Check for
updates

Functional Correlation Bounds and Optimal Iterated Moment Bounds for Slowly-Mixing Nonuniformly Hyperbolic Maps

Nicholas Fleming-Vázquez

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.
E-mail: n.fleming.1@warwick.ac.uk

Abstract: Consider a nonuniformly hyperbolic map $T : M \to M$ modelled by a Young tower with tails of the form $O(n^{-\beta})$, $\beta > 2$. We prove optimal moment bounds for Birkhoff sums $\sum_{i=0}^{n-1} v \circ T^i$ and iterated sums $\sum_{0 \le i < j < n} v \circ T^i \, w \circ T^j$, where $v, w : M \to \mathbb{R}$ are (dynamically) Hölder observables. Previously iterated moment bounds were only known for $\beta > 5$. Our method of proof is as follows; (i) prove that T satisfies an abstract functional correlation bound, (ii) use a weak dependence argument to show that the functional correlation bound implies moment estimates. Such iterated moment bounds arise when using rough path theory to prove deterministic homogenisation results. Indeed, by a recent result of Chevyrev, Friz, Korepanov, Melbourne & Zhang we have convergence to an Itô diffusion for fast-slow systems of the form

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

in the optimal range $\beta > 2$.

1. Introduction

Let $T : M \to M$ be an ergodic, measure-preserving transformation defined on a bounded metric space (M, d) with Borel probability measure μ. Consider a fast-slow system on $\mathbb{R}^d \times M$ of the form

$$x_{k+1}^{(n)} = x_k^{(n)} + n^{-1} a(x_k^{(n)}, y_k) + n^{-1/2} b(x_k^{(n)}, y_k), \quad y_{k+1} = T y_k \qquad (1.1)$$

where the initial condition $x_0^{(n)} \equiv \xi$ is fixed and y_0 is picked randomly from (M, μ). When the fast dynamics $T : M \to M$ is chaotic enough, it is expected that the stochastic process X_n defined by $X_n(t) = x_{[nt]}^{(n)}$ will weakly converge to the solution of a stochastic differential equation driven by Brownian motion. This is referred to as *deterministic homogenisation* and has been of great interest recently [8–10,12,13,18,21,27]. See [7] for a survey of the topic.

In [18], Kelly and Melbourne considered the special case where $a(x, y) \equiv a(x)$ and $b(x, y) = h(x)v(y)$. By using rough path theory, they showed that deterministic homogenisation reduces to proving two statistical properties for $T : M \to M$. In [8] this result was extended to general a, b satisfying mild regularity assumptions.

One of the assumed statistical properties is an "iterated weak invariance principle". In [18,29] it was shown that this property is satisfied by nonuniformly expanding/hyperbolic maps modelled by Young towers, provided that the tails of the return time decay at rate $O(n^{-\beta})$ for some $\beta > 2$ (which is the optimal range for such results).

The second assumed statistical property is control of "iterated moments", which gives tightness in the rough path topology used for proving convergence. This condition has proved much more problematic. Advances in rough path theory [7,8] significantly weakened the moment requirements from [18] and these weakened moment requirements were eventually proved for nonuniformly *expanding* maps in the optimal range (i.e. $\beta > 2$) in [21].

However, for nonuniformly hyperbolic maps modelled by Young towers previously it was only possible to show iterated moment bounds for $\beta > 5$ [11]. In this article, we extend iterated moment bounds to the optimal range $\beta > 2$.

1.1. Illustrative examples. Many examples of invertible dynamical systems are modelled by Young towers [33,34]. For example, Axiom A (uniformly hyperbolic) diffeomorphisms, Henon attractors and the finite-horizon Sinai billiard are modelled by Young towers with exponential tails, so for such systems deterministic homogenisation results follow from [18,19]. 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 $g : [0, 1/2] \to [0, 1]$ by $g(x) = x(1 + 2^{\alpha}x^{\alpha})$. The Liverani-Saussol-Vaienti map $\bar{T} : [0, 1] \to [0, 1]$,

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

is a prototypical example of a slowly-mixing nonuniformly expanding map [25]. As in [29, Exa. 4.1], consider an intermittent Baker's map $T : M \to M$, $M = [0, 1]^2$ defined by

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

There is a unique absolutely continuous invariant probability measure μ. 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 two-sided Young tower with tails of the form $\sim n^{-\beta}$ where $\beta = 1/\alpha$.

For $\beta > 2$ the central limit theorem (CLT) holds for all Hölder observables. For $\beta \leq 2$ the CLT fails for typical Hölder observables [14], so it is natural to restrict to $\beta > 2$ when considering deterministic homogenisation. By [11] 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 $\beta > 2$.

Dispersing billiards provide many examples of slowly-mixing nonuniformly hyperbolic maps. Markarian [26], Chernov and Zhang [5] showed how to model many examples of dispersing billiards by Young towers with polynomial tails.

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

- *Bunimovich flowers* [2]. By [5] 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 [6] Chernov and Zhang introduced a class of billiards modelled by Young towers with tails of the form $O((\log n)^\beta n^{-\beta})$ to any prescribed value of $\beta \in (2, \infty)$.

Notation We endow $\mathbb{R}^k$ with the norm $|y| = \sum_{i=1}^k |y_i|$.

Let $\eta \in (0, 1]$. We say that an observable $v \colon M \to \mathbb{R}$ on a metric space (M, d) is η-Hölder, and write $v \in \mathscr{C}^\eta(M)$, 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.

The rest of this article is structured as follows. In Sect. 2 we state our main results. Our first main result, Theorem 2.3, is that mixing nonuniformly hyperbolic maps modelled by Young towers with polynomial tails satisfy a functional correlation bound. Our second main result, Theorem 2.4, is that this functional correlation bound implies control of iterated moments.

In Sect. 3 we recall background material on Young towers and prove Theorem 2.3. In Sect. 4 we prove that our functional correlation bound implies an elementary weak dependence condition. Finally in Sect. 5 we use this condition to prove Theorem 2.4.

2. Main Results

Let $T \colon M \to M$ be a nonuniformly hyperbolic map modelled by a Young tower. We state our results for the class of dynamically Hölder observables, noting that this includes Hölder observables. We delay the definitions of Young tower and dynamically Hölder until Sect. 3.1. Let $\mathscr{H}(M)$ denote the class of dynamically Hölder observables on M and let $[\cdot]_\mathscr{H}$ denote the dynamically Hölder seminorm.

