some $\beta > 1$. In [MT14] by using ideas from [CG12], it was shown that there exists $C > 0$ such that

$$\Big|\int_M v\ w\circ T^n d\mu - \int_M v d\mu \int_M w d\mu\Big| \le C n^{-(\beta-1)} \|v\|_\eta \|w\|_\eta$$

for all $n \ge 1$, $v,\ w \in \mathcal{C}^\eta(M)$. Our first main result is that the Functional Correlation Bound holds with the same rate:

Theorem 3.2.3. *Let* $\beta > 1$*. Let* T *be a mixing transformation modelled by a two-sided Young tower whose return time has tails of the form* $O(n^{-\beta})$*. Then* T *satisfies the Functional Correlation Bound with rate* $n^{-(\beta-1)}$*.*

Given $v, w \in \mathcal{C}^\eta(M)$ mean zero define

$$S_v(n) = \sum_{0\le i<n} v\circ T^i, \quad \mathbb{S}_{v,w}(n) = \sum_{0\le i<j<n} v\circ T^i\ w\circ T^j.$$

Our second main result is that the Functional Correlation Bound implies moment estimates for $S_v(n)$ and $\mathbb{S}_{v,w}(n)$. Let $\|\cdot\|_\eta = |\cdot|_\infty + [\cdot]_\eta$ denote the η-Hölder norm.

Theorem 3.2.4. *Let $\gamma > 1$. Suppose that T satisfies the Functional Correlation Bound with rate $n^{-\gamma}$. Then there exists a constant $C > 0$ such that for all $n \geq 1$, for any mean zero $v, w \in \mathcal{C}^\eta(M)$,*

(a) $|S_v(n)|_{2\gamma} \leq Cn^{1/2} \|v\|_\eta$.

(b) $|\mathbb{S}_{v,w}(n)|_\gamma \leq Cn \|v\|_\eta \|w\|_\eta$.

Remark 3.2.5. By Theorem 1.1.3 to obtain deterministic homogenisation results it suffices to prove the iterated WIP and iterated moment bounds. Let T be a mixing transformation modelled by a two-sided Young tower with tails of the form $O(n^{-\beta})$ for some $\beta > 2$. By [MV16, Corollary 2.3], the Iterated WIP holds for all Hölder observables. Together Theorem 3.2.3 and Theorem 3.2.4 give the iterated moment bounds required in Theorem 1.1.3.

Remark 3.2.6. In this thesis we derive moment bounds for Hölder observables from a functional correlation bound for separately Hölder functions. For systems modelled by Young towers, one can also consider a wider class of dynamically Hölder observables (see Subsection 3.3.1).

In [FV22], the author considered a functional correlation bound for separately dynamically Hölder functions and showed that it implies moment bounds for dynamically Hölder observables. The arguments used are essentially the same as those in the proof of Theorem 3.2.4.

3.3 Young Towers

3.3.1 Prerequisites

Young towers were first introduced by L.-S. Young in [You98, You99], as a broad framework to prove decay of correlations for nonuniformly hyperbolic maps. Our presentation follows [BMT21]. In particular, this framework does not assume uniform contraction along stable manifolds and hence covers examples such as slowly mixing billiard maps.

Gibbs-Markov maps: Let $(\bar{Y}, \bar{\mu}_Y)$ be a probability space and let $\bar{F} \colon \bar{Y} \to \bar{Y}$ be ergodic and measure-preserving. Let α be an at most countable, measurable partition of $\bar{Y}$. We assume that there exist constants $D_0 > 0$, $\theta \in (0, 1)$ such that for all elements $a \in \alpha$:

- (Full-branch condition) The map $\bar{F}|_a \colon a \to \bar{Y}$ is a measurable bijection.

- For all distinct $y, y' \in \bar{Y}$ the separation time

$$s(y, y') = \inf\{n \geq 0 : \bar{F}^n y,\ \bar{F}^n y' \text{ lie in distinct elements of } \alpha\} < \infty.$$

- Define $\zeta\colon a \to \mathbb{R}^+$ by $\zeta = d\bar{\mu}_Y/(d\,(F|_a^{-1})_*\bar{\mu}_Y)$. We have $|\log\zeta(y) - \log\zeta(y')| \leq D_0\theta^{s(y,y')}$ for all $y, y' \in a$.

Then we call $\bar{F}\colon \bar{Y} \to \bar{Y}$ a full-branch Gibbs-Markov map.

Two-sided Gibbs-Markov maps Let (Y, d) be a bounded metric space with Borel probability measure μ_Y and let $F\colon Y \to Y$ be ergodic and measure-preserving. Let $\bar{F}\colon \bar{Y} \to \bar{Y}$ be a full-branch Gibbs-Markov map with associated measure $\bar{\mu}_Y$.

We suppose that there exists a measure-preserving semi-conjugacy $\bar{\pi}\colon Y \to \bar{Y}$, so $\bar{\pi} \circ F = \bar{F} \circ \bar{\pi}$ and $\bar{\pi}_*\mu_Y = \bar{\mu}_Y$. The separation time $s(\cdot,\cdot)$ on $\bar{Y}$ lifts to a separation time on Y given by $s(y, y') = s(\bar{\pi}y, \bar{\pi}y')$. Suppose that there exist constants $K > 0, \theta \in (0, 1)$ such that

$$d(F^n y, F^n y') \leq D_0(\theta^n + \theta^{s(y,y')-n}) \text{ for all } y, y' \in Y, n \geq 0. \tag{3.3.1}$$

Then we call $F\colon Y \to Y$ a *two-sided Gibbs-Markov map*.

One-sided Young towers: Let $\bar{\phi}\colon \bar{Y} \to \mathbb{Z}^+$ be integrable and constant on partition elements of α. We define the one-sided Young tower $\bar{\Delta} = \bar{Y}^{\bar{\phi}}$ and tower map $\bar{f}\colon \bar{\Delta} \to \bar{\Delta}$ by

$$\bar{\Delta} = \{(\bar{y}, \ell) \in \bar{Y} \times \mathbb{Z} : 0 \leq \ell < \bar{\phi}(y)\}, \quad \bar{f}(\bar{y}, \ell) = \begin{cases} (\bar{y}, \ell+1), & \ell < \bar{\phi}(y) - 1, \\ (\bar{F}\bar{y}, 0), & \ell = \bar{\phi}(y) - 1. \end{cases} \tag{3.3.2}$$

We extend the separation time $s(\cdot,\cdot)$ to $\bar{\Delta}$ by defining

$$s((\bar{y}, \ell), (\bar{y}', \ell')) = \begin{cases} s(\bar{y}, \bar{y}'), & \ell = \ell', \\ 0, & \ell \neq \ell'. \end{cases}$$

Note that for $\theta \in (0, 1)$ we can define a metric by $d_\theta(\bar{p}, \bar{q}) = \theta^{s(\bar{p},\bar{q})}$.

Now, $\bar{\mu}_\Delta = (\bar{\mu}_Y \times \text{counting})/\int_{\bar{Y}} \bar{\phi}\,d\bar{\mu}_Y$ is an ergodic $\bar{f}$-invariant probability measure on $\bar{\Delta}$.

Two-sided Young towers Let $F\colon Y \to Y$ be a two-sided Gibbs-Markov map and let $\phi\colon Y \to \mathbb{Z}^+$ be an integrable function that is constant on $\bar{\pi}^{-1}a$ for each $a \in \alpha$. In particular, ϕ projects to a function $\bar{\phi}\colon \bar{Y} \to M$ that is constant on partition elements of α.

Define the one-sided Young tower $\bar{\Delta} = \bar{Y}^{\bar{\phi}}$ as in (3.3.2). Using ϕ in place of $\bar{\phi}$ and $F\colon Y \to Y$ in place of $\bar{F}\colon \bar{Y} \to \bar{Y}$, we define the *two-sided Young tower* $\Delta = Y^{\phi}$ and tower map $f\colon \Delta \to \Delta$ in the same way. Likewise, we define an ergodic f-invariant probability measure on Δ by $\mu_\Delta = (\mu_Y \times \text{counting})/\int_Y \phi\, d\mu_Y$.

We extend $\bar{\pi}\colon Y \to \bar{Y}$ to a map $\bar{\pi}\colon \Delta \to \bar{\Delta}$ by setting $\bar{\pi}(y,\ell) = (\bar{\pi}y, \ell)$ for all $(y,\ell) \in \Delta$. Note that $\bar{\pi}$ is a measure-preserving semi-conjugacy; $\bar{\pi} \circ f = \bar{f} \circ \bar{\pi}$ and $\bar{\pi}_* \mu_\Delta = \bar{\mu}_\Delta$. The separation time s on $\bar{\Delta}$ lifts to Δ by defining $s(y,y) = s(\bar{\pi}y, \bar{\pi}y')$.

We are now finally ready to say what it means for a map to be modelled by a Young tower:

Definition 3.3.1. *Let $T\colon M \to M$ be a measure-preserving transformation on a probability space (M,μ). Suppose that there exists $Y \subset M$ measurable with $\mu(Y) > 0$ such that:*

- *$F = T^{\phi}\colon Y \to Y$ is a two-sided Gibbs-Markov map with respect to some probability measure μ_Y.*
- *ϕ is constant on partition elements of $\bar{\pi}^{-1}\alpha$, so we can define Young towers $\Delta = Y^{\phi}$ and $\bar{\Delta} = \bar{Y}^{\bar{\phi}}$.*
- *There exist constants $D_0 > 0$ and $\theta \in (0,1)$ such that for all $y, y' \in Y$, $0 \le \ell < \phi(y)$ we have*

$$d(T^\ell y, T^\ell y') \le D_0(d(y,y') + \theta^{s(y,y')}) \tag{3.3.3}$$

- *The map $\pi_M\colon \Delta \to M$, $\pi_M(y,\ell) = T^\ell y$ is a measure-preserving semiconjugacy.*

Then we say that $T\colon M \to M$ is modelled by a (two-sided) Young tower.

Remark 3.3.2. Here we have not assumed that the tower map $f\colon \Delta \to \Delta$ is mixing. However, as in [Che99, Theorem 2.1, Proposition 10.1] and [BMT21] the a priori knowledge that μ is mixing ensures that this is irrelevant.

Remark 3.3.3. (i) The map $\bar{\pi} : Y \to \bar{Y}$ is usually obtained by quotienting along stable leaves. In particular, this is how $\bar{\pi}$ is defined in [You98].
(ii) In [You98] it is assumed that the underlying dynamics T is exponentially contracting along stable leaves. Under this assumption, control of iterated moments follows easily from the nonuniformly expanding case in [KKM22]. Indeed, let $v : M \to \mathbb{R}$ be Hölder. Then the lifted observable $\tilde{v} = v \circ \pi_M$ can be written

in the form $\tilde{v} = w \circ \bar{\pi} + \chi \circ f - \chi$, where $\chi \in L^\infty(\Delta)$ and $w : \bar{\Delta} \to \mathbb{R}$ is d_δ-Lipschitz for some $\delta \in (0,1)$ (see [KKM18, Section 5.1] for more details). Hence it is straightforward to show that iterated moment bounds for Hölder observables on M follow from the corresponding bounds for d_δ-Lipschitz observables on $\bar{\Delta}$.

From now on we suppose that $T\colon M \to M$ is a mixing transformation modelled by a Young tower and that there exist constants $\beta > 1$, $C_\phi > 0$ such that $\mu_Y(\phi \geq n) \leq C_\phi n^{-\beta}$ for all $n \geq 1$. Note that there exist $\delta > 0$ and a finite set $I \subset \mathbb{N}$ with $\gcd\{I\} = \gcd\{\phi(y) : y \in Y\}$ and $\mu_Y(\phi = k) \geq \delta$ for all $k \in I$.

We prove the following more refined version of Theorem 3.2.3:

Theorem 3.3.4. *T satisfies the Functional Correlation Bound with rate $n^{-(\beta-1)}$. Moreover, the implicit constant depends continuously on the system constants D_0, θ, δ, $\max\{I\}$, β and C_ϕ.*

The fact that the implicit constant in the Functional Correlation Bound depends continuously on the above system constants will be important in Chapter 4, where we consider families of nonuniformly hyperbolic maps. Throughout the remainder of this subsection, we use $C > 0$ to denote various constants that depend continuously on the system constants D_0, θ, δ, $\max\{I\}$, β and C_ϕ.

Before proceeding further, we recall and prove some standard facts about Young towers. We quote results from [KKM19] because all of the estimates proved in that article depend continuously on the above system constants.

Let $\psi_n(x) = \#\{j = 1, \ldots, n\colon f^j x \in \Delta_0\}$ denote the number of returns to $\Delta_0 = \{(y,\ell) \in \Delta : \ell = 0\}$ by time n. The following bound is standard, see for example [KKM19, Lemma 5.5].

Lemma 3.3.5. *For all $n \geq 1$, we have $\int_\Delta \theta^{\psi_n}\, d\mu_\Delta \leq C n^{-(\beta-1)}$*

The transfer operator L corresponding to $\bar{f}\colon \bar{\Delta} \to \bar{\Delta}$ and $\bar{\mu}_\Delta$ is given pointwise by

$$(Lv)(x) = \sum_{\bar{f}z = x} g(z)v(z), \text{ where } g(y,\ell) = \begin{cases} \zeta(y), & \ell = \phi(y) - 1, \\ 1, & \ell < \phi(y) - 1 \end{cases}.$$

It follows that for $n \geq 1$, the operator L^n is of the form $(L^n v)(x) = \sum_{\bar{f}^n z = x} g_n(z) v(z)$, where $g_n = \prod_{i=0}^{n-1} g \circ \bar{f}^i$.

We say that $z, z' \in \bar{\Delta}$ are in the same cylinder set of length n if $\bar{f}^k z$ and $\bar{f}^k z'$ lie in the same partition element of $\bar{\Delta}$ for $0 \leq k \leq n-1$. We use the following distortion bound (see e.g. [KKM19, Proposition 5.2]):

Lemma 3.3.6. *For all $n \geq 1$, for all points $z, z' \in \bar{\Delta}$ which belong to the same cylinder set of length n,*

$$|g_n(z) - g_n(z')| \leq C g_n(z) d_\theta(\bar{f}^n z, \bar{f}^n z'). \qquad \square$$

We say that $v \colon \bar{\Delta} \to \mathbb{R}$ is d_θ-Lipschitz if $\|v\|_\theta = |v|_\infty + \mathrm{Lip}(v) < \infty$, where $\mathrm{Lip}(v) = \sup_{x \neq y} |v(x) - v(y)|/d_\theta(x, y)$. If $f \colon \Delta \to \Delta$ is mixing then by [You99],

$$\left| L^n v - \int v \, d\bar{\mu}_\Delta \right|_1 = O(n^{-(\beta-1)} \|v\|_\theta).$$

The same bound holds pointwise on $\bar{\Delta}_0$:

Proposition 3.3.7. *Suppose that $f \colon \Delta \to \Delta$ is mixing. Then for all d_θ-Lipschitz $v : \bar{\Delta} \to \mathbb{R}$, for any $n \geq 1$,*

$$\left| \mathbb{1}_{\bar{\Delta}_0} \left(L^n v - \int_{\bar{\Delta}} v \, d\bar{\mu}_\Delta \right) \right|_\infty \leq C n^{-(\beta-1)} \|v\|_\theta .$$

Proof. We first show that for all d_θ-Lipschitz $w \colon \bar{\Delta} \to \mathbb{R}$ and all $n \geq 0$,

$$\sup_{\bar{\Delta}_0} |L^n w| \leq C(\mathrm{Lip}(w) |\theta^{\psi_n}|_1 + |w|_1). \tag{3.3.4}$$

Let $x, x' \in \bar{\Delta}_0$. Then we can pair preimages z, z' of x, x' so that z, z' are in the same cylinder set of length n. It follows that $s(z, z') = \psi_n(z') + s(x, x')$. Write $(L^n w)(x) - (L^n w)(x') = I_1 + I_2$, where

$$I_1 = \sum_{\bar{f}^n z = x} g_n(z)(w(z) - w(z')), \quad I_2 = \sum_{\bar{f}^n z' = x'} w(z')(g_n(z) - g_n(z')).$$

Note that

$$|I_1| \leq \mathrm{Lip}(w) \sum_{\bar{f}^n z = x} g_n(z) \theta^{\psi_n(z')} d_\theta(x, x') = \mathrm{Lip}(w)(L^n \theta^{\psi_n})(x') d_\theta(x, x').$$

By bounded distortion (Lemma 3.3.6),

$$|I_2| \leq C \sum_{\bar{f}^n z' = x'} |w(z')| g_n(z') d_\theta(x, x') = C(L^n |w|)(x') d_\theta(x, x').$$

It follows that $|(L^n w)(x)| \leq |(L^n w)(x')| + \mathrm{Lip}(w)(L^n \theta^{\psi_n})(x') + C(L^n |w|)(x')$. Hence integrating over $x' \in \bar{\Delta}_0$ gives

$$\begin{aligned}|(L^n w)(x)| &\le \bar{\mu}_\Delta(\bar{\Delta}_0)^{-1} \int_{\bar{\Delta}_0} (|L^n w| + \mathrm{Lip}(w)|L^n \theta^{\psi_n}| + CL^n|w|) d\bar{\mu}_\Delta \\ &\le \bar{\mu}_\Delta(\bar{\Delta}_0)^{-1}\big((1+C)\,|w|_1 + \mathrm{Lip}(w)\left|\theta^{\psi_n}\right|_1\big).\end{aligned}$$

The proof of (3.3.4) follows by noting that $\bar{\mu}_\Delta(\bar{\Delta}_0)^{-1} = \int_{\bar{Y}} \phi\, d\bar{\mu}_Y \le C_\phi \sum_{k\ge 1} k^{-\beta}$.

Finally, let $v\colon \bar{\Delta} \to \mathbb{R}$ be d_θ-Lipschitz and let $k \ge 1$. Without loss of generality take $\int v\, d\bar{\mu}_\Delta = 0$ and set $w = L^{k-[k/2]}v$. By [DP09, Lemma 2.2], $\mathrm{Lip}(w) \le C\,\|v\|_\theta$. By (3.3.4), it follows that

$$\begin{aligned}\sup_{\bar{\Delta}_0} |L^k v| = \sup_{\bar{\Delta}_0} |L^{[k/2]} w| &\le C(\mathrm{Lip}(w)|\theta^{\psi_{[k/2]}}|_1 + |w|_1) \\ &\le C(\|v\|_\theta\,|\theta^{\psi_{[k/2]}}|_1 + |L^{k-[k/2]}v|_1).\end{aligned}$$

Now by [KKM19, Theorem 2.7], $\left|L^{k-[k/2]}v\right|_1 \le C(k-[k/2])^{-(\beta-1)}\,\|v\|_\theta$. By Lemma 3.3.5, we have $|\theta^{\psi_{[k/2]}}|_1 \le C[k/2]^{-(\beta-1)}$. It follows that $\sup_{\bar{\Delta}_0} |L^k v| \le Ck^{-(\beta-1)}\,\|v\|_\theta$, as required. □

When studying systems modelled by Young towers, it is natural to consider the following class of observables:

Dynamically Hölder observables Fix $\theta \in (0,1)$. For $v\colon M \to \mathbb{R}$, define

$$\|v\|_{\mathcal{H}} = |v|_\infty + [v]_{\mathcal{H}}, \quad [v]_{\mathcal{H}} = \sup_{y,y'\in Y, y\ne y'} \sup_{0\le \ell<\phi(y)} \frac{|v(T^\ell y) - v(T^\ell y')|}{d(y,y') + \theta^{s(y,y')}}.$$

We say that v is *dynamically Hölder* if $\|v\|_{\mathcal{H}} < \infty$ and denote by $\mathcal{H}(M)$ the space of all such observables.

Lemma 3.3.8. *Let $v\colon M \to \mathbb{R}$ be Lipschitz. Then $[v]_{\mathcal{H}} \le D_0 \mathrm{Lip}(v)$.*

Proof. Let $y, y' \in Y$, $0 \le \ell < \phi(y)$. By (3.3.3),

$$\begin{aligned}|v(T^\ell y) - v(T^\ell y')| &\le \mathrm{Lip}(v) d(T^\ell y, T^\ell y') \\ &\le D_0 \mathrm{Lip}(v)(d(y,y') + \theta^{s(y,y')}).\end{aligned}$$

□

Let $q \ge 1$. Given a function $G\colon M^q \to \mathbb{R}$ and $0 \le i < q$ we denote

$$[G]_{\mathcal{H},i} = \sup_{x_0,\ldots,x_{q-1}\in M} [G(x_0,\ldots,x_{i-1},\cdot,x_{i+1},\ldots,x_{q-1})]_{\mathcal{H}}.$$

We call the function G *separately dynamically Hölder*, and write $G \in \mathcal{SH}_q(M)$, if

$|G|_\infty + \sum_{i=0}^{q-1} [G]_{\mathcal{H},i} < \infty$.

3.3.2 Reduction to the case of a mixing Young tower

In proofs involving Young towers it is often useful to assume that the Young tower is mixing, i.e. $gcd\{\phi(y) : y \in Y\} = 1$. It is also natural to consider dynamically Hölder observables. Hence in subsequent subsections we focus on the case where the Young tower is mixing and prove a functional correlation for separately dynamically Hölder functions:

Proposition 3.3.9. *Suppose that T is modelled by a mixing two-sided Young tower such that $\mu(\phi \geq k) \leq C_\phi n^{-\beta}$ for all $n \geq 1$. Then for all integers $0 \leq p < q$, $0 \leq n_0 \leq \cdots \leq n_{q-1}$, we have*

$$\Big| \int_M G(T^{n_0}x, \ldots, T^{n_{q-1}}x) d\mu(x) \\ - \int_{M^2} G(T^{n_0}x_0, \ldots, T^{n_{p-1}}x_0, T^{n_p}x_1, \ldots, T^{n_{q-1}}x_1) d\mu(x_0) d\mu(x_1) \Big| \\ \leq C(n_p - n_{p-1})^{-\gamma} \bigg(|G|_\infty + \sum_{i=0}^{q-1} [G]_{\eta,i} \bigg) \tag{3.3.5}$$

for any $G \in \mathcal{SH}_q(M)$.

Proof of Theorem 3.3.4. We can take $\eta = 1$ without loss of generality. Indeed, let $d_\eta(x,y) = d(x,y)^\eta$. Let T be a transformation on (M,d) modelled by a Young tower. Then all the assumptions in the definition of being modelled by a Young tower are satisfied on (M, d_η) with slightly different constants. Moreover, separately Hölder functions with respect to d are separately Lipschitz with respect to d_η.

Let $d = \gcd\{\phi(y) : y \in Y\}$. Set $T' = T^d$ and $\phi' = \phi/d$. Construct a mixing two-sided Young tower $\Delta' = Y^{\phi'}$, with tower measure μ'_Δ. Define $\pi'_M \colon \Delta' \to M$ by $\pi'_M(y,\ell) = (T')^\ell y$. Then T' is modelled by Δ' with ergodic, T'-invariant measure $(\pi'_M)_*\mu'_\Delta$. Next we show that $\mu = (\pi'_M)_*\mu'_\Delta$ by adapting an argument from [BMT21, Section 4.1].

By assumption $\mu = (\pi_M)_*\mu_\Delta$ so μ_Y is absolutely continuous with respect to μ. Let $B \subset M$ be a Borel set such that $\mu(B) = 0$. Then for all $\ell \geq 0$ we have $\mu((T')^{-\ell}B) = \mu(B) = 0$ so $\mu_Y((T')^{-\ell}B) = 0$. Now

$$(\pi'_M)^{-1}B \subset \bigcup_{\ell \geq 0} [(T')^{-\ell}B \cap Y] \times \{\ell\}$$

so

$$(\pi_M)_*\mu'_\Delta(B) \le \sum_{\ell\ge 0} \mu'_\Delta([(T')^{-\ell}B\cap Y]\times\{\ell\})$$
$$= \frac{1}{\int_Y \phi'\, d\mu_Y} \sum_{\ell\ge 0} \mu_Y((T')^{-\ell}B) = 0.$$

Hence $(\pi_M)_*\mu'$ is absolutely continuous with respect to μ. Now μ is mixing for T by assumption and thus ergodic for T'. Since distinct ergodic measures are mutually singular, it follows that $\mu = (\pi'_M)_*\mu'_\Delta$.

Let $G \in \mathcal{H}^\eta_q(M)$ and fix integers $0 \le n_0 \le \cdots \le n_{q-1}$. Define $n'_i = [n_i/d], r_i = n_i \mod d$. We need to bound

$$\nabla G = \int_M G(T^{n_0}x,\dots,T^{n_{q-1}}x)d\mu(x)$$
$$- \int_{M^2} G(T^{n_0}x_0,\dots,T^{n_{p-1}}x_0,T^{n_p}x_1,\dots,T^{n_{q-1}}x_1)d\mu(x_0)d\mu(x_1).$$

Define $G' \colon M^q \to \mathbb{R}$ by $G'(x_0,\dots,x_{q-1}) = G(T^{r_0}x_0,\dots,T^{r_{q-1}}x_{q-1})$. Then

$$\nabla G = \int_M G'((T')^{n'_0}x,\dots,(T')^{n'_{q-1}}x)d\mu(x)$$
$$- \int_{M^2} G'((T')^{n'_0}x_0,\dots,(T')^{n'_{p-1}}x_0,(T')^{n'_p}x_1,\dots,(T')^{n'_{q-1}}x_1)d\mu(x_0)d\mu(x_1).$$

Let $[\cdot]_{\mathcal{H}'}$ denote the dynamically Hölder seminorm as defined with T', ϕ' in place of T, ϕ. Then by Proposition 3.3.9,

$$|\nabla G| \le C(n'_p - n'_{p-1})^{-\gamma}\Big(\big|G'\big|_\infty + \sum_{i=0}^{q-1}[G']_{\mathcal{H}',i}\Big)$$
$$\le Cd^\gamma(n_p - n_{p-1} - d)^{-\gamma}\Big(|G|_\infty + \sum_{i=0}^{q-1}[G']_{\mathcal{H}',i}\Big).$$

Now fix $0 \le i < q$. Let $x_0,\dots,x_{q-1} \in M$ and write

$$v'(y) = G'(x_0,\dots,x_{i-1},y,x_{i+1},\dots,x_{q-1})$$
$$= G(T^{r_0}x_0,\dots,T^{r_{i-1}}x_{i-1},T^{r_i}y,T^{r_{i+1}}x_{i+1},\dots,T^{r_{q-1}}x_{q-1}) = v(T^{r_i}y).$$

Let $y, y' \in Y$ and $0 \le \phi'(y) < \ell$. Then

$$|v'((T')^\ell y) - v'((T')^\ell y')| = |v(T^{d\ell+r_i}y) - v(T^{d\ell+r_i}y')| \le [G]_{\mathcal{H},i}(d(y,y') + \theta^{s(y,y')}),$$

so $[G']_{\mathcal{H}',i} \leq [G]_{\mathcal{H},i}$. Finally, by Lemma 3.3.8 we have $[G]_{\mathcal{H},i} \leq D_0[G]_{1,i}$. □

3.3.3 Approximation by one-sided functions

Let $0 \leq p < q$ and $0 \leq n_0 \leq \cdots \leq n_{q-1}$ be integers and consider a function $G \in \mathcal{SH}_q(M)$. In order to prove Proposition 3.3.9, we wish to bound

$$\nabla G = \int_M G(T^{n_0}x, \ldots, T^{n_{q-1}}x)d\mu(x) \\ - \int_{M^2} G(T^{n_0}x_0, \ldots, T^{n_{p-1}}x_0, T^{n_p}x_1, \ldots, T^{n_{q-1}}x_1)d\mu^2(x_0, x_1).$$

Now since $\pi_M \colon \Delta \to M$ is a measure-preserving semiconjugacy

$$\nabla G = \int_\Delta \widetilde{H}(x, f^{n_p}x)d\mu_\Delta(x) - \int_{\Delta^2} \widetilde{H}(x_0, x_1)d\mu_\Delta^2(x_0, x_1) = \nabla\widetilde{H} \tag{3.3.6}$$

where $\widetilde{H} \colon \Delta^2 \to \mathbb{R}$ is given by

$$\widetilde{H}(x, y) = \widetilde{G}(f^{n_0}x, f^{n_1}x, \ldots, f^{n_{p-1}}x, f^{k_p}y, f^{k_{p+2}}y, \ldots, f^{k_{q-1}}y),$$

where $\widetilde{G} = G \circ \pi_M$ and $k_i = n_i - n_p$.

Let $R \geq 1$. We approximate $\widetilde{H}(f^R\cdot, f^R\cdot)$ by a function $\widetilde{H}_R$ that projects down onto $\bar{\Delta}$. Our approach is based on ideas from Appendix B of [MT14].

