Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2022, Vol. 58, No. 2, 1284–1304
https://doi.org/10.1214/21-AIHP1190

Iterated invariance principle for slowly mixing dynamical systems

Matt Galton[a] and Ian Melbourne[b]

Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK. [a]*Matt.Galton@outlook.com,* [b]*i.melbourne@warwick.ac.uk*

Received 28 April 2020; revised 4 March 2021; accepted 12 May 2021

Abstract. We give sufficient Gordin-type criteria for the iterated (enhanced) weak invariance principle to hold for deterministic dynamical systems. Such an invariance principle is intrinsically related to the interpretation of stochastic integrals. We illustrate this with examples of deterministic fast-slow systems where our iterated invariance principle yields convergence to a stochastic differential equation.

Résumé. Nous donnons des critères suffisants de type Gordin assurant qu'un principe d'invariance itéré (renforcé) faible est satisfait pour des systèmes dynamiques déterministes. Un tel principe d'invariance est intrinsèquement relié à l'interprétation des intégrales stochastiques. Nous illustrons ceci par des exemples de systèmes lents-rapides déterministes où notre principe d'invariance itéré implique la convergence vers une équation différentielle stochastique.

MSC2020 subject classifications: Primary 60F17; 37D25; secondary 37A50; 60H05

Keywords: Iterated weak invariance principle; Gordin criterion; Stochastic integrals

1. Introduction

Recently, there has been a great deal of interest in homogenisation of deterministic systems with multiple timescales [3,4,8,9,16,23,24,26,27,38]; the aim is to prove convergence to a stochastic differential equation (SDE) as the separation of timescales increases. The papers [16,38] considered some simplified situations where it sufficed that the fast dynamics satisfies the weak invariance principle (WIP). In general, however, there are issues regarding the correct interpretation of stochastic integrals (Itô, Stratonovich, ...) in the limiting SDE that are not resolved by the WIP. According to rough path theory [10,11,31], it is necessary to consider an iterated (or enhanced) WIP in order to determine the stochastic integrals. Kelly & Melbourne [23,24] applied rough path theory in the deterministic setting and reduced homogenisation theorems to establishing the iterated WIP and suitable moment control. The conditions on moments were optimized in Chevyrev *et al.* [3,4].

The current paper is based on results of the first author in his Ph. D. thesis [12] and aims to extend the class of dynamical systems for which the iterated WIP holds. There is already a wealth of literature on the central limit theorem (CLT) and WIP for large classes of dynamical systems in both the dynamical systems and probability theory literature [5,13,18,19,22,29,36,46,47]. We slightly extend the class of systems for which the WIP holds, and greatly extend the class of systems for which the iterated WIP holds.

Our approach is based on Gordin's method [13] for proving limit theorems via martingale approximation. It is well-known that the L^2-criterion of Gordin [13] leads to the CLT and WIP, and it follows from [23] that the iterated WIP holds under this criterion (see [6, Proposition 2.5]). Proving the same results under the L^1 version of this criterion (hypotheses (2.1) and (3.1) in this paper) is more delicate. The CLT was obtained by [14] and much later the WIP was obtained by [5]. The WIP in [5] is not quite in the right form for dynamical systems; in this paper we modify it so that it applies to dynamical systems by extending a time-reversal argument from [23].

Previously there were no results on the iterated WIP under L^p Gordin criteria for $p < 2$ (except where there is additional Young tower structure, see [40] and [23, Section 10]). Addressing this is the main aim of this work. In the noninvertible setting (Section 2), we prove the iterated WIP under the L^1 Gordin criterion. In the invertible setting (Section 3), the validity of the iterated WIP under the L^1 Gordin criterion remains unresolved. However, we prove the iterated WIP under a hybrid L^1–L^2 criterion (3.2) which is still a significant improvement on existing results.

The remainder of the paper is organised as follows. In Section 2, we present our main results in the noninvertible setting. Section 3 deals with the invertible setting. In Section 4, we consider some illustrative examples and in Section 5 we give an application to homogenisation of fast-slow systems.

Notation. For $a, b \in \mathbb{R}^d$, we define the outer product $a \otimes b = ab^T \in \mathbb{R}^{d\times d}$. For $J \in \mathbb{R}^{d\times d}$, we write $|J| = (\sum_{i,j=1}^d J_{ij}^2)^{1/2}$.

For real-valued functions f, g, the integral $\int f\,dg$ denotes the Itô integral (where defined). Similarly, for $\mathbb{R}^d$-valued functions, $\int f \otimes dg$ denotes matrices of Itô integrals.

We use "big O" and $\ll$ notation interchangeably, writing $a_n = O(b_n)$ or $a_n \ll b_n$ if there are constants $C > 0$, $n_0 \ge 1$ such that $a_n \le Cb_n$ for all $n \ge n_0$.

2. Noninvertible setting

Let $(\Lambda, \mathcal{F}, \mu)$ be a probability space and $T : \Lambda \to \Lambda$ be an ergodic measure-preserving map. Let $P : L^1 \to L^1$ be the associated transfer operator (so $\int_\Lambda Pv\,w\,d\mu = \int_\Lambda v\,w\circ T\,d\mu$ for $v \in L^1$, $w \in L^\infty$). Also define the Koopman operator $Uv = v \circ T$. We recall that

$$PU = I \quad \text{and} \quad UP = \mathbb{E}\big(\cdot|T^{-1}\mathcal{F}\big).$$

Let $v \in L^\infty(\Lambda, \mathbb{R}^d)$ with $\int_\Lambda v\,d\mu = 0$. Our underlying hypothesis throughout this section is the L^1 Gordin criterion

$$\sum_{n=1}^\infty \big|P^n v\big|_1 < \infty. \tag{2.1}$$

Under this hypothesis, it is well-known that the CLT and WIP hold. We mention [5,29] for this and related results. Our aim is to prove the iterated (or enhanced) version of the WIP. Previously, this was proved in [23, Section 4] under the more restrictive assumption $\sum_{n=1}^\infty |P^n v|_2 < \infty$.

Define the sequences of càdlàg processes

$$W_n \in D\big([0,\infty), \mathbb{R}^d\big), \qquad \mathbb{W}_n \in D\big([0,\infty), \mathbb{R}^{d\times d}\big),$$

by

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{0\le j\le [nt]-1} v\circ T^j, \qquad \mathbb{W}_n(t) = \frac{1}{n} \sum_{0\le i<j\le [nt]-1} \big(v\circ T^i\big)\otimes\big(v\circ T^j\big). \tag{2.2}$$

Theorem 2.1. *Let $v \in L^\infty(\Lambda, \mathbb{R}^d)$ with $\int_\Lambda v\,d\mu = 0$, and suppose that (2.1) holds. Then*

(a) *The limit $\Sigma = \lim_{n\to\infty} \int_\Lambda W_n(1)\otimes W_n(1)\,d\mu \in \mathbb{R}^{d\times d}$ exists.*
(b) *$\det\Sigma = 0$ if and only if there exists $c \in \mathbb{R}^d$ nonzero and $h \in L^1$ such that $c\cdot v = h\circ T - h$.*
(c) *Let ν be any probability measure on Λ absolutely continuous with respect to μ and regard $(W_n, \mathbb{W}_n)$ as a sequence of processes in $D([0,\infty), \mathbb{R}^d \times \mathbb{R}^{d\times d})$ on the probability space (Λ, ν).*

Then $(W_n, \mathbb{W}_n) \to_w (W, \mathbb{W})$ as $n \to \infty$, where W is a d-dimensional Brownian motion with covariance Σ and

$$\mathbb{W}(t) = \int_0^t W\otimes dW + t\sum_{j=1}^\infty \int_\Lambda v\otimes\big(v\circ T^j\big)\,d\mu.$$

Remark 2.2. A standard calculation using (2.1) and Theorem 2.1(a) yields the Green-Kubo formula

$$\Sigma = \int_\Lambda v\otimes v\,d\mu + \sum_{j=1}^\infty \int_\Lambda \big\{v\otimes\big(v\circ T^j\big) + \big(v\circ T^j\big)\otimes v\big\}\,d\mu.$$

Remark 2.3. The assumption that T is noninvertible is not assumed explicitly in Theorem 2.1, but hypothesis (2.1) implies that $v \equiv 0$ when T is invertible.

Remark 2.4. There are various possible extensions to Theorem 2.1:

(1) Let $1 \le p, q \le \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Dedecker & Rio [5] consider unbounded functions $v \in L^p(\Lambda, \mathbb{R})$ and prove that the ordinary WIP $W_n \to_w W$ holds provided the 1-norm in (2.1) is replaced by the q-norm. (In fact it suffices that $\sum_{n=1}^{\infty} P^n v$ converges in L^q in [5].) A natural question is to prove the iterated WIP $(W_n, \mathbb{W}_n) \to_w (W, \mathbb{W})$ under such assumptions. However, the main motivation for studying the iterated WIP is its fundamental role in the theory of fast-slow systems (considered further in Section 5) where it is standard to consider bounded v. Also, considering unbounded v would exacerbate the issues regarding hypotheses (3.1) and (3.2) in the invertible setting. Hence we restrict in this paper to the case of bounded v.

(2) Dedecker & Rio [5] prove a nonergodic version of the WIP following Volný [47]. It seems likely that a nonergodic version of the iterated WIP holds (paying due attention to the limit of $\mathbb{W}_n(1) - \mathbb{M}_n(1)$ in the proof of Theorem 2.1(c)). Again, ergodicity of μ is a standard assumption in the motivating setting of fast-slow systems, and is assumed throughout this paper.

(3) A third possible extension is to consider limits of $(W_n(s), \mathbb{W}_n(t))$ in the space $D([0,\infty) \times [0,\infty), \mathbb{R}^d \times \mathbb{R}^{d\times d})$. This seems to involve a nontrivial extension of [21,28] and hence is beyond the scope of this paper.