Definition 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

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

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

Fix $\gamma > 0$. We consider dynamical systems which satisfy the following property:

Definition 2.2. Suppose that there exists a constant $C > 0$ such that for all integers $0 \leq p < q$, $0 \leq n_0 \leq \cdots \leq n_{q-1}$,

$$\left| \int_M G(T^{n_0}x, \ldots, T^{n_{q-1}}x)d\mu(x) \right.$$

$$\left. - \int_{M^2} G(T^{n_0}x_0, \ldots, T^{n_{p-1}}x_0, T^{n_p}x_1, \ldots, T^{n_{q-1}}x_1)d\mu(x_0)d\mu(x_1) \right|$$

$$\leq C(n_p - n_{p-1})^{-\gamma} \left(|G|_\infty + \sum_{i=0}^{q-1} [G]_{\mathscr{H},i} \right) \tag{2.1}$$

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

A similar condition was introduced by Leppänen in [22] and further studied by Leppänen and Stenlund in [23, 24]. In particular, [22] showed that functional correlation decay implies a multi-dimensional CLT with bounds on the rate of decay. We are now ready to state the main results which we prove in this paper.

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 [28] by using a method due to S. Gouëzel (privately communicated based on ideas from [3]), it was shown that there exists $C > 0$ such that

$$\left| \int_M v \, w \circ T^n d\mu - \int_M v d\mu \int_M w d\mu \right| \leq C n^{-(\beta-1)} \|v\|_{\mathscr{H}} \|w\|_{\mathscr{H}}$$

for all $n \geq 1$, $v, w \in \mathscr{H}(M)$. Our first main result is that the Functional Correlation Bound holds with the same rate:

Theorem 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 $v, w \in \mathscr{H}(M)$ mean zero define

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

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

Theorem 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 $v, w \in \mathscr{H}(M)$,*

(a) $|S_v(n)|_{2\gamma} \leq C n^{1/2} \|v\|_{\mathscr{H}}$.
(b) $\left| \mathbb{S}_{v,w}(n) \right|_\gamma \leq C n \|v\|_{\mathscr{H}} \|w\|_{\mathscr{H}}$.

Remark 2.5. As mentioned above, by [8, Theorem 2.10] 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 [29], the Iterated WIP holds for all Hölder observables. Together Theorem 2.3 and Theorem 2.4 imply that for all $\eta \in (0, 1]$ there exists $C > 0$ such that

(a) $|S_v(n)|_{2(\beta-1)} \leq C n^{1/2} \|v\|_\eta$.
(b) $\left| \mathbb{S}_{v,w}(n) \right|_{\beta-1} \leq C n \|v\|_\eta \|w\|_\eta$.

for all mean zero $v, w \in \mathscr{C}^\eta(M)$, giving the required control of iterated moments.

3. Young Towers

3.1. Prerequisites. Young towers were first introduced by Young in [33,34], as a broad framework to prove decay of correlations for nonuniformly hyperbolic maps. Our presentation follows [1]. In particular, this framework does not assume uniform contraction along stable manifolds and hence covers examples such as billiards.

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 $K > 0$, $\theta \in (0, 1)$ 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 $|\log \zeta(y) - \log \zeta(y')| \leq K\theta^{s(y,y')}$ 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

$$d(F^n y, F^n y') \leq K(\theta^n + \theta^{s(y,y')-n}) \text{ for all } y, y' \in Y, n \geq 0. \tag{3.1}$$

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

$$\bar{\Delta} = \{(\bar{y}, \ell) \in \bar{Y} \times \mathbb{Z} : 0 \leq \ell < \bar{\phi}(y)\}, \quad \bar{f}(\bar{y}, \ell) = \begin{cases} (\bar{y}, \ell + 1), & \ell < \bar{\phi}(y) - 1, \\ (\bar{F}\bar{y}, 0), & \ell = \bar{\phi}(y) - 1. \end{cases} \tag{3.2}$$

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 (3.2). 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 $f: \Delta \to \Delta$ 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 $\bar{\pi}: Y \to \bar{Y}$ to a map $\bar{\pi}: \Delta \to \bar{\Delta}$ 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:

Let $T: M \to M$ 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:

- $F = T^\phi: Y \to Y$ 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 $\bar{\Delta} = \bar{Y}^{\bar{\phi}}$.
- The map $\pi_M: \Delta \to M, \pi_M(y, \ell) = T^\ell y$ is a measure-preserving semiconjugacy.

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

From now on we fix $\beta > 1$ and suppose that $T: M \to M$ is a mixing transformation modelled by a Young tower Δ with tails of the form $\mu_Y(\phi \geq n) = O(n^{-\beta})$.

Remark 3.1. Here we have not assumed that the tower map $f: \Delta \to \Delta$ is mixing. However, as in [4, Theorem 2.1, Proposition 10.1] and [1] the a priori knowledge that μ is mixing ensures that this is irrelevant.

Let $\psi_n(x) = \#\{j = 1, \ldots, n: f^j x \in \Delta_0\}$ 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 [20, Lemma 5.5].

Lemma 3.2. *Let $\theta \in (0, 1)$. Then there exists a constant $D_1 > 0$ such that*

$$\int_\Delta \theta^{\psi_n} d\mu_\Delta \leq D_1 n^{-(\beta-1)} \text{ for } n \geq 1.$$

$\square$

The transfer operator L corresponding to $\bar{f}: \bar{\Delta} \to \bar{\Delta}$ 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. [20, Proposition 5.2]):

Proposition 3.3. *There exists a constant $K_1 > 0$ such that for all $n \geq 1$, for all points $z, z' \in \bar{\Delta}$ which belong to the same cylinder set of length n,*

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

$\square$

Let $\theta \in (0, 1)$. We say that $v \colon \bar{\Delta} \to \mathbb{R}$ is d_θ-Lipschitz if $\|v\|_\theta = |v|_\infty + \sup_{x \neq y} |v(x) - v(y)|/d_\theta(x, y) < \infty$. If $f \colon \Delta \to \Delta$ is mixing then by [34],

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

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

Lemma 3.4. *Suppose that* $f \colon \Delta \to \Delta$ *is mixing. Then there exists* $D_2 > 0$ *such that for all* d_θ*-Lipschitz* $v \colon \bar{\Delta} \to \mathbb{R}$*, for any* $n \geq 1$,

$$\left| \mathbb{1}_{\bar{\Delta}_0} L^n v - \int_{\bar{\Delta}} v \, d\bar{\mu}_\Delta \right|_\infty \leq D_2 n^{-(\beta-1)} \|v\|_\theta .$$

This is a straightforward application of operator renewal theory developed by Sarig [31] and Gouëzel [15,16]. However, we could not find a reference to this result in the literature so we provide a proof.

