Rate of Convergence in the Weak Invariance Principle for Deterministic Systems

Marios Antoniou * Ian Melbourne†

4 June 2018, revised 15 October 2018

Abstract

We obtain the first results on convergence rates in the Prokhorov metric for the weak invariance principle (functional central limit theorem) for deterministic dynamical systems. Our results hold for uniformly expanding/hyperbolic (Axiom A) systems, as well as nonuniformly expanding/hyperbolic systems such as dispersing billiards, Hénon-like attractors, Viana maps and intermittent maps. As an application, we obtain convergence rates for deterministic homogenization in multiscale systems.

1 Introduction

There is considerable interest in proving statistical properties for large classes of dynamical systems. The central limit theorem (CLT) was proved for uniformly hyperbolic (Axiom A) diffeomorphisms and flows in [42] and for various nonuniformly expanding/hyperbolic maps in [27].[1] The latter reference also established the weak invariance principle (WIP), otherwise known as the functional CLT, generalizing the classical result of Donsker [15] for independent and identically distributed random variables. See also [14] for the WIP for uniformly hyperbolic flows. More recently, the WIP was obtained for large classes of nonuniformly hyperbolic maps and flows [5, 19, 35, 39, 40]. The WIP was applied in [18, 37] to obtain results on homogenization for deterministic fast-slow systems.

An important question, especially bearing in mind applications to fast-slow systems, is to obtain convergence rates for these statistical limit laws. In the case of the CLT, these convergence rates are called Berry-Esseen estimates; sharp results have

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

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

[1]For the remainder of the introduction, we write (non)uniformly hyperbolic rather than (non)uniformly expanding/hyperbolic.

been obtained for uniformly hyperbolic diffeomorphisms [11] and nonuniformly hyperbolic systems [20]. However, there are few results on convergence rates in the WIP for dependent random variables in the probability theory literature and apparently none in the dynamical systems literature. In this paper, we obtain convergence rates in the WIP for uniformly and nonuniformly hyperbolic dynamical systems.

For uniformly hyperbolic dynamical systems (including Axiom A diffeomorphisms but also maps with infinitely many branches such as the Gauss map), we obtain the convergence rate $n^{-(\frac{1}{4}-\delta)}$ in the WIP for δ arbitrarily small. This result also applies to certain nonuniformly hyperbolic systems (those modelled by a Young tower with exponential tails [49]) such as dispersing planar periodic billiards, unimodal maps, and Hénon-like maps [8].

More generally, for nonuniformly hyperbolic systems the rate depends on the degree of nonuniformity. As an indicative example, we consider Pomeau-Manneville intermittent maps [41], specifically the map

$$T : [0,1] \to [0,1], \qquad T(x) = \begin{cases} x(1+2^\gamma x^\gamma) & x \in [0, \frac{1}{2}) \\ 2x-1 & x \in [\frac{1}{2}, 1] \end{cases}, \tag{1.1}$$

studied in [34]. Here $\gamma > 0$ is a real parameter and there is a unique ergodic absolutely continuous invariant probability measure μ for $\gamma < 1$. Moreover, Hölder observables satisfy the CLT and WIP provided $\gamma < \frac{1}{2}$. By [20], the convergence rate in the CLT is $\begin{cases} n^{-\frac{1}{2}} & \gamma \in (0, \frac{1}{3}) \\ n^{-(\frac{1}{2\gamma}-1)} & \gamma \in (\frac{1}{3}, \frac{1}{2}) \end{cases}$. For the WIP, we obtain the rate $n^{-(\frac{1}{4}(1-2\gamma)-\delta)}$, $\gamma \in (0, \frac{1}{2})$.

Remark 1.1 (a) The closest previous result that we could find for dynamical systems is due to Grama *et al.* [23]. Their method, is based on a result of Gouëzel [22] which applies to dynamical systems with spectral gaps, so it is plausible that [23] yields the convergence rate $n^{-(\frac{1}{4}-\delta)}$ in the WIP for uniformly hyperbolic maps (though no such claim is explicitly made in their paper). The results in [23] do not apply to nonuniformly hyperbolic systems such as (1.1).

(b) The convergence rates that we obtain are certainly not optimal, but this is the typical situation even in the probability literature as soon as one moves outside of the iid setting, see Remark 2.3 and references therein.

Next, we consider applications to fast-slow systems of the form

$$x_\epsilon(n+1) = x_\epsilon(n) + \epsilon^2 a_\epsilon(x_\epsilon(n), y(n)) + \epsilon b(x_\epsilon(n))v(y(n)), \quad x_\epsilon(0) = \xi, \tag{1.2}$$

where $a_\epsilon : \mathbb{R} \times \Lambda \to \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$, $v : \Lambda \to \mathbb{R}$ satisfy mild conditions, $\xi \in \mathbb{R}$, and the fast variables $y(n) \in \Lambda$ are generated by iterating a nonuniformly hyperbolic dynamical system. In [18], it was shown that the slow variables $x_\epsilon(n)$, suitably scaled in time, converge to the solution of the stochastic differential equation (SDE) given in (2.2). In this paper, we obtain the first estimates of the rate of convergence. For uniformly hyperbolic fast dynamics, we obtain the convergence rate $\epsilon^{\frac{1}{3}-\delta}$ (note that ϵ is identified with $n^{-\frac{1}{2}}$ so this corresponds to $n^{-(\frac{1}{6}-\delta)}$). We also obtain convergence rates for nonuniformly hyperbolic fast dynamics including the intermittent map (1.1) for all $\gamma \in (0, \frac{1}{2})$. Moreover, for $\gamma \in [\frac{1}{12}(11 - \sqrt{73}), \frac{1}{2})$ we obtain the same rate $\epsilon^{\frac{1}{2}(1-2\gamma)-\delta}$ as in the WIP.

Remark 1.2 This paper is based on results of the first author in his Ph. D. thesis [4] which concentrates on uniformly expanding maps but via a method which, as demonstrated in this paper, generalizes to nonuniformly hyperbolic transformations.

The remainder of the paper is organized as follows. In Section 2, we define a class of nonuniformly expanding maps and state precisely our results on convergence rates in the WIP and for homogenization of fast-slow systems. In Section 3, we summarize some recent results of [30] on martingale approximations. The main result for the WIP is proved in Section 4 and the main result for fast-slow systems is proved in Section 5. The extension to nonuniformly hyperbolic systems is covered in Section 6. In Section 7, we discuss examples where our results apply.

Notation We use "big O" and $\ll$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \ll b_n$ if there is a constant $C > 0$ such that $a_n \le C b_n$ for all $n \ge 1$.

Recall that $v : \Lambda \to \mathbb{R}$ is a Hölder observable on a metric space (Λ, d) with exponent $\eta \in (0, 1]$, written $v \in C^\eta(\Lambda)$, if $\|v\|_\eta = |v|_\infty + |v|_\eta < \infty$ where $|v|_\infty = \sup_\Lambda |v|$ and $|v|_\eta = \sup_{x \ne x'} \dfrac{|v(x) - v(x')|}{d(x, x')^\eta}$.

We denote by $C[0, 1]$ the Banach space of continuous functions on $[0, 1]$ equipped with the supnorm.

2 Statement of the main results

Let (Λ, d) be a bounded metric space with Borel probability measure ρ and let $T : \Lambda \to \Lambda$ be a nonsingular transformation (so $\rho(T^{-1}E) = 0$ if and only if $\rho(E) = 0$ for all Borel sets $E \subset \Lambda$). We assume that ρ is ergodic ($\rho(E) = 0$ or 1 for all Borel sets $E \subset \Lambda$ with $T^{-1}E = E$).

Suppose that $Y \subset \Lambda$ is a subset of positive measure, and that α is an at most countable measurable partition of Y. Let $\tau : Y \to \mathbb{Z}^+$ be an integrable function, constant on partition elements, and define $F(y) = T^{\tau(y)}(y)$. We assume that $FY \subset Y$; then τ is called a *return time* and $F : Y \to Y$ is the corresponding *induced map*.

We suppose that that there are constants $\lambda > 1$, $\eta \in (0, 1]$, $C > 0$ such that for all $a \in \alpha$, $y, y' \in a$, $0 \le \ell \le \tau(a) - 1$,

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

(ii) $d(Fy, Fy') \ge \lambda d(y, y')$.

(iii) $g = d\rho|_Y / d\rho|_Y \circ F$ satisfies $|\log g(y) - \log g(y')| \le C d(Fy, Fy')^\eta$.

(iv) $d(T^\ell y, T^\ell y') \le C d(Fy, Fy')$.