Hypothesis (2.1) can be viewed as a slow mixing condition: we recall the following elementary result.

Proposition 2.5. *Let $v \in L^\infty(\Lambda, \mathbb{R})$ with $\int_\Lambda v\,d\mu = 0$. Suppose that there exists $a_n > 0$ such that*

$$\left|\int_\Lambda v\, w \circ T^n\, d\mu\right| \le a_n |w|_\infty \quad \text{for all } w \in L^\infty(\Lambda, \mathbb{R}),\ n \ge 1.$$

Then $|P^n v|_p \le |v|_\infty^{1-1/p} a_n^{1/p}$ for all $1 \le p < \infty$. In particular, hypothesis (2.1) holds if $\sum_{n=1}^{\infty} a_n < \infty$.

Proof. See for example [39, Proposition 2.1]. □

Throughout the remainder of this section, L^p is shorthand for $L^p((\Lambda,\mu), \mathbb{R}^d)$ unless stated otherwise.

2.1. *Martingales*

Let $v : \Lambda \to \mathbb{R}^d$ be an L^∞ observable with mean zero satisfying hypothesis (2.1), and define

$$\chi_\ell^k = \sum_{j=\ell}^{k} P^j v, \quad 1 \le \ell \le k < \infty \qquad \text{and} \qquad \chi = \sum_{j=1}^{\infty} P^j v.$$

It follows from our assumptions that $\chi_\ell^k \in L^\infty$ for all $\ell \le k$ and $\chi \in L^1$. Moreover, $\chi_1^k \to \chi$ in L^1 as $k \to \infty$. Following [34], we write

$$v = m^{(k)} + \chi_1^k \circ T - \chi_1^k + P^k v, \quad k \ge 1 \qquad \text{and} \qquad v = m + \chi \circ T - \chi. \tag{2.3}$$

Since $PU = I$, it is easily verified from the definitions in (2.3) that m, $m^{(k)} \in \ker P$ for all k. It is immediate that $m^{(k)} \in L^\infty$ for all k, that $m \in L^1$ and that $m^{(k)} \to m$ in L^1. A somewhat surprising fact due originally to [25], see also [5, 15,29,32,46], is that $m \in L^2$. We begin by recovering this fact using an elementary argument.

Lemma 2.6. *$m \in L^2$ and $m^{(k)} \to m$ in L^2 as $k \to \infty$.*

Proof. Working componentwise, we can suppose without loss that $d = 1$. For $\ell < k$,

$$\begin{aligned} m^{(k)} - m^{(\ell)} &= \big(\chi_1^k - \chi_1^k \circ T - P^k v\big) - \big(\chi_1^\ell - \chi_1^\ell \circ T - P^\ell v\big) \\ &= \chi_{\ell+1}^k - \chi_{\ell+1}^k \circ T + P^\ell v - P^k v = \chi_\ell^{k-1} - \chi_{\ell+1}^k \circ T. \end{aligned} \tag{2.4}$$

Hence

$$\big|m^{(k)} - m^{(\ell)}\big|_2^2 = \int_\Lambda \big(m^{(k)} - m^{(\ell)}\big)\big(\chi_\ell^{k-1} - \chi_{\ell+1}^k \circ T\big)\,d\mu = \int_\Lambda \big(m^{(k)} - m^{(\ell)}\big)\chi_\ell^{k-1}\,d\mu,$$

where we used that $m^{(k)}$, $m^{(\ell)} \in \ker P$. Continuing and using (2.4) once more,

$$\begin{aligned}
|m^{(k)} - m^{(\ell)}|_2^2 &= \int_\Lambda (\chi_\ell^{k-1} - \chi_{\ell+1}^k \circ T)\chi_\ell^{k-1}\,d\mu = \int_\Lambda \{(\chi_\ell^{k-1})^2 - \chi_{\ell+1}^k P\chi_\ell^{k-1}\}\,d\mu \\
&= \int_\Lambda \{(\chi_\ell^{k-1})^2 - (\chi_{\ell+1}^k)^2\}\,d\mu = \int_\Lambda (\chi_\ell^{k-1} - \chi_{\ell+1}^k)(\chi_\ell^{k-1} + \chi_{\ell+1}^k)\,d\mu \\
&= \int_\Lambda (P^\ell v - P^k v)(\chi_\ell^{k-1} + \chi_{\ell+1}^k)\,d\mu \\
&\le (|P^\ell v|_\infty + |P^k v|_\infty)(|\chi_\ell^{k-1}|_1 + |\chi_{\ell+1}^k|_1) \le 4|v|_\infty \sum_{n=\ell}^\infty |P^n v|_1.
\end{aligned}$$

It follows from (2.1) that $m^{(k)}$ is Cauchy in L^2. By uniqueness of limits in L^1, the L^2 limit of $m^{(k)}$ coincides with m. □

Elements of $\ker P$ enjoy the following martingale structure.

Proposition 2.7. *Let $\phi \in L^1 \cap \ker P$ and fix $n \ge 1$. Define $\mathcal{G}_j = T^{-(n-j)}\mathcal{F}$, $1 \le j \le n$. Then $\{\phi \circ T^{n-j}, \mathcal{G}_j;\ 1 \le j \le n\}$ is a sequence of* martingale differences. *That is, $\mathcal{G}_1 \subset \cdots \subset \mathcal{G}_n$, $\phi \circ T^{n-j}$ is $\mathcal{G}_j$-measurable for each j, and $\mathbb{E}(\phi \circ T^{n-j}|\mathcal{G}_{j-1}) = 0$ for each j.*

Proof. Since $T^{-1}\mathcal{F} \subset \mathcal{F}$, it follows that $\mathcal{G}_j \subset \mathcal{G}_{j+1}$. Measurability of $\phi \circ T^{n-j}$ with respect to $\mathcal{G}_j$ is clear. Finally,

$$\mathbb{E}\big(\phi \circ T^{n-j}|\mathcal{G}_{j-1}\big) = \mathbb{E}\big(\phi|T^{-1}\mathcal{F}\big) \circ T^{n-j} = (UP\phi) \circ T^{n-j} = 0,$$

since $\phi \in \ker P$. □

2.2. *Second moments*

Throughout, we write $v_n = \sum_{j=0}^{n-1} v \circ T^j$, $m_n = \sum_{j=0}^{n-1} m \circ T^j$ and so on for observables v, m, ... defined on Λ.

Corollary 2.8. *Let $\phi \in L^2 \cap \ker P$. Then $|\max_{1\le\ell\le n} |\phi_\ell||_2 \le 4\sqrt{n}|\phi|_2$ for all $n \ge 1$. In particular, $|\max_{1\le\ell\le n} |(m - m^{(k)})_\ell||_2 \le 4\sqrt{n}|m - m^{(k)}|_2$ for all $k, n \ge 1$.*

Proof. Fix $n \ge 1$ and let $X(j) = \phi \circ T^{n-j}$. Since $\phi \in \ker P$, it follows that $|X(1) + \cdots + X(n)|_2 = \sqrt{n}|\phi|_2$. By Proposition 2.7, $\{X(j), \mathcal{G}_j;\ 1 \le j \le n\}$ is a sequence of martingale differences. Hence by Doob's inequality,

$$\Big|\max_{1\le\ell\le n} \big|X(1) + \cdots + X(\ell)\big|\Big|_2 \le 2\big|X(1) + \cdots + X(n)\big|_2 = 2\sqrt{n}|\phi|_2.$$

Finally, $\max_{1\le\ell\le n} |\phi_\ell| \le 2\max_{1\le\ell\le n} |X(1) + \cdots + X(\ell)|$. □

Following [37], we have a similar estimate for v_n.

Proposition 2.9. $|\max_{1\le\ell\le n} |v_\ell||_2^2 \le 128\, n|v|_\infty \sum_{j=0}^\infty |P^j v|_1$.

Proof. Fix $n \ge 1$ and define the random variables $X(j) = v \circ T^{n-j}$, $1 \le j \le n$ which are adapted to the filtration $\mathcal{G}_j = T^{-(n-j)}\mathcal{F}$. The version of Rio's inequality [44] for $p = 2$ in [41, Proposition 7] states that

$$\Big|\max_{1\le\ell\le n} \big|X(1) + \cdots + X(\ell)\big|\Big|_2^2 \le 16 \sum_{j=1}^n b_{j,n}$$

where

$$b_{j,n} = \max_{1\le j\le u\le n} \Big|X(j) \sum_{k=j}^u \mathbb{E}\big(X(k)|\mathcal{G}_j\big)\Big|_1 \le |v|_\infty \max_{1\le j\le u\le n} \Big|\sum_{k=j}^u \mathbb{E}\big(v \circ T^{n-k}|\mathcal{G}_j\big)\Big|_1.$$

By Proposition 2.7, $\mathbb{E}(m\circ T^{n-k}|\mathcal{G}_j)=0$ for all $k>j$. By (2.3),

$$\sum_{k=j}^{u}\mathbb{E}\big(v\circ T^{n-k}|\mathcal{G}_j\big)=\mathbb{E}\big(m\circ T^{n-j}+\chi\circ T^{n+1-j}-\chi\circ T^{n-u}|\mathcal{G}_j\big)=v\circ T^{n-j}+\chi\circ T^{n-j}-\mathbb{E}\big(\chi\circ T^{n-u}|\mathcal{G}_j\big).$$

Hence

$$b_{j,n}\le|v|_\infty\big(|v|_1+2|\chi|_1\big)\le 2|v|_\infty\sum_{j=0}^{\infty}\big|P^jv\big|_1$$

and so

$$\Big|\max_{1\le\ell\le n}\big|X(1)+\cdots+X(\ell)\big|\Big|_2^2\le 32\,|v|_\infty\sum_{j=0}^{\infty}\big|P^jv\big|_1.$$

Finally, $\max_{1\le\ell\le n}|v_\ell|\le 2\max_{1\le\ell\le n}|X(1)+\cdots+X(\ell)|$. □

Lemma 2.10. $\lim_{n\to\infty}\frac{1}{\sqrt{n}}|\max_{1\le\ell\le n}|(v-m)_\ell||_2=0$.