Proof. Define partial transfer operators T_n and B_n as in [17, Section 4]. Then

$$\mathbb{1}_{\bar{\Delta}_0} L^n v = \sum_{k+b=n} T_k B_b v.$$

Define an operator Π by $\Pi v = \int_{\bar{\Delta}_0} v \, d\bar{\mu}_\Delta$. Then as in the proof of [17, Theorem 4.6] we can write $T_k = \Pi + E_k$ where $\|E_k\| = O(k^{-(\beta-1)})$. Moreover, by [17, Theorem 4.6], $\|B_b\| = O(b^{-\beta})$ and

$$\sum_{b=0}^\infty \int_{\bar{\Delta}_0} B_b v \, d\bar{\mu}_\Delta = \int_{\bar{\Delta}} v \, d\bar{\mu}_\Delta.$$

It follows that

$$\begin{aligned}
\mathbb{1}_{\bar{\Delta}_0} L^n v &= \sum_{k+b=n} \Pi B_b v + \sum_{k+b=n} E_k B_b v \\
&= \sum_{b=0}^n \int_{\bar{\Delta}_0} B_b v \, d\bar{\mu}_\Delta + \sum_{k+b=n} E_k B_b v \\
&= \int_{\bar{\Delta}} v \, d\bar{\mu}_\Delta - \sum_{b=n+1}^\infty \int_{\bar{\Delta}_0} B_b v \, d\bar{\mu}_\Delta + \sum_{k+b=n} E_k B_b v.
\end{aligned}$$

The conclusion of the lemma follows by noting that the expressions $\sum_{b=n+1}^\infty b^{-\beta}$ and

$$\sum_{k+b=n} (k + 1)^{-(\beta-1)} (b + 1)^{-\beta}$$

are both $O(n^{-(\beta-1)})$. $\qquad\square$

Finally we recall the class of observables on M that are of interest to us:

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

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

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

It is standard (see e.g. [1, Proposition 7.3]) that Hölder observables are also dynamically Hölder for the classes of dynamical systems that we are interested in:

Proposition 3.5. *Let $\eta \in (0, 1]$ and let d_0 be a bounded metric on M. Let $\mathscr{C}^{\eta}(M)$ be the space of observables that are η-Hölder with respect to d_0. Suppose that there exists $K > 0$, $\gamma_0 \in (0, 1)$ such that $d_0(T^{\ell}y, T^{\ell}y') \leq K(d_0(y, y') + \gamma_0^{s(y, y')})$ for all $y, y' \in Y, 0 \leq \ell < \phi(y)$.*

Then $\mathscr{C}^{\eta}(M)$ is continuously embedded in $\mathscr{H}(M)$ where we may choose any $\theta \in [\gamma_0^{\eta}, 1)$ and $d = d_0^{\eta'}$ for any $\eta' \in (0, \eta]$. $\qquad\qquad\square$

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. $gcd\{\phi(y) : y \in Y\} = 1$. Hence in subsequent subsections we focus on proving the Functional Correlation Bound under this assumption:

Lemma 3.6. *Suppose that T is modelled by a mixing two-sided Young tower whose return time has tails of the form $O(n^{-\beta})$. Then T satisfies the Functional Correlation Bound with rate $n^{-(\beta-1)}$.*

Proof of Theorem 2.3. 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 $(\pi'_M)_* \mu'_{\Delta}$. Now by assumption the measure μ is mixing so by the same argument as in [1, Section 4.1] we must have $\mu = (\pi'_M)_* \mu'_{\Delta}$.

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

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

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

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

Let $[\cdot]_{\mathscr{H}'}$ denote the dynamically Hölder seminorm as defined with T', ϕ' in place of T, ϕ. Then by Lemma 3.6,

$$|\nabla G| \le C(n'_p - n'_{p-1})^{-\gamma}\left(|G'|_\infty + \sum_{i=0}^{q-1}[G']_{\mathscr{H}',i}\right)$$

$$\le Cd^\gamma(n_p - n_{p-1} - d)^{-\gamma}\left(|G|_\infty + \sum_{i=0}^{q-1}[G']_{\mathscr{H}',i}\right)$$

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

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

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

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

so $[G']_{\mathscr{H}',i} \le [G]_{\mathscr{H},i}$. $\qquad\qquad\square$

3.3. Approximation by one-sided functions. Let $0 \le p < q$ and $0 \le n_0 \le \cdots \le n_{q-1}$ be integers and consider a function $G \in \mathscr{S}\mathscr{H}_q(M)$. 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 $\pi_M: \Delta \to M$ is a measure-preserving semiconjugacy

$$\nabla G = \int_\Delta \widetilde{H}(x, f^{n_p}x)d\mu_\Delta(x) - \int_{\Delta^2} \widetilde{H}(x_0, x_1)d\mu_\Delta^2(x_0, x_1) = \nabla\widetilde{H} \qquad (3.3)$$

where $\widetilde{H}: \Delta^2 \to \mathbb{R}$ 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 $R \ge 1$. 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 [28].

Recall that $\psi_R(x) = \#\{j = 1, \ldots, R : f^j x \in \Delta_0\}$ denotes the number of returns to $\Delta_0 = \{(y, \ell) \in \Delta : \ell = 0\}$ by time R. Let $\mathscr{Q}_R$ 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 $\mathscr{Q}_R$. For $x \in \Delta$ let $\hat{x}$ denote the reference point of the element that x belongs to. Define $\widetilde{H}_R: \Delta^2 \to \mathbb{R}$ by

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

Proposition 3.7. *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, there exists a constant $K_2 > 0$ depending only on $T: M \to M$ such that,*

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

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

$$|\widetilde{H}(f^R x, f^R y) - \widetilde{H}_R(x, y)| \le K_2\left(\sum_{i=0}^{p-1}[G]_{\mathscr{H},i}\,\theta^{\psi_R(f^{n_i}x)} + \sum_{i=p}^{q-1}[G]_{\mathscr{H},i}\,\theta^{\psi_R(f^{k_i}y)}\right).$$

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

$$\left\|L^{R+n_{p-1}}\bar{H}_R(\cdot, \bar{y})\right\|_\theta \le K_2\left(|G|_\infty + \sum_{i=0}^{p-1}[G]_{\mathscr{H},i}\right).$$

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

Proof. We follow the proof of Proposition 7.9 in [1].