Recall that $\psi_R(x) = \#\{j = 1, \ldots, R : f^jx \in \Delta_0\}$ denotes the number of returns to $\Delta_0 = \{(y, \ell) \in \Delta : \ell = 0\}$ by time R. Let $\mathcal{Q}_R$ denote the at most countable, measurable partition of Δ with elements of the form $\{x' \in \Delta : s(x, x') > 2\psi_R(x)\}$, $x \in \Delta$. Choose a reference point in each partition element of $\mathcal{Q}_R$. For $x \in \Delta$ let $\hat{x}$ denote the reference point of the element that x belongs to. Define $\widetilde{H}_R \colon \Delta^2 \to \mathbb{R}$ by

$$\widetilde{H}_R(x, y) = \widetilde{G}\Big(f^R(\widehat{f^{n_0}x}), \ldots, f^R(\widehat{f^{n_{p-1}}x}), f^R(\widehat{f^{k_p}y}), \ldots, f^R(\widehat{f^{k_{q-1}}y})\Big).$$

Lemma 3.3.10. *The function $\widetilde{H}_R$ lies in $L^\infty(\Delta^2)$ and projects down to a function $\bar{H}_R \in L^\infty(\bar{\Delta}^2)$. Moreover,*

(i) $\left|\bar{H}_R\right|_\infty = |\widetilde{H}_R|_\infty \leq |G|_\infty$.

(ii) For all $x, y \in \Delta$,

$$|\widetilde{H}(f^Rx, f^Ry) - \widetilde{H}_R(x, y)| \leq C\left(\sum_{i=0}^{p-1}[G]_{\mathcal{H},i}\,\theta^{\psi_R(f^{n_i}x)} + \sum_{i=p}^{q-1}[G]_{\mathcal{H},i}\,\theta^{\psi_R(f^{k_i}y)}\right).$$

(iii) For all $\bar{y} \in \bar{\Delta}$,

$$\left\| L^{R+n_{p-1}} \bar{H}_R(\cdot, \bar{y}) \right\|_\theta \le C\Big(|G|_\infty + \sum_{i=0}^{p-1} [G]_{\mathcal{H},i} \Big).$$

Here we recall that $\|\cdot\|_\theta$ denotes the d_θ-Lipschitz norm, which is given by $\|v\|_\theta = |v|_\infty + \sup_{x \neq y} |v(x) - v(y)|/d_\theta(x,y)$ for $v \colon \bar{\Delta} \to \mathbb{R}$.

Proof. We follow the proof of Proposition 7.9 in [BMT21]. By definition $\widetilde{H}_R$ is piecewise constant on a measurable partition of Δ^2. Moreover, this partition projects down to a measurable partition on $\bar{\Delta}$, since it is defined in terms of s and ψ_R which both project down to $\bar{\Delta}$. It follows that $\bar{H}_R$ is well-defined and measurable. Part (i) is immediate.

Let $x, y \in \Delta$. Write $\widetilde{H}(f^R x, f^R y) - \widetilde{H}_R(x,y) = I_1 + I_2$ where

$$\begin{aligned} I_1 &= \widetilde{G}\Big(f^R(f^{n_0}x), \dots, f^R(f^{n_{p-1}}x), f^R(f^{k_p}y), \dots, f^R(f^{k_{q-1}}y) \Big) \\ &\qquad - \widetilde{G}\Big(f^R(\widehat{f^{n_0}x}), \dots, f^R(\widehat{f^{n_{p-1}}x}), f^R(f^{k_p}y), \dots, f^R(f^{k_{q-1}}y) \Big), \\ I_2 &= \widetilde{G}\Big(f^R(\widehat{f^{n_0}x}), \dots, f^R(\widehat{f^{n_{p-1}}x}), f^R f^{k_p} y, \dots, f^R f^{k_{q-1}} y \Big) \\ &\qquad - \widetilde{G}\Big(f^R(\widehat{f^{n_0}x}), \dots, f^R(\widehat{f^{n_{p-1}}x}), f^R(\widehat{f^{k_p}y}), \dots, f^R(\widehat{f^{k_{q-1}}y}) \Big). \end{aligned}$$

Let $a_i = f^{n_i}x$ and $b_i = f^R f^{k_i} y$. By successively substituting a_i by $\hat{a}_i$,

$$\begin{aligned} I_1 &= \widetilde{G}(f^R a_0, \dots, f^R a_{p-1}, b_p, \dots, b_{q-1}) - \widetilde{G}(f^R \hat{a}_0, \dots, f^R \hat{a}_{p-1}, b_p, \dots, b_{q-1}) \\ &= \sum_{i=0}^{p-1} \big(\widetilde{G}(f^R a_0, \dots, f^R a_{i-1}, f^R a_i, f^R \hat{a}_{i+1}, f^R \hat{a}_{p-1}, b_p, \dots, b_{q-1}) \\ &\qquad - \widetilde{G}(f^R a_0, \dots, f^R a_{i-1}, f^R \hat{a}_i, f^R \hat{a}_{i+1}, f^R \hat{a}_{p-1}, b_p, \dots, b_{q-1}) \big) \\ &= \sum_{i=0}^{p-1} \big(\tilde{v}_i(f^R a_i) - \tilde{v}_i(f^R \hat{a}_i) \big) \end{aligned} \tag{3.3.7}$$

where $\tilde{v}_i(x) = \tilde{G}(f^R a_0, \dots, f^R a_{i-1}, x, f^R \hat{a}_{i+1}, \dots, f^R \hat{a}_{p-1}, b_p, \dots, b_{q-1})$.

Fix $0 \le i < p$. Since a_i and $\hat{a}_i$ are in the same partition element, $s(a_i, \hat{a}_i) > 2\psi_R(a_i)$. Write $a_i = (y, \ell), \hat{a}_i = (\hat{y}, \ell)$. Then $f^R a_i = (F^{\psi_R(a_i)} y, \ell_1)$ and similarly $f^R \hat{a}_i = (F^{\psi_R(a_i)} \hat{y}, \ell_1)$, where $\ell_1 = \ell + R - \Phi_{\psi_R(a_i)}(y)$. (Here, $\Phi_k = \sum_{j=0}^{k-1} \phi \circ F^k$.)

Now by the definition of $[G]_{\mathcal{H},i}$ and (3.3.1),

$$\begin{aligned}|\tilde{v}_i(f^R a_i) - \tilde{v}_i(f^R \hat{a}_i)| &= |\tilde{v}_i(F^{\psi_R(a_i)} y, \ell_1) - \tilde{v}_i(F^{\psi_R(a_i)} \hat{y}, \ell_1)| \\ &\le [G]_{\mathcal{H},i}(d(F^{\psi_R(a_i)} y, F^{\psi_R(a_i)} y') + \theta^{s(F^{\psi_R(a_i)} y, F^{\psi_R(a_i)} y')}) \\ &\le (D_0+1)[G]_{\mathcal{H},i}(\theta^{\psi_R(a_i)} + \theta^{s(a_i,a_i')-\psi_R(a_i)}) \\ &\le 2(D_0+1)[G]_{\mathcal{H},i}\theta^{\psi_R(a_i)}.\end{aligned}$$

Thus

$$|I_1| \le 2(D_0+1)\sum_{i=0}^{p-1}[G]_{\mathcal{H},i}\theta^{\psi_R(f^{n_i}x)}.$$

By a similar argument,

$$|I_2| \le 2(D_0+1)\sum_{i=p}^{q-1}[G]_{\mathcal{H},i}\theta^{\psi_R(f^{k_i}y)},$$

completing the proof of (ii).

Let $\bar{x}, \bar{x}', \bar{y} \in \bar{\Delta}$. Recall that

$$L^{R+n_{p-1}}\bar{H}_R(\cdot,\bar{y})(\bar{x}) = \sum_{\bar{f}^{R+n_{p-1}}\bar{z}=\bar{x}} g_{R+n_{p-1}}(\bar{z})\bar{H}_R(\bar{z},\bar{y}).$$

It follows that $\left|L^{R+n_{p-1}}\bar{H}_R(\cdot,\bar{y})\right|_\infty \le \left|\bar{H}_R\right|_\infty \le |G|_\infty$. If $d_\theta(\bar{x},\bar{x}')=1$, then

$$|L^{R+n_{p-1}}\bar{H}_R(\cdot,\bar{y})(\bar{x}) - L^{R+n_{p-1}}\bar{H}_R(\cdot,\bar{y})(\bar{x}')| \le 2\,|G|_\infty = 2\,|G|_\infty\, d_\theta(\bar{x},\bar{x}').$$

Otherwise, we can write $L^{n_{p-1}+R}\bar{H}_R(\cdot,\bar{y})(\bar{x}) - L^{n_{p-1}+R}\bar{H}_R(\cdot,\bar{y})(\bar{x}') = J_1 + J_2$ where

$$\begin{aligned}J_1 &= \sum_{\bar{f}^{n_{p-1}+R}\bar{z}=\bar{x}} \big(g_{n_{p-1}+R}(\bar{z}) - g_{n_{p-1}+R}(\bar{z}')\big)\bar{H}_R(\bar{z},\bar{y}), \\ J_2 &= \sum_{\bar{f}^{n_{p-1}+R}\bar{z}'=\bar{x}'} g_{n_{p-1}+R}(\bar{z}')\big(\bar{H}_R(\bar{z},\bar{y}) - \bar{H}_R(\bar{z}',\bar{y})\big).\end{aligned}$$

Here, as usual we have paired preimages $\bar{z}, \bar{z}'$ that lie in the same cylinder set of length $n_{p-1}+R$. By bounded distortion (Lemma 3.3.6), $|J_1| \le C\,|G|_\infty\, d_\theta(\bar{x},\bar{x}')$. We claim that $|\bar{H}_R(\bar{z},\bar{y}) - \bar{H}_R(\bar{z}',\bar{y})| \ll \sum_{i=0}^{p-1}[G]_{\mathcal{H},i} d_\theta(\bar{x},\bar{x}')$. It follows that $|J_2| \ll \sum_{i=0}^{p-1}[G]_{\mathcal{H},i} d_\theta(\bar{x},\bar{x}')$.

It remains to prove the claim. Choose points $z, z', y \in \Delta$ that project to

$\bar{z}, \bar{z}', \bar{y}$. Let $a_i = f^{n_i} z, a_i' = f^{n_i} z', b_i = f^{R+n_i} y$. As in part (ii),

$$\bar{H}_R(\bar{z}, \bar{y}) - \bar{H}_R(\bar{z}', \bar{y}) = \widetilde{H}_R(z, y) - \widetilde{H}_R(z', y) = \sum_{i=0}^{p-1} (\tilde{w}_i(f^R \hat{a}_i) - \tilde{w}_i(f^R \hat{a}_i'))$$

where $\tilde{w}_i(x) = \tilde{G}(f^R \hat{a}_0, \ldots, \hat{a}_{i-1}, x, f^R \hat{a}'_{i+1}, \ldots, \hat{a}'_{p-1}, \hat{b}_p, \ldots, \hat{b}_{q-1})$.

Let $0 \leq i < p$. We bound $E_i = \tilde{w}_i(f^R \hat{a}_i) - \tilde{w}_i(f^R \hat{a}_i')$. Without loss of generality suppose that

$$\psi_R(\hat{a}_i') \geq s(\hat{a}_i, \hat{a}_i') - \psi_R(\hat{a}_i),$$

for otherwise $\hat{a}_i$ and $\hat{a}_i'$ are reference points of the same partition element so $\hat{a}_i = \hat{a}_i'$ and $E_i = 0$. Now as in part (ii),

$$E_i \leq (D_0 + 1)(\theta^{\psi_R(\hat{a}_i)} + \theta^{s(\hat{a}_i, \hat{a}_i') - \psi_R(\hat{a}_i)})[G]_{\mathcal{H},i}.$$

Note that

$$s(\hat{a}_i, \hat{a}_i') - \psi_R(\hat{a}_i) \geq \min\{s(\hat{a}_i, a_i), s(a_i, a_i'), s(a_i', \hat{a}_i')\} - \psi_R(\hat{a}_i).$$

Since $\bar{z}, \bar{z}'$ lie in the same cylinder set of length $R + n_{p-1}$, we have $\psi_R(a_i) = \psi_R(a_i')$ and

$$\begin{aligned} s(a_i, a_i') = s(\bar{f}^{n_i} \bar{z}, \bar{f}^{n_i} \bar{z}') &= s(\bar{x}, \bar{x}') + \psi_{R+n_{p-1}-n_i}(\bar{f}^{n_i} \bar{z}) \\ &\geq s(\bar{x}, \bar{x}') + \psi_R(a_i). \end{aligned}$$

Now a_i and $\hat{a}_i$ are contained in the same partition element so $s(\hat{a}_i, a_i) - \psi_R(\hat{a}_i) \geq \psi_R(\hat{a}_i)$ and

$$\psi_R(\hat{a}_i) = \psi_R(a_i) = \psi_R(a_i') = \psi_R(\hat{a}_i').$$

Hence $s(\hat{a}_i, \hat{a}_i') - \psi_R(\hat{a}_i) \geq \min\{s(\bar{x}, \bar{x}'), \psi_R(a_i)\}$. It follows that $E_i \leq 2(D_0 + 1)\theta^{s(\bar{x}, \bar{x}')}[G]_{\mathcal{H},i}$, completing the proof of the claim. □

3.3.4 Proof of Proposition 3.3.9

We continue to assume that $\beta > 1$ and that $\mu_Y(\phi \geq n) \leq C_\phi n^{-\beta}$. We also assume that $\gcd\{\phi(y) : y \in Y\} = 1$ so that $f \colon \Delta \to \Delta$ is mixing. We say that $V \colon \bar{\Delta}^2 \to \mathbb{R}$ is d_θ-Lipschitz in x_0 if

$$\|V\|_{\theta,0} = \sup_{x_1 \in \Delta} \|V(\cdot, x_1)\|_\theta < \infty.$$

Proposition 3.3.11. *For any $V \in L^\infty(\bar{\Delta}^2)$, we have*

$$\left|\int_{\bar{\Delta}} V(x, \bar{f}^n x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1)\right| \le C n^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta$$

for all $n \ge 1$.

Remark 3.3.12. Let $V(x, y) = v(x)w(y)$ where v is d_θ-Lipschitz and $w \in L^\infty(\bar{\Delta})$. Then we obtain that

$$\left|\int_{\bar{\Delta}} v\, w \circ \bar{f}^n d\bar{\mu}_\Delta - \int_{\bar{\Delta}} v\, d\bar{\mu}_\Delta \int_{\bar{\Delta}} w\, d\bar{\mu}_\Delta\right| \le C n^{-(\beta-1)} \|v\|_\theta |w|_\infty ,$$

so Proposition 3.3.11 can be seen as a generalisation of the usual upper bound on decay of correlations for observables on the one-sided tower $\bar{\Delta}$.

Remark 3.3.13. Our proof of Proposition 3.3.11 is based on ideas from [CG12, Section 4]. However, we have chosen to present the proof in full because (i) our assumptions are weaker, in particular we only require $\beta > 1$ instead of $\beta > 2$ and V need not be separately d_θ-Lipschitz and (ii) we avoid introducing Markov chains.

Proof of Proposition 3.3.11. Write $v(x) = V(x, \bar{f}^n x)$ so

$$\begin{aligned}\int_{\bar{\Delta}} V(x, f^n x)\, d\bar{\mu}_\Delta(x) &= \int_{\bar{\Delta}} v\, d\bar{\mu}_\Delta = \int_{\bar{\Delta}} L^n v\, d\bar{\mu}_\Delta \\ &= \int_{\bar{\Delta}} \sum_{\bar{f}^n z = x} g_n(z) V(z, \bar{f}^n z) d\bar{\mu}(x) \\ &= \int_{\bar{\Delta}} \sum_{\bar{f}^n z = x} g_n(z) V(z, x) d\bar{\mu}_\Delta(x) = \int_{\bar{\Delta}} (L^n u_x)(x)\, d\bar{\mu}_\Delta(x).\end{aligned}$$

where $u_x(z) = V(z, x)$. Let $\bar{\Delta}_\ell = \{(y, j) \in \bar{\Delta} : j = \ell\}$ denote the ℓ-th level of $\bar{\Delta}$. It follows that we can decompose

$$\int_{\bar{\Delta}} V(x, \bar{f}^n x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1) = \sum_{\ell \ge 0} A_\ell$$

where

$$A_\ell = \int_{\bar{\Delta}_\ell} \left((L^n u_x)(x) - \int_{\bar{\Delta}} V(z, x) d\bar{\mu}_\Delta(z)\right) d\bar{\mu}_\Delta(x).$$

For all $\ell \ge 0$,

$$|A_\ell| \le 2 |V|_\infty \bar{\mu}_\Delta(\bar{\Delta}_\ell) = 2 |V|_\infty \frac{\bar{\mu}_Y(\phi > \ell)}{\int \phi d\bar{\mu}_Y} \ll |V|_\infty (\ell + 1)^{-\beta}.$$

Hence,

$$\sum_{\ell \geq n/2} |A_\ell| \ll |V|_\infty \, n^{-(\beta-1)}.$$

Let $x \in \bar{\Delta}_\ell$, $\ell \leq n$. Then $(L^n u_x)(x) = (L^{n-\ell} u_x)(x_0)$ where $x_0 \in \bar{\Delta}_0$ is the unique preimage of x under $\bar{f}^\ell$. Thus by Proposition 3.3.7,

$$|A_\ell| \leq \int_{\bar{\Delta}_\ell} C(n-\ell)^{-(\beta-1)} \, \|V(\cdot, x)\|_\theta \, d\bar{\mu}_\Delta \leq C(n-\ell)^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \, \bar{\mu}_\Delta(\bar{\Delta}_\ell).$$

Hence,

$$\begin{aligned}
\sum_{\ell \leq n/2} |A_\ell| &\leq C(n/2)^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \sum_{\ell \leq n/2} \bar{\mu}_\Delta(\bar{\Delta}_\ell) \\
&\leq C(n/2)^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \,,
\end{aligned}$$

completing the proof. □

Proof of Proposition 3.3.9. Recall that we wish to bound

$$\nabla \widetilde{H} = \int_\Delta \widetilde{H}(x, f^{n_p} x) d\mu_\Delta(x) - \int_{\Delta^2} \widetilde{H}(x_0, x_1) d\mu_\Delta^2(x_0, x_1).$$

Without loss of generality assume that $n_p - n_{p-1} \geq 2$. Let $R = [(n_p - n_{p-1})/2]$. Write $\nabla \widetilde{H} = I_1 + I_2 + \nabla \bar{H}_R$ where

$$\begin{aligned}
I_1 &= \int_\Delta \widetilde{H}(x, f^{n_p} x) d\mu_\Delta(x) - \int_\Delta \widetilde{H}_R(x, f^{n_p} x) d\mu_\Delta(x), \\
I_2 &= \int_\Delta \widetilde{H}_R(x_0, x_1) d\mu_\Delta^2(x_0, x_1) - \int_{\Delta^2} \widetilde{H}(x_0, x_1) d\mu_\Delta^2(x_0, x_1), \\
\nabla \bar{H}_R &= \int_\Delta \widetilde{H}_R(x, f^{n_p} x) d\mu_\Delta(x) - \int_{\Delta^2} \widetilde{H}_R(x_0, x_1) d\mu_\Delta^2(x_0, x_1) \\
&= \int_{\bar{\Delta}} \bar{H}_R(x, \bar{f}^{n_p} x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} \bar{H}_R(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1).
\end{aligned}$$

Now by Lemma 3.3.10(ii) and Lemma 3.3.5,

$$
\begin{aligned}
|I_1| &= \left| \int_\Delta \widetilde{H}(f^R x, f^{R+n_p} x) d\mu_\Delta(x) - \int_\Delta \widetilde{H}_R(x, f^{n_p} x) d\mu_\Delta(x) \right| \\
&\ll \int_\Delta \left(\sum_{i=0}^{p-1} [G]_{\mathcal{H},i} \theta^{\psi_R(f^{n_i} x)} + \sum_{i=p}^{q-1} [G]_{\mathcal{H},i} \theta^{\psi_R(f^{n_p+k_i} x)} \right) d\mu_\Delta(x) \\
&= \sum_{i=0}^{q-1} [G]_{\mathcal{H},i} \int_\Delta \theta^{\psi_R} d\mu_\Delta \ll \sum_{i=0}^{q-1} [G]_{\mathcal{H},i} R^{-(\beta-1)}. \qquad (3.3.8)
\end{aligned}
$$

Similarly,

$$
|I_2| \ll \sum_{i=0}^{q-1} [G]_{\mathcal{H},i} R^{-(\beta-1)}. \qquad (3.3.9)
$$

Now let $u_y(z) = \bar{H}_R(z, y)$ and $V(x, y) = (L^{n_{p-1}+R} u_y)(x)$. Then

$$
\int_{\bar{\Delta}^2} V(x_0, x_1)\, d\bar{\mu}^2_\Delta(x_0, x_1) = \int_{\bar{\Delta}^2} \bar{H}_R(x_0, x_1)\, d\bar{\mu}^2_\Delta(x_0, x_1) \qquad (3.3.10)
$$

and

$$
\begin{aligned}
V(x, \bar{f}^{n_p - n_{p-1} - R} x) &= \sum_{\bar{f}^{n_{p-1}+R} z = x} g_{n_{p-1}+R}(z) \bar{H}_R(z, \bar{f}^{n_p - n_{p-1} - R} x) \\
&= \sum_{\bar{f}^{n_{p-1}+R} z = x} g_{n_{p-1}+R}(z) \bar{H}_R(z, \bar{f}^{n_p} z) = (L^{n_{p-1}+R} \hat{u})(x)
\end{aligned}
$$

where $\hat{u}(z) = \bar{H}_R(z, \bar{f}^{n_p} z)$. Hence

$$
\begin{aligned}
\int_{\bar{\Delta}} V(x, \bar{f}^{n_p - n_{p-1} - R} x) d\bar{\mu}_\Delta(x) &= \int_{\bar{\Delta}} L^{n_{p-1}+R} \hat{u}\, d\bar{\mu}_\Delta \\
&= \int_{\bar{\Delta}} \hat{u}\, d\bar{\mu}_\Delta = \int_{\bar{\Delta}} \bar{H}_R(x, \bar{f}^{n_p} x)\, d\bar{\mu}_\Delta(x). \qquad (3.3.11)
\end{aligned}
$$

Now by Lemma 3.3.10(iii), $\sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \ll |G|_\infty + \sum_{i=0}^{p-1} [G]_{\mathcal{H},i}$. By Proposition 3.3.11, (3.3.10) and (3.3.11) it follows that

$$
\begin{aligned}
|\nabla \bar{H}_R| &= \left| \int_{\bar{\Delta}} V(x, \bar{f}^{n_p - n_{p-1} - R} x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}^2_\Delta(x_0, x_1) \right| \\
&\ll \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta\, (n_p - n_{p-1} - R)^{-(\beta-1)} \\
&\ll \left(|G|_\infty + \sum_{i=0}^{p-1} [G]_{\mathcal{H},i} \right) (n_p - n_{p-1} - R)^{-(\beta-1)}. \qquad (3.3.12)
\end{aligned}
$$

Recall that $R = [(n_p - n_{p-1})/2]$. Hence $n_p - n_{p-1} - R \geq R$. By combining (3.3.8), (3.3.9) and (3.3.12) it follows that

$$|\nabla \widetilde{H}| \ll \sum_{i=0}^{q-1} [G]_{\mathcal{H},i}([(n_p - n_{p-1})/2])^{-(\beta-1)},$$

as required. □

3.4 An abstract weak dependence condition

We now give a weak dependence lemma that follows from the Functional Correlation Bound. In most of our applications we use this lemma rather than applying the Functional Correlation Bound directly.

Let $e, q \geq 1$ be integers. For $G = (G_1, \ldots, G_e) : M^q \to \mathbb{R}^e$ and $0 \leq i < q$ we define $[G]_{\eta,i} = \sum_{j=1}^{e} [G_j]_{\eta,i}$. We write $G \in \mathcal{H}_q^\eta(M, \mathbb{R}^e)$ if $|G|_\infty + \sum_{i=0}^{q-1} [G]_{\eta,i} < \infty$.

Let $k \geq 1$ and consider k disjoint blocks of integers $\{\ell_i, \ell_i + 1, \ldots, u_i\}$, $0 \leq i < k$ with $\ell_i \leq u_i < \ell_{i+1}$. Consider $\mathbb{R}^e$-valued random vectors X_i on (M, μ) of the form

$$X_i(x) = \Phi_i(T^{\ell_i} x, \ldots, T^{u_i} x)$$

where $\Phi_i \in \mathcal{H}^\eta_{u_i - \ell_i + 1}(M, \mathbb{R}^e)$, $0 \leq i < k$.

When the gaps $\ell_{i+1} - u_i$ between blocks are large, the random vectors $X_0, \ldots, X_{k-1}$ are weakly dependent. Let $\widehat{X}_0, \ldots, \widehat{X}_{k-1}$ be independent random vectors with $\widehat{X}_i =_d X_i$.

Lemma 3.4.1. *Suppose that T satisfies the Functional Correlation Bound with rate $n^{-\gamma}$ for some $\gamma > 0$. Let $R = \max_i |\Phi_i|_\infty$. Then for all Lipschitz $F : B(0, R)^k \to \mathbb{R}$,*

$$\begin{aligned}&\big|\mathbb{E}_\mu[F(X_0, \ldots, X_{k-1})] - \mathbb{E}[F(\widehat{X}_0, \ldots, \widehat{X}_{k-1})]\,\big| \\ &\qquad\leq C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \Big(|F|_\infty + \operatorname{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\eta,j} \Big),\end{aligned}$$

where $C > 0$ is the constant that appears in the Functional Correlation Bound (3.2.1)*.*

Recall that we have endowed $\mathbb{R}^k$ with the ℓ^1 norm so $|y| = \sum_{i=1}^k |y_i|$.

Proof. We proceed by induction on k. For $k = 1$ the inequality is trivial. Assume that this lemma holds for $k \geq 1$.

Consider an enriched probability space which contains independent copies of $\{X_i\}$ and $\{\widehat{X}_i\}$. Write

$$\mathbb{E}_\mu[F(X_0,\ldots,X_k)] - \mathbb{E}[F(\widehat{X}_0,\ldots,\widehat{X}_k)] = I_1 + I_2$$

where

$$\begin{aligned} I_1 &= \mathbb{E}[F(X_0,\ldots,X_{k-1},\widehat{X}_k)] - \mathbb{E}[F(\widehat{X}_0,\ldots,\widehat{X}_k)]\,, \\ I_2 &= \mathbb{E}_\mu[F(X_0,\ldots,X_k)] - \mathbb{E}[F(X_0,\ldots,X_{k-1},\widehat{X}_k)]\,. \end{aligned}$$

Since $\widehat{X}_k =_d X_k$ and $\widehat{X}_k$ is independent of $X_0,\ldots,X_{k-1}$ and $\widehat{X}_0,\ldots,\widehat{X}_{k-1}$,

$$I_1 = \int_M \big(\mathbb{E}_\mu\big[F\big(X_0,\ldots,X_{k-1},X_k(y)\big)\big] - \mathbb{E}\big[F\big(\widehat{X}_0,\ldots,\widehat{X}_{k-1},X_k(y)\big)\big]\,\big)d\mu(y).$$

Let $y \in M$. The function $F_y = F(\cdot,\ldots,\cdot,X_k(y))\colon M^k \to \mathbb{R}$ satisfies $\mathrm{Lip}(F_y) \le \mathrm{Lip}(F)$. Hence by the inductive hypothesis,

$$\begin{aligned} |I_1| &\le \int \big|\mathbb{E}_\mu[F_y(X_0,\ldots,X_{k-1})] - \mathbb{E}[F_y(\widehat{X}_0,\ldots,\widehat{X}_{k-1})]\,\big|d\mu(y) \\ &\le \int C\sum_{r=0}^{k-2}(\ell_{r+1}-u_r)^{-\gamma}\bigg(|F_y|_\infty + \mathrm{Lip}(F_y)\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j}\bigg)d\mu(y) \\ &\le C\sum_{r=0}^{k-2}(\ell_{r+1}-u_r)^{-\gamma}\bigg(|F|_\infty + \mathrm{Lip}(F)\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j}\bigg). \end{aligned}$$

Now

$$\begin{aligned} I_2 &= \mathbb{E}_\mu[F(X_0,\ldots,X_k)] - \int_M \mathbb{E}_\mu\big[F\big(X_0,\ldots,X_{k-1},X_k(y)\big)\big]\,d\mu(y) \\ &= \int_M F\big(X_0(x),\ldots,X_k(x)\big)d\mu(x) - \int_{M^2} F\big(X_0(x),\ldots,X_{k-1}(x),X_k(y)\big)d\mu^2(x,y). \end{aligned}$$

Write

$$\begin{aligned} &F(X_0(x),\ldots,X_k(x)) \\ &\qquad = F(\Phi_0(T^{\ell_0}x,\ldots,T^{u_0}x);\Phi_1(T^{\ell_1}x,\ldots,T^{u_1}x);\ldots;\Phi_k(T^{\ell_k}x,\ldots,T^{u_k}x)) \\ &\qquad = G(T^{\ell_0}x,\ldots,T^{u_0}x;T^{\ell_1}x,\ldots,T^{u_1}x;\ldots;T^{\ell_k}x,\ldots,T^{u_k}x). \end{aligned}$$

and

$$F(X_0(x), \dots, X_{k-1}(x), X_k(y))$$
$$= G(T^{\ell_0}x, \dots, T^{u_0}x; T^{\ell_1}x, \dots, T^{u_1}x; \dots; T^{\ell_{k-1}}x, \dots, T^{u_{k-1}}x; T^{\ell_k}y, \dots, T^{u_k}y)$$

where $G\colon M^s \to \mathbb{R}$, $s = \sum_{i=0}^{k}(u_i - \ell_i + 1)$. By a straightforward calculation, $G \in \mathcal{H}_s^\eta(M)$ and

$$\sum_{i=0}^{s-1}[G]_{\eta,i} \leq \sum_{i=0}^{k}\sum_{j=0}^{u_i-\ell_i} \operatorname{Lip}(F)[\varPhi_i]_{\eta,j}.$$

Hence by the Functional Correlation Bound,

$$|I_2| = \bigg| \int_M G(T^{\ell_0}x, \dots, T^{u_0}x; \dots; T^{\ell_k}x, \dots, T^{u_k}x)d\mu(x)$$
$$- \int_{M^2} G(T^{\ell_0}x, \dots, T^{u_0}x; \dots; T^{\ell_{k-1}}x, \dots, T^{u_{k-1}}x; T^{\ell_k}y, \dots, T^{u_k}y)d\mu^2(x,y)\bigg|$$
$$\leq C(\ell_k - u_{k-1})^{-\gamma}\bigg(|F|_\infty + \sum_{i=0}^{k}\sum_{j=0}^{u_i-\ell_i} \operatorname{Lip}(F)[\varPhi_i]_{\eta,j}\bigg).$$

This completes the proof. □

3.5 Moment bounds

In this section we prove Theorem 3.2.4. Throughout this section we fix $\gamma > 1$ and assume that $T\colon M \to M$ satisfies the Functional Correlation Bound with rate $n^{-\gamma}$.