Such a dynamical system T is called *nonuniformly expanding*. We say that T is *nonuniformly expanding of order p* if $\tau \in L^p$. A standard consequence of (i)–(iii) is that there is a unique absolutely continuous ergodic T-invariant probability measure μ on Λ.

Let $C_0^\eta(\Lambda) = \{v \in C^\eta(\Lambda) : \int_\Lambda v\,d\mu = 0\}$. Given $v \in C_0^\eta(\Lambda)$, we define $v_n = \sum_{j=0}^{n-1} v \circ T^j$ for $n \ge 1$. Also, define

$$W_n(t) = n^{-\frac{1}{2}} v_{nt}, \quad \text{for } t = \tfrac{j}{n},\ 0 \le j \le n,$$

and linearly interpolate to obtain a process $W_n \in C[0,1]$. The following result is well-known, see for example [19, 30, 35]:

Lemma 2.1 *Suppose that $T : \Lambda \to \Lambda$ is nonuniformly expanding of order 2 and let $v \in C_0^\eta(\Lambda)$. Then*

(a) The limit $\sigma^2 = \lim_{n\to\infty} n^{-1} \int_\Lambda v_n^2\,d\mu$ exists. If in addition $\gcd\{\tau(a) : a \in \alpha\} = 1$ (guaranteeing that T is mixing), then σ^2 is given by the absolutely summable series

$$\sigma^2 = \int_\Lambda v^2\,d\mu + 2\sum_{n=1}^\infty \int_\Lambda v\,v \circ T^n\,d\mu. \tag{2.1}$$

(b) Typically $\sigma^2 > 0$. Indeed, there is a closed subspace $S \subset C_0^\eta(\Lambda)$ of infinite codimension such that $\sigma^2 > 0$ whenever $v \notin S$.

(c) The CLT holds: $n^{-\frac{1}{2}} v_n \to_d N(0, \sigma^2)$ as $n \to \infty$ on the probability space (Λ, μ).

(d) The WIP holds: $W_n \to_w W$ in $C[0,1]$ as $n \to \infty$ on the probability space (Λ, μ), where W is Brownian motion with variance σ^2. ∎

2.1 Rates in the WIP

So far, we assumed only that $\tau \in L^2$. For $\tau \in L^p$, $p > 2$, Gouëzel [20] obtained convergence rates (Berry-Esseen estimates) in the CLT. Our first main result is a convergence rate in the WIP.

The Prokhorov metric π_1 is given by

$$\pi_1(X, Y) = \inf\{\epsilon > 0 : \mathbb{P}(X \in A) \le \mathbb{P}(Y \in A^\epsilon) + \epsilon \text{ for all closed } A \in \mathcal{B}\}.$$

Here, $\mathcal{B}$ is the Borel σ-algebra on $C[0,1]$ and A^ϵ is the ϵ-neighbourhood of A. The WIP in Lemma 2.1(d) can be rewritten as $\lim_{n\to\infty} \pi_1(W_n, W) = 0$.

Theorem 2.2 *Let $T : \Lambda \to \Lambda$ be nonuniformly expanding of order $p > 2$ and suppose that $v \in C_0^\eta(\Lambda)$. Then there is a constant $C > 0$ such that $\pi_1(W_n, W) \le Cn^{-r(p)}$ for all $n \ge 1$, where $r(p) = \frac{p-2}{4p}$.*

Remark 2.3 The exact formula for the function $r : (2, \infty) \to (0, \frac{1}{4})$ can probably be improved slightly using more careful arguments. However, it is known that the main feature, namely $\sup r = \frac{1}{4}$, is essentially optimal under the methods used.

Specifically, we use a result of Kubilius [31] which builds upon [24, 25]. These results use the martingale version of the Skorokhod embedding theorem; by [9, 43], this method cannot yield rates better than $O(n^{-\frac{1}{4}})$.

2.2 Rates for fast-slow systems

Next, we consider fast-slow dynamical systems of the form (1.2). Here $x_\epsilon \in \mathbb{R}$ denotes the slow variables and the fast y-variables are generated by a nonuniformly expanding map $T : \Lambda \to \Lambda$, so $y(n) = T^n y_0$ where y_0 is chosen randomly from (Λ, μ).

We continue to assume that $v \in C_0^\eta(\Lambda)$ and also that $a_\epsilon : \mathbb{R} \times \Lambda \to \mathbb{R}$ and $b : \mathbb{R} \to \mathbb{R}$, satisfy the following regularity conditions:

Regularity conditions: There are constants $C > 0$, $\operatorname{Lip} a_0 > 0$, such that

(i) $\sup_\epsilon |a_\epsilon|_\infty < \infty$. (ii) $\sup_{x,y,\epsilon} |a_\epsilon(x, y) - a_0(x, y)| \le C\epsilon^{\frac{1}{3}}$.

(iii) a_0 is Lipschitz in x uniformly in y. That is, $|a_0(x, y) - a_0(x', y)| \le \operatorname{Lip} a_0 \, |x - x'|$ for all $x, x' \in \mathbb{R}$, $y \in \Lambda$.

Moreover, b is C^2 and nonvanishing with $b, b', b'', 1/b \in L^\infty$. (It is clear from the proof that C^2 can be reduced to $C^{\frac{4}{3}}$.)

Let $\hat{x}_\epsilon(t) = x_\epsilon(t\epsilon^{-2})$ for $t = 0, \epsilon^2, 2\epsilon^2, \dots$ and linearly interpolate to obtain $\hat{x}_\epsilon \in C[0, 1]$. By [18, Theorem 1.3] (see also [30, Section 6]), $\hat{x}_\epsilon \to_w X$ in $C[0, 1]$ for T nonuniformly expanding of order 2, where X is the solution of the Stratonovich SDE

$$dX = \Big\{\bar{a}(X) - \frac{1}{2} b(X)b'(X) \int_\Lambda v^2 \, d\mu\Big\} \, dt + b(X) \circ dW, \quad X(0) = \xi. \tag{2.2}$$

Here W is Brownian motion with variance σ^2 as in Lemma 2.1 and $\bar{a}(x) = \int_\Lambda a(x, y) \, d\mu(y)$. Our second main result gives an estimate for the rate of convergence $\pi_1(\hat{x}_\epsilon, X)$ in the Prokhorov metric.

Theorem 2.4 *Let $T : \Lambda \to \Lambda$ be nonuniformly expanding of order $p > 2$. Suppose that a_ϵ and b satisfy the above regularity conditions. Suppose further that $v \in C_0^\eta(\Lambda)$ and that $\sup_{x\in\mathbb{R}} |a_0(x, \cdot)|_\eta < \infty$. Then there is a constant $C > 0$ such that*

$$\pi_1(\hat{x}_\epsilon, X) \le \begin{cases} C\epsilon^{\frac{p-2}{2p}} & p \le p_* \\ C\epsilon^{\frac{1}{3}\frac{2p-2}{2p-1}}(-\log\epsilon)^{\frac{1}{2}(p-1)} & p > p_* \end{cases},$$

where $p_* = \frac{1}{4}(11 + \sqrt{73}) \approx 4.89$.

Remark 2.5 When T is mixing, we have the alternative representation of the SDE (2.2) in Itô form

$$dX = \Big\{\bar{a}(X) + b(X)b'(X)\sum_{n=1}^{\infty}\int_{\Lambda} v\, v\circ T^n\, d\mu\Big\}\, dt + b(X)\, dW, \quad X(0)=\xi.$$

This is immediate from formula (2.1) and the Itô-Stratonovich conversion formula. The nature of the drift coefficient and the departure from Itô or Stratonovich is discussed further in [16, 18, 32].

3 Martingale approximation

In this section, we recall some results of Gordin-type [17] on martingale approximation from [30]. The key advantage of [30] over other martingale approximation methods [17, 33, 45] is that it gives good control over sums of squares of the approximating martingale, see Proposition 3.5 below.