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

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

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

$$\begin{aligned}
I_1 &= \widetilde{G}(f^R a_0, \ldots, f^R a_{p-1}, b_p, \ldots, b_{q-1}) - \widetilde{G}(f^R \hat{a}_0, \ldots, f^R \hat{a}_{p-1}, b_p, \ldots, b_{q-1})\\
&= \sum_{i=0}^{p-1}\big(\widetilde{G}(f^R a_0, \ldots, f^R a_{i-1}, f^R a_i, f^R \hat{a}_{i+1}, f^R \hat{a}_{p-1}, b_p, \ldots, b_{q-1})\\
&\quad - \widetilde{G}(f^R a_0, \ldots, f^R a_{i-1}, f^R \hat{a}_i, f^R \hat{a}_{i+1}, f^R \hat{a}_{p-1}, b_p, \ldots, b_{q-1})\big)\\
&= \sum_{i=0}^{p-1}\big(\tilde{v}_i(f^R a_i) - \tilde{v}_i(f^R \hat{a}_i)\big) \tag{3.4}
\end{aligned}$$

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

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 $[G]_{\mathscr{H},i}$ and (3.1),

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

$$\leq 2(K+1)[G]_{\mathscr{H},i}\theta^{\psi_R(a_i)}.$$

Thus

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

By a similar argument,

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

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

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

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

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

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

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 (Proposition 3.3), $|J_1| \leq C\,|G|_\infty\, d_\theta(\bar{x},\bar{x}')$. We claim that $|\bar{H}_R(\bar{z},\bar{y}) - \bar{H}_R(\bar{z}',\bar{y})| \leq K_2 \sum_{i=0}^{p-1}[G]_{\mathscr{H},i}d_\theta(\bar{x},\bar{x}')$. It follows that $|J_2| \leq K_2 \sum_{i=0}^{p-1}[G]_{\mathscr{H},i}d_\theta(\bar{x},\bar{x}')$.

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),

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

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

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 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),

$$E_i \leq (K+1)(\theta^{\psi_R(\hat{a}_i)} + \theta^{s(\hat{a}_i,\hat{a}_i')-\psi_R(\hat{a}_i)}).$$

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

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

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 $E_i \leq 2(K + 1)\theta^{s(\bar{x}, \bar{x}')}$, completing the proof of the claim. $\qquad\square$

3.4. Proof of Lemma 3.6. We continue to assume that $\beta > 1$ and that $\mu_Y(\phi \geq n) = O(n^{-\beta})$. We also assume that $\gcd\{\phi(y) : y \in Y\} = 1$ so that $f : \Delta \to \Delta$ is mixing.

Lemma 3.8. *Let $\theta \in (0, 1)$. There exists $D_3 > 0$ such that for any $V \in L^\infty(\bar{\Delta}^2)$,*

$$\left| \int_{\bar{\Delta}} V(x, \bar{f}^n x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1) \right| \leq D_3 n^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta$$

for all $n \geq 1$.

Remark 3.9. 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

$$\left| \int_{\bar{\Delta}} v \, w \circ \bar{f}^n d\bar{\mu}_\Delta - \int_{\bar{\Delta}} v \, d\bar{\mu}_\Delta \int_{\bar{\Delta}} w \, d\bar{\mu}_\Delta \right| \leq D_3 n^{-(\beta-1)} \|v\|_\theta |w|_\infty,$$

so Lemma 3.8 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 3.10. Our proof of Lemma 3.8 is based on ideas from [3, 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 Lemma 3.8. Write $v(x) = V(x, \bar{f}^n x)$ so

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

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

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

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 $\ell \geq 0$,

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

Hence,

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

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 Lemma 3.4,

$$\begin{aligned} |A_\ell| &\leq \int_{\bar{\Delta}_\ell} D_2(n - \ell)^{-(\beta-1)} \|V(\cdot, x)\|_\theta \, d\bar{\mu}_\Delta \\ &\leq D_2(n - \ell)^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \, \bar{\mu}_\Delta(\bar{\Delta}_\ell). \end{aligned}$$

Hence,

$$\sum_{\ell \leq n/2} |A_\ell| \leq D_2(n/2)^{-(\beta-1)} \sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta,$$

completing the proof. $\qquad\square$

Proof of Lemma 3.6. 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 take $n_p - n_{p-1} \geq 2$. 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 Proposition 3.7(ii) and Lemma 3.2,

$$|I_1| = \left| \int_\Delta \widetilde{H}(f^R x, f^{R+n_p} x) d\mu_\Delta(x) - \int_\Delta \widetilde{H}_R(x, f^{n_p} x) d\mu_\Delta(x) \right|$$

$$\leq K_2 \int_\Delta \left(\sum_{i=0}^{p-1} [G]_{\mathscr{H},i} \theta^{\psi_R(f^{n_i} x)} + \sum_{i=p}^{q-1} [G]_{\mathscr{H},i} \theta^{\psi_R(f^{n_p + k_i} x)} \right) d\mu_\Delta(x)$$

$$= K_2 \sum_{i=0}^{q-1} [G]_{\mathscr{H},i} \int_\Delta \theta^{\psi_R} d\mu_\Delta \leq K_2 D_1 \sum_{i=0}^{q-1} [G]_{\mathscr{H},i} R^{-(\beta-1)}. \tag{3.5}$$

Similarly,

$$|I_2| \leq K_2 D_1 \sum_{i=0}^{q-1} [G]_{\mathscr{H},i} R^{-(\beta-1)}. \tag{3.6}$$

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

$$\int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1) = \int_{\bar{\Delta}^2} \bar{H}_R(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1) \tag{3.7}$$

and

$$V(x, \bar{f}^{n_p - n_{p-1} - R} x) = \sum_{\bar{f}^{n_{p-1}+R} z = x} g_{n_{p-1}+R}(z) \bar{H}_R(z, \bar{f}^{n_p - n_{p-1} - R} x)$$

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

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

$$\int_{\bar{\Delta}} V(x, \bar{f}^{n_p - n_{p-1} - R} x) d\bar{\mu}_\Delta(x) = \int_{\bar{\Delta}} L^{n_{p-1}+R} \hat{u} \, d\bar{\mu}_\Delta$$

$$= \int_{\bar{\Delta}} \hat{u} \, d\bar{\mu}_\Delta = \int_{\bar{\Delta}} \bar{H}_R(x, \bar{f}^{n_p} x) \, d\bar{\mu}_\Delta(x). \tag{3.8}$$

Now by Proposition 3.7(iii), $\sup_{y \in \bar{\Delta}} \|V(\cdot, y)\|_\theta \leq K_2(|G|_\infty + \sum_{i=0}^{p-1} [G]_{\mathscr{H},i})$. By Lemma 3.8, (3.7) and (3.8) it follows that

$$|\nabla \bar{H}_R| = \left| \int_{\bar{\Delta}} V(x, \bar{f}^{n_p - n_{p-1} - R} x) d\bar{\mu}_\Delta(x) - \int_{\bar{\Delta}^2} V(x_0, x_1) d\bar{\mu}_\Delta^2(x_0, x_1) \right|$$

$$\leq K_2 D_3 \left(|G|_\infty + \sum_{i=0}^{p-1} [G]_{\mathscr{H},i} \right) (n_p - n_{p-1} - R)^{-(\beta-1)}. \tag{3.9}$$

Recall that $R = [(n_p - n_{p-1})/2]$. Hence $n_p - n_{p-1} - R \geq R$. By combining (3.5), (3.6) and (3.9) it follows that

$$|\nabla \widetilde{H}| \leq K_2(2D_1 + D_3) \sum_{i=0}^{q-1} [G]_{\mathscr{H},i} ([(n_p - n_{p-1})/2])^{-(\beta-1)},$$

as required. $\qquad\qquad\square$

4. An Abstract Weak Dependence Condition

The Functional Correlation Bound can be seen as a weak dependence condition. 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 random variables X_i on (M, μ) of the form

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

where $\Phi_i \in \mathscr{S}\mathscr{H}_{u_i - \ell_i + 1}(M)$, $0 \leq i < k$.