In both parts of Theorem 3.2.4 we use the following moment bounds for independent, mean zero random variables, which are due to von Bahr, Esseen [BE65] and Rosenthal [Ros70], respectively:

Proposition 3.5.1. *Fix $p \geq 1$. There exists a constant $C > 0$ such that for all $k \geq 1$, for all independent, mean zero random variables $\widehat{X}_0, \dots, \widehat{X}_{k-1} \in L^p$:*

(i) If $1 \leq p \leq 2$, then

$$\mathbb{E}\bigg[\bigg|\sum_{i=0}^{k-1}\widehat{X}_i\bigg|^p\bigg] \leq C\sum_{i=0}^{k-1}\mathbb{E}\Big[|\widehat{X}_i|^p\Big].$$

(ii) If $p > 2$, then

$$\mathbb{E}\bigg[\bigg|\sum_{i=0}^{k-1}\widehat{X}_i\bigg|^p\bigg] \leq C\Bigg(\bigg(\sum_{i=0}^{k-1}\mathbb{E}\Big[\widehat{X}_i^2\Big]\bigg)^{p/2} + \sum_{i=0}^{k-1}\mathbb{E}\Big[|\widehat{X}_i|^p\Big]\Bigg).$$

□

Let $v, w \in \mathcal{C}^\eta(M)$ be mean zero. For $b \ge a \ge 0$ we denote

$$S_v(a,b) = \sum_{a \le i < b} v \circ T^i, \quad \mathbb{S}_{v,w}(a,b) = \sum_{a \le i < j < b} v \circ T^i w \circ T^j.$$

Note that $S_v(n) = S_v(0,n)$ and $\mathbb{S}_{v,w}(n) = \mathbb{S}_{v,w}(0,n)$. Some straightforward algebra yields the following lemma.

Lemma 3.5.2. *Fix $\ell \ge 1$ and $0 = a_0 \le a_1 \le \cdots \le a_\ell$. Then,*

(i) $S_v(a_\ell) = \sum_{i=0}^{\ell-1} S_v(a_i, a_{i+1})$.

(ii) $\mathbb{S}_{v,w}(a_\ell) = \sum_{i=0}^{\ell-1} \mathbb{S}_{v,w}(a_i, a_{i+1}) + \sum_{0 \le i < j < \ell} S_v(a_i, a_{i+1}) S_w(a_j, a_{j+1})$. □

We also need the following elementary lemma:

Lemma 3.5.3. *Fix $R > 0$, $p \ge 1$ and an integer $k \ge 1$. Define $F\colon [-R,R]^k \to \mathbb{R}$ by $F(y_0, \dots, y_{k-1}) = |y_0 + \cdots + y_{k-1}|^p$. Then $|F|_\infty \le (kR)^p$ and $\operatorname{Lip}(F) \le p(kR)^{p-1}$.*

Proof. Note that $|F|_\infty \le (kR)^p$. Fix $y = (y_0, \dots, y_{k-1}), y' = (y'_0, \dots, y'_{k-1}) \in [-R,R]^k$ and set $a = |y_0 + \cdots + y_{k-1}|, b = |y'_0 + \cdots + y'_{k-1}|$. By the Mean Value Theorem,

$$\begin{aligned} |F(y_0, \dots, y_{k-1}) - F(y'_0, \dots, y'_{k-1})| &= |a^p - b^p| \\ &\le p \max\{a^{p-1}, b^{p-1}\} |a - b| \\ &\le p(kR)^{p-1} \sum_{i=0}^{k-1} |y_i - y'_i| = p(kR)^{p-1} |y - y'|, \end{aligned}$$

so $\operatorname{Lip}(F) \le p(kR)^{p-1}$. □

Let $k \ge 1, n \ge 2k$ and define $a_i = [\frac{in}{2k}]$ for $0 \le i \le 2k$. Note that

$$\tfrac{n}{2k} - 1 \le a_{i+1} - a_i \le \tfrac{n}{2k} + 1 \le \tfrac{n}{k}. \tag{3.5.1}$$

For $0 \le i < k$ let $X_i = S_v(a_{2i}, a_{2i+1})$. Let $\widehat{X}_0, \dots, \widehat{X}_{k-1}$ be independent random variables with $\widehat{X}_i =_d X_i$.

Proposition 3.5.4. *There exists a constant $C > 0$ such that*

$$\mathbb{E}_\mu \left[\Big| \sum_{i=0}^{k-1} X_i \Big|^{2\gamma} \right] \le C k^{1+\gamma} n^\gamma \, \|v\|_\eta^{2\gamma} + \mathbb{E} \left[\Big| \sum_{i=0}^{k-1} \widehat{X}_i \Big|^{2\gamma} \right],$$

for all $n \geq 2k, k \geq 1$, for any $v \in \mathcal{C}^\eta(M)$.

Proof. Note that

$$X_i(x) = \sum_{q=a_{2i}}^{a_{2i+1}-1} v(T^q x) = \Phi_i(T^{\ell_i}x, \ldots, T^{u_i}x),$$

where $\ell_i = a_{2i}, u_i = a_{2i+1} - 1$ and

$$\Phi_i(x_0, \ldots, x_{u_i-\ell_i}) = \sum_{j=0}^{u_i-\ell_i} v(x_j).$$

Let $R = \max_i |\Phi_i|_\infty$. Then

$$\mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^{2\gamma}\right] = \mathbb{E}_\mu[F(X_0, \ldots, X_{k-1})]$$

where $F\colon [-R, R]^k \to \mathbb{R}$ is given by $F(y_0, \ldots, y_{k-1}) = |y_0 + \cdots + y_{k-1}|^{2\gamma}$. Hence by Proposition 3.4.1,

$$\mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^{2\gamma}\right] \leq A + \mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^{2\gamma}\right]$$

where

$$|A| \leq C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \Big(|F|_\infty + \mathrm{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i-\ell_i} [\Phi_i]_{\eta,j} \Big). \tag{3.5.2}$$

It remains to bound A. First we bound the expressions $[\Phi_i]_{\eta,j}$. Fix $0 \leq i < k$ and $0 \leq j \leq u_i - \ell_i$. For $x_0, \ldots, x_{k-1}, x'_j \in M$,

$$|\Phi_i(x_0, \ldots, x_{u_i-\ell_i}) - \Phi_i(x_0, \ldots, x_{j-1}, x'_j, x_{j+1} \ldots, x_{u_i-\ell_i})| = |v(x_j) - v(x'_j)|$$

so $[\Phi_i]_{\eta,j} \leq [v]_\eta$. Note that by (3.5.1), $|\Phi_i|_\infty \leq (a_{2i+1} - a_{2i}) |v|_\infty \leq \frac{n}{k} |v|_\infty$. Hence by Lemma 3.5.3,

$$|F|_\infty \leq 2\gamma(n |v|_\infty)^{2\gamma} \tag{3.5.3}$$

and $\mathrm{Lip}(F) \leq 2\gamma(n |v|_\infty)^{2\gamma-1}$.

Thus

$$\begin{aligned}
\mathrm{Lip}(F)\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\varPhi_i]_{\eta,j} \le & 2\gamma(n\,|v|_\infty)^{2\gamma-1}\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\varPhi_i]_{\eta,j} \\
\le & 2\gamma(n\,|v|_\infty)^{2\gamma-1}\sum_{i=0}^{k-1}(u_i-\ell_i+1)[v]_\eta \\
\le & 2\gamma(n\,|v|_\infty)^{2\gamma-1}n[v]_\eta. \qquad (3.5.4)
\end{aligned}$$

Now by (3.5.1), $\ell_{r+1}-u_r = a_{2r+2}-(a_{2r+1}-1) \ge \frac{n}{2k}$ for each $0 \le r \le k-2$. Hence

$$\sum_{r=0}^{k-2}(\ell_{r+1}-u_r)^{-\gamma} \le k(\tfrac{n}{2k})^{-\gamma} = 2^\gamma k^{1+\gamma}n^{-\gamma}. \qquad (3.5.5)$$

Substituting (3.5.3), (3.5.4) and (3.5.5) into (3.5.2) gives

$$\begin{aligned}
|A| &\le 2^\gamma k^{1+\gamma}n^{-\gamma}\Big(2\gamma(n\,|v|_\infty)^{2\gamma}+2\gamma(n\,|v|_\infty)^{2\gamma-1}n[v]_\eta\Big) \\
&\le 2^{1+\gamma}\gamma C k^{1+\gamma}n^{-\gamma}\left(n\,\|v\|_\eta\right)^{2\gamma} = 2^{1+\gamma}\gamma C k^{1+\gamma}n^{\gamma}\,\|v\|_\eta^{2\gamma},
\end{aligned}$$

as required. □

We are now ready to prove the moment bound for $S_v(n)$ (Theorem 3.2.4(a)).

Proof of Theorem 3.2.4(a). We prove by induction that there exists $D > 0$ such that

$$|S_v(m)|_{2\gamma} \le D m^{1/2}\,\|v\|_\eta \qquad (3.5.6)$$

for all $m \ge 1$, for any mean zero $v \in \mathcal{C}^\eta(M)$.

<u>Claim.</u> There exists $C > 0$ such that for all mean zero $v \in \mathcal{C}^\eta(M)$, for any $D > 0$, for any $k \ge 1$ and any $n \ge 2k$ such that (3.5.6) holds for all $m < n$, we have

$$|S_v(n)|_{2\gamma}^{2\gamma} \le C(k^{1+\gamma}+k^{1-\gamma}D^{2\gamma})n^\gamma\,\|v\|_\eta^{2\gamma}.$$

Now fix $k \ge 1$ such that $Ck^{1-\gamma} \le \frac{1}{2}$. Fix $D > 0$ such that $Ck^{1+\gamma} \le \frac{1}{2}D^{2\gamma}$ and (3.5.6) holds for all $m < 2k$ and any mean zero $v \in \mathcal{C}^\eta(M)$. Then the claim shows that for any $n \ge 2k$ such that (3.5.6) holds for all $m < n$, we have $|S_v(n)|_{2\gamma}^{2\gamma} \le D^{2\gamma}n^\gamma\,\|v\|_\eta^{2\gamma}$. Hence by induction, (3.5.6) holds for all $m \ge 1$.

It remains to prove the claim. Note that in the following the constant $C > 0$ may vary from line to line.

Fix $n \geq 2k$ and assume that (3.5.6) holds for all $m < n$. By Lemma 3.5.2(i),

$$S_v(n) = \sum_{i=0}^{2k-1} S_v(a_i, a_{i+1}) = I_1 + I_2,$$

where

$$I_1 = \sum_{i=0}^{k-1} S_v(a_{2i}, a_{2i+1}), \quad I_2 = \sum_{i=0}^{k-1} S_v(a_{2i+1}, a_{2i+2}).$$

We first bound $|I_1|_{2\gamma}$. Write $X_i = S_v(a_{2i}, a_{2i+1})$ so that $I_1 = \sum_{i=0}^{k-1} X_i$. By Proposition 3.5.4,

$$|I_1|_{2\gamma}^{2\gamma} = \mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^{2\gamma}\right] \leq Ck^{1+\gamma}n^\gamma \|v\|_\eta^{2\gamma} + \mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^{2\gamma}\right]. \tag{3.5.7}$$

We now bound $\mathbb{E}\Big[|\sum_{i=0}^{k-1} \widehat{X}_i|^{2\gamma}\Big]$ by using Proposition 3.5.1 and the inductive hypothesis.

Fix $0 \leq i < k$. By stationarity, $X_i = S_v(a_{2i}, a_{2i+1}) =_d S_v(a_{2i+1} - a_{2i})$. Thus by the inductive hypothesis (3.5.6), $\mathbb{E}_\mu\big[|X_i|^{2\gamma}\big] \leq D^{2\gamma}(a_{2i+1} - a_{2i})^\gamma \|v\|_\eta^{2\gamma}$. Hence by (3.5.1),

$$\begin{aligned}\sum_{i=0}^{k-1} \mathbb{E}\big[|\widehat{X}_i|^{2\gamma}\big] &\leq \sum_{i=0}^{k-1} D^{2\gamma}(a_{2i+1} - a_{2i})^\gamma \|v\|_\eta^{2\gamma} \\ &\leq \sum_{i=0}^{k-1} D^{2\gamma}(n/k)^\gamma \|v\|_\eta^{2\gamma} = D^{2\gamma}k^{1-\gamma}n^\gamma \|v\|_\eta^{2\gamma}.\end{aligned}$$

Now by the Functional Correlation Bound, $|\mathbb{E}_\mu[v\, v \circ T^n]| \leq Cn^{-\gamma} \|v\|_\eta^2$. By a standard calculation, it follows that $\mathbb{E}_\mu\big[S_v(n)^2\big] \leq Cn \|v\|_\eta^2$. Thus

$$\begin{aligned}\sum_{i=0}^{k-1} \mathbb{E}\Big[\widehat{X}_i^2\Big] &= \sum_{i=0}^{k-1} \mathbb{E}_\mu\big[S_v(a_{2i+1} - a_{2i})^2\big] \\ &\leq \sum_{i=0}^{k-1} C(a_{2i+1} - a_{2i}) \|v\|_\eta^2 \leq C(a_{2k-1} - a_0) \|v\|_\eta^2 \\ &\leq Cn \|v\|_\eta^2.\end{aligned}$$

By Proposition 3.5.1(ii), it follows that

$$\mathbb{E}\left[\left|\sum_{i=0}^{k-1}\widehat{X}_i\right|^{2\gamma}\right] \le C\big((Cn\,\|v\|_\eta^2)^\gamma + D^{2\gamma}k^{1-\gamma}n^\gamma\,\|v\|_\eta^{2\gamma}\big)$$
$$\le C(1+D^{2\gamma}k^{1-\gamma})n^\gamma\,\|v\|_\eta^{2\gamma}\,.$$

Hence by (3.5.7), overall

$$|I_1|_{2\gamma}^{2\gamma} \le C(k^{1+\gamma}+D^{2\gamma}k^{1-\gamma})n^\gamma\,\|v\|_\eta^{2\gamma}\,.$$

Exactly the same argument applies to $|I_2|_{2\gamma}^{2\gamma}$. The conclusion of the claim follows by noting that

$$|S_v(n)|_{2\gamma}^{2\gamma} = |I_1+I_2|_{2\gamma}^{2\gamma} \le 2^{2\gamma}(|I_1|_{2\gamma}+|I_2|_{2\gamma}).$$ □

We now prove Theorem 3.2.4(b). Our proof follows the same lines as that of part (a).

Let $n,k\ge 1$. Recall that $a_i = \left[\frac{in}{2k}\right]$. For $0\le i<k$ define mean zero random variables X_i on (M,μ) by

$$X_i = \mathbb{S}_{v,w}(a_{2i},a_{2i+1}) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i},a_{2i+1})]\,.$$

Let $\widehat{X}_0,\dots,\widehat{X}_{k-1}$ be independent random variables with $\widehat{X}_i =_d X_i$.

The following proposition plays the same role that Proposition 3.5.4 played in the proof of Theorem 3.2.4(a).

Proposition 3.5.5. *There exists a constant $C>0$ such that for any $v,w\in\mathcal{C}^\eta(M)$,*

$$\mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1}X_i\right|^{\gamma}\right] \le Ckn^\gamma\,\|v\|_\eta^\gamma\,\|w\|_\eta^\gamma + \mathbb{E}\left[\left|\sum_{i=0}^{k-1}\widehat{X}_i\right|^{\gamma}\right]$$

for all $n\ge 2k, k\ge 1$.

Proof. Note that

$$X_i(x) = \sum_{a_{2i}\le q<r\le a_{2i+1}-1} v(T^q x)w(T^r x) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i},a_{2i+1})]$$
$$= \Phi_i(T^{\ell_i}x,\dots,T^{u_i}x),$$

where $\ell_i = a_{2i}, u_i = a_{2i+1}-1$ and

$$\Phi_i(x_0,\dots,x_{u_i-\ell_i}) = \sum_{0\le q<r\le u_i-\ell_i} v(x_q)w(x_r) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i},a_{2i+1})]\,.$$

Let $R = \max_i |\Phi_i|_\infty$. Observe that

$$\mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^\gamma\right] = \mathbb{E}_\mu[F(X_0, \dots, X_{k-1})]\,,$$

where $F\colon [-R,R]^k \to \mathbb{R}$ is given by $F(y_0,\dots,y_{k-1}) = |y_0 + \cdots + y_{k-1}|^\gamma$. Hence by Lemma 3.4.1,

$$\mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^\gamma\right] \le A + \mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^{2\gamma}\right]$$

where

$$|A| \le C\sum_{r=0}^{k-2}(\ell_{r+1} - u_r)^{-\gamma}\bigg(|F|_\infty + \operatorname{Lip}(F)\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j}\bigg). \tag{3.5.8}$$

It remains to bound A. The first step is to bound the expressions $[\Phi_i]_{\eta,j}$. Fix $0 \le i < k, 0 \le j \le u_i - \ell_i$. Let $x_0, \dots, x_{k-1}, x_j' \in M$. Note that

$$\Phi_i(x_0,\dots,x_{u_i-\ell_i}) - \Phi_i(x_0,\dots,x_{j-1},x_j',x_{j+1}\dots,x_{u_i-\ell_i}) = J_1 + J_2,$$

where

$$\begin{aligned} J_1 &= \sum_{j<r\le u_i-\ell_i}(v(x_j)w(x_r) - v(x_j')w(x_r)),\\ J_2 &= \sum_{0\le q<j}(v(x_q)w(x_j) - v(x_q)w(x_j')). \end{aligned}$$

Now,

$$|J_1| \le \sum_{j<r\le u_i-\ell_i}|v(x_j) - v(x_j')||w(x_r)| \le |w|_\infty \sum_{j<r\le u_i-\ell_i}|v(x_j) - v(x_j')|$$

and similarly $|J_2| \le |v|_\infty \sum_{0\le q<j}|w(x_j) - w(x_j')|$, so

$$[\Phi_i]_{\eta,j} \le (u_i - \ell_i)\,\|v\|_\eta\,\|w\|_\eta\,.$$

Now recall from (3.5.1) that $u_i - \ell_i + 1 = a_{2i+1} - a_{2i} \le n/k$ so

$$\sum_{i=0}^{k-1}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j} \le \sum_{i=0}^{k-1}(u_i - \ell_i + 1)^2\,\|v\|_\eta\,\|w\|_\eta \le \tfrac{n^2}{k}\,\|v\|_\eta\,\|w\|_\eta\,. \tag{3.5.9}$$

Next note that

$$\begin{aligned}|\Phi_i|_\infty &\le \sum_{0\le q<r\le u_i-\ell_i} |v|_\infty |w|_\infty + |\mathbb{S}_{v,w}(a_{2i},a_{2i+1})|_\infty \\ &\le 2(n/k)^2 |v|_\infty |w|_\infty\end{aligned}$$

so by Lemma 3.5.3, $|F|_\infty \le \big(\frac{2n^2}{k}|v|_\infty |w|_\infty\big)^\gamma$ and $\mathrm{Lip}(F) \le \gamma\big(\frac{2n^2}{k}|v|_\infty |w|_\infty\big)^{\gamma-1}$. Combining these bounds with (3.5.5), (3.5.8) and (3.5.9) yields that

$$\begin{aligned}|A| &\le C2^\gamma k^{1+\gamma} n^{-\gamma}\Big((\tfrac{2n^2}{k}|v|_\infty |w|_\infty)^\gamma + \gamma(\tfrac{2n^2}{k}|v|_\infty |w|_\infty)^{\gamma-1}\tfrac{n^2}{k}\|v\|_\eta \|w\|_\eta\Big) \\ &\le 2^{2\gamma}(1+\gamma/2)Ckn^\gamma \|v\|_\eta^\gamma \|w\|_\eta^\gamma,\end{aligned}$$

as required. □

We are now ready to prove Theorem 3.2.4(b).

Proof of Theorem 3.2.4(b). We prove by induction that there exists $D > 0$ such that

$$|\mathbb{S}_{v,w}(m)|_\gamma \le Dm \|v\|_\eta \|w\|_\eta \tag{3.5.10}$$

for all $m \ge 1$, for any $v, w \in \mathcal{C}^\eta(M)$ mean zero.

Claim. There exists $C > 0$ such that for all $v, w \in \mathcal{C}^\eta(M)$ mean zero, for any $D > 0$, any $k \ge 1$ and any $n \ge 2k$ such that (3.5.10) holds for all $m < n$, we have

$$|\mathbb{S}_{v,w}(n)|_\gamma^\gamma \le C\Big(k^\gamma + (k^{1-\gamma} + k^{-\gamma/2})D^\gamma\Big)(n\|v\|_\eta \|w\|_\eta)^\gamma.$$

Now fix $k \ge 1$ such that $C(k^{1-\gamma} + k^{-\gamma/2}) \le \frac{1}{2}$. Fix $D > 0$ such that $Ck^\gamma \le \frac{1}{2}D^\gamma$ and (3.5.10) holds for all $m < 2k$ and any mean zero $v, w \in \mathcal{C}^\eta(M)$. Then the claim shows that if $n \ge 2k$ and (3.5.10) holds for all $m < n$, then $|\mathbb{S}_{v,w}(n)|_\gamma^\gamma \le D^\gamma(n\|v\|_\eta \|w\|_\eta)^\gamma$. Hence by induction, (3.5.10) holds for all $m \ge 1$.

It remains to prove the claim. Note that in the following the constant $C > 0$ may vary from line to line.

Fix $n \ge 2k$ and assume that (3.5.10) holds for all $m < n$. Recall that $a_i = \big[\frac{in}{2k}\big]$ for $0 \le i \le 2k$. By Lemma 3.5.2(ii),

$$\mathbb{S}_{v,w}(n) = \sum_{0\le i<j<2k} S_v(a_i,a_{i+1})S_w(a_j,a_{j+1}) + \sum_{i=0}^{2k-1}\mathbb{S}_{v,w}(a_i,a_{i+1}) = I_1 + I_2 + I_3 + I_4,$$

where

$$I_1 = \sum_{0 \le i < j < 2k} S_v(a_i, a_{i+1}) S_w(a_j, a_{j+1}), \quad I_2 = \sum_{i=0}^{2k-1} \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_i, a_{i+1})]\,,$$

$$I_3 = \sum_{i=0}^{k-1} \big(\mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i}, a_{2i+1})]\,\big),$$

$$I_4 = \sum_{i=0}^{k-1} \big(\mathbb{S}_{v,w}(a_{2i+1}, a_{2i+2}) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i+1}, a_{2i+2})]\,\big).$$

Recall from (3.5.1) that $a_{i+1} - a_i \le n/k$. Hence by Theorem 3.2.4(a),

$$\begin{aligned}
|I_1|_\gamma &\le \sum_{0 \le i < j < 2k} |S_v(a_i, a_{i+1}) S_w(a_j, a_{j+1})|_\gamma \\
&\le \sum_{0 \le i < j < 2k} |S_v(a_i, a_{i+1})|_{2\gamma} \, |S_w(a_j, a_{j+1})|_{2\gamma} \\
&\le \sum_{0 \le i < j < 2k} C^2 (a_{i+1} - a_i)^{1/2} \|v\|_\eta \, (a_{j+1} - a_j)^{1/2} \|w\|_\eta \\
&\le \sum_{0 \le i < j < 2k} C^2 (n/k)^{1/2} \|v\|_\eta \, (n/k)^{1/2} \|w\|_\eta \le Ckn \|v\|_\eta \, \|w\|\,.
\end{aligned}$$

Now by the Functional Correlation Bound, $|\mathbb{E}_\mu[v\, w \circ T^n]\,| \le Cn^{-\gamma} \|v\|_\eta \, \|w\|_\eta$. By a standard calculation, it follows that $|\mathbb{E}_\mu[\mathbb{S}_{v,w}(n)]| \le Cn \|v\|_\eta \, \|w\|_\eta$. Thus

$$\begin{aligned}
|I_2| &\le \sum_{i=0}^{2k-1} |\mathbb{E}_\mu[\mathbb{S}_{v,w}(a_i, a_{i+1})]\,| \\
&\le \sum_{i=0}^{2k-1} C(a_{i+1} - a_i) \|v\|_\eta \, \|w\|_\eta = C(a_{2k} - a_0) \|v\|_\eta \, \|w\|_\eta \\
&= Cn \|v\|_\eta \, \|w\|_\eta\,.
\end{aligned}$$

We now bound $|I_3|_\gamma^\gamma$. Note that $I_3 = \sum_{i=0}^{k-1} X_i$, where $X_i = \mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i}, a_{2i+1})]$. Hence by Proposition 3.5.5,

$$|I_3|_\gamma^\gamma = \mathbb{E}_\mu\left[\left|\sum_{i=0}^{k-1} X_i\right|^\gamma\right] \le Ckn^\gamma \|v\|_\eta^\gamma \, \|w\|_\eta^\gamma + \mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^\gamma\right]. \tag{3.5.11}$$

Fix $0 \le i < k$. By stationarity, $X_i =_d \mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) - \mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i+1} - a_{2i})]$. Now by the inductive hypothesis (3.5.10), $|\mathbb{S}_{v,w}(a_{2i+1} - a_{2i})|_\gamma \le D(a_{2i+1} - a_{2i}) \|v\|_\eta \, \|w\|_\eta$, so

$$\begin{aligned} |X_i|_\gamma &\leq |\mathbb{S}_{v,w}(a_{2i+1} - a_{2i})|_\gamma + |\mathbb{E}_\mu[\mathbb{S}_{v,w}(a_{2i+1} - a_{2i})]\,| \\ &\leq 2D(a_{2i+1} - a_{2i})\left\|v\right\|_\eta \left\|w\right\|_\eta . \end{aligned}$$

It follows that

$$\begin{aligned} \sum_{i=0}^{k-1} \mathbb{E}\Big[|\widehat{X}_i|^\gamma\Big] &\leq \sum_{i=0}^{k-1} 2^\gamma D^\gamma (a_{2i+1} - a_{2i})^\gamma (\left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma \\ &\leq \sum_{i=0}^{k-1} 2^\gamma D^\gamma (n/k)^\gamma (\left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma = 2^\gamma D^\gamma k^{1-\gamma} (n \left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma. \end{aligned}$$

If $1 < \gamma \leq 2$, then by Proposition 3.5.1(i),

$$\mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^\gamma\right] \leq 2^\gamma C D^\gamma k^{1-\gamma} (n \left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma.$$

Suppose on the other hand that $\gamma > 2$. Note that

$$|\widehat{X}_i|_2 \leq |\widehat{X}_i|_\gamma \leq 2D(a_{2i+1} - a_{2i}) \left\|v\right\|_\eta \left\|w\right\|_\eta$$

so

$$\begin{aligned} \sum_{i=0}^{k-1} \mathbb{E}\Big[\widehat{X}_i^2\Big] &\leq \sum_{i=0}^{k-1} 4D^2 (a_{2i+1} - a_{2i})^2 (\left\|v\right\|_\eta \left\|w\right\|_\eta)^2 \\ &\leq \sum_{i=0}^{k-1} 4D^2 (n/k)^2 (\left\|v\right\|_\eta \left\|w\right\|_\eta)^2 = 4D^2 k^{-1} (n \left\|v\right\|_\eta \left\|w\right\|_\eta)^2. \end{aligned}$$

Hence by Proposition 3.5.1(ii),

$$\begin{aligned} \mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^\gamma\right] &\leq C\Big(\big(4D^2 k^{-1} (n \left\|v\right\|_\eta \left\|w\right\|_\eta)^2\big)^{\gamma/2} + 2^\gamma D^\gamma k^{1-\gamma} \big(n \left\|v\right\|_\eta \left\|w\right\|_\eta\big)^\gamma\Big) \\ &= 2^\gamma C D^\gamma (k^{-\gamma/2} + k^{1-\gamma})(n \left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma. \end{aligned}$$

Hence for any $\gamma > 1$,

$$\mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^\gamma\right] \leq C D^\gamma (k^{-\gamma/2} + k^{1-\gamma})(n \left\|v\right\|_\eta \left\|w\right\|_\eta)^\gamma.$$

By (3.5.11), it follows that

$$|I_3|_\gamma^\gamma \leq C\Big(k + D^\gamma(k^{-\gamma/2} + k^{1-\gamma})\Big)(n\,\|v\|_\eta\,\|w\|_\eta)^\gamma.$$

Exactly the same argument applies to $|I_4|_\gamma^\gamma$. The conclusion of the claim follows by noting that

$$\begin{aligned}|\mathbb{S}_{v,w}(n)|_\gamma^\gamma = |I_1 + I_2 + I_3 + I_4|_\gamma^\gamma &\leq 4^\gamma(|I_1|_\gamma^\gamma + |I_2|_\gamma^\gamma + |I_3|_\gamma^\gamma + |I_4|_\gamma^\gamma) \\ &\leq C\Big[k^\gamma + 1 + 2\Big(k + D^\gamma(k^{-\gamma/2} + k^{1-\gamma})\Big)\Big](n\,\|v\|_\eta\,\|w\|_\eta)^\gamma \\ &\leq C\Big[k^\gamma + D^\gamma(k^{-\gamma/2} + k^{1-\gamma})\Big](n\,\|v\|_\eta\,\|w\|_\eta)^\gamma,\end{aligned}$$

as required. □

Chapter 4

Deterministic homogenisation for families

4.1 Introduction and statement of main results

We now return to the deterministic homogenisation problem that motivated us in Chapter 3. Recall that μ is a Borel probability measure on a metric space M and $T : M \to M$ is an ergodic μ-preserving transformation.