Proof. By hypothesis (2.1) and Lemma 2.6, for each $\epsilon>0$, there exists $k\ge 1$ such that $\sum_{j=k}^{\infty}|P^jv|_1<\epsilon^2$ and $|m-m^{(k)}|_2<\epsilon$.

Recall that $m,\ m^{(k)}\in\ker P$. By Corollary 2.8,

$$\frac{1}{\sqrt{n}}\Big|\max_{1\le\ell\le n}\big|\big(m-m^{(k)}\big)_\ell\big|\Big|_2<4\epsilon. \tag{2.5}$$

Next, $v=m^{(k)}+\chi_1^k\circ T-\chi_1^k+P^kv$, so

$$\big|\big(v-m^{(k)}\big)_n\big|\le 2\big|\chi_1^k\big|_\infty+\big|\big(P^kv\big)_n\big|\le 2k|v|_\infty+\big|\big(P^kv\big)_n\big|.$$

Note that P^kv satisfies our underlying hypotheses, namely $P^kv\in L^\infty$, $\int_\Lambda P^kv\,d\mu=0$, $\sum_{n=1}^{\infty}|P^n(P^kv)|_1<\infty$. Hence by Proposition 2.9,

$$\begin{aligned}\Big|\max_{1\le\ell\le n}\big|\big(v-m^{(k)}\big)_\ell\big|\Big|_2&\le 2k|v|_\infty+\Big|\max_{1\le\ell\le n}\big|\big(P^kv\big)_\ell\big|\Big|_2\le 2k|v|_\infty+\left\{128\,n\big|P^kv\big|_\infty\left(\sum_{j=0}^{\infty}\big|P^{j+k}v\big|_1\right)\right\}^{1/2}\\&\le 2k|v|_\infty+\left\{128\,n|v|_\infty\left(\sum_{j=k}^{\infty}\big|P^jv\big|_1\right)\right\}^{1/2}\ll k+\epsilon\sqrt{n}.\end{aligned}$$

Combining this with (2.5), $\frac{1}{\sqrt{n}}|\max_{1\le\ell\le n}|(v-m)_\ell||_2\ll\frac{1}{\sqrt{n}}k+\epsilon$. Hence $\limsup_{n\to\infty}\frac{1}{\sqrt{n}}|\max_{1\le\ell\le n}|(v-m)_\ell||_2\ll\epsilon$ and the result follows since ϵ is arbitrary. □

Proof of parts (a) and (b) of Theorem 2.1. Since $m\in\ker P$, it holds that $\int_\Lambda m_n\otimes m_n\,d\mu=n\int_\Lambda m\otimes m\,d\mu$ for all n. By Proposition 2.9, $|v_n|_2\ll n^{1/2}$. Hence,

$$\begin{aligned}\left|n^{-1}\int_\Lambda v_n\otimes v_n\,d\mu-\int_\Lambda m\otimes m\,d\mu\right|&=n^{-1}\left|\int_\Lambda(v_n\otimes v_n-m_n\otimes m_n)\,d\mu\right|\\&\le n^{-1}\big(|v_n|_2+|m_n|_2\big)|v_n-m_n|_2\ll n^{-1/2}|v_n-m_n|_2\to 0\end{aligned}$$

by Lemma 2.10. This proves part (a) and shows in addition that

$$\Sigma=\int_\Lambda m\otimes m\,d\mu. \tag{2.6}$$

It follows that $c^T\Sigma c=\int_\Lambda(c\cdot m)^2\,d\mu$ for all $c\in\mathbb{R}^d$.

Next we prove part (b). If $\det \Sigma = 0$, then there exists $c \in \mathbb{R}^d$ nonzero such that $\Sigma c = 0$ and hence $\int_\Lambda (c \cdot m)^2 \, d\mu = c^T \Sigma c = 0$, so $c \cdot m = 0$. By (2.3), $c \cdot v = h \circ T - h$ where $h = c \cdot \chi \in L^1$.

Conversely, suppose that $c \cdot v = h \circ T - h$ for $c \in \mathbb{R}^d$ nonzero and $h \in L^1$. Then $c \cdot Pv = h - Ph$. Also, $Pv = \chi - P\chi$, hence $c \cdot \chi - h \in \ker(P - I)$. By ergodicity, $c \cdot \chi = h + aI$ for some $a \in \mathbb{R}$. Substituting into (2.3),

$$c \cdot v = c \cdot m + c \cdot \chi \circ T - c \cdot \chi = c \cdot m + h \circ T - h = c \cdot m + c \cdot v,$$

and so $c \cdot m = 0$. Hence $c^T \Sigma c = \int_\Lambda (c \cdot m)^2 \, d\mu = 0$. It follows that $\det \Sigma = 0$. □

2.3. *Iterated WIP*

In this subsection, we prove Theorem 2.1(c). First, we prove the ordinary WIP.

Lemma 2.11. *$W_n \to_w W$ in $D([0,\infty), \mathbb{R}^d)$ as $n \to \infty$ on the probability space (Λ, μ).*

Proof. It suffices to prove that $W_n \to_w W$ in $D([0,K], \mathbb{R}^d)$ for each fixed integer $K \geq 1$. Define $M_n(t) = \frac{1}{\sqrt{n}} \times \sum_{0 \leq j \leq [nt]-1} m \circ T^j$. Recall that $m \in L^2 \cap \ker P$. By the pointwise ergodic theorem and (2.6),

$$n^{-1} \sum_{j=0}^{n-1} \big\{ UP(m \otimes m) \big\} \circ T^j \to \int_\Lambda UP(m \otimes m) \, d\mu = \int_\Lambda m \otimes m \, d\mu = \Sigma \quad a.e.$$

It follows from [26, Theorem A.1] that $M_n \to_w W$ in $D([0,K], \mathbb{R}^d)$. Also,

$$\sup_{t \in [0,K]} \big| W_n(t) - M_n(t) \big| = \frac{1}{\sqrt{n}} \max_{1 \leq \ell \leq nK} \big| (v - m)_\ell \big| \to_p 0$$

by Lemma 2.10. Hence $W_n \to_w W$ in $D([0,K], \mathbb{R}^d)$. □

Define the sequence of processes

$$\mathbb{M}_n \in D\big([0,\infty), \mathbb{R}^{d \times d}\big), \qquad \mathbb{M}_n(t) = \frac{1}{n} \sum_{0 \leq i < j \leq [nt]-1} \big(m \circ T^i\big) \otimes \big(v \circ T^j\big).$$

Lemma 2.12. *$(W_n, \mathbb{M}_n) \to_w (W, \mathbb{M})$ in $D([0,\infty), \mathbb{R}^d \times \mathbb{R}^{d \times d})$ as $n \to \infty$ on the probability space (Λ, μ), where $\mathbb{M}(t) = \int_0^t W \otimes dW$.*

Proof. It suffices to prove that $(W_n, \mathbb{M}_n) \to_w (W, \mathbb{M})$ in $D([0,K], \mathbb{R}^d \times \mathbb{R}^{d \times d})$ for each fixed integer $K \geq 1$. Define for $t \in [0,K]$,

$$W_n^-(t) = \frac{1}{\sqrt{n}} \sum_{1 \leq j \leq [nt]} v \circ T^{nK-j}, \qquad M_n^-(t) = \frac{1}{\sqrt{n}} \sum_{1 \leq j \leq [nt]} m \circ T^{nK-j}, \tag{2.7}$$

$$\mathbb{M}_n^-(t) = \frac{1}{n} \sum_{1 \leq i < j \leq [nt]} \big(v \circ T^{nK-i}\big) \otimes \big(m \circ T^{nK-j}\big).$$

There are three main steps:

Step 1 Transfer convergence of W_n in Lemma 2.11 to convergence of W_n^- and M_n^-, showing that $(W_n^-, M_n^-) \to_w (W, W)$ in $D([0,K], \mathbb{R}^d \times \mathbb{R}^d)$.
Step 2 Apply [21,28] to show that $(W_n^-, M_n^-, \mathbb{M}_n^-) \to_w (W, W, \mathbb{M})$ in $D([0,K], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d \times d})$.
Step 3 Transfer convergence of $(W_n^-, M_n^-, \mathbb{M}_n^-)$ in Step 2 back to convergence of $(W_n, \mathbb{M}_n)$, yielding the desired result.