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

Lemma 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 $F : [-R, R]^k \to \mathbb{R}$,*

$$\left| \mathbb{E}_\mu \left[F(X_0, \ldots, X_{k-1}) \right] - \mathbb{E} \left[F(\widehat{X}_0, \ldots, \widehat{X}_{k-1}) \right] \right|$$

$$\leq C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \left(|F|_\infty + \mathrm{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathscr{H}, j} \right),$$

where $C > 0$ only depends on $T : M \to M$.

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

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

where

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

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

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

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

$$|I_1| \leq \int \left| \mathbb{E}_\mu \left[F_y(X_0, \ldots, X_{k-1}) \right] - \mathbb{E} \left[F_y(\widehat{X}_0, \ldots, \widehat{X}_{k-1}) \right] \right| d\mu(y)$$

$$\leq \int C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \left(|F_y|_\infty + \mathrm{Lip}(F_y) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathscr{H}, j} \right) d\mu(y)$$

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

Now

$$I_2 = \mathbb{E}_\mu \left[F(X_0, \ldots, X_k)\right] - \int_M \mathbb{E}_\mu \left[F(X_0, \ldots, X_{k-1}, X_k(y))\right] d\mu(y)$$

$$= \int_M F\big(X_0(x), \ldots, X_k(x)\big) d\mu(x) - \int_{M^2} F\big(X_0(x), \ldots, X_{k-1}(x), X_k(y)\big) d\mu^2(x, y).$$

Write

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

and

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

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

$$\sum_{i=0}^{s-1} [G]_{\mathscr{H},i} \le \sum_{i=0}^k \sum_{j=0}^{u_i - \ell_i} \mathrm{Lip}(F)[\Phi_i]_{\mathscr{H},j}.$$

Hence by the Functional Correlation Bound,

$$|I_2| = \left| \int_M G(T^{\ell_0}x, \ldots, T^{u_0}x; \ldots; T^{\ell_k}x, \ldots, T^{u_k}x) d\mu(x) \right.$$
$$\left. - \int_{M^2} G(T^{\ell_0}x, \ldots, T^{u_0}x; \ldots; T^{\ell_{k-1}}x, \ldots, T^{u_{k-1}}x; T^{\ell_k}y, \ldots, T^{u_k}y) d\mu^2(x, y) \right|$$
$$\le C(\ell_k - u_{k-1})^{-\gamma} \left(|F|_\infty + \sum_{i=0}^k \sum_{j=0}^{u_i - \ell_i} \mathrm{Lip}(F)[\Phi_i]_{\mathscr{H},j} \right).$$

This completes the proof. $\qquad\square$

5. Moment Bounds

In this section we prove Theorem 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 2.4 we use the following moment bounds for independent, mean zero random variables, which are due to von Bahr, Esseen [32] and Rosenthal [30], respectively:

Lemma 5.1. *Fix $p \ge 1$. There exists a constant $C > 0$ such that for all $k \ge 1$, for all independent, mean zero random variables $\widehat{X}_0, \ldots, \widehat{X}_{k-1} \in L^p$:*

(i) *If $1 \le p \le 2$, then*

$$\mathbb{E}\left[\left| \sum_{i=0}^{k-1} \widehat{X}_i \right|^p \right] \le C \sum_{i=0}^{k-1} \mathbb{E}\left[|\widehat{X}_i|^p \right].$$

(ii) *If $p > 2$, then*

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

□

Let $v, w \in \mathscr{H}(M)$ 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 proposition.

Proposition 5.2. *Fix $\ell \ge 1$ and $0 = a_0 \le a_1 \le \cdots \le a_\ell$. Then,*

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

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

□

We also need the following elementary proposition:

Proposition 5.3. *Fix $R > 0$, $p \ge 1$ and an integer $k \ge 1$. Define $F: [-R, R]^k \to \mathbb{R}$ by $F(y_0, \ldots, y_{k-1}) = |y_0 + \cdots + y_{k-1}|^p$. Then $|F|_\infty \le (kR)^p$ and $\mathrm{Lip}(F) \le p(kR)^{p-1}$.*

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

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

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

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

$$\tfrac{n}{2k} - 1 \le a_{i+1} - a_i \le \tfrac{n}{2k} + 1 \le \tfrac{n}{k}. \tag{5.1}$$

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

Lemma 5.4. *There exists a constant $C > 0$ such that*

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

for all $n \ge 2k$, $k \ge 1$, for any $v \in \mathscr{H}(M)$.

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

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

where $F: [-R, R]^k \to \mathbb{R}$ is given by $F(y_0, \ldots, y_{k-1}) = |y_0 + \cdots + y_{k-1}|^{2\gamma}$. Hence by Lemma 4.1,

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

where

$$|A| \le C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma}\left(|F|_\infty + \mathrm{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathscr{H}, j}\right). \tag{5.2}$$

It remains to bound A. First we bound the expressions $[\Phi_i]_{\mathscr{H}, j}$. Fix $0 \le i < k$ and $0 \le j \le u_i - \ell_i$. 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 $[\Phi_i]_{\mathscr{H}, j} \le [v]_{\mathscr{H}}$. Note that by (5.1), $|\Phi_i|_\infty \le (a_{2i+1} - a_{2i}) |v|_\infty \le \frac{n}{k} |v|_\infty$. Hence by Proposition 5.3,

$$|F|_\infty \le 2\gamma (n |v|_\infty)^{2\gamma} \tag{5.3}$$

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

Thus

$$\mathrm{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathscr{H}, j} \le 2\gamma (n |v|_\infty)^{2\gamma - 1} \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathscr{H}, j}$$

$$\le 2\gamma (n |v|_\infty)^{2\gamma - 1} \sum_{i=0}^{k-1} (u_i - \ell_i + 1)[v]_{\mathscr{H}}$$

$$\le 2\gamma (n |v|_\infty)^{2\gamma - 1} n [v]_{\mathscr{H}}. \tag{5.4}$$

Now by (5.1), $\ell_{r+1} - u_r = a_{2r+2} - (a_{2r+1} - 1) \geq \frac{n}{2k}$ for each $0 \leq r \leq k - 2$. Hence

$$\sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \leq k(\tfrac{n}{2k})^{-\gamma} = 2^\gamma k^{1+\gamma} n^{-\gamma}. \tag{5.5}$$