By Theorem 1.1.3 ([Che$^+$22, Theorem 2.10]), deterministic homogenisation reduces to proving two statistical properties for $T : M \to M$. In Theorem 3.2.4 we showed that one of these statistical properties, namely iterated moment bounds, follows from an abstract functional correlation bound. Thus it is natural to ask whether the Functional Correlation Bound is a sufficient condition for homogenisation.

Indeed, homogenisation results have many interesting physical applications (cf. [PS08, Sect. 11.8]), particularly in stochastic climate theory [GCF17]. As such, it is desirable to find a sufficient condition for homogenisation that is accessible to a broad audience.

Before proceeding further, let us discuss other possible sufficient conditions for homogenisation. For many classes of chaotic dynamical systems (particularly ones with some hyperbolicity), it is possible to prove bounds on the correlations

$$C_n(v, w) = \int_M v\, w \circ T^n d\mu - \int_M v\, d\mu \int_M w\, d\mu$$

for all Hölder v, $w : M \to \mathbb{R}$. Moreover, correlation bounds play a crucial role in most standard proofs of the central limit theorem, so it is natural to seek a sufficient condition for homogenisation in terms of decay of correlations.

Fast decay of the autocorrelations $C_n(v,v)$ is not enough to guarantee that $v : M \to \mathbb{R}$ satisfies the central limit theorem, see [GM14] for a counterexample. More generally, fast decay of correlations for Hölder observables is not thought to be sufficient to prove the central limit theorem.

Let $v = (v_1, \dots, v_d) \in L^\infty(M, \mathbb{R}^d)$ with $\int v\, d\mu = 0$. Suppose that there exists a sequence $a_n > 0$ such that $\sum_{n \geq 1} a_n^{1/3} < \infty$ and

$$\left|C_n(v_i, w)\right| \leq a_n \left|w\right|_\infty \text{ for all } w \in L^\infty(M), 1 \leq i \leq d. \tag{4.1.1}$$

Then v satisfies the statistical properties required by [KM16]. However, if T is invertible, then this condition fails whenever $v \neq 0$.

Our first main result is that homogenisation holds whenever the Functional Correlation Bound is satisfied with a fast enough rate. Before stating this result, let us recall the homogenisation problem that we are interested in. Consider a fast-slow system on $\mathbb{R}^d \times M$ of the form

$$x_{k+1}^{(n)} = x_k^{(n)} + n^{-1} a(x_k^{(n)}, y_k) + n^{-1/2} b(x_k^{(n)}, y_k), \qquad y_{k+1} = T y^{(n)}, \tag{4.1.2}$$

where $x_0^{(n)} \equiv \xi$ is fixed and y_0 is drawn randomly from (M, μ). Assume that $\int_M b(x,y) d\mu(y) = 0$ for all $x \in \mathbb{R}^d$. We are interested in the limiting behaviour of the random process X_n defined by $X_n(t) = x_{[nt]}^{(n)}$ for $t \in [0,1]$. We view X_n as a random element of the space $D([0,1], \mathbb{R}^d)$, that is the space of functions $h : [0,1] \to \mathbb{R}^d$ that are right continuous with left limits. We endow $D([0,1], \mathbb{R}^d)$ with the sup-norm topology.

Let us now describe our regularity assumptions on the coefficients $a, b : \mathbb{R}^d \times M \to \mathbb{R}^d$. For $\alpha \geq 0$, $\kappa \in [0,1]$ define $C^{\alpha,\kappa}(\mathbb{R}^d \times M, M)$ to be the space of functions $f : \mathbb{R}^d \times M \to \mathbb{R}^d$ such that

$$\|f\|_{C^{\alpha,\kappa}} = \sum_{|k| \leq [\alpha]} \sup_{x \in \mathbb{R}^d} \left\| D^k f(x, \cdot) \right\|_\kappa + \sum_{|k| = [\alpha]} \frac{\left\| D^k f(x, \cdot) - D^k f(x', \cdot) \right\|_\kappa}{|x - x'|^{\alpha - [\alpha]}} < \infty.$$

Here D^k is the differential operator acting in the x component.

Fix parameters $\alpha > 2 + \frac{d}{\gamma}$ and $\kappa \in (0,1)$. Let $a \in C^{1+\kappa,0}(\mathbb{R}^d \times M, \mathbb{R}^d)$ and $b \in C^{\alpha,\eta}(\mathbb{R}^d \times M, \mathbb{R}^d)$. Our first main result is as follows:

Theorem 4.1.1. *Suppose that T satisfies the Functional Correlation Bound with rate $k^{-\gamma}$, $\gamma > 1$. Then $X_n \xrightarrow{w} X$ in $D([0,1], \mathbb{R}^d)$, where X is the solution of a stochastic differential equation driven by Brownian motion.*

We refer to [Che$^+$22, Sect. 2] for explicit expressions for the drift and diffusion coefficients of the stochastic differential equation satisfied by X.

The condition $\gamma > 1$ in Theorem 4.1.1 is sharp. For $\gamma \le 1$ there are examples where the central limit theorem fails for generic Hölder observables (see Example 4.1.4 for more details).

4.1.1 More general fast-slow systems

Until now we have assumed that the fast direction y_k in (4.1.2) is independent of n. Let $T_n : M \to M$ be a family of maps with invariant probability measures μ_n. We generalise (4.1.2) by considering fast-slow systems on $\mathbb{R}^d \times M$ of the form

$$x_{k+1}^{(n)} = x_k^{(n)} + n^{-1} a_n(x_k^{(n)}, y_k^{(n)}) + n^{-1/2} b_n(x_k^{(n)}, y_k^{(n)}), \quad y_{k+1}^{(n)} = T_n y_k^{(n)}. \qquad (4.1.3)$$

Here $x_0^{(n)} \equiv \xi$ is fixed and $y_0^{(n)}$ is drawn randomly from (M, μ_n). The homogenisation problem we are interested in is the same as before, only now X_n is a stochastic process on (M, μ_n). This homogenisation problem was previously considered by [KKM18, KKM22, Che$^+$22]. In particular, in [KKM22, Che$^+$22] this problem was settled for families of nonuniformly expanding maps.

We now give a generalisation of Theorem 4.1.1 for families T_n that satisfy the Functional Correlation Bound in the following uniform sense.

Let $T_n \colon M \to M$, $n \in \mathbb{N} \cup \{\infty\}$, be a family of dynamical systems with invariant probability measures μ_n. Suppose that T_n satisfies the Functional Correlation Bound with rate $k^{-\gamma}$ for each n, and that the constant $C > 0$ can be chosen independently of n. Then we say that the family T_n *satisfies the Functional Correlation Bound uniformly with rate* $k^{-\gamma}$.

As before, fix parameters $\kappa \in (0,1)$ and $\alpha > 2 + \frac{d}{\gamma}$. We assume that $a_n \in C^{1+\kappa,0}(\mathbb{R}^d \times M, \mathbb{R}^d)$, $b_n \in C^{\alpha,\eta}(\mathbb{R}^d \times M, \mathbb{R}^d)$ for all $n \ge 1$ and that $\sup_n \|a_n\|_{C^{1+\kappa,0}} < \infty$. Moreover, we assume that there exist functions $\bar{a} \in C^{1+\kappa}(\mathbb{R}^d, \mathbb{R}^d)$ and $b_\infty \in C^{\alpha,\eta}(\mathbb{R}^d \times M, \mathbb{R}^d)$ such that $\lim_{n\to\infty} \int a_n(x,y) d\mu_n(y) = \bar{a}(x)$ for all $x \in \mathbb{R}^d$ and $\lim_{n\to\infty} \|b_n - b_\infty\|_{C^{\alpha,\eta}} = 0$.

Our main result on homogenisation for fast-slow systems of the form (4.1.3) is as follows:

Theorem 4.1.2. *Let $T_n : M \to M$ be a family of dynamical systems that satisfies the Functional Correlation Bound uniformly with rate $k^{-\gamma}$, $\gamma > 1$. Suppose that*

$$\lim_{n\to\infty} \int_M vw \circ T_n^j \, d\mu_n = \int_M vw \circ T_\infty^j \, d\mu_\infty \qquad (4.1.4)$$

for all $j \geq 0$ and all η-Hölder v, $w : M \to \mathbb{R}$. Then $X_n \xrightarrow{w} X$ in $D([0,1],\mathbb{R}^d)$, where X is the solution of a stochastic differential equation driven by Brownian motion.

Note that Theorem 4.1.1 follows from Theorem 4.1.2 by taking $T_n \equiv T$, $\mu_n \equiv \mu$.

Remark 4.1.3. It is straightforward to check that condition (4.1.4) holds provided that

(A1) T_∞ is continuous μ_∞-a.e. and μ_n is statistically stable, i.e. $\mu_n \xrightarrow{w} \mu_\infty$.

(A2) $\lim_{n\to\infty} \mu_n(x \in M : d(T_n^j x, T_\infty^j x) > a) = 0$ for all $a > 0$ and $j \geq 0$.

Conditions (A1)-(A2) are similar to condition (7.2) in [KKM18]. Many examples of families of nonuniformly expanding maps are strongly statistically stable: the measures μ_n are all absolutely continuous with respect to some reference measure m and $d\mu_n/dm \to d\mu_\infty/dm$ in $L^1(m)$. For such examples, it is fairly straightforward to check conditions (A1) and (A2). In contrast, for many natural examples of families of nonuniformly hyperbolic diffeomorphisms the measures μ_n are mutually singular so strong statistical stability fails.

Condition (4.1.4) is easier to check in certain situations. In particular, in Subsection 4.4.2 we show that it is also possible to verify condition (4.1.4) for an example where statistical stability is proved by Keller-Liverani perturbation theory.

We end this introduction by explaining how Theorem 4.1.2 applies to intermittent Baker's maps, which we previously saw in Subsection 3.1.1.

Example 4.1.4. *Let $\alpha \in (0,1)$. Define $g_\alpha : [0,\frac{1}{2}] \to [0,1]$ by $g_\alpha(x) = x(1+2^\alpha x^\alpha)$. Recall that the Liverani-Saussol-Vaienti [LSV99] map is given by*

$$\bar{T}_\alpha : [0,1] \to [0,1], \quad \bar{T}_\alpha(x) = \begin{cases} g_\alpha(x), & x < \frac{1}{2}, \\ 2x-1, & x \geq \frac{1}{2}. \end{cases}$$

The map $\bar{T}_\alpha$ is an archetypal example of a nonuniformly expanding map with slow decay of correlations. Let $M = [0,1]^2$. Consider an intermittent Baker's map [MV16] given by

$$T_\alpha : M \to M, \quad T_\alpha(x,y) = \begin{cases} (\bar{T}_\alpha(x), g_\alpha^{-1}(y)), & x < \frac{1}{2}, \\ (\bar{T}_\alpha(x), \frac{1}{2}(y+1)), & x \geq \frac{1}{2}. \end{cases} \tag{4.1.5}$$

The map T_α is invertible and there is a unique probability measure μ_α such that $\pi_*\mu_\alpha = \bar{\mu}_\alpha$, where π denotes the projection onto the first coordinate. By Theorem 3.2.3, both $\bar{T}_\alpha$ and T_α satisfies the Functional Correlation Bound with rate

$k^{1-1/\alpha}$. *For* $\alpha \geq \frac{1}{2}$, *the central limit theorem fails for generic Hölder observables even for the* $\bar{T}_\alpha$ *dynamics [Gou04]. Hence it is natural to restrict to the range* $\alpha < \frac{1}{2}$ *when considering homogenisation.*

Let $\alpha_n \in (0, \frac{1}{2})$ *satisfy* $\alpha_n \to \alpha_\infty \in (0, \frac{1}{2})$. *Let* $\bar{T}_n = \bar{T}_{\alpha_n}$, $T_n = T_{\alpha_n}$ *and* $\mu_n = \mu_{\alpha_n}$. *In [KKM22] homogenisation is obtained for the family* $\bar{T}_n$. *By Theorem 4.1.2, we obtain homogenisation for the family* T_n, *see Subsection 4.4.1 for more details.*

Notation: Throughout this chapter, 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 $X_n \to_{\mu_n} X$ means that X_n is a family of random variables on (M, μ_n) and $X_n \xrightarrow{w} X$.

4.2 Proof of Theorem 4.1.2

This section is dedicated to the proof of Theorem 4.1.2. We proceed by applying [Che+22, Theorem 2.17]. In Subsection 4.2.1, we prove the iterated WIP, which is one of the main hypotheses of [Che+22, Theorem 2.17]. Then in Subsection 4.2.2, we complete the proof of Theorem 4.1.2.

4.2.1 The Iterated WIP

Let $\gamma > 1$. Throughout this subsection, T_n, $n \geq 1$ is a family of maps that satisfies the Functional Correlation Bound uniformly with rate $k^{-\gamma}$.

Fix $d \geq 1$. Let $v_n \colon M \to \mathbb{R}^d$, $n \geq 1$ be a family of observables with $\sup_{n \geq 1} \|v_n\|_\eta < \infty$ and $\int v_n \, d\mu_n = 0$. Recall that for $a, b \in \mathbb{R}^d$ we denote $a \otimes b = ab^T$. For $t \geq 0$ define

$$W_n(t) = n^{-1/2} \sum_{0 \leq r < [nt]} v_n \circ T_n^r, \qquad \mathbb{W}_n(t) = n^{-1} \sum_{0 \leq r < s < [nt]} v_n \circ T_n^r \otimes v_n \circ T_n^s.$$

Let $v \colon M \to \mathbb{R}^d$ and $k, n \geq 1$. Define $S_v(k, n) = \sum_{0 \leq r < k} v \circ T_n^r$ and $\mathbb{S}_v(k, n) = \sum_{0 \leq r < s < k} v \circ T_n^r \otimes v \circ T_n^s$.

Lemma 4.2.1. *For each* $n \geq 1$, *the limits*

$$\Sigma_n = \lim_{k \to \infty} k^{-1} \mathbb{E}_{\mu_n}[S_{v_n}(k, n) \otimes S_{v_n}(k, n)], \quad E_n = \lim_{k \to \infty} k^{-1} \mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(k, n)]$$

exist and are given by

$$\begin{aligned}\Sigma_n &= \mathbb{E}_{\mu_n}[v_n \otimes v_n] + \sum_{\ell\geq 1}\Big(\mathbb{E}_{\mu_n}[v_n \otimes v_n \circ T_n^\ell] + \mathbb{E}_{\mu_n}[v_n \circ T_n^\ell \otimes v_n]\Big),\\ E_n &= \sum_{\ell\geq 1}\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big].\end{aligned} \tag{4.2.1}$$

Moreover, the convergence is uniform in n.

Proof. We first prove the existence of the limit E_n. Note that

$$\mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(k,n)] = \sum_{\ell=1}^{k-1}\sum_{r=0}^{k-\ell-1}\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big] = \sum_{\ell=1}^{k-1}(k-\ell)\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big].$$

Let $1 \leq i,j \leq d$ and $\ell \geq 1$. Define $G\colon M^2 \to \mathbb{R}$ by $G(x,y) = v_n^i(x)v_n^j(y)$. By the Functional Correlation Bound,

$$\begin{aligned}|\mathbb{E}_{\mu_n}\Big[v_n^i v_n^j \circ T_n^\ell\Big]| &= \left|\int G(x, T_n^\ell x)d\mu_n(x)\right|\\ &\ll \ell^{-\gamma}\left\|v_n^i\right\|_\eta \left\|v_n^j\right\|_\eta + \left|\int v_n^i d\mu_n \int v_n^j d\mu_n\right| = \ell^{-\gamma}\left\|v_n^i\right\|_\eta \left\|v_n^j\right\|_\eta.\end{aligned}$$

It follows that for all $n \geq 1$,

$$\begin{aligned}&\left|\sum_{\ell\geq 1}\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big] - k^{-1}\mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(k,n)]\right|\\ &\qquad \leq k^{-1}\sum_{\ell=1}^{k-1}\ell\Big|\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big]\Big| + \sum_{\ell\geq k}\Big|\mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell\Big]\Big|\\ &\qquad \ll k^{-1}\sum_{\ell=1}^{k-1}\ell^{1-\gamma} + \sum_{\ell\geq k}\ell^{-\gamma} \ll k^{-1}(1+k^{2-\gamma}) + k^{1-\gamma} \ll k^{1-\gamma} + k^{-1} = o(1),\end{aligned}$$

which proves the existence of the limit E_n. Since

$$\mathbb{E}_{\mu_n}[S_{v_n}(k,n) \otimes S_{v_n}(k,n)] = k\mathbb{E}_{\mu_n}[v_n \otimes v_n] + \mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(k,n)] + \mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(k,n)]^T$$

and $\Sigma_n = \mathbb{E}_{\mu_n}[v_n \otimes v_n] + E_n + E_n^T$ it follows that

$$\Big|\Sigma_n - k^{-1}\mathbb{E}_{\mu_n}[S_{v_n}(k,n) \otimes S_{v_n}(k,n)]\Big| \ll k^{1-\gamma} + k^{-1} = o(1), \tag{4.2.2}$$

as required. □

We are now ready to state the main result of this subsection:

Theorem 4.2.2 (Iterated WIP)**.** *Suppose that* $\lim_{n\to\infty}\Sigma_n = \Sigma$ *and* $\lim_{n\to\infty} E_n = E$. *Then* $(W_n, \mathbb{W}_n) \to (W, \mathbb{W})$ *in the sense of finite-dimensional distributions, where* W *is a Brownian motion with covariance* Σ *and* $\mathbb{W}(t) = \int_0^t W \otimes dW + Et$. *This means that for all* $\ell \geq 1$ *and* $0 \leq t_1, \ldots, t_\ell \leq 1$,

$$((W_n, \mathbb{W}_n)(t_1), \ldots, (W_n, \mathbb{W}_n)(t_\ell)) \to_{\mu_n} ((W, \mathbb{W})(t_1), \ldots, (W, \mathbb{W})(t_\ell)).$$

Here the stochastic integral $\int W \otimes dW$ is interpreted in the Itô sense.

Our proof of the Iterated WIP (Theorem 4.2.2) is inspired by the proof of the central limit theorem in [CM06, Chap. 7], which is based on Bernstein's 'big block-small block' technique. Let $0 < b < a < 1$. We split $\{0, \ldots, n-1\}$ into alternating big blocks of length $p = [n^a]$ and small blocks of length $q = [n^b]$. Let k denote the number of big blocks, which is equal to the number of small blocks. Then $k = [n/(p+q)] = O(n^{1-a})$. The last remaining block is of length at most $p+q$.

Let $\mathcal{B} \subset \{0, \ldots, n-1\}$ denote the set of terms contained in big blocks. Let $t \in [0,1]$. Then $W_n(t) = I_1(t) + I_2(t)$ and $\mathbb{W}_n(t) = J_1(t) + J_2(t) + J_3(t)$, where

$$\begin{aligned}
I_1(t) &= \frac{1}{n^{1/2}} \sum_{0\leq r<[nt]:\, r\in\mathcal{B}} v_n \circ T_n^r, \qquad I_2(t) = \frac{1}{n^{1/2}} \sum_{0\leq r<[nt]:\, r\notin\mathcal{B}} v_n \circ T_n^r, \\
J_1(t) &= \frac{1}{n} \sum_{0\leq r<s<[nt]:\, r,s\in\mathcal{B}} v_n \circ T_n^r \otimes v_n \circ T_n^s, \\
J_2(t) &= \frac{1}{n} \sum_{0\leq r<s<[nt]:\, r\notin\mathcal{B}, s\in\mathcal{B}} v_n \circ T_n^r \otimes v_n \circ T_n^s, \\
J_3(t) &= \frac{1}{n} \sum_{0\leq r<s<[nt]:\, s\notin\mathcal{B}} v_n \circ T_n^r \otimes v_n \circ T_n^s.
\end{aligned} \tag{4.2.3}$$

Remark 4.2.3. In [CM06, Chap. 7] the central limit theorem is proved under a hypothesis on decay of multiple correlations, where the Functional Correlation Bound (3.2.1) is only assumed for functions $G : M^q \to \mathbb{R}$ of the form $G(x_0, \ldots, x_{q-1}) = \prod_{i=0}^{q-1} v_i(x_i)$. This hypothesis is strong enough to control the characteristic function of $I_1(t)$. However, functions G which are not of the above form arise naturally when we consider the characteristic function of $J_1(t)$.

We first show that the terms $I_2(t)$, $J_2(t)$, $J_3(t)$ that involve small blocks can be neglected.

Lemma 4.2.4. *Suppose that* $a > \frac{b+1}{2}$. *Let* $t \in [0,1]$. *Then* $I_2(t) \to_{\mu_n} 0$, $J_2(t) \to_{\mu_n} 0$ *and* $J_3(t) \to_{\mu_n} 0$ *as* $n \to \infty$.

Proof. We show that $|J_3(t)|_{L^1(\mu_n)} \to 0$. By the same line of argument,

$|I_2(t)|_{L^1(\mu_n)} \to 0$ and $|J_2(t)|_{L^1(\mu_n)} \to 0$.

Write $\{0,\dots,[nt]-1\} \setminus \mathcal{B} = \bigcup_{i=1}^{k+1} C_i$ where C_i denotes the intersection of $\{0,\dots,[nt]-1\}$ with the ith small block for $1 \le i \le k$. Also, C_{k+1} denotes the intersection of $\{0,\dots,[nt]-1\}$ with the last remaining block. Write $C_i = \{\ell, \ell+1,\dots,u\}$. Then

$$\sum_{0\le r<s<[nt]:\, s\in C_i} v_n \circ T_n^r \otimes v_n \circ T_n^s = \sum_{r=0}^{\ell-1}\sum_{s=\ell}^{u} v_n \circ T_n^r \otimes v_n \circ T_n^s + \sum_{\ell\le r<s\le u} v_n \circ T_n^r \otimes v_n \circ T_n^s.$$

Hence by Theorem 3.2.4,

$$\begin{aligned}\Big|\sum_{0\le r<s<[nt]:\, s\in C_i} v_n \circ T_n^r \otimes v_n \circ T_n^s\Big|_{L^1(\mu_n)} &\le \Big|\sum_{r=0}^{\ell-1} v_n \circ T_n^r\Big|_{L^2(\mu_n)} \Big|\sum_{s=\ell}^{u} v_n \circ T_n^s\Big|_{L^2(\mu_n)} \\ &\quad + \Big|\sum_{\ell\le r\le s\le u} v_n \circ T_n^r \otimes v_n \circ T_n^s\Big|_{L^1(\mu_n)} \\ &\ll \ell^{1/2}\#C_i^{1/2} + \#C_i \ll \#C_i^{1/2} n^{1/2}. \quad (4.2.4)\end{aligned}$$

Let $1 \le i \le k$. Then $\#C_i \le q = [n^b]$. Also, $k = O(n^{1-a})$ and $\#C_{k+1} = O(n^a)$. Thus

$$\begin{aligned}|J_3(t)|_{L^1(\mu_n)} &\le \frac{1}{n}\sum_{i=1}^{k+1}\Big|\sum_{0\le r<s<[nt]:\, s\in C_i} v_n \circ T_n^r \otimes v_n \circ T_n^s\Big|_{L^1(\mu_n)} \\ &\ll \frac{1}{n}(n^{1-a}n^{\frac12(b+1)} + n^{\frac12(a+1)}) \ll n^{\frac12(b+1)-a} + n^{\frac12(a-1)} = o(1),\end{aligned}$$

as required. □

For $1 \le i \le k$ let

$$X_i = n^{-1/2}\sum_{0\le r<p} v_n \circ T_n^{r+(i-1)(p+q)},$$

$$\mathbb{X}_i = n^{-1}\sum_{0\le r<s<p} (v_n \circ T_n^r \otimes v_n \circ T_n^s) \circ T_n^{(i-1)(p+q)}.$$

For $0 \le t \le 1$ define

$$\widetilde{W}_n(t) = \sum_{1\le i\le [kt]} X_i, \quad \widetilde{\mathbb{W}}_n(t) = \sum_{1\le i<j\le [kt]} X_i \otimes X_j.$$

Lemma 4.2.5. *Suppose that $a > \frac{b+1}{2}$. Let $t \in [0,1]$. Then*

$$W_n(t) - \widetilde{W}_n(t) \to_{\mu_n} 0, \qquad \mathbb{W}_n(t) - \widetilde{\mathbb{W}}_n(t) - \sum_{i=1}^{[kt]} \mathbb{X}_i \to_{\mu_n} 0.$$

Proof. Recall the definitions of I_1 and J_1 from (4.2.3). Let $t \in [0,1]$. Since $[kt](p+q)$ is the first term of the $([kt]+1)$th big block,

$$I_1\left(\frac{[kt](p+q)}{n}\right) = \sum_{1\le i\le [kt]} X_i, \quad J_1\left(\frac{[kt](p+q)}{n}\right) = \sum_{1\le i\le [kt]} \mathbb{X}_i + \sum_{1\le i<j\le [kt]} X_i \otimes X_j.$$

Hence

$$W_n([kt]\tfrac{p+q}{n}) = \sum_{1\le i\le [kt]} X_i + I_2([kt]\tfrac{p+q}{n}) = \widetilde{W_n}(t) + I_2([kt]\tfrac{p+q}{n})$$

and similarly

$$\mathbb{W}_n([kt]\tfrac{p+q}{n}) = \widetilde{\mathbb{W}}_n(t) + \sum_{1\le i\le [kt]} \mathbb{X}_i + J_2([kt]\tfrac{p+q}{n}) + J_3([kt]\tfrac{p+q}{n}).$$

By Lemma 4.2.4, it follows that $W_n([kt]\frac{p+q}{n}) - \widetilde{W}_n(t) \to_{\mu_n} 0$ and $\mathbb{W}_n([kt]\frac{p+q}{n}) - \widetilde{\mathbb{W}}_n(t) - \sum_{i=1}^{[kt]} \mathbb{X}_i \to_{\mu_n} 0$. It remains to show that $\mathbb{W}_n(t) - \mathbb{W}_n([kt]\frac{p+q}{n}) \to_{\mu_n} 0$ and $W_n(t) - W_n([kt]\frac{p+q}{n}) \to_{\mu_n} 0$. Let $0 \le t' \le t$. Let $C = \{[nt'], \dots, [nt]-1\}$. By (4.2.4),

$$\begin{aligned}\left|\mathbb{W}_n(t) - \mathbb{W}_n(t')\right|_{L^1(\mu_n)} &= \left| n^{-1} \sum_{0\le r<s<[nt]:\ s\in C} v_n \circ T_n^r \otimes v_n \circ T_n^s \right|_{L^1(\mu_n)} \\ &\ll n^{-1/2} \# C^{1/2} \ll n^{-1/2}([nt]-[nt'])^{1/2} \\ &\ll (n^{-1} + t - t')^{1/2}.\end{aligned}$$

Now, $[kt]\frac{p+q}{n} = \left[\left[\frac{n}{p+q}\right]t\right]\frac{p+q}{n} \to t$ as $n \to \infty$ so $\left|\mathbb{W}_n(t) - \mathbb{W}_n([kt]\frac{p+q}{n})\right|_{L^1(\mu_n)} \to 0$. By a similar argument, $W_n(t) - W_n([kt]\frac{p+q}{n}) \to_{\mu_n} 0$. ∎

Proposition 4.2.6. *Suppose that $b > \gamma^{-1}$. Let $t \in [0,1]$. Then $\sum_{i=1}^{[kt]} \mathbb{X}_i \to_{\mu_n} tE$ as $n \to \infty$.*

Proof. First note that $[kt]/n \sim t/p$. Hence by Lemma 4.2.1 and the fact that

$\lim_{n\to\infty} E_n = E$,

$$\lim_{n\to\infty} \sum_{i=1}^{[kt]} \mathbb{E}_{\mu_n}[\mathbb{X}_i] = \lim_{n\to\infty} [kt]\mathbb{E}_{\mu_n}[\mathbb{X}_1] = \lim_{n\to\infty} \mathbb{E}_{\mu_n}\left[\frac{t}{p} \sum_{0\le r<s<p} v_n \circ T_n^r \otimes v_n \circ T_n^s\right]$$
$$= \lim_{n\to\infty} tp^{-1}\mathbb{E}_{\mu_n}[\mathbb{S}_{v_n}(p,n)] = tE.$$

It remains to show that

$$\lim_{n\to\infty} \mathbb{E}_{\mu_n}\left[\left|\sum_{i=1}^{[kt]}(\mathbb{X}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|\right] = 0.$$

Write $\mathbb{X}_i(x) = \Phi_i(T_n^{\ell_i}x, \dots, T_n^{u_i}x)$ where $\ell_i = (i-1)(p+q)$, $u_i = \ell_i + p - 1$ and

$$\Phi_i(y_0, \dots, y_{u_i-\ell_i}) = \frac{1}{n} \sum_{0\le r<s\le u_i-\ell_i} v_n(y_r) \otimes v_n(y_s).$$

Let $R = \max_i |\Phi_i|_\infty$ and define $F\colon B_{\mathbb{R}^{d\times d}}(0,R)^{[kt]} \to \mathbb{R}$ by $F(z_1, \dots, z_{[kt]}) = |\sum_{i=1}^{[kt]}(z_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])|$.