Recall that a measure-preserving transformation $f:\Delta\to\Delta$ on a probability space $(\Delta,\mathcal{M},\mu_\Delta)$ is called an *extension* of $T:\Lambda\to\Lambda$ if there is a measure-preserving map $\pi_\Delta:\Delta\to\Lambda$ such that $\pi_\Delta\circ f = T\circ\pi_\Delta$. The map π_Δ is called a *semiconjugacy.*

Throughout this section, we suppose that $T:\Lambda\to\Lambda$ is a nonuniformly expanding map of order $p\ge 2$ and that $\eta\in(0,1]$. By [30, Propositions 2.4 and 2.5, and Corollary 2.8], we can form the following "martingale-coboundary decomposition":

Proposition 3.1 *There is an extension $f:\Delta\to\Delta$ of $T:\Lambda\to\Lambda$ such that for any $v\in C_0^\eta(\Lambda)$ there exists $m\in L^p(\Delta)$ and $\chi\in L^{p-1}(\Delta)$ with*

$$v\circ\pi_\Delta = m + \chi\circ f - \chi, \qquad \mathbb{E}(m|f^{-1}\mathcal{M}) = 0.$$

Moreover, there is a constant $C>0$ such that for all $v\in C_0^\eta(\Lambda)$, $n\ge 1$,

$$|m|_p \le C\|v\|_\eta, \quad \text{and} \quad \Big|\max_{j\le n}|\chi\circ f^j - \chi|\Big|_p \le C\|v\|_\eta\, n^{\frac{1}{p}}.\ [2]$$

∎

Remark 3.2 The extension Δ is called the Young tower [50] associated to the nonuniformly expanding map T. The definition and properties of Δ are not required in this paper; the current section summarizes all the results that we need.

Let

$$\mathcal{G}_{n,j} = f^{-(n-j)}\mathcal{M}, \quad \xi_{n,j} = n^{-\frac{1}{2}}\sigma^{-1} m\circ f^{n-j}.$$

Proposition 3.3 $\{\xi_{n,j},\mathcal{G}_{n,j};\ 1\le j\le n\}$ *is a martingale difference array.*

Proof See for example [30, Proposition 2.9]. ∎

[2] By [30], $\chi\circ f^j - \chi\in L^p(\Delta)$ for $j\ge 1$ even though χ is generally only L^{p-1}.

Proposition 3.4 *There is a constant $C>0$ such that*

$$\big|\max_{j\le n}|v_j|\big|_{2(p-1)}\le C\|v\|_\eta\, n^{\frac12}\quad \textit{for all } v\in C_0^\eta(\Lambda),\ n\ge 1.$$

Proof This result is due to [36, 38]. For the formulation stated here, see for example [30, Corollary 2.10]. ∎

Proposition 3.5 *There is a constant $C>0$ such that*

$$\Big|n^{-1}\sum_{j=0}^{k-1}\mathbb{E}(m^2-\sigma^2|f^{-1}\mathcal{M})\circ f^j\Big|_{2(p-1)}\le C\|v\|_\eta^2\, n^{-\frac12}\quad \textit{for all } v\in C_0^\eta(\Lambda),\ n\ge 1.$$

Proof This is [30, Corollary 3.2]. For some reason, the result is stated there only with an L^p bound, but the superior $L^{2(p-1)}$ bound here is immediate from the argument used in [30] (specifically [30, Corollary 2.10, Remark 2.16 and Proposition 3.1]). ∎

4 Convergence rates in the WIP

In this section, we prove Theorem 2.2. Recall that $T:\Lambda\to\Lambda$ is a nonuniformly expanding map of order $p>2$ and $v\in C_0^\eta(\Lambda)$. Let $r(p)=\frac{p-2}{4p}$.

We let C denote constants that may depend on p and $\|v\|_\eta$. Also, C' denotes constants that may depend on p but not v.

Define $v\circ\pi_\Delta=m+\chi\circ f-\chi$ and the martingale difference array $\{\xi_{n,j},f^{-(n-j)}\mathcal{M}\}$ as in Section 3.

For $\ell\ge1$, define $V_{n,\ell}=\sum_{j=1}^{\ell}\mathbb{E}(\xi_{n,j}^2|\mathcal{G}_{n,j-1})$.

Proposition 4.1 $\big|\max_{k\le n}|V_{n,k}-\frac{k}{n}|\big|_{2(p-1)}\le C'\|v\|_\eta^2\, n^{-\frac12}$.

Proof We have

$$\begin{aligned}V_{n,k}-\frac{k}{n}&=\sigma^{-2}n^{-1}\sum_{j=1}^{k}\mathbb{E}(m^2\circ f^{n-j}|f^{-(n-j+1)}\mathcal{M})-\frac{k}{n}\\&=\sigma^{-2}n^{-1}\sum_{j=0}^{k-1}\mathbb{E}(m^2-\sigma^2|f^{-1}\mathcal{M})\circ f^j.\end{aligned}$$

The result follows from Proposition 3.5. ∎

For each $n\ge1$, define

$$X_n(t)=\sum_{j=1}^{k}\xi_{n,j},\quad \text{for } t=V_{n,k}/V_{n,n},\ 0\le k\le n, \tag{4.1}$$

and linearly interpolate to obtain a process $X_n \in C[0,1]$. We use the following result of [31] for martingale difference arrays to estimate the rate of convergence of X_n to the unit Brownian motion B.

Theorem 4.2 (Kubilius [31, Theorem 1]) *Let $\delta \in [0,\frac{3}{4}] \cup \{1\}$. There is a constant $C>0$ such that $\pi_1(X_n,B) \le C\lambda|\log\lambda|$ where $\lambda = \lambda_1+\lambda_2$ and*

$$\lambda_1 = \inf_{0\le\epsilon\le 1}\big\{\epsilon^{\frac12} + \big(\mathbb{E}\sum_{j=1}^n |\xi_{n,j}|^{2+2\delta}1_{\{|\xi_{n,j}|>\epsilon\}}\big)^{1/(3+2\delta)}\big\},$$
$$\lambda_2 = \inf_{0\le\epsilon\le 1}\big\{\epsilon + \mathbb{P}(|V_{n,n}-1|>\epsilon^2)\big\}.$$
■

Lemma 4.3 $\pi_1(X_n,B) \le Cn^{-r(p)}$.

Proof Let $\lambda=\lambda_1+\lambda_2$ be as in Theorem 4.2. We claim that

$$\lambda_1 \ll \|v\|_\eta^{r'} n^{-r_1(p)}, \qquad \lambda_2 \ll \|v\|_\eta^{(4p-4)/(4p-3)} n^{-(p-1)/(4p-3)},$$

where $r_1(p) = \begin{cases} \frac{p-2}{2p+2} & 2<p\le\frac72 \\ \frac{p-2}{4p-5} & \frac72\le p<4 \\ \frac{p-2}{4p-6} & p\ge 4\end{cases}$. Then $\lambda_2 \ll \lambda_1$ and it follows from Theorem 4.2 that $\pi_1(X_n,B) \ll n^{-r_1(p)}\log n$. In all cases, $r_1(p)>r(p)$ so the result follows.

First, we verify the estimate for λ_1. Choose $\delta \in [0,\frac34]\cup\{1\}$ greatest such that $2+2\delta\le p$. In other words, $\delta = \begin{cases}\frac{p-2}{2} & 2\le p\le\frac72 \\ \frac34 & \frac72\le p<4 \\ 1 & p\ge 4\end{cases}$. By stationarity,

$$\mathbb{E}\sum_{j=1}^n |\xi_{n,j}|^{2+2\delta}1_{\{|\xi_{n,j}|\ge\epsilon\}} = \sigma^{-(2+2\delta)}n^{-\delta}\mathbb{E}(|m|^{2+2\delta}1_{\{|m|\ge\epsilon\sigma n^{1/2}\}}).$$

By Hölder's inequality, and then Markov's inequality,

$$\begin{aligned}\sigma^{2+2\delta}\mathbb{E}\sum_{j=1}^n |\xi_{n,j}|^{2+2\delta}1_{\{|\xi_{n,j}|\ge\epsilon\}} &\le n^{-\delta}|m|_p^{2+2\delta}\,\mu(|m|\ge\epsilon\sigma n^{1/2})^{(p-2-2\delta)/p} \\ &\le n^{-\delta}|m|_p^{2+2\delta}\Big(\frac{|m|_p^p}{\epsilon^p\sigma^p n^{p/2}}\Big)^{(p-2-2\delta)/p} \\ &= \sigma^{-(p-2-2\delta)}|m|_p^p\,\epsilon^{-(p-2-2\delta)}n^{-(p-2)/2}.\end{aligned}$$

Hence

$$\begin{aligned}\lambda_1 &\ll \inf_{0\le\epsilon\le1}\{\epsilon^{1/2} + \|v\|_\eta^{p/(3+2\delta)}\epsilon^{-(p-2-2\delta)/(3+2\delta)}n^{-(p-2)/(6+4\delta)}\} \\ &\le 2\|v\|_\eta^{p/(2p-2\delta-1)}n^{-(p-2)/(4p-4\delta-2)} = \|v\|_\eta^{p/(2p-2\delta-1)}n^{-r_1(p)}.\end{aligned}$$

Second, we verify the estimate for λ_2. By Proposition 4.1 and Markov's inequality, $\mathbb{P}(|V_{n,n}-1|>\epsilon^2) \ll \|v\|_\eta^{4(p-1)}\epsilon^{-4(p-1)}n^{-(p-1)}$. Hence

$$\lambda_2 \ll \inf_{0\le\epsilon\le 1}\{\epsilon+\|v\|_\eta^{4(p-1)}\epsilon^{-4(p-1)}n^{-(p-1)}\} \le 2\|v\|_\eta^{(4p-4)/(4p-3)}n^{-(p-1)/(4p-3)},$$

completing the proof. ∎

The integer k in (4.1) is a random variable $k=k_{n,t}:\Delta\to\{0,\dots,n\}$ given by

$$V_{n,k}\le tV_{n,n}<V_{n,k+1}.$$

Proposition 4.4 $\big|\sup_{[0,1]}|k-[nt]|\big|_{2(p-1)}\le Cn^{\frac12}$.