Let $\widetilde{D}$ denote càglàd functions. Following [23], we define

$$g : D\big([0,K], \mathbb{R}^d\big) \to \widetilde{D}\big([0,K], \mathbb{R}^d\big), \qquad g(r)(t) = r(K) - r(K-t).$$

Then

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{j=nK-[nt]+1}^{nK} v \circ T^{nK-j} = \frac{1}{\sqrt{n}} \sum_{j=[n(K-t)]+1}^{nK} v \circ T^{nK-j} - F_n^1(t)$$
$$= W_n^-(K) - W_n^-(K-t) - F_n^1(t) = g\big(W_n^-\big)(t) - F_n^1(t),$$

where $F_n^1(t)$ is either 0 or $n^{-1/2} v \circ T^{nK-[n(K-t)]-1}$. In particular,

$$\sup_{t\in[0,K]} \big|F_n^1(t)\big| \le n^{-1/2}|v|_\infty \to 0.$$

By Lemma 2.11 and the continuous mapping theorem,

$$W_n^- = g^{-1}\big(W_n + F_n^1\big) \to_w g^{-1}(W) \quad \text{in } \widetilde{D}([0,K],\mathbb{R}^d).$$

Using the fact that the limiting process has continuous sample paths, it follows (see [23, Proposition 4.9]) that $W_n^- \to_w g^{-1}(W)$ in $D([0,K],\mathbb{R}^d)$. By [23, Lemma 4.11], the processes $g^{-1}(W)$ and W are equal in distribution, so $W_n^- \to_w W$ in $D([0,K],\mathbb{R}^d)$. By the continuous mapping theorem, $(W_n^-, W_n^-) \to_w (W,W)$ in $D([0,K],\mathbb{R}^d\times\mathbb{R}^d)$. Also,

$$\sup_{t\in[0,K]} \big|W_n^-(t) - M_n^-(t)\big| \le 2n^{-1/2} \max_{1\le\ell\le nK} \big|(v-m)_\ell\big| \tag{2.8}$$

so $|\sup_{t\in[0,K]}|W_n^-(t) - M_n^-(t)||_2 \to 0$ by Lemma 2.10. Hence $(W_n^-, M_n^-) \to_w (W,W)$ in $D([0,K],\mathbb{R}^d\times\mathbb{R}^d)$ completing Step 1.

By Proposition 2.7, $\{m\circ T^{nK-j};\ 1\le j\le nK\}$ is a martingale difference sequence with respect to the filtration $\mathcal{G}_{n,j} = T^{-(nK-j)}\mathcal{F}$ for each $n\ge 1$. Moreover, W_n^- is adapted (i.e. $v\circ T^{nK-j}$ is $\mathcal{G}_{n,j}$-measurable for all j, n). Also $\int_\Lambda |M_n^-(t)|^2\,d\mu = n^{-1}[nt]\int_\Lambda |m|^2\,d\mu \le K|m|_2^2$, so condition C2.2(i) in [28, Theorem 2.2] is satisfied. Applying [28, Theorem 2.2] (or alternatively [21]) we deduce that $(W_n^-, M_n^-, \mathbb{M}_n^-) \to_w (W,W,\mathbb{M})$ in $D([0,K],\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d})$ completing Step 2.

Adapting [23], we define $h : D([0,K],\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}) \to \widetilde{D}([0,K],\mathbb{R}^d\times\mathbb{R}^{d\times d})$,

$$h(r,u,v)(t) = \big(r(K) - r(K-t)\,,\ \big\{v(K) - v(K-t) - r(K-t)\otimes\big(u(K)-u(K-t)\big)\big\}^*\big),$$

where * denotes matrix transpose.

We claim that

$$(W_n,\mathbb{M}_n) = h\big(W_n^-, M_n^-, \mathbb{M}_n^-\big) - F_n \quad \text{where } \sup_{t\in[0,K]} \big|F_n(t)\big| \to_p 0.$$

Suppose that the claim is true. By the continuous mapping theorem and [23, Proposition 4.9], $(W_n,\mathbb{M}_n) \to_w h(W,W,\mathbb{M})$ in $D([0,K],\mathbb{R}^d\times\mathbb{R}^{d\times d})$. By [23, Lemma 4.11], the processes $h(W,W,\mathbb{M})$ and $(W,\mathbb{M})$ are equal in distribution so $(W_n,\mathbb{M}_n) \to_w (W,\mathbb{M})$ in $D([0,K],\mathbb{R}^d\times\mathbb{R}^{d\times d})$.

It remains to prove the claim. Write $h = (h^1, h^2)$ where $h^1 : D([0,K],\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}) \to \widetilde{D}([0,K],\mathbb{R}^d)$ and $h^2 : D([0,K],\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{R}^{d\times d}) \to \widetilde{D}([0,K],\mathbb{R}^{d\times d})$.

By Step 1,

$$W_n(t) = h^1\big(W_n^-, M_n^-, \mathbb{M}_n^-\big)(t) - F_n^1(t) \quad \text{where } \sup_{t\in[0,K]} \big|F_n^1(t)\big| \le n^{-1/2}|v|_\infty \to 0.$$

Also,

$$\mathbb{M}_n(t) = \frac{1}{n} \sum_{nK-[nt]<j<i\le nK} \big(m\circ T^{nK-i}\big)\otimes\big(v\circ T^{nK-j}\big)$$
$$= \frac{1}{n} \sum_{nK-[nt]<i<j\le nK} \big\{\big(v\circ T^{nK-i}\big)\otimes\big(m\circ T^{nK-j}\big)\big\}^*$$
$$= \frac{1}{n} \sum_{[n(K-t)]<i<j\le nK} \big\{\big(v\circ T^{nK-i}\big)\otimes\big(m\circ T^{nK-j}\big)\big\}^* - F_n^2(t)^*$$

$$= \big\{\mathbb{M}_n^-(K) - \mathbb{M}_n^-(K-t) - W_n^-(K-t) \otimes \big(M_n^-(K) - M_n^-(K-t)\big)\big\}^* - F_n^2(t)^*$$
$$= h^2\big(W_n^-, M_n^-, \mathbb{M}_n^-\big)(t) - F_n^2(t)^*,$$

where $F_n^2(t)$ is either 0 or $n^{-1}\sum_{[n(K-t)]+1<j\le nK}(v\circ T^{nK-[n(K-t)]-1})\otimes(m\circ T^{nK-j})$. In particular, $|F_n^2(t)|\le n^{-1}|v|_\infty \max_{1\le\ell\le nK}|m_\ell|$, so by Corollary 2.8,

$$\Big|\sup_{t\in[0,K]}\big|F_n^2(t)\big|\Big|_2 \ll n^{-1/2}|v|_\infty|m|_2 \to 0.$$

This completes the proof of the claim. □

Proof of Theorem 2.1(c). First we consider the case $\nu=\mu$. It follows from the definition of χ that

$$\int_\Lambda \chi\otimes v\,d\mu = \sum_{j=1}^\infty \int_\Lambda \big(P^j v\big)\otimes v\,d\mu = \sum_{j=1}^\infty \int_\Lambda v\otimes\big(v\circ T^j\big)\,d\mu.$$

By Lemma 2.12, it suffices to show for all $K>0$ that

$$\Big|\sup_{t\in[0,K]}\Big(\mathbb{W}_n(t) - \mathbb{M}_n(t) - t\int_\Lambda \chi\otimes v\,d\mu\Big)\Big| \to_p 0 \quad \text{as } n\to\infty.$$

Now,

$$\sum_{j=1}^n\sum_{i=0}^{j-1}\big((v-m)\circ T^i\big)\otimes\big(v\circ T^j\big) = \sum_{j=1}^n\big(\chi\circ T^j-\chi\big)\otimes\big(v\circ T^j\big) = \sum_{j=1}^n(\chi\otimes v)\circ T^j - \chi\otimes\sum_{j=1}^n v\circ T^j.$$

Since $v\in L^\infty$, $\chi\in L^1$, and $\int_\Lambda v\,d\mu=0$, it follows from the pointwise ergodic theorem that

$$\mathbb{W}_n(1)-\mathbb{M}_n(1) = n^{-1}\sum_{j=1}^n\sum_{i=0}^{j-1}\big((v-m)\circ T^i\big)\otimes\big(v\circ T^j\big) \to \int_\Lambda \chi\otimes v\,d\mu \quad \text{a.e.}$$

as $n\to\infty$. Hence for any $K>0$,

$$\Big|\sup_{t\in[0,K]}\Big(\mathbb{W}_n(t) - \mathbb{M}_n(t) - t\int_\Lambda \chi\otimes v\,d\mu\Big)\Big| \to 0 \quad \text{a.e.}$$

The iterated WIP on (Λ,μ) follows.

Now we consider the case where ν is a general probability measure absolutely continuous with respect to μ. Since μ is ergodic, it suffices by [52, Theorem 1] to show that

$$\lim_{n\to\infty}\mu\Big(\sup_{t\in[0,K]}\big|\mathbf{W}_n(t)\circ T - \mathbf{W}_n(t)\big|>\epsilon\Big)=0 \tag{2.9}$$

for all $\epsilon>0$, where $\mathbf{W}_n=(W_n,\mathbb{W}_n)$.

Now, $W_n(t)\circ T - W_n(t) = n^{-1/2}(v\circ T^{[nt]}-v)$ so

$$\big|W_n(t)\circ T - W_n(t)\big| \le 2n^{-1/2}\max_{0\le k\le nK}\big|v\circ T^k\big| \le 2n^{-1/2}|v|_\infty$$

for all $t\in[0,K]$. Similarly,

$$\big|\mathbb{W}_n(t)\circ T - \mathbb{W}_n(t)\big| \le 2n^{-1}|v|_\infty \max_{1\le k\le nK}|v_k|$$

for all $t\in[0,K]$. By Proposition 2.9 and (2.1), $|\max_{1\le k\le nK}|v_k||_2 \ll (n|v|_\infty\sum_{j\ge0}|P^j v|_1)^{1/2} \ll n^{1/2}$. Hence $|\sup_{t\in[0,K]}|\mathbf{W}_n(t)\circ T-\mathbf{W}_n(t)||_2 \ll n^{-1/2}$, and (2.9) follows. □

3. Invertible setting

Let $(\Lambda, \mathcal{F}, \mu)$ be a probability space and $T : \Lambda \to \Lambda$ be an invertible ergodic measure-preserving map. We suppose that there is a sub-sigma-algebra $\mathcal{F}_0 \subset \mathcal{F}$ such that $T^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$. Then $\mathcal{F}_j = T^j \mathcal{F}_0$ defines a nondecreasing filtration $\{\mathcal{F}_j : j \in \mathbb{Z}\}$.