Let $(\widehat{\mathbb{X}}_i)$ be independent copies of $(\mathbb{X}_i)$. By Lemma 3.4.1,

$$\mathbb{E}_{\mu_n}\left[\left|\sum_{i=1}^{[kt]}(\mathbb{X}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|\right] \le A + \mathbb{E}\left[\left|\sum_{i=1}^{[kt]}(\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|\right]$$

where

$$|A| \le C \sum_{r=1}^{[kt]-1} (\ell_{r+1} - u_r)^{-\gamma}\left(|F|_\infty + \mathrm{Lip}(F) \sum_{i=1}^{[kt]} \sum_{j=0}^{u_i-\ell_i} [\Phi_i]_{\eta,j}\right).$$

Note that $|\Phi_i|_\infty \le \frac{p^2}{n} |v_n|_\infty^2$. By a similar calculation to the bound on $[\Phi_i]_{\eta,j}$ in the proof of Proposition 3.5.5, $[\Phi_i]_{\eta,j} \le \frac{1}{n}(u_i - \ell_i) \|v_n\|_\eta^2 = \frac{p-1}{n} \|v_n\|_\eta^2$.

Let $z = (z_1, \dots, z_{[kt]}), z' = (z'_1, \dots, z'_{[kt]}) \in (\mathbb{R}^{d\times d})^{[kt]}$. Then

$$|F(z) - F(z')| = \left|\sum_{i=1}^{[kt]} z_i - z'_i\right| \le \sum_{i=1}^{[kt]} |z_i - z'_i| = |z - z'|$$

so $\mathrm{Lip}(F) \le 1$. Moreover,

$$|F|_\infty \le \sum_{i=1}^{[kt]}(R + |\mathbb{E}_{\mu_n}[\mathbb{X}_i]\,|) \le \sum_{i=1}^{[kt]}(R + |\Phi_i|_\infty) \le \frac{2kp^2}{n} |v_n|_\infty^2.$$

Now $pk \le n$ and $\ell_{r+1} - u_r = q + 1 \ge n^b$ so

$$
\begin{aligned}
|A| &\le Ckq^{-\gamma}\left(\frac{2kp^2}{n}\,|v_n|_\infty^2 + \frac{kp^2}{n}\,\|v_n\|_\eta^2\right) \ll \frac{k^2p^2}{n}q^{-\gamma} \le nq^{-\gamma} \\
&\ll n^{1-b\gamma} = o(1).
\end{aligned}
$$

It remains to prove that $\mathbb{E}\Big[\big|\sum_{i=1}^{[kt]}(\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\big|\Big] \to 0$. Without loss of generality take $\gamma \le 2$. By von Bahr-Esseen's inequality (Proposition 3.5.1(i)),

$$
\mathbb{E}\left[\left|\sum_{i=1}^{[kt]}(\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|\right] \le \left|\sum_{i=1}^{[kt]}(\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|_\gamma \ll \left(\sum_{i=1}^{[kt]}\left|\widehat{\mathbb{X}}_i - \mathbb{E}[\mathbb{X}_i]\right|_\gamma^\gamma\right)^{1/\gamma}.
$$

Now by Theorem 3.2.4,

$$
\begin{aligned}
\left|\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i]\right|_\gamma &= |\mathbb{X}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i]|_\gamma \le 2\,|\mathbb{X}_i|_\gamma \\
&= \frac{2}{n}\left|\sum_{0 \le r < s < p} v_n \circ T_n^r \otimes v_n \circ T_n^s\right|_\gamma = O(p/n)
\end{aligned}
$$

so

$$
\mathbb{E}\left[\left|\sum_{i=1}^{[kt]}(\widehat{\mathbb{X}}_i - \mathbb{E}_{\mu_n}[\mathbb{X}_i])\right|\right] \ll ([kt](p/n)^\gamma)^{1/\gamma} \ll k^{(1-\gamma)/\gamma} = o(1),
$$

as required. □

Proposition 4.2.7. *Suppose that $a + \gamma b > 2$. Then*

$$
(\widetilde{W}_n, \widetilde{\mathbb{W}}_n) \to_{\mu_n} \left(W, \int W \otimes dW\right) \text{ in } D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}).
$$

Before proving this proposition, we record a result that will be useful both here and in Chapter 5. Let $(\widehat{X}_i)$ be independent copies of (X_i) and define $(W_n, \mathbb{W}_n) \in D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d})$ by

$$
(\widehat{W}_n, \widehat{\mathbb{W}}_n)(t) = \left(\sum_{1 \le i \le [kt]} \widehat{X}_i, \sum_{1 \le i < j \le [kt]} \widehat{X}_i \otimes \widehat{X}_j\right)
$$

for $t \in [0,1]$.

Lemma 4.2.8. *There exists a constant $C > 0$ such that for any $n \ge 1$,*

(i) $|\mathbb{E}_{\mu_n}[G(\widetilde{W}_n)] - \mathbb{E}[G(\widehat{W}_n)]| \le Cn^{3/2-a-b\gamma}\mathrm{Lip}(G)$ *for all Lipschitz functions*

$G : D([0,1],\mathbb{R}^d) \to \mathbb{R}$.

(ii) $|\mathbb{E}_{\mu_n}[H(\widetilde{W}_n, \widetilde{\mathbb{W}}_n)] - \mathbb{E}[H(\widehat{W}_n, \widehat{\mathbb{W}}_n)]\,| \le Cn^{2-a-b\gamma}\mathrm{Lip}(H)$ *for all Lipschitz functions* $H : D([0,1],\mathbb{R}^d \times \mathbb{R}^{d\times d}) \to \mathbb{R}$.

Proof. We proceed by applying the weak dependence lemma (Lemma 3.4.1). Note that

$$X_i(x) = \Phi_i(T_n^{\ell_i}x, \ldots, T_n^{u_i}x)$$

where $\ell_i = (i-1)(p+q)$, $u_i = \ell_i + p - 1$ and

$$\Phi_i(y_0, \ldots, y_{u_i-\ell_i}) = n^{-1/2}\sum_{r=0}^{u_i-\ell_i} v_n(y_r).$$

Let $0 \le r \le u_i - \ell_i$. Then $[\Phi_i]_{\eta,r} \le n^{-1/2}[v_n]_\eta$. Let $R = \max_i |\Phi_i|_\infty \le pn^{-1/2}|v_n|_\infty$. Define $\pi_k : B(0,R)^k \to D([0,1],\mathbb{R}^d \times \mathbb{R}^{d\times d})$ and $\rho_k : B(0,R)^k \to D([0,1],\mathbb{R}^d \times \mathbb{R}^{d\times d})$ by

$$\pi_k(x_1,\ldots,x_k)(t) = \sum_{i=1}^{[kt]} x_i, \qquad \rho_k(x_1,\ldots,x_k)(t) = \left(\sum_{i=1}^{[kt]} x_i, \sum_{1\le i<j\le [kt]} x_i \otimes x_j\right)$$

for $t \in [0,1]$. Then $\widetilde{W}_n = \pi_k(X_1,\ldots,X_k)$ and $\widehat{W}_n = \pi_k(\widehat{X}_1,\ldots,\widehat{X}_k)$. Similarly, $(\widetilde{W}_n, \widetilde{\mathbb{W}}_n) = \rho_k(X_1,\ldots,X_k)$ and $(\widehat{W}_n, \widehat{\mathbb{W}}_n) = \rho_k(\widehat{X}_1,\ldots,\widehat{X}_k)$.

Now for all $(x_1,\ldots,x_k), (x_1',\ldots,x_k') \in B(0,R)^k$,

$$\sup_{t\in[0,1]}\left|\sum_{i=1}^{[kt]} x_i - \sum_{i=1}^{[kt]} x_i'\right| \le \sum_{i=1}^{k} |x_i - x_i'|,$$

so $\mathrm{Lip}(\pi_k) \le 1$. Also,

$$\begin{aligned}\sup_{t\in[0,1]}\left|\sum_{1\le i<j\le [kt]} (x_i\otimes x_j - x_i'\otimes x_j')\right| &\le \sum_{1\le i<j\le k} |x_i\otimes x_j - x_i'\otimes x_j'| \\ &\le \sum_{1\le i<j\le k} |x_i\otimes (x_j - x_j')| + |(x_i - x_i')\otimes x_j'| \\ &\le 2kR\sum_{j=1}^{k} |x_j - x_j'|.\end{aligned}$$

Thus $\mathrm{Lip}(\rho_k) \le 1 + 2kR$.

Let $G : D([0,1],\mathbb{R}^d) \to \mathbb{R}$ be Lipschitz. Without loss of generality assume that $G(0) = 0$, for otherwise we can consider $G' = G - G(0)$. Thus

$$|G \circ \pi_k|_\infty \le \mathrm{Lip}(G)\,|\pi_k|_\infty \le kR\mathrm{Lip}(G) \ll n^{1/2}\mathrm{Lip}(G). \tag{4.2.5}$$

By applying Lemma 3.4.1 with $F = G \circ \pi_k$, we obtain that $|\mathbb{E}_{\mu_n}[G(\widetilde{W}_n)] - \mathbb{E}[G(\widehat{W}_n)]\,| \le A$, where

$$A = C\sum_{r=1}^{k-1}(\ell_{r+1} - u_r)^{-\gamma}\bigg(|G \circ \pi_k|_\infty + \mathrm{Lip}(G \circ \pi_k)\sum_{i=1}^{k}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j}\bigg).$$

Note that $u_i - \ell_i = p - 1$ and $\ell_{i+1} - u_i = q + 1 \ge n^b$ for $1 \le i \le k$. Thus

$$\sum_{r=1}^{k-1}(\ell_{r+1} - u_r)^{-\gamma} \le kn^{-b\gamma} \ll n^{1-a-b\gamma} \tag{4.2.6}$$

and

$$\sum_{i=1}^{k}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j} \le kpn^{-1/2}[v_n]_\eta \le n^{1/2}[v_n]_\eta. \tag{4.2.7}$$

By combining these bounds with (4.2.5) it follows that

$$\begin{aligned} A &\ll n^{1-a-b\gamma}\Big(|G \circ \pi_k|_\infty + \mathrm{Lip}(G \circ \pi_k)n^{1/2}\Big) \\ &\ll n^{1-a-b\gamma}\Big(\mathrm{Lip}(G)\,n^{1/2} + \mathrm{Lip}(G)\mathrm{Lip}(\pi_k)n^{1/2}\Big) \\ &\le 2n^{3/2-a-b\gamma}\mathrm{Lip}(G). \end{aligned}$$

This completes the proof of (i).

Let $H : D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d}) \to \mathbb{R}$ be Lipschitz. Without loss of generality assume that $H(0) = 0$. Thus

$$|H \circ \rho_k|_\infty \le \mathrm{Lip}(H)\,|\rho_k|_\infty \le kR\mathrm{Lip}(H)\mathrm{Lip}(\rho_k) \ll n^{1/2}\mathrm{Lip}(H)\mathrm{Lip}(\rho_k).$$

Now applying Lemma 3.4.1 with $F = H \circ \rho_k$ yields that

$$|\mathbb{E}_{\mu_n}[H(\widetilde{W}_n, \widetilde{\mathbb{W}}_n)] - \mathbb{E}[G(\widehat{W}_n, \widehat{\mathbb{W}}_n)]\,| \le A,$$

where

$$A = C\sum_{r=1}^{k-1}(\ell_{r+1} - u_r)^{-\gamma}\bigg(|H \circ \rho_k|_\infty + \mathrm{Lip}(H \circ \pi_k)\sum_{i=1}^{k}\sum_{j=0}^{u_i-\ell_i}[\Phi_i]_{\eta,j}\bigg).$$

Now by (4.2.6) and (4.2.7),

$$\begin{aligned}
A &\ll n^{1-a-b\gamma}\left(|H\circ\rho_k|_\infty + \mathrm{Lip}(H\circ\rho_k)n^{1/2}\right)\\
&\ll n^{3/2-a-b\gamma}\mathrm{Lip}(H)\mathrm{Lip}(\rho_k)\\
&\le n^{3/2-a-b\gamma}\mathrm{Lip}(H)(1+2kR) \ll n^{2-a-b\gamma}\mathrm{Lip}(H).
\end{aligned}$$

This completes the proof of (ii). □

We are now ready to complete the proof of Proposition 4.2.7.

Proof of Lemma 4.2.7. By the portmanteau theorem (Lemma 2.1.1), it suffices to prove that

$$\lim_{n\to\infty}\mathbb{E}_{\mu_n}\Big[H(\widetilde{W}_n,\widetilde{\mathbb{W}}_n)\Big] = \mathbb{E}\Big[H\left(W,\int W\otimes dW\right)\Big]$$

for all bounded Lipschitz $H : D([0,1],\mathbb{R}^d\times\mathbb{R}^{d\times d})\to\mathbb{R}$. Since $a+b\gamma>2$, by Lemma 4.2.8(ii), we have $\lim_{n\to\infty}\big|\mathbb{E}_{\mu_n}[H(\widetilde{W}_n,\widetilde{\mathbb{W}}_n)] - \mathbb{E}[H(\widehat{W}_n,\widehat{\mathbb{W}}_n)]\big| = 0$. It remains to show that

$$(\widehat{W}_n,\widehat{\mathbb{W}}_n)\to_{\mu_n}\left(W,\int W\otimes dW\right) \text{ in } D([0,1],\mathbb{R}^d\times\mathbb{R}^{d\times d}).$$

Indeed, once we have proved this it follows that

$$\lim_{n\to\infty}\mathbb{E}_{\mu_n}[H(\widetilde{W}_n,\widetilde{\mathbb{W}}_n)] = \lim_{n\to\infty}\mathbb{E}[H(\widehat{W}_n,\widehat{\mathbb{W}}_n)] = \mathbb{E}\Big[H\left(W,\int W\otimes dW\right)\Big],$$

completing the proof of this proposition.

Note that $n/k\sim p$. Hence, by Lemma 4.2.1 and the fact that $\lim_{n\to\infty}\Sigma_n = \Sigma$,

$$\begin{aligned}
\lim_{n\to\infty}\sum_{i=1}^{[kt]}\mathbb{E}[\widehat{X}_i\otimes\widehat{X}_i] &= \lim_{n\to\infty}[kt]\mathbb{E}[\widehat{X}_1\otimes\widehat{X}_1]\\
&= \lim_{n\to\infty}\frac{[kt]}{n}\mathbb{E}_{\mu_n}\left[\sum_{r=0}^{p-1}v_n\circ T_n^r\otimes\sum_{r=0}^{p-1}v_n\circ T_n^r\right]\\
&= \lim_{n\to\infty}\frac{t}{p}\mathbb{E}_{\mu_n}[S_{v_n}(p,n)\otimes S_{v_n}(p,n)] = \Sigma t.
\end{aligned}$$

Hence hypothesis (i) of Proposition A.1 is satisfied with $\chi_{n,i}=\widehat{X}_i$. Now by Theorem 3.2.4,

$$\begin{aligned}\sum_{i=1}^{k} \mathbb{E}\big[|\widehat{X}_i|^{2\gamma}\big] &= kn^{-\gamma}\mathbb{E}_{\mu_n}\left[\Big|\sum_{r=0}^{p-1} v_n \circ T_n^r\Big|^{2\gamma}\right] \ll kn^{-\gamma}p^{\gamma} \\ &\le k(kp)^{-\gamma}p^{\gamma} = k^{1-\gamma} = o(1)\end{aligned} \tag{4.2.8}$$

so hypothesis (ii) of Proposition A.1 is satisfied with $p = \gamma$. This completes the proof. □

We are now ready to prove the iterated WIP (Theorem 4.2.2).

Proof of Theorem 4.2.2. Since $\gamma > 1$, we can choose $0 < b < a < 1$ such that $b > \gamma^{-1}$, $a > \frac{b+1}{2}$ and $a + \gamma b > 2$. Then the conditions of Lemma 4.2.5 and Propositions 4.2.6 and 4.2.7 are satisfied. Let $0 \le t_1, t_2, \ldots, t_\ell \le 1$, $\ell \ge 1$. Write

$$\big((W_n, \mathbb{W}_n)(t_1), (W_n, \mathbb{W}_n)(t_2), \ldots, (W_n, \mathbb{W}_n)(t_\ell)\big) = K_1 + K_2 + K_3,$$

where

$$\begin{aligned}K_1 &= (A(t_1), A(t_2), \ldots, A(t_\ell)),\\ K_2 &= \left(\Big(0, \sum_{i=1}^{[kt_1]} \mathbb{X}_i\Big), \Big(0, \sum_{i=1}^{[kt_2]} \mathbb{X}_i\Big), \ldots, \Big(0, \sum_{i=1}^{[kt_\ell]} \mathbb{X}_i\Big)\right),\\ K_3 &= \big((\widetilde{W}_n, \widetilde{\mathbb{W}}_n)(t_1), (\widetilde{W}_n, \widetilde{\mathbb{W}}_n)(t_2), \ldots, (\widetilde{W}_n, \widetilde{\mathbb{W}}_n)(t_\ell)\big).\end{aligned}$$

Here $A(t) = (W_n(t) - \widetilde{W}_n(t), \mathbb{W}_n(t) - \widetilde{\mathbb{W}}_n(t) - \sum_{i=1}^{[kt]} \mathbb{X}_i)$.

By Lemma 4.2.5, $K_1 \to_{\mu_n} 0$ and by Proposition 4.2.6,

$$K_2 \to_{\mu_n} ((0, t_1 E), (0, t_2 E), \ldots, (0, t_\ell E)).$$

Moreover, by Proposition 4.2.7,

$$K_3 \to_{\mu_n} \left(\Big(W(t_1), \int_0^{t_1} W \otimes dW\Big), \Big(W(t_2), \int_0^{t_2} W \otimes dW\Big), \ldots, \Big(W(t_\ell), \int_0^{t_\ell} W \otimes dW\Big)\right).$$

Hence by Slutsky's theorem (Lemma 2.1.3),

$$K_1 + K_2 + K_3 \to_{\mu_n} \big((W, \mathbb{W})(t_1), (W, \mathbb{W})(t_2), \ldots, (W, \mathbb{W})(t_\ell)\big),$$

as required. □

4.2.2 Completing the proof of Theorem 4.1.2

We now have all the ingredients needed to prove Theorem 4.1.2.

Proof of Theorem 4.1.2. We proceed by applying [Che+22, Theorem 2.17], so we need to check Assumptions 2.11 and 2.12 from [Che+22].

By the arguments in the proof of [KKM22, Proposition 3.9], Assumptions 2.11 and 2.12(ii)(a) follow from (4.1.4) and Theorem 3.2.4. Since μ_n is T_n-invariant for all n, Assumption 2.12(i) also follows from Theorem 3.2.4 (cf. [Che+22, Remark 2.13]).

It remains to verify Assumption 2.12(ii)(b). Let $v_n \in \mathcal{C}^\eta(M, \mathbb{R}^d)$, $n \in \mathbb{N} \cup \{\infty\}$, with $\mathbb{E}_{\mu_n}[v_n] = 0$ and $\sup_{n \geq 1} \|v_n\|_\eta < \infty$. We assume that $\lim_{n\to\infty} |v_n - v_\infty|_\infty = 0$. Define Σ_n and E_n as in (4.2.1). It suffices to prove that $\lim_{n\to\infty} \Sigma_n = \Sigma_\infty$ and $\lim_{n\to\infty} E_n = E_\infty$. Assumption 2.12(ii)(b) then follows from Theorem 4.2.2, with

$$\mathfrak{B}_1(v, w) = \mathbb{E}_{\mu_\infty}[vw]\,, \quad \mathfrak{B}_2(v, w) = \sum_{\ell \geq 1}^{\infty} \mathbb{E}_{\mu_\infty}\Big[vw \circ T_\infty^\ell\Big]\,.$$

We show that $\lim_{n\to\infty} E_n = E_\infty$; the proof that $\lim_{n\to\infty} \Sigma_n = \Sigma_\infty$ is similar. By Lemma 4.2.1, the series $E_n = \sum_{\ell \geq 1} \mathbb{E}_{\mu_n}\big[v_n \otimes v_n \circ T_n^\ell\big]$ is convergent uniformly in n. Hence it suffices to show that $\lim_{n\to\infty} \mathbb{E}_{\mu_n}\big[v_n \otimes v_n \circ T_n^\ell\big] = \mathbb{E}_{\mu_\infty}\big[v_\infty \otimes v_\infty \circ T_\infty^\ell\big]$ for each fixed $\ell \geq 1$. Write $\mathbb{E}_{\mu_n}\big[v_n \otimes v_n \circ T_n^\ell\big] - \mathbb{E}_{\mu_\infty}\big[v_\infty \otimes v_\infty \circ T_\infty^\ell\big] = A(n) + B(n)$, where

$$\begin{aligned} A(n) &= \mathbb{E}_{\mu_n}\Big[v_n \otimes v_n \circ T_n^\ell - v_\infty \otimes v_\infty \circ T_n^\ell\Big]\,, \\ B(n) &= \mathbb{E}_{\mu_n}\Big[v_\infty \otimes v_\infty \circ T_n^\ell\Big] - \mathbb{E}_{\mu_\infty}\Big[v_\infty \otimes v_\infty \circ T_\infty^\ell\Big]\,. \end{aligned}$$

Now,

$$\Big|v_n \otimes v_n \circ T_n^\ell - v_\infty \otimes v_\infty \circ T_n^\ell\Big|_\infty \leq |v_n - v_\infty|_\infty |v_n|_\infty + |v_\infty|_\infty |v_n - v_\infty|_\infty$$

so $A(n) \to 0$ as $n \to \infty$. Finally, by (4.1.4), we have $B(n) \to 0$. □

4.3 Uniform families of nonuniformly hyperbolic maps

In Theorem 3.3.4 we showed that mixing transformations modelled by a Young tower with $O(n^{-\beta})$, $\beta > 1$ tails satisfy the Functional Correlation Bound. Moreover, we showed that the implicit constant in the Functional Correlation Bound depends continuously on various system constants associated with T. Consequently, it is straightforward to give necessary conditions for a family of nonuniformly hyperbolic transformations to satisfy the Functional Correlation Bound with a uniform rate.

Let $T_n \colon M \to M$, $n \geq 1$ be a family of mixing transformations with invariant probability measures μ_n. Suppose that T_n is modelled by a Young tower for each n (see Definition 3.3.1).

Definition 4.3.1. *Let $\beta > 1$. We call $T_n \colon M \to M$ a uniform family of nonuniformly hyperbolic maps with $O(k^{-\beta})$ tails if:*

(i) The constants $D_0 > 0$, $\theta \in (0,1)$ in Subsection 3.3.1 can be chosen independent of $n \geq 1$.

(ii) There exists $C_\phi > 0$ such that the return time functions $\phi_n \colon Y_n \to \mathbb{Z}^+$ satisfy $\mu_{Y_n}(\phi_n \geq k) \leq C_\phi k^{-\beta}$ for all $n, k \geq 1$.

(iii) There exist $\delta > 0$ and $K > 0$ such that for all n there exists $I_n \subset [1, K]$ with $\mu_{Y_n}(\phi_n = k) \geq \delta$ for $k \in I_n$ and $\gcd\{I_n\} = \gcd\{\phi_n(y) \colon y \in Y_n\}$.

We are now ready to state the main result of this section, which follows immediately from Theorem 3.3.4:

Theorem 4.3.2. *Let T_n be a uniform family of nonuniformly hyperbolic maps with $O(k^{-\beta})$ tails. Then the family T_n satisfies the Functional Correlation Bound uniformly with rate $k^{-(\beta-1)}$.*

4.4 Examples of families of dynamical systems

In this section we consider examples of families of dynamical systems for which we can verify the hypotheses of Theorem 4.1.2.

4.4.1 Intermittent Baker's maps

Let $I = [0,1]$, $M = I^2$. Fix a family of intermittent Baker's maps $T_n : M \to M$, $n \in \mathbb{N} \cup \{\infty\}$, as in (4.1.5) with parameters $\alpha_n \in (0, \frac{1}{2})$ such that $\lim_{n\to\infty} \alpha_n = \alpha_\infty \in (0, \frac{1}{2})$. Recall that T_n is a skew product map of the form

$$T_n(x,z) = (\bar{T}_n(x), h_n(x,z)), \quad h_n(x,z) = \begin{cases} g_{n,0}(z), & x < \frac{1}{2}, \\ g_{n,1}(z), & x \geq \frac{1}{2}, \end{cases}$$

where $\bar{T}_n : I \to I$ is the Liverani-Saussol-Vaienti map with parameter α_n and $g_{n,0}$, $g_{n,1}$ are the inverse branches of $\bar{T}_n$. The projection $\pi : M \to I$, $\pi(x,z) = x$

defines a semiconjugacy between T_n and $\bar{T}_n$. By [LSV99], there is a unique $\bar{T}_n$-invariant ergodic probability measure $\bar{\mu}_n$ which is absolutely continuous with respect to Lebesgue. Let μ_n be the unique T_n-invariant probability measure such that $\pi_*\mu_n = \bar{\mu}_n$ (we construct this measure in Lemma 4.4.3).