Substituting (5.3), (5.4) and (5.5) into (5.2) gives

$$|A| \leq 2^\gamma k^{1+\gamma} n^{-\gamma} (2\gamma (n |v|_\infty)^{2\gamma} + 2\gamma (n |v|_\infty)^{2\gamma - 1} n[v]_{\mathcal{H}})$$
$$\leq 2^{1+\gamma} \gamma C k^{1+\gamma} n^{-\gamma} (n \|v\|_{\mathcal{H}})^{2\gamma} = 2^{1+\gamma} \gamma C k^{1+\gamma} n^\gamma \|v\|_{\mathcal{H}}^{2\gamma},$$

as required. $\square$

We are now ready to prove the moment bound for $S_v(n)$ (Theorem 2.4(a)).

Proof of Theorem 2.4(a). We prove by induction that there exists $D > 0$ such that

$$|S_v(m)|_{2\gamma} \leq Dm^{1/2} \|v\|_{\mathcal{H}} \tag{5.6}$$

for all $m \geq 1$, for any mean zero $v \in \mathcal{H}(M)$.

Claim There exists $C > 0$ such that for all mean zero $v \in \mathcal{H}(M)$, for any $D > 0$, for any $k \geq 1$ and any $n \geq 2k$ such that (5.6) holds for all $m < n$, we have

$$|S_v(n)|_{2\gamma}^{2\gamma} \leq C(k^{1+\gamma} + k^{1-\gamma} D^{2\gamma}) n^\gamma \|v\|_{\mathcal{H}}^{2\gamma}.$$

Now fix $k \geq 1$ such that $Ck^{1-\gamma} \leq \frac{1}{2}$. Fix $D > 0$ such that $Ck^{1+\gamma} \leq \frac{1}{2} D^{2\gamma}$ and (5.6) holds for all $m < 2k$ and any mean zero $v \in \mathcal{H}(M)$. Then the claim shows that for any $n \geq 2k$ such that (5.6) holds for all $m < n$, we have $|S_v(n)|_{2\gamma}^{2\gamma} \leq D^{2\gamma} n^\gamma \|v\|_{\mathcal{H}}^{2\gamma}$. Hence by induction, (5.6) holds for all $m \geq 1$.

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 (5.6) holds for all $m < n$. By Proposition 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 Lemma 5.4,

$$|I_1|_{2\gamma}^{2\gamma} = \mathbb{E}_\mu \left[\left| \sum_{i=0}^{k-1} X_i \right|^{2\gamma} \right] \leq Ck^{1+\gamma} n^\gamma \|v\|_{\mathcal{H}}^{2\gamma} + \mathbb{E} \left[\left| \sum_{i=0}^{k-1} \widehat{X}_i \right|^{2\gamma} \right]. \tag{5.7}$$

We now bound $\mathbb{E} \left[| \sum_{i=0}^{k-1} \widehat{X}_i |^{2\gamma} \right]$ by using Lemma 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 (5.6), $\mathbb{E}_\mu \left[|X_i|^{2\gamma} \right] \leq D^{2\gamma} (a_{2i+1} - a_{2i})^\gamma \|v\|_{\mathscr{H}}^{2\gamma}$. Hence by (5.1),

$$\sum_{i=0}^{k-1} \mathbb{E}\left[|\widehat{X}_i|^{2\gamma} \right] \leq \sum_{i=0}^{k-1} D^{2\gamma} (a_{2i+1} - a_{2i})^\gamma \|v\|_{\mathscr{H}}^{2\gamma}$$

$$\leq \sum_{i=0}^{k-1} D^{2\gamma} (n/k)^\gamma \|v\|_{\mathscr{H}}^{2\gamma} = D^{2\gamma} k^{1-\gamma} n^\gamma \|v\|_{\mathscr{H}}^{2\gamma}.$$

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

$$\sum_{i=0}^{k-1} \mathbb{E}\left[\widehat{X}_i^2 \right] = \sum_{i=0}^{k-1} \mathbb{E}_\mu \left[S_v(a_{2i+1} - a_{2i})^2 \right]$$

$$\leq \sum_{i=0}^{k-1} C(a_{2i+1} - a_{2i}) \|v\|_{\mathscr{H}}^2 \leq C(a_{2k-1} - a_0) \|v\|_{\mathscr{H}}^2$$

$$\leq Cn \|v\|_{\mathscr{H}}^2.$$

By Lemma 5.1(ii), it follows that

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

$$\leq C(1 + D^{2\gamma} k^{1-\gamma}) n^\gamma \|v\|_{\mathscr{H}}^{2\gamma}.$$

Hence by (5.7), overall

$$|I_1|_{2\gamma}^{2\gamma} \leq C(k^{1+\gamma} + D^{2\gamma} k^{1-\gamma}) n^\gamma \|v\|_{\mathscr{H}}^{2\gamma}.$$

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

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

$\square$

We now prove Theorem 2.4(b). Our proof follows the same lines as that of part (a).

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

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

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

The following lemma plays the same role that Lemma 5.4 played in the proof of Theorem 2.4(a).

Lemma 5.5. *There exists a constant $C > 0$ such that for any $v, w \in \mathcal{H}(M)$,*

$$\mathbb{E}_\mu \left[\left| \sum_{i=0}^{k-1} X_i \right|^\gamma \right] \leq C k n^\gamma \|v\|_{\mathcal{H}}^\gamma \|w\|_{\mathcal{H}}^\gamma + \mathbb{E} \left[\left| \sum_{i=0}^{k-1} \widehat{X}_i \right|^\gamma \right]$$

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

Proof. Note that

$$X_i(x) = \sum_{a_{2i} \leq q < r \leq a_{2i+1}-1} v(T^q x) w(T^r x) - \mathbb{E}_\mu \left[\mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) \right]$$

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

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

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

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

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

where $F : [-R, R]^k \to \mathbb{R}$ is given by $F(y_0, \ldots, y_{k-1}) = |y_0 + \cdots + y_{k-1}|^\gamma$. Hence by Lemma 4.1,

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

where

$$|A| \leq C \sum_{r=0}^{k-2} (\ell_{r+1} - u_r)^{-\gamma} \left(|F|_\infty + \text{Lip}(F) \sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathcal{H},j} \right). \tag{5.8}$$