Fix $d \geq 1$ and let $v \in L^\infty(\Lambda, \mathbb{R}^d)$ with $\int_\Lambda v\, d\mu = 0$. The L^1 Gordin criterion now takes the form

$$\sum_{n=1}^{\infty} \big|\mathbb{E}_0\big(v \circ T^{-n}\big)\big|_1 + \sum_{n=0}^{\infty} \big|\mathbb{E}_0\big(v \circ T^n\big) - v \circ T^n\big|_1 < \infty, \tag{3.1}$$

where $\mathbb{E}_j = \mathbb{E}(\cdot\,|\mathcal{F}_j)$.

Under hypotheses similar to (3.1), the CLT and WIP have been proved by various authors, including [5,42,46]. In Section 3.1, we recover the WIP under hypothesis (3.1) using techniques similar to those in Section 2 combined with ideas from [5].

The iterated WIP holds under the L^2 Gordin criterion $\sum_{n=1}^\infty |\mathbb{E}_0(v \circ T^{-n})|_2 + \sum_{n=0}^\infty |\mathbb{E}_0(v \circ T^n) - v \circ T^n|_2 < \infty$ by [23, Section 4] (see [6, Proposition 2.5]). An interesting open question is to prove the iterated WIP under the L^1 criterion (3.1), but this seems currently out of reach. In Section 3.2, we prove the iterated WIP under a hybrid L^1–L^2 Gordin criterion

$$\sum_{n=1}^{\infty} \big|\mathbb{E}_0\big(v \circ T^{-n}\big)\big|_1 + \sum_{n=0}^{\infty} \big|\mathbb{E}_0\big(v \circ T^n\big) - v \circ T^n\big|_2 < \infty. \tag{3.2}$$

The same argument works if the roles of $||_1$ and $||_2$ are reversed in (3.2).

We note that the existence of a suitable sub-sigma-algebra $\mathcal{F}_0$ is very natural in the dynamical setting. Indeed it is often the case that Λ is covered by a collection $\mathcal{W}^s$ of disjoint measurable sets, called "stable leaves", such that $T W_x^s \subset W_{Tx}^s$ for all $x \in \Lambda$, where W_x^s is the stable leaf containing x. In this situation, let $\mathcal{F}_0$ denote the sigma-algebra generated by $\mathcal{W}^s$. Then $T^{-1}\mathcal{F}_0 \subset \mathcal{F}_0$. The following result gives sufficient conditions for hypotheses (3.1) and (3.2) to hold.

Proposition 3.1. *Let $p \geq 1$ and let $v \in L^\infty(\Lambda, \mathbb{R})$ with $\int_\Lambda v\, d\mu = 0$.*

(a) *Suppose that there exists $C > 0$, $\epsilon > 0$ such that*

$$\left|\int_\Lambda v\, w \circ T^n\, d\mu\right| \leq C|w|_\infty n^{-(p+\epsilon)}$$

for all $\mathcal{F}_0$-measurable $w \in L^\infty(\Lambda, \mathbb{R})$, $n \geq 1$. Then $\sum_{n=1}^\infty |\mathbb{E}_0(v \circ T^{-n})|_p < \infty$.

(b) *Suppose that there exists $C > 0$, $\epsilon > 0$ such that*

$$\int_\Lambda \operatorname{diam}\big(v\big(T^n W^s\big)\big)\, d\mu \leq C|w|_\infty n^{-(p+\epsilon)}$$

for all $n \geq 1$. Then $\sum_{n=0}^\infty |\mathbb{E}_0(v \circ T^n) - v \circ T^n|_p < \infty$.

Proof. The arguments are standard. See for example [6, Theorem 3.1]. □

Throughout the remainder of this section, L^p is shorthand for $L^p((\Lambda, \mu), \mathbb{R}^d)$ unless stated otherwise.

3.1. *WIP in the invertible setting*

Define $W_n \in D([0, \infty), \mathbb{R}^d)$ as in (2.2). Let ν be any probability measure on Λ absolutely continuous with respect to μ. In this subsection, we prove:

Theorem 3.2. *Let $v \in L^\infty$ with $\int_\Lambda v\, d\mu = 0$, and suppose that (3.1) holds. Then conclusions* (a) *and* (b) *of Theorem 2.1 hold, and $W_n \to_w W$ in $D([0, \infty), \mathbb{R}^d)$ as $n \to \infty$ on (Λ, ν), where W is a d-dimensional Brownian motion with covariance Σ.*

For $-\infty < \ell \le k < \infty$, define

$$\chi_\ell^k = \sum_{j=\ell}^{k} a_j, \quad a_j = \begin{cases} \mathbb{E}_0(v \circ T^j) & j \le -1 \\ \mathbb{E}_0(v \circ T^j) - v \circ T^j & j \ge 0. \end{cases}$$

Also define $\chi = \sum_{j=-\infty}^{\infty} a_j$. It follows from our assumptions that $\chi_\ell^k \in L^\infty$ for all $\ell \le k$ and $\chi \in L^1$. Moreover, $\chi_{-k}^k \to \chi$ in L^1 as $k \to \infty$.

Proposition 3.3.

(a) $\mathbb{E}_{-1}(\chi_{-k}^{-\ell}) = \chi_{-k-1}^{-\ell-1} \circ T$ *for all* $k \ge \ell > 0$.
(b) $\mathbb{E}(\chi_{\ell+1}^{k+1}(\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1})) = 0$ *for all* $k \ge \ell \ge 0$.

Proof. (a) Since $\mathbb{E}_{-1}\mathbb{E}_0 = \mathbb{E}_{-1}$ and $\mathbb{E}_{-1}(g \circ T) = (\mathbb{E}_0 g) \circ T$,

$$\mathbb{E}_{-1}\big(\chi_{-k}^{-\ell}\big) = \sum_{j=-k}^{-\ell} \mathbb{E}_{-1}\big(v \circ T^j\big) = \sum_{j=-k}^{-\ell} \big(\mathbb{E}_0\big(v \circ T^{j-1}\big)\big) \circ T = \chi_{-k-1}^{-\ell-1} \circ T.$$

(b) Note that

$$\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1} = \sum_{j=\ell}^{k} \big\{\mathbb{E}_0\big(v \circ T^j\big) \circ T - \mathbb{E}_0\big(v \circ T^{j+1}\big)\big\},$$

so $\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1}$ is $\mathcal{F}_0$-measurable. Also $\mathbb{E}_0 \chi_{\ell+1}^{k+1} = 0$. Hence

$$\mathbb{E}\big(\chi_{\ell+1}^{k+1}\big(\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1}\big)\big) = \mathbb{E}\mathbb{E}_0\big(\chi_{\ell+1}^{k+1}\big(\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1}\big)\big) = \mathbb{E}\big(\big(\chi_\ell^k \circ T - \chi_{\ell+1}^{k+1}\big)\mathbb{E}_0 \chi_{\ell+1}^{k+1}\big) = 0$$

as required. □

Write

$$v = m^{(k)} + \chi_{-k}^k \circ T - \chi_{-k}^k + a_{-k} - a_{k+1}, \quad k \ge 1 \quad \text{and} \quad v = m + \chi \circ T - \chi. \tag{3.3}$$

It is immediate from the definitions that $m^{(k)} \in L^\infty$ for all k, that $m \in L^1$ and that $m^{(k)} \to m$ in L^1. Moreover, we have the following result corresponding to Lemma 2.6:

Lemma 3.4. $m \in L^2$ *and* $m^{(k)} \to m$ *in* L^2 *as* $k \to \infty$.

Proof. For $k \ge \ell \ge 0$,

$$\begin{aligned} m^{(k)} - m^{(\ell)} &= \big(\chi_{-k}^k - \chi_{-\ell}^\ell\big) - \big(\chi_{-k}^k - \chi_{-\ell}^\ell\big) \circ T + (a_{k+1} - a_{-k}) - (a_{\ell+1} - a_{-\ell}) \\ &= \big(\chi_{-k}^{-\ell-1} + \chi_{\ell+1}^k\big) - \big(\chi_{-k}^{-\ell-1} + \chi_{\ell+1}^k\big) \circ T + (a_{k+1} - a_{-k}) - (a_{\ell+1} - a_{-\ell}) \\ &= \big(\chi_{-k+1}^{-\ell} + \chi_{\ell+2}^{k+1}\big) - \big(\chi_{-k}^{-\ell-1} + \chi_{\ell+1}^k\big) \circ T. \end{aligned}$$

Hence $|m^{(k)} - m^{(\ell)}|_2 \le A + B$ where

$$A = \big|\chi_{-k+1}^{-\ell} - \chi_{-k}^{-\ell-1} \circ T\big|_2, \qquad B = \big|\chi_{\ell+2}^{k+1} - \chi_{\ell+1}^k \circ T\big|_2.$$

Now,

$$A^2 = \mathbb{E}\big(\big(\chi_{-k+1}^{-\ell}\big)^2 - 2\chi_{-k+1}^{-\ell}\,\chi_{-k}^{-\ell-1} \circ T + \big(\chi_{-k}^{-\ell-1}\big)^2\big).$$

By Proposition 3.3(a),

$$\mathbb{E}\big(\chi_{-k+1}^{-\ell}\,\chi_{-k}^{-\ell-1} \circ T\big) = \mathbb{E}\mathbb{E}_{-1}\big(\chi_{-k+1}^{-\ell}\,\chi_{-k}^{-\ell-1} \circ T\big) = \mathbb{E}\big(\chi_{-k}^{-\ell-1} \circ T\,\mathbb{E}_{-1}\big(\chi_{-k+1}^{-\ell}\big)\big) = \mathbb{E}\big(\big(\chi_{-k}^{-\ell-1} \circ T\big)^2\big) = \mathbb{E}\big(\big(\chi_{-k}^{-\ell-1}\big)^2\big).$$