Proposition 4.4.1. *T_n, $n \in \mathbb{N} \cup \{\infty\}$, is a uniform family of nonuniformly hyperbolic maps with $O(k^{-1/\alpha})$ tails, where $\alpha = \sup_n \alpha_n$.*

Proof. For each n, we take $\bar{Y} = [1/2, 1]$ and let $\bar{\phi}_n : \bar{Y} \to \mathbb{Z}^+$ be the first return time to $\bar{Y}$, i.e. $\bar{\phi}_n(y) = \inf\{k \geq 1 : \bar{T}_n^k(y) \in \bar{Y}\}$. Then by [You99, Section 6], the first return map $\bar{F}_n = \bar{T}_n^{\bar{\phi}_n} : \bar{Y} \to \bar{Y}$ is a Gibbs-Markov map and there exists a constant $C > 0$ such that $\bar{\mu}_n(\bar{\phi}_n > k) \leq Ck^{-1/\alpha_n}$ for all $k \geq 1$. Moreover, by [KKM17, Example 5.1] both the constant C and the constants that appear in the definition of Gibbs-Markov map can be chosen independently of n. It follows that condition (ii) in Definition 4.3.1 is satisfied.

Note that $\{\bar{\phi}_n = 1\} = [3/4, 1]$. By [LSV99, Lemma 2.4], $\inf_n \inf_I d\bar{\mu}_n/d\text{Leb} > 0$. Hence $\inf_n \bar{\mu}_n|_{\bar{Y}}(\bar{\phi}_n = 1) > 0$, and so condition (iii) in Definition 4.3.1 is satisfied.

Let $Y = \bar{Y} \times I$ and let $\phi_n : Y \to \mathbb{Z}^+$ be the first return time to Y. Then $\phi_n = \bar{\phi}_n \circ \pi|_Y$ and $\pi|_Y$ defines a semiconjugacy between $F_n = T_n^{\phi_n} : Y \to Y$ and $\bar{F}_n$.

We now complete the proof of condition (i) in Definition 4.3.1 by verifying that (3.3.1) and (3.3.3) hold with constants D_0, θ that are uniform in n. Denote $\psi_{n,k}(x) = \#\{j = 0, \ldots, k-1 : \bar{T}_n^j x \in \bar{Y}\}$. We claim that

$$d(T_n^k(x_1, z_1), T_n^k(x_2, z_2)) \leq 2\left(\tfrac{1}{2}\right)^{s(x_1,x_2)-\psi_{n,k}(x_1)} + \left(\tfrac{1}{2}\right)^{\psi_{n,k}(x_1)} d(z_1, z_2) \qquad (4.4.1)$$

for all $k \geq 1$, $n \in \mathbb{N} \cup \{\infty\}$, $(x_1, z_1), (x_2, z_2) \in Y$. It is straightforward to check that both (3.3.1) and (3.3.3) follow from this claim with $D_0 = 2$, $\theta = \frac{1}{2}$.

Let $x_1, z_1, z_2 \in I$. Then for $i = 1, 2$,

$$T_n^k(x_1, z_i) = (\bar{T}_n^k x_1, g_{n,a(k-1)} \circ \cdots \circ g_{n,a(0)}(z_i)), \qquad (4.4.2)$$

where $a(j) = \mathbb{1}\{\bar{T}_n^j x_1 \in \bar{Y}\}$. Since $\left|g'_{n,0}\right|_\infty \geq 1$ and $\left|g'_{n,1}\right|_\infty = \frac{1}{2}$, by the mean value theorem it follows that

$$d(T_n^k(x_1, z_1), T_n^k(x_1, z_2)) \leq \prod_{j=0}^{k-1} \left|g'_{n,a(j)}\right|_\infty d(z_1, z_2) \leq \left(\tfrac{1}{2}\right)^{\psi_{n,k}(x_1)} d(z_1, z_2). \qquad (4.4.3)$$

Let $x_1, x_2 \in \bar{Y}$. Without loss of generality assume that $s(x_1, x_2) > \psi_{n,k}(x_1)$, for otherwise (4.4.1) is satisfied trivially. Since ϕ_n is the first return time to $\bar{Y}$, it follows that for all $0 \leq j < k$ we have $\bar{T}_n^j x_1 \in \bar{Y}$ if and only if $\bar{T}_n^j x_2 \in \bar{Y}$. Hence by (4.4.2), $d(T_n^k(x_1,z_1), T_n^k(x_2,z_1)) = d(\bar{T}_n^k x_1, \bar{T}_n^k x_2)$. Note that $\bar{F}_n^{\psi_{n,k}+1} = \bar{T}_n^r$, where $r = \sum_{\ell=0}^{\psi_{n,k}} \bar{\phi}_n \circ \bar{F}_n^\ell$. Since $r(x_1) = r(x_2) > k$ and $\bar{T}_n' \geq 1$, it follows that

$$d(\bar{T}_n^k x_1, \bar{T}_n^k x_2) \leq d(\bar{T}_n^{r(x_1)} x_1, \bar{T}_n^{r(x_1)} x_2) = d(\bar{F}_n^{\psi_{n,k}+1} x_1, \bar{F}_n^{\psi_{n,k}+1} x_2). \tag{4.4.4}$$

Now $\bar{T}_n' \geq 1$ and $\bar{T}_n' = 2$ on $\bar{Y}$ so

$$d(\bar{F}_n y, \bar{F}_n y') \geq 2d(y, y') \tag{4.4.5}$$

whenever $y, y' \in \bar{Y}$ belong to the same partition element. By [Alv20, Lemma 3.2], it follows that

$$d(\bar{F}_n^{\psi_{n,k}+1} x_1, \bar{F}_n^{\psi_{n,k}+1} x_2) \leq \left(\tfrac{1}{2}\right)^{s(x_1,x_2) - (\psi_{n,k}(x_1)+1)}. \tag{4.4.6}$$

For completeness, let us recall the proof of (4.4.6). By the definition of the separation time s, the points $\bar{F}^j x_1$ and $\bar{F}^j x_2$ belong to the same partition element for all $0 \leq j < s(x_1, x_2)$. Hence by backward induction on j, the bound (4.4.5) implies that

$$d(\bar{F}_n^j x_1, \bar{F}_n^j x_2) \leq \left(\tfrac{1}{2}\right)^{s(x_1,x_2)-j} d(\bar{F}_n^{s(x_1,x_2)} x_1, \bar{F}_n^{s(x_1,x_2)} x_2) \leq \left(\tfrac{1}{2}\right)^{s(x_1,x_2)-j}$$

for all $0 \leq j \leq s(x_1, x_2)$. Now by assumption, $\psi_{n,k}(x_1) = \psi_{n,k}(x_2) < s(x_1, x_2)$ and so (4.4.6) follows by taking $j = \psi_{n,k}(x_1) + 1$. Finally, we obtain the claim (4.4.1) by combining (4.4.3) with (4.4.4) and (4.4.6). □

By Theorem 4.3.2, it follows that the family T_n satisfies the Functional Correlation Bound uniformly with rate $k^{-(1/\alpha - 1)}$. In the remainder of this subsection we verify that condition (4.1.4) is satisfied by verifying the conditions (A1), (A2) from Remark 4.1.3. It follows that our main result, Theorem 4.1.2, applies to the family T_n.

By [FT09] (see also [Kor16, BT16]), $\bar{\mu}_n$ is *strongly statistically stable*, i.e. $d\bar{\mu}_n/d\mathrm{Leb} \to d\bar{\mu}_\infty/d\mathrm{Leb}$ in $L^1(\mathrm{Leb})$. The following lemma immediately implies that (A2) holds and will also be useful in the proof that μ_n is statistically stable, as required by (A1).

Lemma 4.4.2. *For all $a > 0$, for all $j \in \mathbb{N}$,*

$$\lim_{n\to\infty} \bar{\mu}_n\Big(x : \sup_{z\in I} d(T_n^j(x,z), T_\infty^j(x,z)) > a\Big) = 0. \tag{4.4.7}$$

Proof. Let $\varepsilon > 0$. Choose $K \subset I$ compact such that $\frac{1}{2} \notin \bar{T}_\infty^i(K)$ for $0 \leq i \leq j$ and

$\bar{\mu}_\infty(K) \geq 1 - \varepsilon$. Then $T_n^j \to T_\infty^j$ uniformly on $K \times I$ so for all $a > 0$,

$$\limsup_{n\to\infty} \bar{\mu}_\infty\Big(x : \sup_{z\in I} d(T_n^j(x,z), T_\infty^j(x,z)) > a\Big) \leq \bar{\mu}_\infty(I \setminus K) < \varepsilon.$$

It follows that (4.4.7) holds with $\bar{\mu}_\infty$ in place of $\bar{\mu}_n$. The inequality (4.4.7) follows by strong statistical stability. $\square$

Note that T_∞ is continuous on $I^2 \setminus (\{\frac{1}{2}\} \times I)$ so T_∞ is continuous μ_∞-a.e. In the remainder of this subsection, we complete the verification of condition (A2) by showing that μ_n is statistically stable. We closely follow the strategy that Alves & Soufi [AS14] used to prove statistical stability for the Poincaré maps of geometric Lorenz attractors.

First let us recall the standard procedure for constructing invariant measures for skew products with contracting fibres. Given a bounded, measurable function $\phi : M \to \mathbb{R}$, define $\phi^+ : I \to \mathbb{R}$ by $\phi^+(x) = \sup_{z\in I} \phi(x,z)$.

Lemma 4.4.3. *Let $n \in \mathbb{N} \cup \{\infty\}$. There exists a unique probability measure μ_n such that for any continuous function $\phi : M \to \mathbb{R}$,*

$$\int_M \phi \, d\mu_n = \lim_{m\to\infty} \int_I (\phi \circ T_n^m)^+ d\bar{\mu}_n. \tag{4.4.8}$$

Moreover, the convergence is uniform in n. Besides, μ_n is ergodic and is the unique T_n-invariant probability measure such that $\pi_* \mu_n = \bar{\mu}_n$.

Proof. We first show that the maps T_n uniformly contract fibres in the sense that

$$\operatorname{diam}\big(T_n^m \pi^{-1}(x)\big) \to 0 \text{ as } m \to \infty, \text{ uniformly in } x \text{ and } n. \tag{4.4.9}$$

Fix x and n. By (4.4.2) and the fact that $g_{n,0}$ and $g_{n,1}$ are inverse branches of $\bar{T}_n$,

$$T_n^m \pi^{-1}(x) = \{\bar{T}_\infty^m(x)\} \times H(I)$$

where $H : I \to I$ is an inverse branch of $\bar{T}_n^m$. By [Lep17, equation (5)], there exists $C > 0$ such that for all $m, n \geq 1$, for any inverse branch H of $\bar{T}_n^m$ we have $\operatorname{diam}(H(I)) \leq C m^{-1/\sup_n \alpha_n}$. This proves (4.4.9). The rest of the proof that the limit (4.4.8) exists and the convergence is uniform in n proceeds exactly as in [AS14, Proposition 3.3] (with P_n and f_n changed to T_n and $\bar{T}_n$). In [Ara$^+$09, Corollary 6.4] it is shown that μ_n indeed defines a T_n-invariant probability measure and that the ergodicity of μ_n follows from the ergodicity of $\bar{\mu}_n$.

Let ν be a T_n-invariant probability measure such that $\pi_*\nu = \bar{\mu}_n$. Let $\phi : M \to \mathbb{R}$ be continuous. Then for all m,

$$\left| \int_M \phi \, d\nu - \int_I (\phi \circ T_n^m)^+ d\bar{\mu}_n \right| = \left| \int_M \big(\phi \circ T_n^m - (\phi \circ T_n^m)^+ \circ \pi\big) d\nu \right|$$
$$\leq \int_M |\phi \circ T_n^m - (\phi \circ T_n^m)^+ \circ \pi| d\nu.$$

Let $(x,y) \in M$. Note that $(\phi \circ T_n^m)^+(\pi(x,y)) = \sup_{z \in I} \phi \circ T_n^m(x,z)$. Hence by (4.4.9) and uniform continuity of ϕ,

$$\sup_{(x,y)\in M} |\phi \circ T_n^m(x,y) - (\phi \circ T_m^n)^+(\pi(x,y))| \leq \sup_x \sup_{p_1,p_2 \in T_n^m \pi^{-1}(x)} |\phi(p_1) - \phi(p_2)| \to 0$$

as $m \to \infty$. Hence $\int \phi \, d\nu = \lim_{m\to\infty} \int (\phi \circ T_n^m)^+ d\bar{\mu}_n = \int \phi \, d\mu_n$. Since ϕ is an arbitrary continuous function, $\nu = \mu_n$ as required. □

Lemma 4.4.4. *For all $m \geq 1$,*

$$\lim_{n\to\infty} \int (\phi \circ T_n^m)^+ d\bar{\mu}_n = \int (\phi \circ T_\infty^m)^+ d\bar{\mu}_\infty.$$

Proof. We proceed as in [AS14, Lemma 3.2].

Write $\int (\phi \circ T_n^m)^+ d\bar{\mu}_n - \int (\phi \circ T_\infty^m)^+ d\bar{\mu}_\infty = I_n + J_n$, where

$$I_n = \int (\phi \circ T_n^m)^+ - (\phi \circ T_\infty^m)^+ \, d\bar{\mu}_n, \quad J_n = \int (\phi \circ T_\infty^m)^+ d\bar{\mu}_n - \int (\phi \circ T_\infty^m)^+ d\bar{\mu}_\infty.$$

Now for all $x \in I$,

$$|(\phi \circ T_n^m)^+(x) - (\phi \circ T_\infty^m)^+(x)| \leq \sup_{y \in I} |\phi \circ T_n^m(x,z) - \phi \circ T_\infty^m(x,z)|. \tag{4.4.10}$$

Since M is compact, ϕ is uniformly continuous on M. Hence for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|\phi(z) - \phi(z')| < \varepsilon$ for all $z, z' \in M$ with $d(z,z') < \delta$. Let

$$S_n = \Big\{ x \in I : \sup_{y \in I} d(T_n^m(x,y), T_\infty^m(x,y)) \geq \delta \Big\}.$$

Then by (4.4.10),

$$|I_n| \leq \int_{S_n} |(\phi \circ T_n^m)^+ - (\phi \circ T_\infty^m)^+| \, d\bar{\mu}_n + \int_{I \setminus S_n} |(\phi \circ T_n^m)^+ - (\phi \circ T_\infty^m)^+| \, d\bar{\mu}_n$$
$$\leq 2 \, |\phi|_\infty \, \bar{\mu}_n(S_n) + \varepsilon.$$

By Lemma 4.4.2, $\bar{\mu}_n(S_n) \to 0$ as $n \to \infty$. Since $\varepsilon > 0$ is arbitrary, it follows that $I_n \to 0$.

Finally, note that

$$|J_n| = \left| \int (\phi \circ T_\infty^m)^+ \left(\frac{d\bar{\mu}_n}{d\text{Leb}} - \frac{d\bar{\mu}_\infty}{d\text{Leb}} \right) d\text{Leb} \right| \le |\phi|_\infty \left| \frac{d\bar{\mu}_n}{d\text{Leb}} - \frac{d\bar{\mu}_\infty}{d\text{Leb}} \right|_{L^1(\text{Leb})}.$$

Hence by strong statistical stability, $J_n \to 0$ as $n \to \infty$. □

We can now complete the proof that μ_n is statistically stable, i.e. $\mu_n \xrightarrow{w} \mu_\infty$. Let $\phi : M \to \mathbb{R}$ be continuous. Then by Lemmas 4.4.3 and 4.4.4,

$$\int_M \phi \, d\mu_\infty = \lim_{m \to \infty} \int_I (\phi \circ \bar{T}_\infty^m)^+ d\bar{\mu}_\infty = \lim_{m \to \infty} \lim_{n \to \infty} \int_I (\phi \circ \bar{T}_n^m)^+ d\bar{\mu}_n.$$

Since $\int_I (\phi \circ \bar{T}_n^m) d\bar{\mu}_n \to \int_M \phi \, d\mu_n$ as $m \to \infty$ uniformly in n, we can swap the limits as $m \to \infty$ and $n \to \infty$ in the above expression. Thus

$$\int_M \phi \, d\mu_\infty = \lim_{n \to \infty} \lim_{m \to \infty} \int_I (\phi \circ \bar{T}_n^m)^+ d\bar{\mu}_n = \lim_{n \to \infty} \int_M \phi \, d\mu_n,$$

as required.

4.4.2 Externally forced dispersing billiards

A Sinai billiard table on the two-torus $\mathbb{T}^2$ is a set of the form $Q = \mathbb{T}^2 \setminus \cup_i B_i$ where $\{B_i\}$ is a finite collection of open sets such that $\bar{B}_i \cap \bar{B}_j = \emptyset$ for $i \neq j$. It is assumed that the sets B_i have C^3 boundaries with positive curvature. The billiard flow on $Q \times S^1$ is induced by the motion of a particle that moves in straight lines at unit speed on Q and collides elastically with the boundary ∂Q. We say that the table has finite horizon if there exists a constant $L > 0$ such that any line of length L in $\mathbb{T}^2$ intersects ∂Q.

In [Che01, Che08] Chernov studied perturbations of the finite horizon Sinai billiard flow where a small stationary force F acts on the particle between its collisions with ∂Q. We refer to [Che01, Section 2] for the precise details of the model. In particular, it is assumed that the force preserves an additional integral of motion and that the phase space obtained by restricting to one of its level sets is a compact 3-dimensional manifold with boundary.

Consider the flow obtained upon restricting to one of these level sets. The assumptions then guarantee that the collision map T_F with the table can be parametrised on the same space $M = \partial Q \times [-\pi/2, \pi/2]$ as the collision map of the unperturbed Sinai billiard flow. Let $(F_n)_{n\in\mathbb{N}}$ be a sequence of admissible forces such that $F_n \to F_\infty = 0$ in C^2 and define $T_n = T_{F_n} : M \to M$. By [Che01, Theorem 2.1], the map T_n admits a unique SRB measure μ_n for all $n \in \mathbb{N} \cup \{\infty\}$.

Lemma 4.4.5. *For all $\gamma > 1$ the family $T_n : M \to M$, $n \in \mathbb{N} \cup \{\infty\}$, satisfies the Functional Correlation Bound uniformly with rate $k^{-\gamma}$.*

Remark 4.4.6. In principle, it should be possible to prove this lemma by verifying that the family T_n is a uniform family of nonuniformly hyperbolic maps and applying Theorem 4.3.2. Indeed, for each n, the system T_n is modelled by a Young tower with exponential tails [Che01]. However, the construction of the base of the tower in [Che01] is quite intricate so it seems difficult to check condition (iii) in Definition 4.3.1.

Proof. In [LS17], Leppänen & Stenlund considered the finite horizon Sinai billiard map and proved a functional correlation bound for separately dynamically Hölder functions. Recall the definition of the past/future separation times $s_\pm$ and the dynamically Hölder function classes $\mathcal{H}_\pm$ from [LS17, Section 2].

We first show that separately Hölder functions are separately dynamically Hölder with parameters independent of n. By [Che08, p.95], there exist constants $C > 0$, $\Lambda > 1$ independent of n such that $d(x, y) \le C\Lambda^{-s_+(x,y)}$ whenever $x, y \in M$ belong to the same local unstable manifold. Similarly, $d(x, y) \le C\Lambda^{-s_-(x,y)}$ whenever $x, y \in M$ belong to the same local stable manifold. Let $v \in C^\eta(M)$. It follows that

$$|v(x) - v(y)| \le [v]_\eta d(x, y)^\eta \le C^\eta [v]_\eta (\Lambda^{-\eta})^{s_+(x,y)}$$

whenever x and y belong to the same local unstable manifold. Hence $v \in \mathcal{H}_+(C^\eta[v]_\eta, \Lambda^{-\eta})$. Similarly, $v \in \mathcal{H}_-(C^\eta[v]_\eta, \Lambda^{-\eta})$. Let $G : M^q \to \mathbb{R}$, $q \ge 1$ be separately η-Hölder and set $c = C^\eta \max_i [G]_{\eta,i}$, $\vartheta = \Lambda^{-\eta}$. Then $G(x_0, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_{q-1}) \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$ for all $x_0, \dots, x_{q-1} \in M$, $0 \le i < q$.

It remains to explain why the arguments used in the proof of [LS17, Theorem 2.4] go through with system constants M_0, M_1 and θ_0, θ_1 uniform in n. The result then follows by applying [LS17, Theorem 2.4] with $K = 2$, $F = G$ and c, ϑ as defined above.

Note that [LS17, Lemma 4.1] merely gives the usual decomposition of μ into a standard family. Let $\{(\xi_q, \nu_q) : q \in \mathcal{Q}\}$ be as defined in that lemma. By [Che08, p. 96], $\{(\xi_q, \nu_q) : q \in \mathcal{Q}\}$ is a proper standard family. In particular, there exists a constant $M_1 > 0$ independent of n such that $\lambda(\{q \in \mathcal{Q} : |\xi_q| \le \varepsilon\}) \le M_1\varepsilon$ for all $\varepsilon >$

0. By [CM06, p.171], it follows that there exists a constant M_1' independent of n such that $\int_{\mathcal{Q}} |\xi_q|^{-1} d\lambda(q) \le M_1'$, so [Lep17, Lemma 4.3] goes through. Finally, it follows from the growth lemma ([Che01, Proposition 5.3]) and the equidistribution property ([Che08, Proposition 2.2]) that [LS17, Lemma 4.2] goes through with constants a_0, M_0 and θ_0 that are uniform in n. The rest of the proof of [LS17, Theorem 2.4] proceeds exactly as in [LS17]. □

We finish this subsection by showing that condition (4.1.4) is satisfied. It follows that Theorem 4.1.2 applies to the family T_n.

Lemma 4.4.7. *Condition* (4.1.4) *is satisfied.*

Proof. For $n \in \mathbb{N} \cup \{\infty\}$, Demers & Zhang [DZ13] considered the action of the transfer operator $\mathcal{L}_n$ associated with T_n on certain spaces of distributions. In particular, if ν is a finite signed measure, then $\mathcal{L}_n \nu = (T_n)_* \nu$.

Fix $\eta \in (0, 1]$. The article [DZ13] constructs Banach spaces $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ and $(\mathcal{B}_w, \|\cdot\|_{\mathcal{B}_w})$ with the following properties:

(i) There is a sequence of continuous embeddings $\mathcal{B} \hookrightarrow \mathcal{B}_w \hookrightarrow (C^\eta(M))'$.

(ii) For each n, $\mathcal{L}_n$ is a well-defined bounded linear operator on both $\mathcal{B}$ and $\mathcal{B}_w$. Moreover, $\sup_n \|\mathcal{L}_n\|_{\mathcal{B}_w} < \infty$.

(iii) ([DZ13, Theorem 2.2]) For each n, we have $\mu_n \in \mathcal{B}$ and μ_n is the unique element of $\mathcal{B}$ such that $\mathcal{L}_n \mu_n = \mu_n$ and $\mu_n(1) = 1$.

(iv) ([DZ13, Theorem 2.11]) $\|\mathcal{L}_n - \mathcal{L}_\infty\|_{\mathcal{B} \to \mathcal{B}_w} \to 0$ as $n \to \infty$. By [DZ13, Theorem 2.1], it follows that $\mu_n \to \mu_\infty$ in $\mathcal{B}_w$.

(v) ([DZ13, Lemma 5.3]) Let $v \in C^\eta(M)$. Then v is a bounded multiplier on $\mathcal{B}$ (that is, $h \mapsto vh$ is a well-defined bounded operator on $\mathcal{B}$). Moreover, v is a bounded multiplier on $\mathcal{B}_w$.[1]

Fix $v \in C^\eta(M)$ and $k \ge 0$. For $n \in \mathbb{N} \cup \{\infty\}$, define a signed probability measure ν_n by $\nu_n = v\mu_n$. Then $\mathcal{L}_n^k \nu_n(w) = \nu_n(w \circ T_n^k) = \mu_n(vw \circ T_n^k)$ for all $w \in C^\eta(M)$, so it sufficient to prove that $\mathcal{L}_n^k \nu_n \to \mathcal{L}_\infty^k \nu_\infty$ in $(C^\eta(M))'$. Now $\mu_n \in \mathcal{B}$ so by properties (v) and (ii) we have $\nu_n \in \mathcal{B}$ and $\mathcal{L}_n^k \nu_n \in \mathcal{B}$. Hence by (i), it suffices to show that $\mathcal{L}_n^k \nu_n \to \mathcal{L}_\infty^k \nu_\infty$ in $\mathcal{B}_w$.

[1]Since the definitions of the weak norm and the strong stable norm are similar, this follows easily from the arguments used to bound the strong stable norm at the beginning of the proof of [DZ13, Lemma 5.3].

Write $\mathcal{L}_n^k \nu_n - \mathcal{L}_\infty^k \nu_\infty = I_n + J_n$, where $I_n = \mathcal{L}_n^k(\nu_n - \nu_\infty)$ and

$$J_n = \mathcal{L}_n^k \nu_\infty - \mathcal{L}_\infty^k \nu_\infty = \sum_{j=0}^{k-1} \mathcal{L}_n^j (\mathcal{L}_n - \mathcal{L}_\infty) \mathcal{L}_\infty^{k-j-1} \nu_\infty.$$

By properties (ii) and (iv),

$$\|I_n\|_{\mathcal{B}_w} \le \sup_n \left\| \mathcal{L}_n^k \right\|_{\mathcal{B}_w} \|v\mu_n - v\mu_\infty\|_{\mathcal{B}_w} \le C_v \sup_n \left\| \mathcal{L}_n^k \right\|_{\mathcal{B}_w} \|\mu_n - \mu_\infty\|_{\mathcal{B}_w} \to 0$$

as $n \to \infty$. Using properties (ii) and (iv) again, we have

$$\|J_n\|_{\mathcal{B}_w} \le \sum_{j=0}^{k-1} \sup_n \left\| \mathcal{L}_n^j \right\|_{\mathcal{B}_w} \|\mathcal{L}_n - \mathcal{L}_\infty\|_{\mathcal{B} \to \mathcal{B}_w} \left\| \mathcal{L}_\infty^{k-j-1} \right\|_{\mathcal{B}} \|\nu_\infty\|_{\mathcal{B}} \to 0$$

as $n \to \infty$. □

Remark 4.4.8. Note that we have not used any facts about the anisotropic Banach spaces in [DZ13] apart from properties (i)-(v). These properties arise naturally in situations where statistical stability is proved by Keller-Liverani perturbation theory (see e.g. [GL06, DL08]).

Chapter 5

Rates of convergence in the multidimensional weak invariance principle

5.1 Statement of results

Let $T : M \to M$ be an ergodic, measure-preserving transformation on a probability space (M, μ). Let $v \in L^2(M, \mathbb{R}^d)$ be mean zero. Define a random process $W_n \in D([0,1], \mathbb{R}^d)$ by

$$W_n(t) = n^{-1/2} \sum_{r=0}^{[nt]-1} v \circ T^r \qquad \text{for } t \in [0,1].$$

Recall that v is said to satisfy the *weak invariance principle (WIP)* if $W_n \xrightarrow{w} W$ where W is a Brownian motion. As before, we endow $D([0,1], \mathbb{R}^d)$ with the topology induced by the sup-norm $|\cdot|_\infty$.

Throughout this chapter we assume that $T : M \to M$ is a map that satisfies the Functional Correlation Bound with rate $n^{-\gamma}$, $\gamma > 1$. Fix $d \geq 1$ and let $v \in \mathcal{C}^\eta(M, \mathbb{R}^d)$ be mean zero. By our main homogenisation result (Theorem 4.1.1), we know that $W_n \xrightarrow{w} W$ where W is a Brownian motion with covariance

$$\Sigma = \mathbb{E}_\mu[v \otimes v] + \sum_{\ell \geq 1} \left(\mathbb{E}_\mu[v \otimes v \circ T^\ell] + \mathbb{E}_\mu[v \circ T^\ell \otimes v] \right)$$

so v satisfies the WIP.

In this chapter, we consider the rate at which W_n weakly converges to W. In order to make sense of this question, we need a way to metrise weak convergence of probability measures on $D = D([0,1], \mathbb{R}^d)$. Let $\mathcal{M}^1(D)$ denote the set of Borel probability measures on D such that $x \mapsto |x|_\infty$ is integrable. The *Kantorovich metric* on $\mathcal{M}^1(D)$ is defined by

$$\kappa(\mu, \nu) = \sup \left\{ \int_D f \, d\mu - \int_D f \, d\nu \middle| f : D \to \mathbb{R} \text{ is Lipschitz and } \operatorname{Lip}(f) \le 1 \right\}.$$

It is straightforward to see that κ is indeed a metric on $\mathcal{M}^1(D)$. By the portmanteau theorem (Lemma 2.1.1), if μ, $(\mu_n)_{n\ge1} \in \mathcal{M}^1(D)$ with $\kappa(\mu_n, \mu) \to 0$, then $\mu_n \xrightarrow{w} \mu$.