Proof Set $\widetilde{V}_{n,j}=nV_{n,j}-j$. Then $\widetilde{V}_{n,k}+k\le t\widetilde{V}_{n,n}+nt<\widetilde{V}_{n,k+1}+k+1$, and it follows that $k-nt$ satisfies the inequalities

$$k-nt\le t\widetilde{V}_{n,n}-\widetilde{V}_{n,k}\le \widetilde{V}_{n,n},$$

and

$$k-nt>t\widetilde{V}_{n,n}-\widetilde{V}_{n,k+1}-1\ge -\widetilde{V}_{n,k+1}-1.$$

Hence

$$|k-[nt]|\le|k-nt|+1\le\max_{j\le n+1}|\widetilde{V}_{n,j}|+2=n\max_{j\le n+1}|V_{n,j}-\tfrac{j}{n}|+2,$$

and so the result follows from Proposition 4.1. ∎

4.1 Passing from X_n to W_n

Proposition 4.5 *Let Y, $Y'\in C[0,1]$ be random elements defined on a common probability space, and let ϵ_0, $\epsilon_1>0$, $q\ge 1$.*
(a) If $\mathbb{P}\big(\sup_{[0,1]}|Y-Y'|\ge\epsilon_0\big)\le\epsilon_1$, then $\pi_1(Y,Y')\le\max\{\epsilon_0,\epsilon_1\}$.
(b) If $\big|\sup_{[0,1]}|Y-Y'|\big|_q\le\epsilon_0$, then $\pi_1(Y,Y')\le\epsilon_0^{q/(q+1)}$.

Proof (a) This is immediate from the definition of π_1.
(b) By Markov's inequality, $\mathbb{P}(\sup_{[0,1]}|Y-Y'|\ge\epsilon)\le\epsilon^{-q}\epsilon_0^q$. In particular, $\mathbb{P}(\sup_{[0,1]}|Y-Y'|\ge\epsilon_1)\le\epsilon_1$ if ϵ_0 satisfies $\epsilon_1^{-q}\epsilon_0^q=\epsilon_1$. In other words, $\epsilon_1=\epsilon_0^{q/(q+1)}$. Hence the result follows from part (a). ∎

Proposition 4.6 [3] *For $n\ge1$, define $Z_n=\max_{0\le i,\ell\le n^{\frac12}}|v_\ell|\circ T^{i[n^{\frac12}]}$. Then*

[3] This estimate was suggested to us by Alexey Korepanov.

(a) $|\sum_{j=a}^{b-1} v\circ T^j| \le Z_n((b-a)(n^{\frac12}-1)^{-1}+3)$ *for all* $0\le a<b\le n$.

(b) $|Z_n|_{2(p-1)} \le C'\|v\|_\eta\, n^{\frac14+\frac{1}{4(p-1)}}$ *for all* $n\ge 1$.

Proof (a) Choose $0\le \ell_1\le\ell_2\le\sqrt{n}$ greatest such that $\ell_1[\sqrt{n}]\le a$ and $\ell_2[\sqrt{n}]\le b$. Then $\ell_2-\ell_1\le \frac{b-a}{\sqrt{n}-1}+1$. It follows that

$$\Big|\sum_{j=a}^{b-1} v\circ T^j\Big| \le (\ell_2-\ell_1+2)Z_n \le \Big(\frac{b-a}{\sqrt{n}-1}+3\Big)Z_n,$$

as required.
(b) We have

$$\int_\Delta |Z_n|^{2(p-1)}\,d\mu_\Delta \le \sum_{i\le n^{1/2}} \int_\Delta \max_{\ell\le n^{1/2}} |v_\ell|^{2(p-1)}\,d\mu_\Delta \ll n^{\frac12}\int_\Delta \max_{\ell\le n^{1/2}} |v_\ell|^{2(p-1)}\,\mu_\Delta.$$

By Proposition 3.4,

$$|Z_n|_{2(p-1)} \ll n^{\frac{1}{4(p-1)}}\Big|\max_{\ell\le n^{1/2}}|v_\ell|\Big|_{2(p-1)} \ll \|v\|_\eta\, n^{\frac14+\frac{1}{4(p-1)}}.$$

∎

Following [29, Lemma 4.8], we define the linear functional

$$g: C[0,1]\to C[0,1], \qquad g(u)(t)=u(1)-u(1-t).$$

Lemma 4.7 $\pi_1(g\circ W_n\circ\pi_\Delta, \sigma X_n)\le Cn^{-r(p)}$.

Proof Define the piecewise constant process $V_n'(t)=n^{-\frac12}\sum_{n-[nt]}^{n-k-1} v\circ T^j$. Then

$$\begin{aligned} g\circ W_n(t)\circ\pi_\Delta-\sigma X_n(t) &= n^{-\frac12}\sum_{j=n-[nt]}^{n-1} v\circ T^j\circ\pi_\Delta - n^{-\frac12}\sum_{j=1}^{k} m\circ f^{n-j}+E_n(t)\\ &= n^{-\frac12}\sum_{j=n-[nt]}^{n-1} v\circ T^j\circ\pi_\Delta - n^{-\frac12}\Big(\sum_{j=n-k}^{n-1} v\circ\pi_\Delta\circ f^j-\chi\circ f^n+\chi\circ f^{n-k}\Big)+E_n(t)\\ &= V_n'(t)\circ\pi_\Delta + n^{-\frac12}(\chi\circ f^{n-k}-\chi\circ f^n)+E_n(t),\end{aligned}$$

where $|E_n(t)|_\infty\le n^{-\frac12}|v|_\infty$. By Proposition 3.1,

$$\big|n^{-1/2}\sup_{[0,1]}|\chi\circ f^{n-k}-\chi\circ f^n|\big|_p \le 2n^{-\frac12}\big|\max_{j\le n}|\chi\circ f^j-\chi|\big|_p \ll n^{-(\frac12-\frac1p)}.$$

By Propositions 4.4 and 4.6, and Cauchy-Schwarz,

$$\begin{aligned}\big|\sup_{[0,1]}|V_n'(t)\circ\pi_\Delta|\big|_{p-1} = \big|\sup_{[0,1]}|V_n'(t)|\big|_{p-1} &\le n^{-\frac12}\big|Z_n\big(n^{-\frac12}\sup_{[0,1]}|[nt]-k|+3\big)\big|_{p-1} \\ &\le n^{-\frac12}|Z_n|_{2(p-1)}\,\big(n^{-\frac12}\big|\sup_{[0,1]}|[nt]-k|\big|_{2(p-1)}+3\big) \\ &\ll n^{-\frac12}|Z_n|_{2(p-1)} \ll n^{-(\frac14-\frac{1}{4(p-1)})} = n^{-\frac14\frac{p-2}{p-1}}.\end{aligned}$$

Combining these estimates, we obtain $\big|\sup_{[0,1]}|g\circ W_n\circ\pi_\Delta-\sigma X_n|\big|_{p-1} \ll n^{-\frac14\frac{p-2}{p-1}}$. Now apply Proposition 4.5(b). ∎

Proof of Theorem 2.2 It is easy to see that $g(W) =_d W = \sigma B$. Since π_Δ is a semiconjugacy, $W_n\circ\pi_\Delta =_d W_n$. Hence combining Lemma 4.3 and Lemma 4.7,

$$\begin{aligned}\pi_1(g\circ W_n, g\circ W) &= \pi_1(g\circ W_n\circ\pi_\Delta, \sigma B) \\ &\le \pi_1(g\circ W_n\circ\pi_\Delta, \sigma X_n) + \pi_1(\sigma X_n, \sigma B) \ll n^{-r(p)}.\end{aligned}$$

Now $g\circ g = \mathrm{Id}$. Also, $g: C[0,1]\to C[0,1]$ is Lipschitz with $\operatorname{Lip} g\le 2$. That is $\sup_{[0,1]}|g(u)-g(v)| \le 2\sup_{[0,1]}|u-v|$ for all $u,v\in C[0,1]$. Hence it follows from the Lipschitz mapping theorem [47, Theorem 3.2] that $\pi_1(W_n,W) = \pi_1(g(g\circ W_n), g(g\circ W)) \le 2\pi_1(g\circ W_n, g\circ W) \ll n^{-r(p)}$. ∎

5 Convergence rates for homogenization

In this section, we prove Theorem 2.4. We begin by proving an abstract result, Theorem 5.1 below, before specializing to the case where T is nonuniformly expanding.