It remains to bound A. The first step is to bound the expressions $[\Phi_i]_{\mathcal{H},j}$. Fix $0 \leq i < k, 0 \leq j \leq u_i - \ell_i$. Let $x_0, \ldots, x_{k-1}, x'_j \in M$. Note that

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

where

$$J_1 = \sum_{j < r \leq u_i - \ell_i} (v(x_j) w(x_r) - v(x'_j) w(x_r)),$$

$$J_2 = \sum_{0 \leq q < j} (v(x_q) w(x_j) - v(x_q) w(x'_j)).$$

Now,

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

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

$$[\Phi_i]_{\mathcal{H}, j} \leq (u_i - \ell_i) \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}}.$$

Now recall from (5.1) that $u_i - \ell_i + 1 = a_{2i+1} - a_{2i} \leq n/k$ so

$$\sum_{i=0}^{k-1} \sum_{j=0}^{u_i - \ell_i} [\Phi_i]_{\mathcal{H}, j} \leq \sum_{i=0}^{k-1} (u_i - \ell_i + 1)^2 \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}} \leq \frac{n^2}{k} \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}}. \quad (5.9)$$

Next note that

$$|\Phi_i|_\infty \leq \sum_{0 \leq q < r \leq u_i - \ell_i} |v|_\infty |w|_\infty + \left| \mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) \right|_\infty$$

$$\leq 2(n/k)^2 |v|_\infty |w|_\infty$$

so by Proposition 5.3, $|F|_\infty \leq \left(\frac{2n^2}{k} |v|_\infty |w|_\infty \right)^\gamma$ and $\mathrm{Lip}(F) \leq \gamma \left(\frac{2n^2}{k} |v|_\infty |w|_\infty \right)^{\gamma - 1}$. Combining these bounds with (5.5), (5.8) and (5.9) yields that

$$|A| \leq C 2^\gamma k^{1+\gamma} n^{-\gamma} \left(\left(\frac{2n^2}{k} |v|_\infty |w|_\infty \right)^\gamma + \gamma \left(\frac{2n^2}{k} |v|_\infty |w|_\infty \right)^{\gamma - 1} \frac{n^2}{k} \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}} \right)$$

$$\leq 2^{2\gamma} (1 + \gamma/2) C k n^\gamma \|v\|_{\mathcal{H}}^\gamma \|w\|_{\mathcal{H}}^\gamma,$$

as required. $\qquad \square$

We are now ready to prove Theorem 2.4(b).

Proof of Theorem 2.4(b). We prove by induction that there exists $D > 0$ such that

$$\left| \mathbb{S}_{v,w}(m) \right|_\gamma \leq Dm \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}} \quad (5.10)$$

for all $m \geq 1$, for any $v, w \in \mathcal{H}(M)$ mean zero.

Claim There exists $C > 0$ such that for all $v, w \in \mathcal{H}(M)$ mean zero, for any $D > 0$, any $k \geq 1$ and any $n \geq 2k$ such that (5.10) holds for all $m < n$, we have

$$\left| \mathbb{S}_{v,w}(n) \right|_\gamma^\gamma \leq C(k^\gamma + (k^{1-\gamma} + k^{-\gamma/2}) D^\gamma)(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma.$$

Now fix $k \geq 1$ such that $C(k^{1-\gamma} + k^{-\gamma/2}) \leq \frac{1}{2}$. Fix $D > 0$ such that $Ck^\gamma \leq \frac{1}{2} D^\gamma$ and (5.10) holds for all $m < 2k$ and any mean zero $v, w \in \mathcal{H}(M)$. Then the claim shows that if $n \geq 2k$ and (5.10) holds for all $m < n$, then $\left| \mathbb{S}_{v,w}(n) \right|_\gamma^\gamma \leq D^\gamma (n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma$. Hence by induction, (5.10) holds for all $m \geq 1$.

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 (5.10) holds for all $m < n$. Recall that $a_i = \left[\frac{in}{2k} \right]$ for $0 \leq i \leq 2k$. By Proposition 5.2(ii),

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

where

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

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

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

Recall from (5.1) that $a_{i+1} - a_i \le n/k$. Hence by Theorem 2.4(a),

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

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

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

We now bound $|I_3|_\gamma^\gamma$. Note that $I_3 = \sum_{i=0}^{k-1} X_i$, where $X_i = \mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) - \mathbb{E}_\mu \big[\mathbb{S}_{v,w}(a_{2i}, a_{2i+1}) \big]$. Hence by Lemma 5.5,

$$|I_3|_\gamma^\gamma = \mathbb{E}_\mu \left[\left| \sum_{i=0}^{k-1} X_i \right|^\gamma \right] \le Ckn^\gamma \|v\|_{\mathscr{H}}^\gamma \|w\|_{\mathscr{H}}^\gamma + \mathbb{E} \left[\left| \sum_{i=0}^{k-1} \widehat{X}_i \right|^\gamma \right]. \quad (5.11)$$

Fix $0 \le i < k$. By stationarity, $X_i =_d \mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) - \mathbb{E}_\mu \big[\mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) \big]$. Now by the inductive hypothesis (5.10), $\big| \mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) \big|_\gamma \le D(a_{2i+1} - a_{2i}) \|v\|_{\mathscr{H}} \|w\|_{\mathscr{H}}$, so

$$
\begin{aligned}
|X_i|_\gamma &\le \big| \mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) \big|_\gamma + |\mathbb{E}_\mu \big[\mathbb{S}_{v,w}(a_{2i+1} - a_{2i}) \big]| \\
&\le 2D(a_{2i+1} - a_{2i}) \|v\|_{\mathscr{H}} \|w\|_{\mathscr{H}} .
\end{aligned}
$$

It follows that

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

$$\leq \sum_{i=0}^{k-1} 2^\gamma D^\gamma (n/k)^\gamma (\|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma = 2^\gamma D^\gamma k^{1-\gamma} (n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma.$$

If $1 < \gamma \leq 2$, then by Lemma 5.1(i),

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

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

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

so

$$\sum_{i=0}^{k-1} \mathbb{E}\left[\widehat{X}_i^2\right] \leq \sum_{i=0}^{k-1} 4D^2 (a_{2i+1} - a_{2i})^2 (\|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^2$$

$$\leq \sum_{i=0}^{k-1} 4D^2 (n/k)^2 (\|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^2 = 4D^2 k^{-1} (n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^2.$$

Hence by Lemma 5.1(ii),

$$\mathbb{E}\left[\left|\sum_{i=0}^{k-1} \widehat{X}_i\right|^\gamma\right] \leq C\left(\left(4D^2 k^{-1} (n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^2\right)^{\gamma/2} + 2^\gamma D^\gamma k^{1-\gamma} \left(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}}\right)^\gamma\right)$$

$$= 2^\gamma C D^\gamma (k^{-\gamma/2} + k^{1-\gamma})(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma.$$