For random processes $X_1, X_2 \in D$ we write $\kappa(X_1, X_2) = \kappa(\mathbb{P}_{X_1}, \mathbb{P}_{X_2})$, where $\mathbb{P}_{X_i}$ is the distribution of X_i. We are now ready to state the main result of this chapter:

Theorem 5.1.1. *Let $T : M \to M$ satisfy the Functional Correlation Bound with rate $n^{-\gamma}$, $\gamma > 1$ and suppose that $v : M \to \mathbb{R}^d$ is η-Hölder and mean zero. Assume that Σ is nondegenerate. Then there is a constant $C > 0$ such that $\kappa(W_n, W) \le Cn^{-r(\gamma)}$ for all $n \ge 1$, where*

$$r(\gamma) = \begin{cases} \frac{1}{8} - \frac{1}{8\gamma}, & 1 < \gamma < 4 \\ \frac{1}{4} - \frac{5}{8\gamma}, & \gamma \ge 4 \end{cases}.$$

The exact formula for $r : (1, \infty) \to (0, \frac{1}{4})$ could be improved by more careful arguments. However, since our proof is based on Bernstein's 'big block-small block' technique, it does not seem possible to obtain a rate better than $O(n^{-1/4})$.

Remark 5.1.2. (i) The matrix Σ is typically nondegenerate. Indeed, let $\gamma > 2$ and suppose that Σ is degenerate. Then there exists a nonzero vector $c \in \mathbb{R}^d$ such that $c^T \Sigma c = 0$. Let $w = c^T v$. Then $\sum_n n |\mathbb{E}_\mu[w w \circ T^n]| \ll \sum_n n^{1-\gamma} < \infty$. By [CLB01, Lemme 2.2], it follows that w is an L^2 coboundary, i.e. there exists $u \in L^2$ such that $w = u - u \circ T$.
(ii) Rather than considering W_n, one often considers the continuous process X_n defined by setting $X_n(j/n) = W_n(j/n)$ for $j = 0, \dots, n$ and linearly interpolating in between. That is, $X_n(t) = W_n(t) + n^{-1/2}(nt - [nt]) v \circ T^{[nt]}$ for $t \in [0,1]$. Since $\left| \sup_{t\in[0,1]} |X_n(t) - W_n(t)| \right|_\infty \le n^{-1/2} |v|_\infty$, we have $\kappa(W_n, X_n) \le n^{-1/2} |v|_\infty$. By Theorem 5.1.1, it follows that $\kappa(X_n, W) = O(n^{-r(\gamma)})$.

5.1.1 Comparison with existing results

Most results in the literature on rates of convergence in the weak invariance principle consider bounds in the Prokhorov metric or the Wasserstein-p metric. Before discussing existing results, let us recall the definition of these metrics and their relation to the Kantorovich metric. A more complete account of metrics on spaces of measures can be found in Chapter 3 of [Bog18].

The *Prokhorov metric* on the space of Borel probability measures is defined by

$$\pi(\mu,\nu) = \inf\{\varepsilon > 0 : \nu(B) \le \mu(B^\varepsilon)+\varepsilon, \mu(B) \le \nu(B^\varepsilon)+\varepsilon \text{ for all Borel sets } B \subset D\}.$$

Here B^ε denotes the open ε-neighbourhood of the set B.

Let μ and ν be Radon probability measures on D. A *coupling* between μ and ν is a Radon measure on D^2 for which the projections on the first and second factors coincide with μ and ν, respectively. Let $\Pi(\mu, \nu)$ denote the set of couplings between μ and ν.

For $p \ge 1$, let $\mathcal{P}_r^p(D)$ denote the set of Radon probability measures on D such that $x \mapsto |x|_\infty^p$ is integrable. The *Wasserstein-p metric* on $\mathcal{P}_r^p(D)$ is defined by

$$\mathcal{W}_p(\mu,\nu) = \inf_{\sigma\in\Pi(\mu,\nu)} \left(\int_{D^2} |x-y|_\infty^p \, d\sigma(x,y)\right)^{1/p}.$$

Note that $\mathcal{W}_p$ is indeed a metric on $\mathcal{P}_r^p(D)$, see [Bog18, p.118] for more details.

The metrics that we have just introduced are related as follows:

Lemma 5.1.3. *Fix $p \ge 1$. Let μ, $\nu \in \mathcal{P}_r^p(D)$. Then*

(i) $\kappa(\mu,\nu) = \mathcal{W}_1(\mu,\nu)$.

(ii) $\pi(\mu,\nu) \le \mathcal{W}_p(\mu,\nu)^{\frac{p}{p+1}}$.

Proof. Part (i) is Theorem 3.2.7 in [Bog18]. The proof of (ii) is similar to the proof of Proposition 2.6 in [LW22], however, we give the details because our setting is somewhat different. Let $\varepsilon > 0$ and let $B \subset D$ be Borel. Let $\sigma \in \Pi(\mu,\nu)$. Since we have endowed D with the sup-norm metric, we have that

$$\begin{aligned}
\nu(B) &= \sigma((x,y) \in D^2 : y \in B) \\
&\le \sigma((x,y) \in D^2 : x \in B^\varepsilon) + \sigma((x,y) \in D^2 : |x-y|_\infty^p \ge \varepsilon) \\
&\le \mu(B^\varepsilon) + \frac{1}{\varepsilon^p}\int_{D^2} |x-y|_\infty^p \, d\sigma(x,y).
\end{aligned}$$

Taking the infimum over $\sigma \in \Pi(\mu,\nu)$ yields that $\nu(B) \le \mu(B^\varepsilon) + \varepsilon^{-p}\mathcal{W}_p(\mu,\nu)^p$. By interchanging μ and ν, we also have $\mu(B) \le \nu(B^\varepsilon) + \varepsilon^{-p}\mathcal{W}_p(\mu,\nu)^p$.

Let $\varepsilon = \mathcal{W}_p(\mu,\nu)^{\frac{p}{p+1}}$. Then $\varepsilon = \varepsilon^{-p}\mathcal{W}_p(\mu,\nu)^p$ so $\nu(B) \le \mu(B^\varepsilon) + \varepsilon$ and $\mu(B) \le \nu(B^\varepsilon) + \varepsilon$. Since $B \subset D$ was an arbitrary Borel set, it follows that $\pi(\mu,\nu) \le \varepsilon$. □

Let X_n be as defined in Remark 5.1.2. Then X_n and W both have continuous sample paths. Since $C([0,1],\mathbb{R}^d)$ is complete and separable, any probability measure on $C([0,1],\mathbb{R}^d)$ is Radon. It follows that $\mathcal{W}_1(X_n,W) = \kappa(X_n,W)$.

To the author's knowledge, Antoniou & Melbourne [AM19] were the first to study the problem of obtaining rates in the WIP for dynamical systems. They obtained rates in the Prokhorov metric for $\mathbb{R}$-valued Hölder observables on nonuniformly expanding maps modelled by Young towers with $O(n^{-\beta})$ tails with $\beta > 2$. In particular, for nonuniformly expanding maps with superpolynomial tails the rates are of the form $O(n^{-(1/4-\delta)})$ for any $\delta > 0$. More recently, Liu & Wang [LW22] have obtained similar rates in the Wasserstein-p metric in the same setting but they require the map to have $O(n^{-\beta})$ tails with $\beta > 4$. The arguments in both papers are based on the martingale approximation method.

Recall that by Theorem 3.2.3, mixing nonuniformly expanding/hyperbolic maps modelled by Young towers with $O(n^{-\beta})$ tails satisfy the Functional Correlation Bound with rate $n^{-(\beta-1)}$. Hence our results also apply in this setting, however the rates that we obtain in the Wasserstein-1 metric are worse than those in [LW22]. By Lemma 5.1.3(ii) these results also yield rates in the Prokhorov metric, however, these rates are worse than the ones obtained in [AM19].

In his PhD thesis, Paviato [Pav23] has obtained rates in the Wasserstein-1 metric for $\mathbb{R}^d$-valued Hölder observables on nonuniformly expanding maps. By the same method, he has also obtained rates in the multidimensional WIP for nonuniformly expanding semiflows. The rates that he obtains are independent of d but are at best $O(n^{-1/6})$ for $d > 1$ because obtaining optimal rates in the multidimensional WIP is still an open problem for martingales (see the introduction of [CDM21] for more details).

Hence our main theorem improves on existing results for $d > 1$. Moreover, even for $d = 1$ it is the first result on rates that applies to slowly-mixing invertible maps such as the examples given in Subsection 3.1.1.

In our main theorem we only prove bounds in the Kantorovich metric. Nevertheless, in some of our arguments we work with the Wasserstein-p metric because this is equally convenient. In future work, the author plans to consider bounds in the Wasserstein-p metric and the Prokhorov metric. We also plan to remove the assumption that Σ is nondegenerate.

5.2 Preliminaries

In this section we state and prove some probabilistic results that are useful in the proof of our main theorem.

Lemma 5.2.1. *Fix $p \geq 2$. There exists a constant $C > 0$ such that for all $k \geq 1$, for all independent, mean zero random variables $\widehat{X}_1, \dots, \widehat{X}_k \in L^p$,*

$$\mathbb{E}\Big[\Big|\max_{1\leq j\leq k}\Big|\sum_{i=1}^{j}\widehat{X}_i\Big|\Big|^p\Big] \leq C\left(\left(\sum_{i=1}^{k}\mathbb{E}\Big[|\widehat{X}_i|^2\Big]\right)^{p/2} + \sum_{i=1}^{k}\mathbb{E}\Big[|\widehat{X}_i|^p\Big]\right).$$

Proof. Define $M(t) = \sum_{i=1}^{[kt]}\widehat{X}_i$. Then M is a martingale (see Appendix A). Hence by Doob's maximal inequality,

$$\Big|\max_{1\leq j\leq k}\Big|\sum_{i=1}^{j}\widehat{X}_i\Big|\Big|_p = \Big|\sup_{0\leq t\leq 1}|M(t)|\Big|_p \leq \frac{p}{p-1}\Big|\sum_{i=1}^{k}\widehat{X}_i\Big|_p.$$

The result follows by Rosenthal's inequality (Proposition 3.5.1). □

Lemma 5.2.2 ([Ser70, Corollary B.2])**.** *Let $p > 2$. Let $(X_n)_{n\geq 1}$ be a sequence of mean zero random variables defined on the same probability space. Suppose that there exists a constant $M > 0$ such that*

$$\Big|\sum_{i=a+1}^{a+n}X_i\Big|_p \leq Mn^{1/2} \text{ for all } a \geq 1, n \geq 1.$$

Then there exists a constant $C > 0$ such that

$$\Big|\max_{1\leq j\leq n}\Big|\sum_{i=a+1}^{a+j}X_i\Big|\Big|_p \leq Cn^{1/2} \text{ for all } a \geq 1, n \geq 1.$$

Corollary 5.2.3. *Let $v \in \mathcal{C}^\eta(M, \mathbb{R}^d)$ be mean zero. Then there exists $C > 0$ such that*

$$\Big|\max_{1\leq j\leq n}\Big|\sum_{i=0}^{j-1}v\circ T^i\Big|\Big|_{2\gamma} \leq Cn^{1/2} \text{ for all } n \geq 1.$$

Proof. Since μ is T-invariant,

$$\Big|\sum_{i=a}^{a+n-1}v\circ T^i\Big|_{2\gamma} = \Big|\sum_{i=0}^{n-1}v\circ T^i\Big|_{2\gamma}$$

for all $a \geq 0$. Now by Theorem 3.2.4, $\big|\sum_{i=0}^{n-1} v \circ T^i\big|_{2\gamma} = O(n^{1/2})$. The result follows by Lemma 5.2.2. $\square$

It is well-known that Brownian motion is a.s. α-Hölder continuous for all $\alpha < 1/2$. Let us recall the following standard bound on the α-Hölder norm of Brownian motion (cf. [RY99, p. 28]).

Lemma 5.2.4. *Let B be a standard d-dimensional Brownian motion. Let $\alpha \in (0, 1/2)$. Then*

$$\Big| \sup_{0 \leq s < t \leq 1} \frac{|B(t) - B(s)|}{|t-s|^\alpha} \Big|_p < \infty$$

for all $1 \leq p < \infty$.

Lemma 5.2.5 ([GZ09, Theorem 3]). *Let $p \in (2, \infty)$. Then there exists a constant $C > 0$ such that the following holds. Let $(\widehat{X}_i)_{i=1}^k$ be iid mean zero random vectors in L^p. Then there exists a probability space supporting random vectors $(\tilde{X}_i)_{i=1}^k$ with the same joint distribution as $(\widehat{X}_i)_{i=1}^k$ and iid random vectors $(Z_i)_{i=1}^k$ with distribution $\mathcal{N}(0, \mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1])$ such that*

$$\Big| \max_{1 \leq j \leq k} \Big| \sum_{i=1}^j \tilde{X}_i - \sum_{i=1}^j Z_i \Big| \Big|_p \leq C k^{1/p} \big|\widehat{X}_1\big|_p \frac{\sigma_{\max}(V)}{\sigma_{\min}(V)},$$

where $\sigma^2_{\min}(V)$ and $\sigma^2_{\max}(V)$ are the minimal and maximal positive eigenvalues of $V = k\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]$, respectively.

We now use this lemma to prove a weak invariance principle with rates for iid random vectors. Define $Y_k \in D([0,1], \mathbb{R}^d)$ by $Y_k(t) = \sum_{i=0}^{[kt]} \widehat{X}_i$ for $t \in [0, 1]$.

Corollary 5.2.6. *Let $p \in (2, \infty)$. Then there exists $C > 0$ such that for all $k \geq 1$, for any sequence of iid mean zero random vectors $(\widehat{X}_i)_{i=1}^k$ in L^p we have*

$$\mathcal{W}_p(Y_k, W) \leq C k^{1/p} \Big(\big|\widehat{X}_1\big|_p \frac{\sigma_{\max}(V)}{\sigma_{\min}(V)} + |\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]^{1/2}| \Big),$$

where W is a Brownian motion with covariance $V = k\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]$.

Proof. Define $A_k, \tilde{Y}_k \in D([0,1], \mathbb{R}^d)$ by $A_k(t) = \sum_{i=1}^{[kt]} Z_i$ and $\tilde{Y}_k(t) = \sum_{i=1}^{[kt]} \tilde{X}_i$. Then by Lemma 5.2.5,

$$\Big| \sup_{0 \leq t \leq 1} |\tilde{Y}_k(t) - A_k(t)| \Big|_p = \Big| \max_{1 \leq j \leq k} \Big| \sum_{i=1}^j \tilde{X}_i - \sum_{i=1}^j Z_i \Big| \Big|_p \leq C k^{1/p} \big|\widehat{X}_1\big|_p \frac{\sigma_{\max}(V)}{\sigma_{\min}(V)}.$$

It follows that

$$\mathcal{W}_p(Y_k, A_k) = \mathcal{W}_p(\tilde{Y}_k, A_k) \le Ck^{1/p} \big|\widehat{X}_1\big|_p \sigma_{\max}(V)/\sigma_{\min}(V). \tag{5.2.1}$$

Now let W be a Brownian motion with covariance V. Then

$$\left(W\left(\frac{j}{k}\right) - W\left(\frac{j-1}{k}\right)\right)_{1\le j\le k} =_d (Z_j)_{1\le j\le k}.$$

Define $\tilde{A}_k \in D([0,1],\mathbb{R}^d)$ by

$$\tilde{A}_k(t) = \sum_{i=1}^{[kt]} \left(W\left(\frac{j}{k}\right) - W\left(\frac{j-1}{k}\right)\right) = W\left(\frac{[kt]}{k}\right).$$

It follows that $\tilde{A}_k =_d A_k$. Write $W = V^{1/2}B$, where B is a standard Brownian motion. Then

$$\begin{aligned} |\tilde{A}(t) - W(t)| &= \left|V^{1/2}\left(B(t) - B\left(\frac{[kt]}{k}\right)\right)\right| \\ &\le k^{1/2} |\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]^{1/2}| \left|B(t) - B\left(\frac{[kt]}{k}\right)\right|. \end{aligned}$$

Since $|t - \frac{1}{k}[kt]| \le \frac{1}{k}$, it follows from Lemma 5.2.4 that

$$\left|\sup_{0\le t\le 1}\left|B(t) - B\left(\frac{[kt]}{k}\right)\right|\right|_p \le k^{-(1/2-1/p)} \left|\sup_{0\le s<t\le 1} \frac{|B(t)-B(s)|}{|t-s|^{1/2-1/p}}\right|_p \le Ck^{-(1/2-1/p)}.$$

Hence

$$\mathcal{W}_p(A_k, W) = \mathcal{W}_p(\tilde{A}_k, W) \le Ck^{1/p} |\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]^{1/2}|.$$

The result follows by combining this bound with (5.2.1). □

5.3 Proof of Theorem 5.1.1

We decompose $\{0,\ldots,n-1\}$ into big blocks and small blocks in the same manner as in the proof of the iterated WIP (Theorem 4.2.2). Let $0 < b < a < 1$. Recall that we split $\{0,\ldots,n-1\}$ into alternating big blocks of length $p = [n^a]$ and small blocks of length $q = [n^b]$. Again we let k denote the number of big blocks, which is equal to the number of small blocks. For $n \ge 1$, we denote $S_n = \sum_{r=0}^{n-1} v \circ T^r$.

Let $1 \le i \le k$. Recall that

$$X_i = n^{-1/2} S_p \circ T^{(i-1)(p+q)} \quad \text{and} \quad \widetilde{W}_n(t) = \sum_{i=1}^{[kt]} X_i$$

for $0 \le t \le 1$. Hence X_i is the sum of $n^{-1/2} v \circ T^r$ over the ith big block.

We also define

$$Y_i = n^{-1/2} S_q \circ T^{p+(i-1)(p+q)} \quad \text{and} \quad R_n(t) = \sum_{i=1}^{[kt]} Y_i.$$

Thus Y_i is the sum of $n^{-1/2} v \circ T^r$ over the ith small block.

Lemma 5.3.1. *There exists a constant $C > 0$ such that $\mathcal{W}_{2\gamma}(W_n, \widetilde{W}_n + R_n) \le Cn^{(1-a)\frac{1-\gamma}{2\gamma}}$ for all $n \ge 1$.*

Proof. Let $t \in [0,1]$. Since $\{0, \ldots, [kt](p+q)-1\}$ is the union of first $[kt]$ big blocks and the first $[kt]$ small blocks,

$$n^{-1/2} S_{[kt](p+q)} = \sum_{i=1}^{[kt]} (X_i + Y_i) = \widehat{W}_n(t) + R_n(t).$$

Now since $n \ge k(p+q)$, we have $[nt] \ge [kt(p+q)] \ge [kt](p+q)$. Let $h(t) = [nt] - [kt](p+q)$. Then

$$\begin{aligned} W_n(t) = n^{-1/2} S_{[nt]} &= n^{-1/2} S_{[kt](p+q)} + n^{-1/2} S_{h(t)} \circ T^{[kt](p+q)} \\ &= \widetilde{W}_n(t) + R_n(t) + n^{-1/2} S_{h(t)} \circ T^{[kt](p+q)}. \end{aligned}$$

Recall that $n - k(p+q) \le p+q$. Hence

$$h(t) \le nt - (kt-1)(p+q) = t(n - k(p+q)) + p + q \le 2(p+q).$$

It follows that

$$\left| W_n(t) - \big(\widetilde{W}_n(t) + R_n(t)\big) \right| = n^{-1/2} |S_{h(t)}| \circ T^{[kt](p+q)} \le A \circ T^{[kt](p+q)},$$

where $A = n^{-1/2}\max_{j\le 2(p+q)}|S_j|$. Hence

$$\left|\sup_{0\le t\le 1}\left|W_n(t) - \big(\widetilde{W}_n(t) + R_n(t)\big)\right|\right|_{2\gamma}^{2\gamma} \le \mathbb{E}_\mu\left[\left|\sup_{0\le t\le 1} A\circ T^{[kt](p+q)}\right|^{2\gamma}\right]$$
$$= \mathbb{E}_\mu\left[\left|\max_{0\le i\le k} A\circ T^{i(p+q)}\right|^{2\gamma}\right].$$

Now by Corollary 5.2.3, we have $|A|_{2\gamma} \ll n^{-1/2}(p+q)^{1/2} \ll n^{(a-1)/2}$. It follows that

$$\mathbb{E}_\mu\left[\left|\max_{0\le i\le k} A\circ T^{i(p+q)}\right|^{2\gamma}\right] \ll \mathbb{E}_\mu\left[\sum_{i=0}^{k}|A\circ T^{i(p+q)}|^{2\gamma}\right]$$
$$= (k+1)\mathbb{E}_\mu\big[|A|^{2\gamma}\big]$$
$$\ll n^{1-a}n^{\gamma(a-1)} = n^{(1-a)(1-\gamma)},$$

which completes the proof. □

Let $(\widehat{X}_i)$ be independent copies of (X_i) and define $\widehat{W}_n(t) = \sum_{i=1}^{[kt]}\widehat{X}_i$ for $t \in [0,1]$.

Lemma 5.3.2. *There exists a constant* $C > 0$ *such that* $\kappa(\widetilde{W}_n + R_n, \widehat{W}_n) \le C(n^{3/2-a-b\gamma} + n^{(b-a)/2})$ *for all* $n \ge 1$.

Proof. By Lemma 4.2.8(i), $\kappa(\widetilde{W}_n, \widehat{W}_n) \ll n^{3/2-a-b\gamma}$.

Let $(\widehat{Y}_i)$ be independent copies of (Y_i) and define $\widehat{R}_n(t) = \sum_{i=1}^{[kt]}\widehat{Y}_i$ for $t \in [0,1]$. By making some minor modifications to the proof of Lemma 4.2.8(i), it follows that $\kappa(R_n, \widehat{R}_n) \ll n^{3/2-b-a\gamma} \le n^{3/2-a-b\gamma}$.

Next we bound $\left|\sup_{0\le t\le 1}|\widehat{R}_n(t)|\right|_1$. By working componentwise, without loss of generality we can assume that $(\widehat{Y}_i)$ are $\mathbb{R}$-valued. Now, $\widehat{Y}_i =_d Y_i =_d n^{-1/2}S_q$ so by Theorem 3.2.4, $|\widehat{Y}_i|_2 \ll n^{-1/2}q^{1/2} \ll n^{(b-1)/2}$. Hence by Lemma 5.2.1,

$$\left|\sup_{0\le t\le 1}|\widehat{R}_n(t)|\right|_2 = \left|\max_{1\le j\le k}\left|\sum_{i=1}^{j}\widehat{Y}_i\right|\right|_2 \ll \left(\sum_{i=1}^{k}\mathbb{E}\Big[|\widehat{Y}_i|^2\Big]\right)^{1/2}$$
$$\ll k^{1/2}n^{(b-1)/2} \ll n^{(b-a)/2}.$$

Since $|\cdot|_\infty$ is 1-Lipschitz it follows that

$$\mathbb{E}_\mu\left[\sup_{0\le t\le 1}|R_n(t)|\right] \le \kappa(R_n, \widehat{R}_n) + \mathbb{E}\left[\sup_{0\le t\le 1}|\widehat{R}_n(t)|\right] \ll n^{3/2-a-b\gamma} + n^{(b-a)/2}.$$

Hence $\kappa(\widetilde{W}_n + R_n, \widetilde{W}_n) \le \left|\sup_{0\le t\le 1}|R_n(t)|\right|_1 \ll n^{3/2-a-b\gamma} + n^{(b-a)/2}$. □

Let $V_n = k\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]$, so $V_n = \mathbb{E}[\widehat{W}_n(1) \otimes \widehat{W}_n(1)]$ is the covariance of $\widehat{W}_n(1)$. Before proceeding further, we need a bound on the difference between V_n and the limiting covariance Σ.

Lemma 5.3.3. *There exists a constant $C > 0$ such that for all $n \geq 1$,*

$$|V_n - \Sigma| \leq C(n^{a(1-\gamma)} + n^{b-a} + n^{a-1}).$$

Proof. Since $\widehat{X}_1 =_d X_1$, we have $V_n = k\mathbb{E}_\mu[X_1 \otimes X_1] = kn^{-1}\mathbb{E}_\mu[S_p \otimes S_p]$. Recall that by (4.2.2),

$$|p^{-1}\mathbb{E}[S_p \otimes S_p] - \Sigma| \ll p^{-1} + p^{1-\gamma}.$$

Since $kp \leq n$ and $p = [n^a] \geq \frac{1}{2}n^a$, it follows that

$$\left|V_n - \tfrac{kp}{n}\Sigma\right| = \tfrac{kp}{n}\left|p^{-1}\mathbb{E}[S_p \otimes S_p] - \Sigma\right| \ll n^{-a} + n^{a(1-\gamma)}.$$

Now $n - k(p+q) \leq p + q$ so

$$\left|\left(\tfrac{kp}{n} - 1\right)\Sigma\right| \leq \tfrac{1}{n}(kq + p + q)|\Sigma| \ll \tfrac{1}{n}(n^{1-a}n^b + n^a) = n^{b-a} + n^{a-1}.$$

Hence

$$|V_n - \Sigma| \ll n^{-a} + n^{a(1-\gamma)} + n^{b-a} + n^{a-1} \ll n^{a(1-\gamma)} + n^{b-a} + n^{a-1},$$

as required. □

Lemma 5.3.4. *There exists a constant $C > 0$ such that for all $n \geq 1$,*

$$\mathcal{W}_{2\gamma}(\widehat{W}_n, W) \leq C(n^{(1-a)\frac{1-\gamma}{2\gamma}} + n^{a(1-\gamma)} + n^{b-a}).$$

Proof. By Corollary 5.2.6,

$$\mathcal{W}_{2\gamma}(\widehat{W}_n, W^{(V_n)}) \leq Ck^{1/(2\gamma)}\Big(\big|\widehat{X}_1\big|_{2\gamma}\frac{\sigma_{\max}(V_n)}{\sigma_{\min}(V_n)} + |\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]^{1/2}|\Big),$$

where $W^{(V_n)}$ is a Brownian motion with covariance V_n. Now $\widehat{X}_1 =_d X_1 = n^{-1/2}S_p$ so by Theorem 3.2.4, we have $|\widehat{X}_1|_{2\gamma} \ll n^{-1/2}p^{1/2}$. By Lemma 5.3.3, $V_n \to \Sigma$. Since Σ is positive definite, it follows that $\sigma_{\max}(V_n)/\sigma_{\min}(V_n) \to \sigma_{\max}(\Sigma)/\sigma_{\min}(\Sigma)$. Moreover, $V_n^{1/2} \to \Sigma^{1/2}$ so

$$|\mathbb{E}[\widehat{X}_1 \otimes \widehat{X}_1]^{1/2}| = k^{-1/2}|V_n^{1/2}| \ll k^{-1/2}.$$

Since $k = \left[\frac{n}{p+q}\right] \sim n/p \sim n^{1-a}$, it follows that

$$\begin{aligned}\mathcal{W}_{2\gamma}(\widehat{W}_n, W^{(V_n)}) &\ll k^{1/(2\gamma)}(n^{-1/2}p^{1/2} + k^{-1/2}) \\ &\ll (n/p)^{\frac{1-\gamma}{2\gamma}} \ll n^{(1-a)\frac{1-\gamma}{2\gamma}}. \end{aligned} \tag{5.3.1}$$

Write $W^{(V_n)} = V_n^{1/2}B$, where B is a standard Brownian motion. Let $W = \Sigma^{1/2}B$, so W is a Brownian motion with covariance Σ. Note that

$$|W^{(V_n)}(t) - W(t)| = |(V_n^{1/2} - \Sigma^{1/2})B(t)| \leq |V_n^{1/2} - \Sigma^{1/2}||B(t)|.$$

Since Σ is positive definite, by [HJ85, equation (7.2.13)],

$$\|V_n^{1/2} - \Sigma^{1/2}\|_2 \leq \|\Sigma^{-1/2}\|_2 \|V_n - \Sigma\|_2,$$

where $\|\cdot\|_2$ denotes the spectral norm. Hence by Lemma 5.3.3,

$$\left|\sup_{0\leq t\leq 1} |W^{(V_n)}(t) - W(t)|\right|_{2\gamma} \ll |V_n - \Sigma| \left|\sup_{0\leq t\leq 1}|B(t)|\right|_{2\gamma} \ll n^{a(1-\gamma)} + n^{b-a} + n^{a-1}.$$

By combining this bound with (5.3.1), it follows that

$$\begin{aligned}\mathcal{W}_{2\gamma}(\widehat{W}_n, W) &\leq \mathcal{W}_{2\gamma}(\widehat{W}_n, W^{(V_n)}) + \mathcal{W}_{2\gamma}(W^{(V_n)}, W) \\ &\ll n^{(1-a)\frac{1-\gamma}{2\gamma}} + n^{a(1-\gamma)} + n^{b-a} + n^{a-1} \\ &\ll n^{(1-a)\frac{1-\gamma}{2\gamma}} + n^{a(1-\gamma)} + n^{b-a},\end{aligned}$$

as required. □

We now have all the estimates required to prove our main theorem.