Let $T:\Lambda\to\Lambda$ be a map with ergodic invariant probability measure μ. Define $y(n) = T^n y_0$ where y_0 is chosen randomly from (Λ,μ), and fix $\xi\in\mathbb{R}$. We consider fast-slow systems of the form (1.2) where $a_\epsilon:\mathbb{R}\times\Lambda\to\mathbb{R}$ and $b:\mathbb{R}\to\mathbb{R}$ satisfy the regularity assumptions listed in Section 2.2 and $v\in L^\infty(\Lambda)$.

Define

$$\bar a(x) = \textstyle\int_\Lambda a_0(x,y)\,d\mu(y), \qquad \tilde a(x,y) = a_0(x,y)-\bar a(x).$$

Let $\tilde a_u(y) = \tilde a(u,y)$. We suppose that there exist $C>0$, $q\ge 1$ such that

$$\big|\textstyle\sum_{j=0}^{n-1}\tilde a_u\circ T^j\big|_q \le Cn^{\frac12} \quad \text{for all } n\ge 1,\ u\in\mathbb{R}. \tag{5.1}$$

$$\big|\textstyle\sum_{j=0}^{n-1} v\circ T^j\big|_q \le Cn^{\frac12} \quad \text{for all } n\ge 1. \tag{5.2}$$

When $b\not\equiv 1$, we require in addition that

$$\big|\textstyle\sum_{j=0}^{n-1} v^2\circ T^j - n\int_\Lambda v^2\,d\mu\big|_q \le Cn^{\frac12} \quad \text{for all } n\ge 1. \tag{5.3}$$

Define $W_\epsilon(t) = \epsilon v_{t\epsilon^{-2}}$ for $t = 0,\epsilon^2,2\epsilon^2,\dots$ and linearly interpolate (and restrict to $[0,1]$) to obtain $W_\epsilon\in C[0,1]$. We assume that $W_\epsilon\to_w W$ in $C[0,1]$ where W

is a one-dimensional Brownian motion with variance σ^2. Let $\hat{x}_\epsilon(t) = x_\epsilon(t\epsilon^{-2})$ for $t = 0, \epsilon^2, 2\epsilon^2, \ldots$ and linearly interpolate to obtain $\hat{x}_\epsilon \in C[0,1]$. By [18, Theorem 1.3], $\hat{x}_\epsilon \to_w X$ in $C[0,1]$ where X is the solution of the Stratonovich SDE (2.2).

Theorem 5.1 *Suppose that $\pi_1(W_\epsilon, W) = O(\epsilon^r)$ for some $r > 0$. Then $\pi_1(\hat{x}_\epsilon, X) \leq C(\epsilon^r + \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}})$.*

Remark 5.2 It is easily seen from the proof that in the special case $a_\epsilon(x,y) \equiv \bar{a}(x)$, $b \equiv 1$, we obtain the same rate $\pi_1(\hat{x}_\epsilon, X) = O(\epsilon^r)$ as in the WIP.

Proof of Theorem 2.4 In the definitions of W_n and W_ϵ, notice that ϵ is identified with $n^{-\frac{1}{2}}$. Hence it follows from Theorem 2.2 that $\pi_1(W_\epsilon, W) = O(\epsilon^{\frac{p-2}{2p}})$.

Next, we verify the moment conditions (5.1)–(5.3) with $q = 2(p-1)$. Since v and $v^2 - \int_\Lambda v^2\,d\mu$ lie in $C_0^\eta(\Lambda)$, conditions (5.2) and (5.3) follow from Proposition 3.4. The assumption on a_0 implies that $\tilde{a}_u \in C_0^\eta(\Lambda)$ for all $u \in \mathbb{R}$ and that $\sup_u \|\tilde{a}_u\|_\eta < \infty$, so (5.1) also follows from Proposition 3.4.

We have now verified all of the hypotheses of Theorem 5.1 and it follows that $\pi_1(\hat{x}_\epsilon, X) \leq C\big\{\epsilon^{\frac{p-2}{2p}} + \epsilon^{\frac{1}{3}\frac{2p-2}{2p-1}}(-\log\epsilon)^{\frac{1}{2}(p-1)}\big\}$. ∎

In Subsection 5.1, we prove Theorem 5.1 when $b \equiv 1$. The general case is proved in Subsection 5.2 by reducing to the case $b \equiv 1$.

5.1 The case $b \equiv 1$

In this subsection, we prove Theorem 5.1 in the special case where $b \equiv 1$. In this case, the limiting SDE takes the form

$$dX = \bar{a}(X)\,dt + dW. \qquad X(0) = \xi. \tag{5.4}$$

Theorem 5.3 *Suppose that $\pi_1(W_\epsilon, W) = O(\epsilon^r)$ for some $r > 0$. Then $\pi_1(\hat{x}_\epsilon, X) \leq C(\epsilon^r + \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}})$.*

We have the following preliminary calculation: Set $M = [\epsilon^{-\frac{4}{3}}]$.

Proposition 5.4 *Suppose that $b(x) \equiv 1$. Then*

$$\hat{x}_\epsilon(t) = \xi + \int_0^t \bar{a}(\hat{x}_\epsilon(s))\,ds + W_\epsilon(t) + D_\epsilon(t) + E_\epsilon(t),$$

where

$$D_\epsilon(t) = \epsilon^{\frac{2}{3}} \sum_{n=0}^{[t\epsilon^{-\frac{2}{3}}]-1} J_\epsilon(n), \qquad J_\epsilon(n) = \epsilon^{\frac{4}{3}} \sum_{j=nM}^{(n+1)M-1} \tilde{a}(x_\epsilon(nM), y(j)),$$

and $\big|\sup_{[0,1]}|E_\epsilon|\big|_q \leq C\epsilon^{\frac{1}{3}}$.

Proof Introduce the step function $\tilde{x}_\epsilon(t) = x_\epsilon([t\epsilon^{-2}])$. Then

$$\tilde{x}_\epsilon(t) = \xi + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} a_\epsilon(x_\epsilon(j), y(j)) + \epsilon \sum_{j=0}^{[t\epsilon^{-2}]-1} v(y(j)).$$

Using the estimates

$$|\hat{x}_\epsilon(t) - \tilde{x}_\epsilon(t)|_\infty \le \epsilon^2 |a_\epsilon|_\infty + \epsilon |v|_\infty, \qquad |W_\epsilon(t) - \epsilon \textstyle\sum_{j=0}^{[t\epsilon^{-2}]-1} v(y(j))|_\infty \le \epsilon |v|_\infty,$$
$$|\epsilon^2 \textstyle\sum_{j=0}^{[t\epsilon^{-2}]-1} \{a_\epsilon(x_\epsilon(j), y(j)) - a_0(x_\epsilon(j), y(j))\}|_\infty = O(\epsilon^{\frac{1}{3}}),$$

we obtain

$$\begin{aligned} \hat{x}_\epsilon(t) &= \xi + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} a_0(x_\epsilon(j), y(j)) + W_\epsilon(t) + O(\epsilon^{\frac{1}{3}}) && (5.5) \\ &= \xi + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} \bar{a}(\hat{x}_\epsilon(\epsilon^2 j)) + W_\epsilon(t) + F_\epsilon(t) + O(\epsilon^{\frac{1}{3}}), \end{aligned}$$

where $F_\epsilon(t) = \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} \tilde{a}(x_\epsilon(j), y(j))$.