Hence

$$A^2 = \mathbb{E}\big((\chi_{-k+1}^{-\ell})^2 - (\chi_{-k}^{-\ell-1})^2\big) = \mathbb{E}\big((\chi_{-k+1}^{-\ell} - \chi_{-k}^{-\ell-1})(\chi_{-k+1}^{-\ell} + \chi_{-k}^{-\ell-1})\big)$$
$$= \mathbb{E}\big((a_{-\ell} - a_{-k})(\chi_{-k+1}^{-\ell} + \chi_{-k}^{-\ell-1})\big) \le |a_{-\ell} - a_{-k}|_\infty \big|\chi_{-k+1}^{-\ell} + \chi_{-k}^{-\ell-1}\big|_1 \le 4|v|_\infty \sum_{j=-\infty}^{-\ell} \big|\mathbb{E}_0\big(v \circ T^j\big)\big|_1.$$

Next,

$$B^2 = \mathbb{E}\big((\chi_{\ell+2}^{k+1})^2 - 2\chi_{\ell+2}^{k+1}\,\chi_{\ell+1}^{k} \circ T + (\chi_{\ell+1}^{k})^2\big).$$

By Proposition 3.3(b),

$$\mathbb{E}\big(\chi_{\ell+2}^{k+1}\,\chi_{\ell+1}^{k} \circ T\big) = \mathbb{E}\big(\chi_{\ell+2}^{k+1}\,(\chi_{\ell+1}^{k} \circ T - \chi_{\ell+2}^{k+1}) + (\chi_{\ell+2}^{k+1})^2\big) = \mathbb{E}\big((\chi_{\ell+2}^{k+1})^2\big).$$

Hence

$$B^2 = \mathbb{E}\big((\chi_{\ell+1}^{k})^2 - (\chi_{\ell+2}^{k+1})^2\big) = \mathbb{E}\big((\chi_{\ell+1}^{k} - \chi_{\ell+2}^{k+1})(\chi_{\ell+1}^{k} + \chi_{\ell+2}^{k+1})\big)$$
$$= \mathbb{E}\big((a_{\ell+1} - a_{k+1})(\chi_{\ell+1}^{k} + \chi_{\ell+2}^{k+1})\big) \le |a_{\ell+1} - a_{k+1}|_\infty \big|\chi_{\ell+1}^{k} + \chi_{\ell+2}^{k+1}\big|_1 \le 8|v|_\infty \sum_{j=\ell+1}^{\infty} \big|\mathbb{E}_0\big(v \circ T^j\big) - v \circ T^j\big|_1.$$

It follows from hypothesis (3.1) together with these estimates for A and B that $m^{(k)}$ is Cauchy in L^2. By uniqueness of limits in L^1, the L^2 limit of $m^{(k)}$ coincides with m. □

Standard calculations (see for example [18,47] or [6, Proposition 2.2]) show that m is $\mathcal{F}_0$-measurable and that $\mathbb{E}_{-1}m = 0$. Hence $\{m \circ T^{-j} : n \in \mathbb{Z}\}$ is a martingale with respect to the filtration $\mathcal{F}_j$. The same is true for

$$m^{(k)} = \sum_{j=-k+1}^{k+1} \mathbb{E}_0\big(v \circ T^j\big) - \sum_{j=-k}^{k} \big(\mathbb{E}_0\big(v \circ T^j\big)\big) \circ T.$$

Maximal inequality for a_{-k}

Proposition 3.5. *Let $w \in L^\infty$ and suppose that w is $\mathcal{F}_0$-measurable. Then*

$$|\max_{1\le\ell\le n} |w_\ell||_2^2 \le 128\, n |w|_\infty \sum_{j=0}^{\infty} |\mathbb{E}_0(w \circ T^{-j})|_1.$$

Proof. Fix $n \ge 1$ and define the random variables $X(j) = w \circ T^{n-j}$ which are adapted to the filtration $\mathcal{F}_{j-n}$. Using Rio's inequality as in the proof of Proposition 2.9, $|\max_{1\le\ell\le n} |X(1) + \cdots + X(\ell)||_2^2 \le 16 \sum_{j=1}^n b_{j,n}$ where

$$b_{j,n} = \max_{1\le j\le u\le n} \left| X(j) \sum_{k=j}^{u} \mathbb{E}\big(X(k)|\mathcal{F}_{j-n}\big) \right|_1 \le |w|_\infty \max_{1\le j\le u\le n} \left| \sum_{k=j}^{u} \mathbb{E}\big(w \circ T^{n-k}|\mathcal{F}_{j-n}\big) \right|_1.$$

Define m_-, $\chi_- \in L^1$,

$$\chi_- = \sum_{j=1}^{\infty} \mathbb{E}_0\big(w \circ T^{-j}\big), \quad w = m_- + \chi_- \circ T - \chi_-.$$

Using that w is $\mathcal{F}_0$-measurable, it is easily verified that m_- is $\mathcal{F}_0$-measurable and $\mathbb{E}_{-1}m_- = 0$. Hence $\mathbb{E}(m_- \circ T^{n-k}|\mathcal{F}_{j-n}) = 0$ for all $k > j$. It follows that

$$\sum_{k=j}^{u} \mathbb{E}\big(w \circ T^{n-k}|\mathcal{F}_{j-n}\big) = \mathbb{E}\big(m_- \circ T^{n-j} + \chi_- \circ T^{n+1-j} - \chi_- \circ T^{n-u}|\mathcal{F}_{j-n}\big).$$

Now continue as in the proof of Proposition 2.9. □

Corollary 3.6. $|\max_{1\le\ell\le n}|(a_{-k})_\ell||_2^2 \le 128\,n|v|_\infty \sum_{j=k}^\infty |\mathbb{E}_0(v\circ T^{-j})|_1$.

Proof. Recall that $a_{-k}=\mathbb{E}_0(v\circ T^{-k})$, so $|a_{-k}|_\infty\le|v|_\infty$ and a_{-k} is $\mathcal{F}_0$-measurable. By Proposition 3.5,

$$\Big|\max_{1\le\ell\le n}\big|(a_{-k})_\ell\big|\Big|_2^2 \le 128\,n|v|_\infty\sum_{j=0}^\infty\big|\mathbb{E}_0\big(a_{-k}\circ T^{-j}\big)\big|_1.$$

Setting $g=v\circ T^{-k}$,

$$\mathbb{E}_0\big(a_{-k}\circ T^{-j}\big)=\mathbb{E}_0\big((\mathbb{E}_0 g)\circ T^{-j}\big)=(\mathbb{E}_{-j}\mathbb{E}_0 g)\circ T^{-j}=(\mathbb{E}_{-j}g)\circ T^{-j}=\mathbb{E}_0\big(g\circ T^{-j}\big)=\mathbb{E}_0\big(v\circ T^{-(j+k)}\big).$$

The result follows. □

Maximal inequality for a_k

Here we rely heavily on ideas from [5]. In particular, we require the following maximal inequality [5, Equation (3.4)]:

Lemma 3.7. *Let $S_n=\sum_{j=1}^n X(j)$ be a sum of L^2 random variables. Then*

$$\mathbb{E}\big(S_n^{*2}\big)\le 4\mathbb{E}\big(S_n^2\big)-4\sum_{j=1}^n\mathbb{E}\big(X(j)S_{j-1}^*\big)$$

where $S_n^*=\max\{0,S_1,\dots,S_n\}$. □

The following elementary estimate is useful:

Proposition 3.8. *Define $h:\mathbb{R}^n\to\mathbb{R}$, $h(b)=\max\{0,b_1,b_1+b_2,\dots,\sum_{j=1}^n b_i\}$. Then $|h(b)-h(b')|\le\sum_{i=1}^n|b_i-b_i'|$.* □

Proposition 3.9. *Let $w\in L^\infty$ with $\mathbb{E}_0 w=0$. Then*

$$\Big|\max_{1\le\ell\le n}|w_\ell|\Big|_2^2\le 96\,n|w|_\infty\sum_{j=0}^\infty\big|\mathbb{E}_0\big(w\circ T^j\big)-w\circ T^j\big|_1.$$

Proof. Define $X(j)=w\circ T^{-j}$ and $S_n=\sum_{j=1}^n X(j)$. Then

$$\mathbb{E}\big(S_n^2\big)=\sum_{i,j=0}^{n-1}\mathbb{E}\big(w\circ T^{-i}\,w\circ T^{-j}\big)=n\mathbb{E}\big(w^2\big)+2\sum_{j=1}^{n-1}(n-j)\mathbb{E}\big(w\,w\circ T^j\big).$$

Also, $\mathbb{E}(w\,\mathbb{E}_0(w\circ T^j))=\mathbb{E}(\mathbb{E}_0(w\circ T^j)\,\mathbb{E}_0 w)=0$ and so

$$\begin{aligned}\mathbb{E}\big(S_n^2\big)&=n\mathbb{E}\big(w^2\big)+2\sum_{j=1}^{n-1}(n-j)\mathbb{E}\big(w\,\big(w\circ T^j-\mathbb{E}_0\big(w\circ T^j\big)\big)\big)\\&\le 2n|w|_\infty\bigg(|w|_1+\sum_{j=1}^\infty\big|\mathbb{E}_0\big(w\circ T^j\big)-w\circ T^j\big|_1\bigg)=2n|w|_\infty\sum_{j=0}^\infty\big|\mathbb{E}_0\big(w\circ T^j\big)-w\circ T^j\big|_1.\end{aligned}$$