Proof of Theorem 5.1.1. By Lemmas 5.3.1 and 5.3.2,

$$\begin{aligned}\kappa(W_n, \widehat{W}_n) &\leq \kappa(W_n, \widetilde{W}_n + R_n) + \kappa(\widetilde{W}_n + R_n, \widehat{W}_n) \\ &\ll n^{(1-a)\frac{1-\gamma}{2\gamma}} + n^{3/2-a-b\gamma} + n^{(b-a)/2}.\end{aligned}$$

By Lemma 5.3.4,

$$\kappa(\widehat{W}_n, W) \leq \mathcal{W}_{2\gamma}(\widehat{W}_n, W) \ll n^{(1-a)\frac{1-\gamma}{2\gamma}} + n^{a(1-\gamma)} + n^{b-a}.$$

Hence $\kappa(W_n, W) \ll n^{-r(\gamma)}$ where

$$\begin{aligned} r(\gamma) &= \min\left\{(1-a)\frac{\gamma-1}{2\gamma},\, a(1-\gamma),\, a+b\gamma-\frac{3}{2},\, a(\gamma-1), a-b,\, \frac{a-b}{2}\right\} \\ &= \min\left\{(1-a)\frac{\gamma-1}{2\gamma},\, a+b\gamma-\frac{3}{2},\, a(\gamma-1),\, \frac{a-b}{2}\right\}. \end{aligned} \tag{5.3.2}$$

Recall that $0 < b < a < 1$ are arbitrary. Finally we show how to choose a and b in order to obtain the closed form expression for $r(\gamma)$ in the statement of this theorem.[1]

Let $1 < \gamma \leq 4$. Choose $a = \frac{3}{4}$ and $b = \frac{9}{8\gamma+4} < a$. Then

$$a + b\gamma - \frac{3}{2} = \frac{a-b}{2} = \frac{3(\gamma-1)}{4(2\gamma+1)}.$$

Hence

$$r(\gamma) = \min\left\{\frac{\gamma-1}{8\gamma}, \frac{3(\gamma-1)}{4(2\gamma+1)}, \frac{3(\gamma-1)}{2}\right\} = \frac{\gamma-1}{8\gamma}.$$

Let $\gamma > 4$ and choose $a = \frac{1}{2}$, $b = \frac{5}{4\gamma} < a$. Then $a + b\gamma - \frac{3}{2} = \frac{1}{4}$ so

$$r(\gamma) = \min\left\{\frac{\gamma-1}{4\gamma}, \frac{1}{4}, \frac{1}{2}(\gamma-1), \frac{1}{4} - \frac{5}{8\gamma}\right\} = \frac{1}{4} - \frac{5}{8\gamma},$$

as required. □

[1] These choices of a and b are suboptimal, however, we have been unable to find a closed form expression for the maximum of (5.3.2) over $0 < b < a < 1$.

Appendix A

A limit theorem for triangular arrays of random vectors

In this appendix, we state and prove an iterated WIP for triangular arrays of random vectors. Our assumptions are similar to those of Lyapunov's classical central limit theorem.

Proposition A.1. *Fix $d \geq 1$. Let $(\chi_{n,i})_{n\geq 1, 1\leq i\leq k_n}$ be an array of mean zero $\mathbb{R}^d$-valued random vectors such that $(\chi_{n,i})_{1\leq i\leq k_n}$ are independent for each $n \geq 1$. Suppose that*

(i) There exists a matrix $\Sigma \in \mathbb{R}^{d\times d}$ such that for all $t \in [0,1]$,

$$\lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{[tk_n]} \chi_{n,i} \otimes \sum_{i=1}^{[tk_n]} \chi_{n,i}\right] = t\Sigma.$$

(ii) There exists $p > 1$ such that $\lim_n \sum_{i=1}^{k_n} \mathbb{E}\big[|\chi_{n,i}|^{2p}\big] = 0$.

Define random processes $\widehat{W}_n \in D([0,1], \mathbb{R}^d)$, $\widehat{\mathbb{W}}_n \in D([0,1], \mathbb{R}^{d\times d})$ by

$$\widehat{W}_n(t) = \sum_{1\leq i\leq [tk_n]} \chi_{n,i} \quad \textit{and} \quad \widehat{\mathbb{W}}_n(t) = \sum_{1\leq i<j\leq [tk_n]} \chi_{n,i} \otimes \chi_{n,j}.$$

Then $(\widehat{W}_n, \widehat{\mathbb{W}}_n) \xrightarrow{w} (W, \int W \otimes dW)$ in $D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d})$ with the sup-norm topology, where W is a Brownian motion with covariance Σ and the stochastic integral $\int W \otimes dW$ is interpreted in the Itô sense.

In the following we make use of convergence results for continuous-time martingales. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $(\mathcal{A}_t)_{t\in[0,1]}$ be a filtration, i.e. a collection of sub-σ-algebras of $\mathcal{A}$ satisfying $\mathcal{A}_s \subseteq \mathcal{A}_t$ for all $0 \le s \le t \le 1$. Let $(M_t)_{t\in[0,1]}$ be a collection of $\mathbb{R}^d$-valued random vectors on $(\Omega, \mathcal{A}, \mathbb{P})$. We say that $(M_t)_{t\in[0,1]}$ is a *martingale* with respect to $(\mathcal{A}_t)_{t\in[0,1]}$ if for all $0 \le s \le t \le 1$,

(i) M_t is $\mathcal{A}_t$-measurable

(ii) $M_t \in L^1$

(iii) $\mathbb{E}[M_t|\mathcal{A}_s] = M_s$ a.s.

For $n \ge 1$, $t \in [0,1]$ let $\mathcal{F}_t^n$ be the σ-algebra generated by $\chi_{n,1}, \ldots, \chi_{n,[tk_n]}$. Then $(\mathcal{F}_t^n)_{t\in[0,1]}$ forms a filtration and $\widehat{W}_n$ is an $(\mathcal{F}_t^n)_{t\in[0,1]}$-martingale. Indeed, conditions (i) and (ii) are immediate. Moreover, for all $0 \le s \le t \le 1$, the random vectors $\chi_{n,[sk_n]+1}, \ldots, \chi_{n,[tk_n]}$ are independent of $\mathcal{F}_s^n$ so

$$\begin{aligned}\mathbb{E}[\widehat{W}_n(t)|\mathcal{F}_s^n] &= \mathbb{E}\left[\sum_{1\le i\le [sk_n]} \chi_{n,i}\middle|\mathcal{F}_s^n\right] + \mathbb{E}\left[\sum_{[sk_n]<i\le[tk_n]} \chi_{n,i}\middle|\mathcal{F}_s^n\right] \\ &= \widehat{W}_n(s) + \sum_{[sk_n]<i\le[tk_n]} \mathbb{E}[\chi_{n,i}] = \widehat{W}_n(s).\end{aligned}$$

Hence condition (iii) is satisfied.

For $X \in D[0,1]$, let $X(t-) = \lim_{s\uparrow t} X(s)$ and define

$$J(X,t) = \sup\{X(s) - X(s-) : 0 < s \le t\}.$$

Let $t \in [0,1]$. Let $0 = t_0^n < t_1^n < \cdots < t_{\ell_n}^n = t$ be a nested sequence of partitions such that the mesh size $\max_i |t_{i+1}^n - t_i^n| \to 0$. For $Y \in D[0,1]$ define the quadratic covariation of X and Y by

$$[X,Y](t) = \lim_{n\to\infty} \sum_{i=0}^{\ell_n-1} (X(t_{i+1}^n) - X(t_i^n))(Y(t_{i+1}^n) - Y(t_i^n))$$

where the limit is taken in probability, whenever this limit exists. Let $\Delta X(t) = X(t) - X(t-)$. If X and Y are piecewise constant then

$$[X,Y](t) = \sum_{0<s\le t} \Delta X(s)\Delta Y(s).$$

Define the Itô integral by

$$\int_0^t XdY = \lim_{n\to\infty} \sum_{i=0}^{\ell_n-1} X(t_i^n)(Y(t_{i+1}^n) - Y(t_i^n))$$

where the limit is taken in probability, whenever this limit exists.

If X and Y are piecewise constant then

$$\int_0^t X dY = \sum_{0<s\leq t} X(s-)\Delta Y(s). \tag{A.1}$$

In the following, we temporarily consider weak convergence with respect to the Skorokhod topology on $D([0,1],\mathbb{R}^d)$. At the end of the proof of Proposition A.1 we will explain why weak convergence also takes place in the sup-norm topology.

Lemma A.2 (Martingale WIP, [Whi07])**.** *For $n\geq 1$, let $M_n=(M_n^1,\dots,M_n^d)$ be a martingale in $D([0,1],\mathbb{R}^d)$ with respect to a filtration $(\mathcal{F}_t^n)_{t\in[0,1]}$ satisfying $M_n(0)=0$. Let $\Sigma\in\mathbb{R}^{d\times d}$ be positive semidefinite. Suppose that*

(i) $\lim_{n\to\infty}\mathbb{E}[J(M_n,1)]=0$.

(ii) For all $t\geq 0$ and $1\leq a,b\leq d$, $[M_n^a,M_n^b](t)\xrightarrow{p}\Sigma_{ab}t$.

Then $M_n\xrightarrow{w}W$ in $D([0,1],\mathbb{R}^d)$ where W is a Brownian motion with covariance Σ. □

Corollary A.3. *Under the assumptions of Proposition A.1, we have $\widehat{W}_n\xrightarrow{w}W$ in $D([0,1],\mathbb{R}^d)$ where W is a Brownian motion with covariance Σ.*

Proof. Note that $\widehat{W}_n(t)-\widehat{W}_n(t-)=\chi_{n,i}$ if $t=\frac{i}{k_n}$ for some $1\leq i\leq k_n$ and 0 otherwise. Hence $J(\widehat{W}_n,1)=\max_{1\leq i\leq k_n}|\chi_{n,i}|$. Thus

$$\mathbb{E}[J(\widehat{W}_n,1)]^2=\mathbb{E}\left[\max_{1\leq i\leq k_n}|\chi_{n,i}|\right]^2\leq\mathbb{E}\left[\max_{1\leq i\leq k_n}|\chi_{n,i}|^2\right] \tag{A.2}$$

Let $\varepsilon>0$. Note that

$$\sum_{i=1}^{k_n}\mathbb{E}\big[|\chi_{n,i}|^2\mathbb{1}\{|\chi_{n,i}|>\varepsilon\}\big]\leq\sum_{i=1}^{k_n}\frac{1}{\varepsilon^{2p-2}}\mathbb{E}\big[|\chi_{n,i}|^{2p}\mathbb{1}\{|\chi_{n,i}|>\varepsilon\}\big]\to 0 \tag{A.3}$$

by Proposition A.1(ii). Now for all $n\geq 1$, $1\leq i\leq k_n$ we have

$$|\chi_{n,i}|^2\leq\varepsilon^2+|\chi_{n,i}|^2\mathbb{1}\left\{|\chi_{n,i}|^2>\varepsilon\right\}.$$

Hence

$$\mathbb{E}\Big[\max_{1\le i\le k_n} |\chi_{n,i}|^2\Big] \le \varepsilon^2 + \mathbb{E}\Big[\max_{1\le i\le k_n} |\chi_{n,i}|^2 \mathbb{1}\left\{|\chi_{n,i}|^2 > \varepsilon\right\}\Big]$$
$$\le \varepsilon^2 + \mathbb{E}\left[\sum_{i=1}^{k_n} |\chi_{n,i}|^2 \mathbb{1}\left\{|\chi_{n,i}| > \varepsilon\right\}\right].$$

By (A.2) and (A.3), it follows that $\limsup_n \mathbb{E}[J(\widehat{W}_n, 1)] \le \varepsilon$. Since $\varepsilon > 0$ was arbitrary, it follows that condition (i) of Lemma A.2 is satisfied.

Fix $t \in [0,1]$ and $1 \le a, b \le d$. By condition (i) of Proposition A.1,

$$\lim_{n\to\infty} \mathbb{E}\Big[[\widehat{W}_n^a, \widehat{W}_n^b](t)\Big] = \lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{[tk_n]} \chi_{n,i}^a \chi_{n,i}^b\right] = t\Sigma_{ab}. \tag{A.4}$$

Let $1 < p < 2$ be as in Proposition A.1. By von Bahr-Esseen's inequality (Proposition 3.5.1(i)),

$$\mathbb{E}\left[\left|\sum_{i=1}^{[tk_n]} \Big(\chi_{n,i}^a\, \chi_{n,i}^b - \mathbb{E}[\chi_{n,i}^a\, \chi_{n,i}^b]\Big)\right|^p\right] \le C \sum_{i=1}^{[tk_n]} \mathbb{E}\Big[\Big|\chi_{n,i}^a\, \chi_{n,i}^b - \mathbb{E}[\chi_{n,i}^a\, \chi_{n,i}^b]\Big|^p\Big].$$

Now,

$$\mathbb{E}\Big[\Big|\chi_{n,i}^a\, \chi_{n,i}^b - \mathbb{E}[\chi_{n,i}^a\, \chi_{n,i}^b]\Big|^p\Big] \le 2^p \mathbb{E}\Big[|\chi_{n,i}^a\, \chi_{n,i}^b|^p\Big]$$
$$\le 2^p \mathbb{E}\big[|\chi_{n,i}|^{2p}\big].$$

By condition (i) of Proposition A.1, it follows that

$$\sum_{i=1}^{[tk_n]} \left(\chi_{n,i}^a\, \chi_{n,i}^b - \mathbb{E}[\chi_{n,i}^a\, \chi_{n,i}^b]\right) \longrightarrow 0 \text{ in } L^p.$$

Hence by (A.4),

$$[\widehat{W}_n^a, \widehat{W}_n^b](t) = \sum_{i=1}^{[tk_n]} \chi_{n,i}^a\, \chi_{n,i}^b \longrightarrow t\Sigma_{ab}$$

in L^p, so condition (ii) of Lemma A.2 is satisfied. □

We use the following result from [JMP89, KP91] to complete the proof of Proposition A.1:

Lemma A.4 ([KP91, Thm 2.2])**.** *For each $n \ge 1$, let $(X_n, Y_n) \in D([0,1], \mathbb{R}^d \times \mathbb{R}^d)$ be an $\mathcal{F}_t^n$-adapted process and let Y_n be an $\mathcal{F}_t^n$-martingale. Suppose that*

$\sup_n \mathbb{E}[[Y_n^a, Y_n^a](t)] < \infty$ *for all* $1 \le a \le d$ *and* $t \in [0,1]$ *and that* $(X_n, Y_n) \xrightarrow{w} (X, Y)$ *in* $D([0,1], \mathbb{R}^d \times \mathbb{R}^d)$. *Then*

$$\left(X_n, Y_n, \int X_n \otimes dY_n\right) \xrightarrow{w} \left(X, Y, \int X \otimes dY\right) \text{ in } D([0,1], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}).$$

Proof of Proposition A.1. Take $X_n = Y_n = \widehat{W}_n$. By (A.1), for all $1 \le a, b \le d$,

$$\int_0^t \widehat{W}_n^a \, d\widehat{W}_n^b = \sum_{j=1}^{[tk_n]} \widehat{W}^a(\tfrac{j}{k_n}-)\chi_{n,i}^b = \sum_{j=1}^{[tk_n]} \sum_{i=1}^{j-1} \chi_{n,i}^a \chi_{n,j}^b = \widehat{\mathbb{W}}_n^{ab}(t).$$

Now by Corollary A.3, $\widehat{W}_n \xrightarrow{w} W$ in $D([0,1], \mathbb{R}^d)$ so by the continuous mapping theorem (Lemma 2.1.2), $(\widehat{W}_n, \widehat{W}_n) \xrightarrow{w} (W, W)$ in $D([0,1], \mathbb{R}^d \times \mathbb{R}^d)$. As shown in the proof of Corollary A.3, for all $t \in [0,1]$ and $1 \le a \le d$ the quadratic covariation $[\widehat{W}_n^a, \widehat{W}_n^a](t)$ converges in L^1, so $\sup_n \mathbb{E}[[\widehat{W}_n^a, \widehat{W}_n^a](t)] < \infty$. Hence Lemma A.4 applies and it follows that $(\widehat{W}_n, \widehat{\mathbb{W}}_n) \xrightarrow{w} (W, \int W \otimes dW)$ in $D([0,1], \mathbb{R}^d \times \mathbb{R}^{d\times d})$ with Skorokhod topology. Since the limit $(W, \int W \otimes dW)$ has continuous sample paths, by [Bil99, Section 15] we also have weak convergence in the sup-norm topology. $\square$

Bibliography

[Alv20] J. F. Alves. *Nonuniformly hyperbolic attractors—geometric and probabilistic aspects*. Springer Monographs in Mathematics. Springer, Cham, [2020] ©2020, pp. xi+259.

[AS14] J. F. Alves and M. Soufi. “Statistical stability of geometric Lorenz attractors”. *Fund. Math.* 224.3 (2014), pp. 219–231.

[AM19] M. Antoniou and I. Melbourne. “Rate of convergence in the weak invariance principle for deterministic systems”. *Comm. Math. Phys.* 369.3 (2019), pp. 1147–1165.

[Ara$^+$09] V. Araujo, M. J. Pacifico, E. R. Pujals, and M. Viana. “Singular-hyperbolic attractors are chaotic”. *Trans. Amer. Math. Soc.* 361.5 (2009), pp. 2431–2485.

[BE65] B. von Bahr and C.-G. Esseen. “Inequalities for the rth absolute moment of a sum of random variables, $1 \le r \le 2$”. *Ann. Math. Statist.* 36 (1965), pp. 299–303.

[BT16] V. Baladi and M. Todd. “Linear response for intermittent maps”. *Comm. Math. Phys.* 347.3 (2016), pp. 857–874.

[BBNV12] P. Bálint, G. Borbély, and A. Némedy Varga. “Statistical properties of the system of two falling balls”. *Chaos* 22.2 (2012), pp. 026104, 30.

[Bas11] R. F. Bass. *Stochastic processes*. Vol. 33. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2011, pp. xvi+390.

[Bil99] P. Billingsley. *Convergence of probability measures*. Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999, pp. x+277.

[Bog18] V. I. Bogachev. *Weak convergence of measures*. Vol. 234. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2018, pp. xii+286.

[BMT21] H. Bruin, I. Melbourne, and D. Terhesiu. "Sharp polynomial bounds on decay of correlations for multidimensional nonuniformly hyperbolic systems and billiards". *Ann. H. Lebesgue* 4 (2021), pp. 407–451.

[Bun73] L. A. Bunimovič. "The ergodic properties of billiards that are nearly scattering". *Dokl. Akad. Nauk SSSR* 211 (1973), pp. 1024–1026.

[CG12] J.-R. Chazottes and S. Gouëzel. "Optimal concentration inequalities for dynamical systems". *Comm. Math. Phys.* 316.3 (2012), pp. 843–889.

[Che99] N. Chernov. "Decay of correlations and dispersing billiards". *J. Statist. Phys.* 94.3-4 (1999), pp. 513–556.

[Che01] N. Chernov. "Sinai billiards under small external forces". *Ann. Henri Poincaré* 2.2 (2001), pp. 197–236.

[Che08] N. Chernov. "Sinai billiards under small external forces. II". *Ann. Henri Poincaré* 9.1 (2008), pp. 91–107.

[CM06] N. Chernov and R. Markarian. *Chaotic billiards*. Vol. 127. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2006, pp. xii+316.

[CZ05a] N. Chernov and H.-K. Zhang. "A family of chaotic billiards with variable mixing rates". *Stochastics and Dynamics* 5.04 (2005), pp. 535–553.

[CZ05b] N. Chernov and H.-K. Zhang. "Billiards with polynomial mixing rates". *Nonlinearity* 18.4 (2005), pp. 1527–1553.

[Che+19] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. "Multiscale systems, homogenization, and rough paths". *Probability and analysis in interacting physical systems*. Vol. 283. Springer Proc. Math. Stat. Springer, Cham, 2019, pp. 17–48.

[Che+22] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne, and H. Zhang. "Deterministic homogenization under optimal moment assumptions for fast–slow systems. Part 2". *Ann. Inst. Henri Poincaré Probab. Stat.* 58.3 (2022), pp. 1328–1350.

[CP90] Z. Coelho and W. Parry. "Central limit asymptotics for shifts of finite type". *Israel J. Math.* 69.2 (1990), pp. 235–249.

[CLB01] J.-P. Conze and S. Le Borgne. “Méthode de martingales et flot géodésique sur une surface de courbure constante négative”. *Ergodic Theory Dynam. Systems* 21.2 (2001), pp. 421–441.

[CDM21] C. Cuny, J. Dedecker, and F. Merlevède. “Rates of convergence in invariance principles for random walks on linear groups via martingale methods”. *Trans. Amer. Math. Soc.* 374.1 (2021), pp. 137–174.

[DP09] J. Dedecker and C. Prieur. “Some unbounded functions of intermittent maps for which the central limit theorem holds”. *ALEA Lat. Am. J. Probab. Math. Stat.* 5 (2009), pp. 29–45.

[DL08] M. F. Demers and C. Liverani. “Stability of statistical properties in two-dimensional piecewise hyperbolic maps”. *Trans. Amer. Math. Soc.* 360.9 (2008), pp. 4777–4814.

[DMN20] M. F. Demers, I. Melbourne, and M. Nicol. “Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas”. *Nonlinearity* 33.8 (2020), pp. 4095–4113.

[DZ13] M. F. Demers and H.-K. Zhang. “A functional analytic approach to perturbations of the Lorentz gas”. *Comm. Math. Phys.* 324.3 (2013), pp. 767–830.

[Dol04] D. Dolgopyat. “Limit theorems for partially hyperbolic systems”. *Trans. Amer. Math. Soc.* 356.4 (2004), pp. 1637–1689.

[EL21] P. Eslami and C. Liverani. “Mixing rates for symplectic almost Anosov maps”. *Nonlinearity* 34.6 (2021), pp. 3709–3731.

[FV22] N. Fleming-Vázquez. “Functional Correlation Bounds and Optimal Iterated Moment Bounds for Slowly-Mixing Nonuniformly Hyperbolic Maps”. *Comm. Math. Phys.* 391.1 (2022), pp. 173–198.

[FT09] J. M. Freitas and M. Todd. “The statistical stability of equilibrium states for interval maps”. *Nonlinearity* 22.2 (2009), pp. 259–281.

[GCF17] G. A. Gottwald, D. T. Crommelin, and C. L. E. Franzke. “Stochastic climate theory”. *Nonlinear and stochastic climate dynamics.* Cambridge Univ. Press, Cambridge, 2017, pp. 209–240.

[GM13] G. A. Gottwald and I. Melbourne. “Homogenization for deterministic maps and multiplicative noise”. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469.2156 (2013), p. 20130201.

[GZ09] F. Götze and A. Y. Zaĭtsev. “The accuracy of approximation in the multidimensional invariance principle for sums of independent identically distributed random vectors with finite moments”. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 368.Veroyatnost’ i Statistika. 15 (2009), pp. 110–121, 283–284.

[Gou04] S. Gouëzel. “Central limit theorem and stable laws for intermittent maps”. *Probab. Theory Related Fields* 128.1 (2004), pp. 82–122.

[Gou05] S. Gouëzel. “Berry-Esseen theorem and local limit theorem for non uniformly expanding maps”. *Ann. Inst. H. Poincaré Probab. Statist.* 41.6 (2005), pp. 997–1024.

[GL06] S. Gouëzel and C. Liverani. “Banach spaces adapted to Anosov systems”. *Ergodic Theory Dynam. Systems* 26.1 (2006), pp. 189–217.

[GM14] J. Griffin and J. Marklof. “Limit theorems for skew translations”. *J. Mod. Dyn.* 8.2 (2014), pp. 177–189.

[HK82] F. Hofbauer and G. Keller. “Ergodic properties of invariant measures for piecewise monotonic transformations”. *Math. Z.* 180.1 (1982), pp. 119–140.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1985, pp. xiii+561.

[JMP89] A. Jakubowski, J. Mémin, and G. Pagès. “Convergence en loi des suites d’intégrales stochastiques sur l’espace $\mathbf{D}^1$ de Skorokhod”. *Probab. Theory Related Fields* 81.1 (1989), pp. 111–137.

[KM16] D. Kelly and I. Melbourne. “Smooth approximation of stochastic differential equations”. *Ann. Probab.* 44.1 (2016), pp. 479–520.

[Kor16] A. Korepanov. “Linear response for intermittent maps with summable and nonsummable decay of correlations”. *Nonlinearity* 29.6 (2016), pp. 1735–1754.

[KKM17] A. Korepanov, Z. Kosloff, and I. Melbourne. “Averaging and rates of averaging for uniform families of deterministic fast-slow skew product systems”. *Studia Math.* 238.1 (2017), pp. 59–89.

[KKM18] A. Korepanov, Z. Kosloff, and I. Melbourne. “Martingale-coboundary decomposition for families of dynamical systems”. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* 35.4 (2018), pp. 859–885.

[KKM19] A. Korepanov, Z. Kosloff, and I. Melbourne. "Explicit coupling argument for non-uniformly hyperbolic transformations". *Proc. Roy. Soc. Edinburgh Sect. A* 149.1 (2019), pp. 101–130.

[KKM22] A. Korepanov, Z. Kosloff, and I. Melbourne. "Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1". *Ann. Inst. Henri Poincaré Probab. Stat.* 58.3 (2022), pp. 1305–1327.

[KP91] T. G. Kurtz and P. Protter. "Weak limit theorems for stochastic integrals and stochastic differential equations". *Ann. Probab.* 19.3 (1991), pp. 1035–1070.

[Lep17] J. Leppänen. "Functional correlation decay and multivariate normal approximation for non-uniformly expanding maps". *Nonlinearity* 30.11 (2017), pp. 4239–4259.

[LS17] J. Leppänen and M. Stenlund. "A note on the finite-dimensional distributions of dispersing billiard processes". *J. Stat. Phys.* 168.1 (2017), pp. 128–145.

[LS20] J. Leppänen and M. Stenlund. "Sunklodas' approach to normal approximation for time-dependent dynamical systems". *J. Stat. Phys.* 181.5 (2020), pp. 1523–1564.

[LW22] Z. Liu and Z. Wang. *Wasserstein convergence rate in the invariance principle for deterministic dynamical systems*. 2022. arXiv: `arXiv:2204.00263`.

[LSV99] C. Liverani, B. Saussol, and S. Vaienti. "A probabilistic approach to intermittency". *Ergodic Theory Dynam. Systems* 19.3 (1999), pp. 671–685.

[Mar04] R. Markarian. "Billiards with polynomial decay of correlations". *Ergodic Theory Dynam. Systems* 24.1 (2004), pp. 177–197.

[MS11] I. Melbourne and A. M. Stuart. "A note on diffusion limits of chaotic skew-product flows". *Nonlinearity* 24.4 (2011), pp. 1361–1367.

[MT14] I. Melbourne and D. Terhesiu. "Decay of correlations for non-uniformly expanding systems with general return times". *Ergodic Theory Dynam. Systems* 34.3 (2014), pp. 893–918.

[MV16] I. Melbourne and P. Varandas. "A note on statistical properties for nonuniformly hyperbolic systems with slow contraction and expansion". *Stoch. Dyn.* 16.3 (2016), pp. 1660012, 13.

[Pav23] N. Paviato. “Rates for maps and flows in a deterministic multidimensional weak invariance principle”. PhD thesis. University of Warwick, 2023.

[PS08] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*. Vol. 53. Texts in Applied Mathematics. Averaging and homogenization. Springer, New York, 2008, pp. xviii+307.

[Rat73] M. Ratner. “The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature”. *Israel J. Math.* 16 (1973), pp. 181–197.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. Third. Vol. 293. Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1999, pp. xiv+602.

[Ros70] H. P. Rosenthal. “On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables”. *Israel Journal of Mathematics* 8.3 (1970), pp. 273–303.

[Ser70] R. J. Serfling. “Moment inequalities for the maximum cumulative sum”. *Ann. Math. Statist.* 41 (1970), pp. 1227–1234.

[Sin60] J. G. Sinaĭ. “The central limit theorem for geodesic flows on manifolds of constant negative curvature”. *Soviet Math. Dokl.* 1 (1960), pp. 983–987.

[Whi07] W. Whitt. “Proofs of the martingale FCLT”. *Probab. Surv.* 4 (2007), pp. 268–302.

[You98] L.-S. Young. “Statistical properties of dynamical systems with some hyperbolicity”. *Ann. of Math. (2)* 147.3 (1998), pp. 585–650.

[You99] L.-S. Young. “Recurrence times and rates of mixing”. *Israel Journal of Mathematics* 110.1 (1999), pp. 153–188.