If $t\epsilon^{-2}$ is an integer, then $\epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} \bar{a}(\hat{x}_\epsilon(\epsilon^2 j)) = \int_0^t \bar{a}(\hat{x}_\epsilon(s))\, ds$, while in general $\left|\epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} \bar{a}(\hat{x}_\epsilon(\epsilon^2 j)) - \int_0^t \bar{a}(\hat{x}_\epsilon(s))\, ds\right|_\infty \le \epsilon^2 |a_0|_\infty$. Hence

$$\hat{x}_\epsilon(t) = \xi + \int_0^t \bar{a}(\hat{x}_\epsilon(s))\, ds + W_\epsilon(t) + F_\epsilon(t) + O(\epsilon^{\frac{1}{3}}).$$

Next, for $nM \le j < (n+1)M$, $y \in \Lambda$,

$$\begin{aligned} |\tilde{a}(x_\epsilon(j), y) - \tilde{a}(x_\epsilon(nM, y)| &\le 2 \operatorname{Lip} a_0\, |x_\epsilon(j) - x_\epsilon(nM)| \\ &\le 2 \operatorname{Lip} a_0\, (\epsilon^2 M |a_\epsilon|_\infty + \epsilon | \textstyle\sum_{i=nM}^{j-1} v(T^i y_0)|). \end{aligned}$$

By (5.2), $|\sum_{i=nM}^{j-1} v \circ T^i|_q = |\sum_{i=0}^{j-nM-1} v \circ T^i|_q \ll (j - nM)^{\frac{1}{2}} \le M^{\frac{1}{2}}$. Hence $|\tilde{a}(x_\epsilon(j), y(j)) - \tilde{a}(x_\epsilon(nM, y(j))|_q = O(\epsilon M^{\frac{1}{2}}) = O(\epsilon^{\frac{1}{3}})$ uniformly in $nM \le j < (n+1)M$. It follows that

$$\epsilon^2 \Big| \sup_{t\in[0,1]} \Big| \sum_{j=0}^{[t\epsilon^{-\frac{2}{3}}]M-1} \tilde{a}(x_\epsilon(j), y(j)) - \sum_{n=0}^{[t\epsilon^{-\frac{2}{3}}]-1} \sum_{j=nM}^{(n+1)M-1} \tilde{a}(x_\epsilon(nM), y(j)) \Big| \Big|_q = O(\epsilon^{\frac{1}{3}}).$$

Also,

$$\Big| F_\epsilon(t) - \epsilon^2 \sum_{j=0}^{[t\epsilon^{-\frac{2}{3}}][\epsilon^{-\frac{4}{3}}]-1} \tilde{a}(x_\epsilon(j), y(j)) \Big|_\infty \le \epsilon^2 (\epsilon^{-\frac{4}{3}} + \epsilon^{-\frac{2}{3}}) |\tilde{a}|_\infty \le 4\epsilon^{\frac{2}{3}} |a_0|_\infty.$$

Combining these last two estimates, we obtain $\left|\sup_{[0,1]} |F_\epsilon - D_\epsilon|\right|_q = O(\epsilon^{\frac{1}{3}})$ and the result follows. ∎

Lemma 5.5 *Let* $Q = (-32\sigma^2 \log \epsilon)^{\frac{1}{2}}$. *Then* $\mu\big(\sup_{[0,1]}|\hat{x}_\epsilon| \ge Q\big) \le C\epsilon^r$.

Proof By the reflection principle,

$$\begin{aligned}\mathbb{P}\big(\sup_{[0,1]}|W| \ge \tfrac{1}{4}Q\big) &\le 2\mathbb{P}\big(\sup_{[0,1]}W \ge \tfrac{1}{4}Q\big) = 4\mathbb{P}(W(1) \ge \tfrac{1}{4}Q) \\ &\ll \textstyle\int_{Q/4}^{\infty} e^{-x^2/(2\sigma^2)}\,dx \ll Q^{-1}e^{-Q^2/(32\sigma^2)} \le \epsilon.\end{aligned}$$

By assumption, $\pi_1(W_\epsilon, W) = O(\epsilon^r)$. In particular, for ϵ sufficiently small, $\mu\big(\sup_{[0,1]}|W_\epsilon| \ge \frac{1}{2}Q\big) \le \mathbb{P}(\sup_{[0,1]}|W| \ge \frac{1}{4}Q) + O(\epsilon^r)$.

By (5.5), $|\hat{x}_\epsilon(t) - W_\epsilon(t)| \le |\xi| + |a_0|_\infty + O(\epsilon^{\frac{1}{3}})$. Hence for ϵ sufficiently small,

$$\mu\big(\sup_{[0,1]}|\hat{x}_\epsilon| \ge Q\big) \le \mu\big(\sup_{[0,1]}|W_\epsilon| \ge \tfrac{1}{2}Q\big) \le \mathbb{P}\big(\sup_{[0,1]}|W| \ge \tfrac{1}{4}Q\big) + O(\epsilon^r) = O(\epsilon^r),$$

as required. ∎

Let $B_\epsilon = \{\sup_{[0,1]}|\hat{x}_\epsilon| \le Q\}$ where $Q = Q_\epsilon$ is as in Lemma 5.5.

Lemma 5.6 $\big|1_{B_\epsilon}\sup_{[0,1]}|D_\epsilon|\big|_q \le C\epsilon^{\frac{1}{3}}(-\log\epsilon)^{\frac{1}{4}}$.

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

$$\tilde{J}_\epsilon(n,u)(y_0) = \epsilon^{\frac{4}{3}} \sum_{j=nM}^{(n+1)M-1} \tilde{a}(u, y(j)) = \epsilon^{\frac{4}{3}} \sum_{j=nM}^{(n+1)M-1} \tilde{a}_u(T^j y_0),$$

where $\tilde{a}_u(y) = \tilde{a}(u,y)$. Note that $\tilde{J}_\epsilon(n,u) = \tilde{J}_\epsilon(0,u) \circ T^{nM}$. By assumption (5.1),

$$|\tilde{J}_\epsilon(n,u)|_q = |\tilde{J}_\epsilon(0,u)|_q \ll \epsilon^{\frac{4}{3}} M^{\frac{1}{2}} \le \epsilon^{\frac{2}{3}},$$

uniformly in n and u.

We can choose a partition $S = S_\epsilon \subset [-Q,Q]$ of finite cardinality $|S|$ such that $\operatorname{dist}(x,S) \le 2Q/|S|$ for all $x \in [-Q,Q]$. For $x \in [-Q,Q]$, there exists $u_x \in S$ such that for all $y \in \Lambda$,

$$|\tilde{a}(x,y) - \tilde{a}(u_x,y)| \le 2\operatorname{Lip} a_0\, 2Q/|S| = 4\operatorname{Lip} a_0\, Q/|S|.$$

It follows that

$$1_{B_\epsilon}|J_\epsilon(n) - \tilde{J}_\epsilon(n, u_{x_\epsilon(nM)})| \le 4\operatorname{Lip} a_0\, Q/|S|,$$

and hence

$$1_{B_\epsilon}|J_\epsilon(n)| \le \textstyle\sum_{u\in S}|\tilde{J}_\epsilon(n,u)| + 4\operatorname{Lip} a_0\, Q/|S|.$$

Choosing $|S| = [\epsilon^{-\frac{1}{3}}(-\log\epsilon)^{\frac{1}{4}}]$,

$$\begin{aligned}\big|1_{B_\epsilon}|J_\epsilon(n)|\big|_q &\le \textstyle\sum_{u\in S}|\tilde{J}_\epsilon(n,u)|_q + 4\operatorname{Lip} a_0\, Q/|S| \\ &\ll |S|\epsilon^{\frac{2}{3}} + (-\log\epsilon)^{\frac{1}{2}}/|S| \ll \epsilon^{\frac{1}{3}}(-\log\epsilon)^{\frac{1}{4}}.\end{aligned}$$

It follows that

$$\big|1_{B_\epsilon}\sup_{[0,1]}|D_\epsilon|\big|_q \le \epsilon^{\frac{2}{3}}\sum_{n=0}^{[\epsilon^{-\frac{2}{3}}]-1}\big|1_{B_\epsilon}|J_\epsilon(n)|\big|_q \ll \epsilon^{\frac{1}{3}}(-\log\epsilon)^{\frac{1}{4}},$$

as required. ∎

Corollary 5.7 $\pi_1(W_\epsilon + D_\epsilon + E_\epsilon, W) \le C(\epsilon^r + \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}})$.