Next, define

$$Y_{i,j}=\mathbb{E}_j\big(w\circ T^{-i}\big),\qquad Z_{p,j}=\sum_{i=1}^p Y_{i,j},\qquad Z_{j-1}^*=\max\{0,Z_{1,j},\dots,Z_{j-1,j}\}.$$

Note that $Y_{i,j}$ is $\mathcal{F}_j$-measurable for all $i<j$, so in particular Z_{j-1}^* is $\mathcal{F}_j$-measurable. Hence $\mathbb{E}(X(j)Z_{j-1}^*)=\mathbb{E}(Z_{j-1}^*\mathbb{E}_j X(j))=0$. It follows that

$$\sum_{j=1}^n\big|\mathbb{E}\big(X(j)S_{j-1}^*\big)\big|=\sum_{j=1}^n\big|\mathbb{E}\big(X(j)\big(S_{j-1}^*-Z_{j-1}^*\big)\big)\big|\le|w|_\infty\sum_{j=1}^n\mathbb{E}\big|S_{j-1}^*-Z_{j-1}^*\big|.$$

By Proposition 3.8,

$$\big|S^*_{j-1} - Z^*_{j-1}\big| \le \sum_{i=1}^{j-1} \big|X(i) - Y_{i,j}\big| = \sum_{i=1}^{j-1} |w \circ T^{-i} - (\mathbb{E}_{j-i} w) \circ T^{-i}|$$

and hence

$$\begin{aligned}\sum_{j=1}^{n} \big|\mathbb{E}\big(X(j) S^*_{j-1}\big)\big| &\le |w|_\infty \sum_{1\le i<j\le n} |w - \mathbb{E}_{j-i} w|_1 = |w|_\infty \sum_{j=1}^{n-1} (n-j)|w - \mathbb{E}_j w|_1 \\ &\le n|w|_\infty \sum_{j=1}^{\infty} |\mathbb{E}_j w - w|_1 = n|w|_\infty \sum_{j=1}^{\infty} |\big(\mathbb{E}_0\big(w \circ T^j\big)\big) \circ T^{-j} - w|_1 \\ &= n|w|_\infty \sum_{j=1}^{\infty} \big|\mathbb{E}_0\big(w \circ T^j\big) - w \circ T^j\big|_1.\end{aligned}$$

Combining this with the estimate for $\mathbb{E}S_n^2$ it follows from Lemma 3.7 that

$$\mathbb{E}\big(S_n^{*2}\big) \le 12n|w|_\infty \sum_{j=0}^{\infty} \big|\mathbb{E}_0\big(w \circ T^j\big) - w \circ T^j\big|_1.$$

The transformation $w \mapsto -w$ sends $S_n^* \mapsto S_{n,*} = \max\{0, -S_1, \dots, -S_n\}$. Hence

$$\mathbb{E}(S_{n,*}^2) \le 12n|w|_\infty \sum_{j=0}^{\infty} |\mathbb{E}_0(w \circ T^j) - w \circ T^j|_1,$$

and so

$$\max_{1\le\ell\le n} |S_\ell|^2 = \max\big\{S_n^{*2}, S_{n,*}^2\big\} \le S_n^{*2} + S_{n,*}^2 \le 24\, n|w|_\infty \sum_{j=0}^{\infty} \big|\mathbb{E}_0\big(w \circ T^j\big) - w \circ T^j\big|_1.$$

Finally, $w_\ell = (S_n - S_{n-\ell}) \circ T^n$, so $|\max_{1\le\ell\le n} |w_\ell||_2 \le 2|\max_{1\le\ell\le n} |S_\ell||_2$ and the result follows. □

Corollary 3.10. $|\max_{1\le\ell\le n} |(a_k)_\ell||_2^2 \le 192\, n|v|_\infty \sum_{j=k}^{\infty} |\mathbb{E}_0(v \circ T^j) - v \circ T^j|_1$.

Proof. Recall that $a_k = \mathbb{E}_0(v \circ T^k) - v \circ T^k$, so $|a_k|_\infty \le 2|v|_\infty$ and $\mathbb{E}_0 a_k = 0$. By Proposition 3.9,

$$\Big|\max_{1\le\ell\le n} \big|(a_k)_\ell\big|\Big|_2^2 \le 192\, n|v|_\infty \sum_{j=0}^{\infty} \big|\mathbb{E}_0\big(a_k \circ T^j\big) - a_k \circ T^j\big|_1.$$

Setting $g = v \circ T^k$, $a_k = \mathbb{E}_0 g - g$,

$$\begin{aligned}\mathbb{E}_0\big(a_k \circ T^j\big) - a_k \circ T^j &= \mathbb{E}_0\big((\mathbb{E}_0 g) \circ T^j\big) - \mathbb{E}_0\big(g \circ T^j\big) - (\mathbb{E}_0 g) \circ T^j + g \circ T^j \\ &= (\mathbb{E}_j \mathbb{E}_0 g) \circ T^j - \mathbb{E}_0\big(g \circ T^j\big) - (\mathbb{E}_0 g) \circ T^j + g \circ T^j \\ &= -\mathbb{E}_0\big(g \circ T^j\big) + g \circ T^j = -\mathbb{E}_0\big(v \circ T^{j+k}\big) + v \circ T^{j+k}.\end{aligned}$$

The result follows. □

Lemma 3.11. $\lim_{n\to\infty} \frac{1}{\sqrt{n}} |\max_{1\le\ell\le n} |(v-m)_\ell||_2 = 0$.

Proof. By Lemma 3.4 and hypothesis (3.1), for each $\epsilon > 0$, there exists $k \ge 1$ such that

$$\big|m - m^{(k)}\big|_2 < \epsilon, \qquad \sum_{j=k}^{\infty} \big|\mathbb{E}_0\big(v \circ T^{-j}\big)\big|_1 < \epsilon^2, \qquad \sum_{j=k+1}^{\infty} \big|\mathbb{E}_0\big(v \circ T^j\big) - v \circ T^j\big|_1 < \epsilon^2.$$

Since $\mathbb{E}_{-1}m = \mathbb{E}_{-1}m^{(k)} = 0$, it follows from Doob's inequality as in Corollary 2.8 that

$$\Big|\max_{1\le\ell\le n}\big|\big(m-m^{(k)}\big)_\ell\big|\Big|_2 < 4\sqrt{n}\,\epsilon. \tag{3.4}$$

By (3.3),

$$\big|\big(v-m^{(k)}\big)_n\big| \le 2\big|\chi_{-k}^k\big|_\infty + \big|(a_{-k})_n\big| + \big|(a_{k+1})_n\big| \le (6k+2)|v|_\infty + \big|(a_{-k})_n\big| + \big|(a_{k+1})_n\big|. \tag{3.5}$$

Substituting the estimates from Corollaries 3.6 and 3.10 into (3.5), $|\max_{1\le\ell\le n}|(v-m^{(k)})_\ell||_2 \ll k+\epsilon n^{1/2}$, and combining this with (3.4),

$$\frac{1}{\sqrt{n}}\Big|\max_{1\le\ell\le n}\big|(v-m)_\ell\big|\Big|_2 \ll kn^{-1/2}+\epsilon.$$

Hence $\limsup_{n\to\infty}\frac{1}{\sqrt{n}}|\max_{1\le\ell\le n}|(v-m)_\ell||_2 \ll \epsilon$ and the result follows since ϵ is arbitrary. □

We require the following standard result from probability theory.

Proposition 3.12. *Let $Y_1, Y_2\ldots$ be identically distributed random variables with finite second moment. Then $|\max_{1\le\ell\le n}|Y_\ell||_2 = o(\sqrt{n})$ as $n\to\infty$.* □

Proof of Theorem 3.2. Conclusions (a) and (b) hold by the same arguments in the proof of Theorem 2.1 (using Lemma 3.11 in place of Lemma 2.10).

Fix $K \ge 1$ to be an integer and define $W_n^-(t)$ and $M_n^-(t)$ for $t\in[0,K]$ as in (2.7). Also, for $t\ge 0$ define

$$\widetilde{M}_n^-(t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{[nt]} m\circ T^{-j}.$$

Note that $\{m\circ T^{-n};\, n\in\mathbb{Z}\}$ is a martingale difference sequence with respect to the filtration $\mathcal{F}_n$. By Proposition 3.12, $\frac{1}{\sqrt{n}}|\max_{1\le j\le n}|m\circ T^{-j}||_2\to 0$. Also, by the ergodic theorem $\frac{1}{n}\sum_{j=1}^n (m\otimes m)\circ T^{-j}\to\int_\Lambda m\otimes m\,d\mu=\Sigma$ almost everywhere. Hence we have verified the hypotheses of [48, Theorem 2.1], yielding $\widetilde{M}_n^-\to_w W$ in $D[0,\infty),\mathbb{R}^d)$. Since $M_n^- = \widetilde{M}_n^-\circ T^{nK}$, it follows that $M_n^-\to_w W$ in $D[0,K],\mathbb{R}^d)$. By (2.8) and Lemma 3.11, $W_n^-\to_w W$ in $D[0,K],\mathbb{R}^d)$ on (Λ,μ).

Defining g as in Step 1 of the proof of Lemma 2.12, we obtain

$$W_n(t) = g\big(W_n^-(t)\big) - F_n^1(t) \quad \text{for } t\in[0,K],$$

where $\sup_{t\in[0,K]}|F_n^1(t)|\le n^{-1/2}|v|_\infty$. Applying [23, Proposition 4.9 and Lemma 4.11], $W_n\to_w g(W)=_w W$ in $D[0,K],\mathbb{R}^d)$ on (Λ,μ).