Hence for any $\gamma > 1$,

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

By (5.11), it follows that

$$|I_3|_\gamma^\gamma \leq C\left(k + D^\gamma (k^{-\gamma/2} + k^{1-\gamma})\right)(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma.$$

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

$$\left|\mathbb{S}_{v,w}(n)\right|_\gamma^\gamma = |I_1 + I_2 + I_3 + I_4|_\gamma^\gamma \leq 4^\gamma (|I_1|_\gamma^\gamma + |I_2|_\gamma^\gamma + |I_3|_\gamma^\gamma + |I_4|_\gamma^\gamma)$$

$$\leq C\left(k^\gamma + 1 + 2(k + D^\gamma (k^{-\gamma/2} + k^{1-\gamma}))\right)(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma$$

$$\leq C\left(k^\gamma + D^\gamma (k^{-\gamma/2} + k^{1-\gamma})\right)(n \|v\|_{\mathcal{H}} \|w\|_{\mathcal{H}})^\gamma,$$

as required. $\qquad\qquad\square$

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Bruin, H., Melbourne, I., Terhesiu, D.: Sharp polynomial bounds on decay of correlations for multidimensional nonuniformly hyperbolic systems and billiards. Ann. H. Lebesgue **4**, 407–451 (2021)
2. Bunimovič, L.A.: The ergodic properties of billiards that are nearly scattering. Dokl. Akad. Nauk SSSR **211**, 1024–1026 (1973)
3. Chazottes, J.-R., Gouëzel, S.: Optimal concentration inequalities for dynamical systems. Commun. Math. Phys. **316**(3), 843–889 (2012)
4. Chernov, N.: Decay of correlations and dispersing billiards. J. Stat. Phys. **94**(3–4), 513–556 (1999)
5. Chernov, N., Zhang, H.-K.: Billiards with polynomial mixing rates. Nonlinearity **18**(4), 1527–1553 (2005)
6. Chernov, N., Zhang, H.-K.: A family of chaotic billiards with variable mixing rates. Stoch. Dyn. **5**(4), 535–553 (2005)
7. Chevyrev, I., Friz, P.K., Korepanov, A., Melbourne, I., Zhang, H.: Multiscale systems, homogenization, and rough paths. In: Probability and Analysis in Interacting Physical Systems, Volume 283 of Springer Proceedings of Mathematical Statistics, pp. 17–48. Springer, Cham (2019)
8. Chevyrev, I., Friz, P.K., Korepanov, A., Melbourne, I., Zhang, H.: Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2. Ann. Inst. H. Poincaré Probab. Statist. to appear
9. De Simoi, J., Liverani, C.: Statistical properties of mostly contracting fast-slow partially hyperbolic systems. Invent. Math. **206**(1), 147–227 (2016)
10. De Simoi, J., Liverani, C.: Limit theorems for fast-slow partially hyperbolic systems. Invent. Math. **213**(3), 811–1016 (2018)
11. Demers, M., Melbourne, I., Nicol, M.: Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas. Nonlinearity **33**(8), 4095–4113 (2020)
12. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. Trans. Am. Math. Soc. **356**(4), 1637–1689 (2004)
13. Gottwald, G.A., Melbourne, I.: Homogenization for deterministic maps and multiplicative noise. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **469**(2156), 20130201 (2013)
14. Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. Probab. Theory Relat. Fields **128**(1), 82–122 (2004)
15. Gouëzel, S.: Sharp polynomial estimates for the decay of correlations. Israel J. Math. **139**, 29–65 (2004)
16. Gouëzel, S.: Vitesse de décorrélation et théoremes limites pour les applications non uniformément dilatantes. PhD thesis, Ecole Normale Supérieure (2004)
17. Gouëzel, S.: Berry–Esseen theorem and local limit theorem for non uniformly expanding maps. Ann. Inst. H. Poincaré Probab. Statist. **41**(6), 997–1024 (2005)
18. Kelly, D., Melbourne, I.: Smooth approximation of stochastic differential equations. Ann. Probab. **44**(1), 479–520 (2016)
19. Kelly, D., Melbourne, I.: Deterministic homogenization for fast–slow systems with chaotic noise. J. Funct. Anal. **272**(10), 4063–4102 (2017)
20. Korepanov, A., Kosloff, Z., Melbourne, I.: Explicit coupling argument for non-uniformly hyperbolic transformations. Proc. Roy. Soc. Edinb. Sect. A **149**(1), 101–130 (2019)
21. Korepanov, A., Kosloff, Z., Melbourne, I.: Deterministic homogenization under optimal moment assumptions for fast–slow systems. Part 1. Ann. Inst. H. Poincaré Probab. Statist., to appear

22. Leppänen, J.: Functional correlation decay and multivariate normal approximation for non-uniformly expanding maps. Nonlinearity **30**(11), 4239–4259 (2017)
23. Leppänen, J., Stenlund, M.: A note on the finite-dimensional distributions of dispersing billiard processes. J. Stat. Phys. **168**(1), 128–145 (2017)
24. Leppänen, J., Stenlund, M.: Sunklodas' approach to normal approximation for time-dependent dynamical systems. J. Stat. Phys. **181**(5), 1523–1564 (2020)
25. Liverani, C., Saussol, B., Vaienti, S.: A probabilistic approach to intermittency. Ergod. Theory Dyn. Syst. **19**(3), 671–685 (1999)
26. Markarian, R.: Billiards with polynomial decay of correlations. Ergod. Theory Dyn. Syst. **24**(1), 177–197 (2004)
27. Melbourne, I., Stuart, A.M.: A note on diffusion limits of chaotic skew-product flows. Nonlinearity **24**(4), 1361–1367 (2011)
28. Melbourne, I., Terhesiu, D.: Decay of correlations for non-uniformly expanding systems with general return times. Ergod. Theory Dyn. Syst. **34**(3), 893–918 (2014)
29. Melbourne, I., Varandas, P.: A note on statistical properties for nonuniformly hyperbolic systems with slow contraction and expansion. Stoch. Dyn. **16**(3), 1660012 (2016)
30. Rosenthal, H.P.: On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables. Israel J. Math. **8**, 273–303 (1970)
31. Sarig, O.: Subexponential decay of correlations. Invent. Math. **150**(3), 629–653 (2002)
32. von Bahr, B., Esseen, C.-G.: Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. Ann. Math. Stat. **36**, 299–303 (1965)
33. Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. (2) **147**(3), 585–650 (1998)
34. Young, L.-S.: Recurrence times and rates of mixing. Israel J. Math. **110**, 153–188 (1999)