Proof By Proposition 5.4 and Lemma 5.6, $\big|\sup_{[0,1]}|1_{B_\epsilon}D_\epsilon + E_\epsilon|\big|_q \ll \epsilon^{\frac{1}{3}}(-\log\epsilon)^{\frac{1}{4}}$. Hence by Markov's inequality,

$$\mu(\sup_{[0,1]}|1_{B_\epsilon}D_\epsilon + E_\epsilon| > \epsilon^{\frac{1}{3}\frac{q}{q+1}}) \ll \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}}.$$

Combining this with Lemma 5.5, we obtain

$$\mu(\sup_{[0,1]}|D_\epsilon + E_\epsilon| > \epsilon^{\frac{1}{3}\frac{q}{q+1}}) \ll \epsilon^r + \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}}.$$

By Proposition 4.5(a),

$$\pi_1(W_\epsilon + D_\epsilon + E_\epsilon, W_\epsilon) \ll \epsilon^r + \epsilon^{\frac{1}{3}\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}}.$$

Now combine this estimate with the assumption $\pi_1(W_\epsilon, W) = O(\epsilon^r)$. ∎

Proof of Theorem 5.3 Consider the functional $\mathcal{G} : C[0,1] \to [0,1]$ given by $\mathcal{G}(u) = v$ where v is the unique solution to $v(t) = \xi + \int_0^t \bar{a}(v(s))\,ds + u(t)$. Since $\bar{a}$ is globally Lipschitz, it follows from existence and uniqueness of solutions for ordinary differential equations that $\mathcal{G}$ is well-defined. By Gronwall's inequality, $\mathcal{G}$ is Lipschitz with $\operatorname{Lip}\mathcal{G} \le e^{\operatorname{Lip}\bar{a}}$.

By definition, $X(t) = \xi + \int_0^t \bar{a}(X(s))\,ds + W(t)$, so $X = \mathcal{G}(W)$. By Proposition 5.4, $\hat{x}_\epsilon = \mathcal{G}(W_\epsilon + D_\epsilon + E_\epsilon)$. Hence by the Lipschitz mapping theorem [47, Theorem 3.2],

$$\pi_1(\hat{x}_\epsilon, X) = \pi_1(\mathcal{G}(W_\epsilon + D_\epsilon + E_\epsilon), \mathcal{G}(W)) \le e^{\operatorname{Lip}\bar{a}}\pi_1(W_\epsilon + D_\epsilon + E_\epsilon, W).$$

Now apply Corollary 5.7. ∎

5.2 The general case

Let $b = 1/\psi'$ and write $z_\epsilon(n) = \psi(x_\epsilon(n))$, $\hat{z}_\epsilon(t) = \psi(\hat{x}_\epsilon(t))$. Define

$$\bar{A}(z) = \psi'(\psi^{-1}(z))\bar{a}(\psi^{-1}(z)) + \frac{1}{2}\psi''(\psi^{-1}(z))b(\psi^{-1}(z))^2\int_\Lambda v^2\,d\mu.$$

Lemma 5.8 $\pi_1(\hat z_\epsilon, Z) \le C(\epsilon^r + \epsilon^{\frac13\frac{q}{q+1}}(-\log\epsilon)^{\frac{q}{4}})$ *where* $dZ = \bar A(Z)\,dt + dW$ *and* $Z(0) = \psi(\xi)$.

Proof The assumptions on b ensure that ψ is C^3 uniformly on $\mathbb{R}$. A calculation as in [18, 30], using the Taylor expansion of ψ, yields

$$z_\epsilon(n+1) = z_\epsilon(n) + \epsilon^2 A_\epsilon(z_\epsilon(n), y(n)) + \epsilon v(y(n)), \quad z_\epsilon(0) = \psi(\xi),$$

where

$$A_\epsilon(z,y) = \psi'(\psi^{-1}(z))a_\epsilon(\psi^{-1}(z),y) + \tfrac12\psi''(\psi^{-1}(z))b(\psi^{-1}(z))^2 v(y)^2 + O(\epsilon),$$

uniformly in z, y. By the inverse function theorem, $(\psi^{-1})' = b\circ\psi^{-1} \in L^\infty$ and hence ψ^{-1} is uniformly Lipschitz on $\mathbb{R}$. It follows easily that A_ϵ inherits the regularity conditions (i)–(iii) from a. By (5.3), A_0 inherits condition (5.1) from a_0. Hence we can apply Theorem 5.3. ∎

Proof of Theorem 5.1 As in [18], a calculation using the definition of X in (2.2) shows that $Z = \psi(X)$ satisfies the SDE in Lemma 5.8. The functional $\chi : C[0,1] \to C[0,1]$, $u \mapsto \psi^{-1}\circ u$, is Lipschitz with $\operatorname{Lip}\chi = \operatorname{Lip}\psi^{-1}$, so by the Lipschitz mapping theorem,

$$\pi_1(\hat x_\epsilon, X) = \pi_1(\chi(\hat z_\epsilon), \chi(Z)) \ll \pi_1(\hat z_\epsilon, Z).$$

Now apply Lemma 5.8. ∎

6 Nonuniformly hyperbolic transformations

In this section, we show how the main results in Section 2 extend from nonuniformly expanding maps to transformations that are nonuniformly hyperbolic in the sense of Young [49, 50]. We focus on the parts necessary for this paper, referring to [49, 50] for further details. (In particular, we do not restrict to systems with physical measures even though this is the case for most of the examples.)

Let (Λ, d) be a bounded metric space with Borel probability measure μ and let $T:\Lambda\to\Lambda$ be an ergodic measure-preserving transformation.

Definition 6.1 Let $p \ge 1$, $\eta\in(0,1]$. The transformation $T:\Lambda\to\Lambda$ is a *nonuniformly hyperbolic transformation of order* p if there exists a nonuniformly expanding map $\bar f:\bar\Delta\to\bar\Delta$ of order p (with ergodic invariant probability measure $\bar\mu$) such that

(a) $T:\Lambda\to\Lambda$ and $\bar f:\bar\Delta\to\bar\Delta$ have a common extension $f:\Delta\to\Delta$ with semiconjugacies $\pi_\Delta:\Delta\to\Lambda$, $\bar\pi:\Delta\to\bar\Delta$,

(b) There exists $\eta'\in(0,1]$ such that for any $v\in C^\eta(\Lambda)$, there exists $\bar v\in C^{\eta'}(\bar\Delta)$ and $\chi\in L^\infty(\Delta)$ such that

$$v\circ\pi_\Delta = \bar v\circ\bar\pi + \psi\circ f - \psi.$$

(c) There is a constant $C > 0$ such that for all $v \in C^\eta(\Lambda)$,

$$\|\bar{v}\|_{\eta'} \le C\|v\|_\eta, \quad \text{and} \quad |\psi|_\infty \le C\|v\|_\eta.$$

Remark 6.2 (a) By [30, Propositions 5.3 and 5.4], nonuniformly hyperbolic transformations modelled by Young towers with exponential tails [49] are nonuniformly hyperbolic of order p (in the sense of Definition 6.1) for all p. Also, transformations that are modelled by Young tails with polynomial tails [50] are nonuniformly hyperbolic of order p provided (i) the inducing time in [50] lies in L^p, and (ii) the transformation T contracts exponentially along stable leaves (see [30, Remark 5.1]).

(b) By [39], the conclusions in Lemma 2.1 hold for transformations that are nonuniformly hyperbolic of order $p \ge 2$ and observables $v \in C_0^\eta(\Lambda)$. This result does not require condition (ii) from part (a) of this remark.

Now define $W_n, W \in C[0,1]$ as in Section 2.

Theorem 6.3 *Let $T : \Lambda \to \Lambda$ be nonuniformly hyperbolic of order $p > 2$ and suppose that $v \in C_0^\eta(\Lambda)$. Then there is a constant $C > 0$ such that $\pi_1(W_n, W) \le Cn^{-r(p)}$ for all $n \ge 1$, where $r(p) = \frac{p-2}{4p}$.*

Proof Define $\overline{W}_n$ using the observable $\bar{v}$ from Definition 6.1(b). By Definition 6.1(b), $|W_n \circ \pi_\Delta - \overline{W}_n \circ \bar{\pi}|_\infty \le 2n^{-1/2}|\psi|_\infty$. Hence by Proposition 4.5(a), $\pi_1(W_n \circ \pi_\Delta, \overline{W}_n \circ \bar{\pi}) \ll n^{-(\frac{1}{2}-\delta)}$ for all $\delta > 0$. In particular, $\pi_1(W_n \circ \pi_\Delta, \overline{W}_n \circ \bar{\pi}) \ll n^{-r(p)}$. Since $\pi_\Delta : \Delta \to \Lambda$ and $\bar{\pi} : \Delta \to \bar{\Delta}$ are semiconjugacies, $\pi_1(W_n, \overline{W}_n) \ll n^{-r(p)}$.

Since $\bar{f}$ is nonuniformly expanding of order p and $\bar{v} \in C_0^{\eta'}(\bar{\Delta})$, it follows from Theorem 2.2 that $\pi_1(\overline{W}_n, W) \ll n^{-r(p)}$. This completes the proof. ∎

Next, consider a fast-slow system (1.2) satisfying the regularity conditions in Section 2.2. Define $\hat{x}_\epsilon$, $X \in C[0,1]$ as in Section 2.2.

Theorem 6.4 *Let $T : \Lambda \to \Lambda$ be nonuniformly hyperbolic of order $p > 2$. Suppose further that $v \in C_0^\eta(\Lambda)$ and that $\sup_{x\in\mathbb{R}} |a_0(x,\cdot)|_\eta < \infty$. Then there is a constant $C > 0$ such that*

$$\pi_1(\hat{x}_\epsilon, X) \le \begin{cases} C\epsilon^{\frac{p-2}{2p}} & p \le p_* \\ C\epsilon^{\frac{1}{3}\frac{2p-2}{2p-1}}(-\log\epsilon)^{\frac{1}{2}(p-1)} & p > p_* \end{cases},$$

where $p_* = \frac{1}{4}(11 + \sqrt{73}) \approx 4.89$.

Proof As in the proof of Theorem 2.4, it suffices to verify the hypotheses of Theorem 5.1. By Theorem 6.3, $\pi_1(W_\epsilon, W) = O(\epsilon^{\frac{p-2}{2p}})$.

Next, v and $v^2 - \int_\Lambda v^2\, d\mu$ lie in $C_0^\eta(\Lambda)$ so conditions (5.2) and (5.3) follow from Proposition 3.4 and Definition 6.1(b,c). Also, $\tilde{a}_u \in C_0^\eta(\Lambda)$ uniformly in $u \in \mathbb{R}$, so (5.1) follows from Proposition 3.4 and Definition 6.1(b,c). ∎

7 Examples

In this section, we list some examples to which our results apply. The simplest class of examples are uniformly expanding maps, which are nonuniformly expanding with $\tau = 1$. A specific example is the Gauss map $T : [0,1] \to [0,1]$ given by $Tx = \frac{1}{x} - [\frac{1}{x}]$. Our results hold for all p, so we obtain rates $n^{-(\frac{1}{4}-\delta)}$ in the WIP and $\epsilon^{-(\frac{1}{3}-\delta)}$ for homogenization.

Similarly for uniformly hyperbolic maps (including nontrivial basic sets for Axiom A diffeomorphism) we can take $\tau = 1$ and p arbitrarily large.

More generally, the rates $n^{-(\frac{1}{4}-\delta)}$ in the WIP and $\epsilon^{-(\frac{1}{3}-\delta)}$ for homogenization hold provided T is nonuniformly expanding/hyperbolic and $\tau \in L^p$ for all p. In particular, this covers all systems that are modelled by Young towers with exponential tails [49], including:

- Planar periodic dispersing billiards with finite horizon [49] and infinite horizon [10], as well as billiards with external forcing and corners [10, 13].
- Unimodal maps $T : [-1,1] \to [-1,1]$ given by $Tx = 1 - ax^2$, $a \in [0,2]$ satisfying the Collet-Eckmann condition [12], namely there are constants $b, c > 0$ such that $|(T^n)'(1)| \geq ce^{bn}$ for all $n \geq 1$. By [28, 6], this condition holds for a positive Lebesgue measure set of parameters a.
- Hénon like attractors. The Hénon map $T : \mathbb{R}^2 \to \mathbb{R}^2$ introduced in [26] is given by $(x,y) = (1 - ax^2 + y, bx)$ where $a, b \in \mathbb{R}$. By [7, 8], there is a positive measure set of parameters $a < 2$, b small, such that T is modelled by a Young tower with exponential tails.

Other class of examples for which these rates hold are Viana maps. These maps, introduced in [46], comprise a C^3 open class of multidimensional nonuniformly expanding maps. For definiteness, we restrict attention to maps on $M = S^1 \times \mathbb{R}$. Let $T_0 : M \to M$ be the map $T_0(\theta, y) = (\lambda\theta \bmod 1, a_0 + a \sin 2\pi\theta - y^2)$, where $\lambda \in \mathbb{N}$ with $\lambda \geq 16$, a_0 is chosen so that 0 is a preperiodic point for the quadratic map $y \mapsto a_0 - y^2$, and a is sufficiently small. It follows from [1, 3] that C^3 maps sufficiently close to T_0 are nonuniformly expanding of order p for all p. (In fact, they are modelled by Young towers with stretched exponential tails [21].)

Finally, we mention that the intermittent maps (1.1), with parameter $\gamma \in (0, \frac{1}{2})$, are nonuniformly expanding of order p for any $p < \frac{1}{\gamma}$, yielding the rates

$$\pi_1(W_n, W) = O(n^{-(\frac{1}{4}(1-2\gamma)-\delta)}), \qquad \pi_1(\hat{x}_\epsilon, X) \leq \begin{cases} C\epsilon^{\frac{1}{2}(1-2\gamma)-\delta} & \gamma \geq \gamma_* \\ C\epsilon^{\frac{1}{3}\frac{2-2\gamma}{2-\gamma}-\delta} & \gamma \leq \gamma_* \end{cases},$$

where $\gamma_* = \frac{1}{12}(11 - \sqrt{73})$. The classical solenoid construction of Smale and Williams [44, 48] can be used as in [2] to construct nonuniformly hyperbolic intermittent solenoids with these rates.

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

References

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

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

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

[4] M. Antoniou. Rates of convergence for statistical limit laws in deterministic systems. Ph. D. Thesis. University of Warwick, 2018.

[5] P. Bálint and I. Melbourne. Statistical properties for flows with unbounded roof function, including the Lorenz attractor. *J. Stat. Phys.* **172** (2018) 1101–1126.

[6] M. Benedicks and L. Carleson. On iterations of $1-ax^2$ on (-1.1). *Ann. of Math.* **122** (1985) 1–25.

[7] M. Benedicks and L. Carleson. The dynamics of the Hénon map. *Ann. of Math.* **133** (1991) 73–169.

[8] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Astérisque* (2000) no. **261**, 13–56.

[9] A. A. Borovkov. The rate of convergence in the invariance principle. *Probab. Theory Appl.* **18** (1973) 217–234.

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

[11] Z. Coehlo and W. Parry. Central limit asymptotics for subshifts of finite type. *Israel J. Math.* **69** (1990) 235–249.

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

[13] J. De Simoi and I. P. Tóth. An expansion estimate for dispersing planar billiards with corner points. *Ann. Henri Poincaré* **15** (2014) 1223–1243.

[14] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems* **4** (1984) 541–552.

[15] M. D. Donsker. An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.,* **No. 6** (1951).

[16] D. Givon and R. Kupferman. White noise limits for discrete dynamical systems driven by fast deterministic dynamics. *Phys. A* **335** (2004) 385–412.

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

[18] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. London A* **469** (2013) 20130201.

[19] S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128** (2004) 82–122.

[20] S. Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 997–1024.

[21] S. Gouëzel. Decay of correlations for nonuniformly expanding systems. *Bull. Soc. Math. France* **134** (2006) 1–31.

[22] S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* **38** (2010) 1639–1671.

[23] I. Grama, E. Le Page and M. Peigné. On the rate of convergence in the weak invariance principle for dependent random variables with applications to Markov chains. *Colloq. Math.* **134** (2014) 1–55.

[24] E. Haeusler. An exact rate of convergence in the functional central limit theorem for special martingale difference arrays. *Z. Wahrsch. Verw. Gebiete* **65** (1984) 523–534.

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

[26] M. Hénon. A two-dimensional mapping with a strange attractor. *Comm. Math. Phys.* **50** (1976) 69–77.

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

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

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

[30] A. Korepanov, Z. Kosloff and I. Melbourne. Martingale-coboundary decomposition for families of dynamical systems. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **35** (2018) 859–885.

[31] K. Kubilius. The rate of convergence in the invariance principle for martingale differences. *Liet. Mat. Rink.* **34** (1994) 482–494.

[32] R. Kupferman, G. A. Pavliotis and A. M. Stuart. Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Phys. Rev. E (3)* **70** (2004) 036120, 9.

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

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

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

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

[37] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity* **24** (2011) 1361–1367.

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

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

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

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

[42] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.* **16** (1973) 181–197.

[43] S. Sawyer. Rates of convergence for some functionals in probability. *Ann. Math. Statist.* **43** (1972) 273–284.

[44] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967) 747–817.

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

[46] M. Viana. Multidimensional nonhyperbolic attractors. *Inst. Hautes Études Sci. Publ. Math.* **85** (1997) 63–96.

[47] W. Whitt. Preservation of rates of convergence under mappings. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **29** (1974) 39–44.

[48] R. F. Williams. One-dimensional non-wandering sets. *Topology* **6** (1967) 473–487.

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

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