Finally, the case where ν is a general probability measure absolutely continuous with respect to μ follows from [52, Corollary 3]. □

3.2. *Iterated WIP in the invertible setting*

Define $W_n\in D([0,\infty),\mathbb{R}^d)$, $\mathbb{W}_n\in D([0,\infty),\mathbb{R}^{d\times d})$ as in (2.2). Let ν be any probability measure on Λ absolutely continuous with respect to μ. In this subsection, we prove:

Theorem 3.13. *Let $v\in L^\infty$ with $\int_\Lambda v\,d\mu=0$, and suppose that (3.2) holds. Assume also that T is mixing. Then $(W_n,\mathbb{W}_n)\to_w(W,\mathbb{W})$ in $D([0,\infty),\mathbb{R}^d\times\mathbb{R}^{d\times d})$ as $n\to\infty$ on (Λ,ν), where W is as in Theorem 3.2 and*

$$\mathbb{W}(t) = \int_0^t W\otimes dW + t\sum_{j=1}^\infty\int_\Lambda v\otimes\big(v\circ T^j\big)\,d\mu.$$

Write

$$\chi = \chi_- + \chi_+, \qquad \chi_- = \sum_{j=1}^{\infty} \mathbb{E}_0\big(v \circ T^{-j}\big), \qquad \chi_+ = \sum_{j=0}^{\infty} \big(\mathbb{E}_0\big(v \circ T^j\big) - v \circ T^j\big).$$

By (3.2), $\chi_- \in L^1$ and $\chi_+ \in L^2$. Define $\hat{v} \in L^2$ by

$$v = \hat{v} + \chi_+ \circ T - \chi_+.$$

Then

$$\hat{v} = \sum_{j=0}^{\infty} \big\{\mathbb{E}_0\big(v \circ T^j\big) - \big(\mathbb{E}_0\big(v \circ T^j\big)\big) \circ T\big\}$$

is $\mathcal{F}_0$-measurable.

Define

$$\widehat{\mathbb{M}}_n \in D\big([0,\infty), \mathbb{R}^{d\times d}\big), \qquad \widehat{\mathbb{M}}_n(t) = \frac{1}{n} \sum_{0\le i<j\le [nt]-1} \big(m \circ T^i\big) \otimes \big(\hat{v} \circ T^j\big),$$

where m is as in (3.3). (This differs from the definition of $\mathbb{M}_n$ in Section 2; we use $\hat{v}$ instead of v since v is not $\mathcal{F}_0$-measurable.)

Lemma 3.14. *$(W_n, \widehat{\mathbb{M}}_n) \to_w (W, \mathbb{M})$ in $D([0,\infty), \mathbb{R}^d \times \mathbb{R}^{d\times d})$ as $n \to \infty$ on (Λ, μ), where $\mathbb{M}(t) = \int_0^t W \otimes dW$.*

Proof. Fix $K \ge 1$ an integer and define W_n^-, M_n^- as in (2.7). As shown in the proof of Theorem 3.2, $M_n^- \to_w W$ in $D([0,[0,K],\mathbb{R}^d)$. By the continuous mapping theorem, $(M_n^-, M_n^-) \to_w (W, W)$ in $D([0,K]), \mathbb{R}^d \times \mathbb{R}^d)$.

Define

$$\widehat{W}_n(t) = \frac{1}{n} \sum_{0\le j\le [nt]-1} \hat{v} \circ T^j, \qquad \widehat{W}_n^-(t) = \frac{1}{n} \sum_{1\le j\le [nt]} \hat{v} \circ T^{nK-j}.$$

By Lemma 3.11, $\frac{1}{\sqrt{n}} |\max_{1\le \ell\le n} |(v-m)_\ell||_2 \to 0$. Also, $\frac{1}{\sqrt{n}} |\max_{1\le \ell\le n} |(v-\hat{v})_\ell||_2 \le \frac{2}{\sqrt{n}} |\max_{1\le \ell\le n} \chi_+ \circ T^\ell|_2 \to 0$ by Proposition 3.12. Hence $\frac{1}{\sqrt{n}} |\max_{1\le \ell\le n} |(\hat{v}-m)_\ell||_2 \to 0$. It follows that

$$\big(\widehat{W}_n^-, M_n^-\big) \to_w (W, W) \quad \text{in } D\big([0,K], \mathbb{R}^d \times \mathbb{R}^d\big).$$

Define

$$\widehat{\mathbb{M}}_n^-(t) = \frac{1}{n} \sum_{1\le j<i\le [nt]} \big(\hat{v} \circ T^{-j}\big) \otimes \big(m \circ T^{-i}\big).$$

We apply [21,28] as in Step 2 of the proof of Lemma 2.12: M_n^- is a martingale and $\widehat{W}_n^-$ is adapted with respect to the filtration $\mathcal{F}_j$. Moreover, $\int_\Lambda |M_n^-(t)|^2\, d\mu = n^{-1}[nt] \int_\Lambda |m|^2\, d\mu \le K|m|_2^2$ for all $t \in [0,K]$ so condition C2.2(i) in [28, Theorem 2.2] is satisfied. Hence

$$\big(\widehat{W}_n^-, M_n^-, \widehat{\mathbb{M}}_n^-\big) \to_w (W, W, \mathbb{M}) \quad \text{in } D\big([0,K], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}\big).$$

Next, define $h : D([0,K], \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d\times d}) \to \widetilde{D}([0,K], \mathbb{R}^d \times \mathbb{R}^{d\times d})$, as in Step 3 of the proof of Lemma 2.12. Then

$$(\widehat{W}_n, \widehat{\mathbb{M}}_n) = h\big(\widehat{W}_n^-, M_n^-, \widehat{\mathbb{M}}_n^-\big) - F_n \quad \text{where } \sup_{t\in[0,K]} \big|F_n(t)\big| \to_p 0,$$

and we deduce that

$$(\widehat{W}_n, \widehat{\mathbb{M}}_n) \to_w (W, \mathbb{M}) \quad \text{in } D\big([0,K], \mathbb{R}^d \times \mathbb{R}^{d\times d}\big).$$

Using once again that $\frac{1}{\sqrt{n}} |\max_{1\le \ell\le n} |(v-\hat{v})_\ell||_2 \to 0$, we obtain the desired result. □

Proposition 3.15. $\sum_{j=1}^{\infty}\int_\Lambda v\otimes(v\circ T^j)\,d\mu=\int_\Lambda(\chi\otimes v-m\otimes(\chi_+\circ T))\,d\mu$.

Proof. Write

$$\begin{aligned}v\otimes\big(v\circ T^j\big)&=(m+\chi\circ T-\chi)\otimes\big(v\circ T^j\big)\\&=m\otimes\big(\hat v\circ T^j+\chi_+\circ T^{j+1}-\chi_+\circ T^j\big)+(\chi\circ T-\chi)\otimes\big(v\circ T^j\big).\end{aligned}$$

Then $\sum_{j=1}^n\int_\Lambda v\otimes(v\circ T^j)\,d\mu=I_1+I_2+I_3$ where

$$I_1=\sum_{j=1}^n\int_\Lambda m\otimes\big(\hat v\circ T^j\big)\,d\mu,\qquad I_2=\int_\Lambda m\otimes\big(\chi_+\circ T^{n+1}-\chi_+\circ T\big)\,d\mu,$$

$$I_3=\int_\Lambda\sum_{j=1}^n(\chi\circ T-\chi)\otimes v\circ T^j)\,d\mu.$$

Now,

$$\mathbb{E}\big(m\otimes\big(\hat v\circ T^j\big)\big)=\mathbb{E}\mathbb{E}_{-j}\big(m\otimes\big(\hat v\circ T^j\big)\big)=\mathbb{E}\big((\mathbb{E}_{-j}m)\otimes\big(\hat v\circ T^j\big)\big)=0,$$

so $I_1=0$. Since T is mixing, $I_2\to-\int_\Lambda m\otimes(\chi_+\circ T)\,d\mu$. Finally,

$$I_3=\int_\Lambda\sum_{j=1}^n\big(\chi\circ T^{-(j-1)}-\chi\circ T^{-j}\big)\otimes v\,d\mu=\int_\Lambda\big(\chi-\chi\circ T^{-n}\big)\otimes v\,d\mu\to\int_\Lambda\chi\otimes v\,d\mu$$

since T is mixing. □

Proof of Theorem 3.13. Write

$$\begin{aligned}\big(v\circ T^i\big)\otimes\big(v\circ T^j\big)&=\big(m\circ T^i\big)\otimes\big(v\circ T^j\big)+\big(\chi\circ T^{i+1}-\chi\circ T^i\big)\otimes\big(v\circ T^j\big)\\&=\big(m\circ T^i\big)\otimes\big(\hat v\circ T^j\big)+\big(m\circ T^i\big)\otimes\big(\chi_+\circ T^{j+1}-\chi_+\circ T^j\big)+\big(\chi\circ T^{i+1}-\chi\circ T^i\big)\otimes\big(v\circ T^j\big).\end{aligned}$$

Then

$$\begin{aligned}\mathbb{W}_n(t)-\widehat{\mathbb{M}}_n(t)&=\frac1n\sum_{0\le i\le[nt]-2}\big(m\circ T^i\big)\otimes\big(\chi_+\circ T^{[nt]}-\chi_+\circ T^{i+1}\big)\\&\quad+\frac1n\sum_{1\le j\le[nt]-1}\big(\chi\circ T^j-\chi\big)\otimes\big(v\circ T^j\big)=\frac1n\big(A_n(t)+B_n(t)\big)\end{aligned}$$

where

$$A_n(t)=\sum_{0\le i\le[nt]-2}\big(m\circ T^i\big)\otimes\big(\chi_+\circ T^{[nt]}\big)$$

$$B_n(t)=-\sum_{0\le i\le[nt]-2}\big(m\otimes(\chi_+\circ T)\big)\circ T^i+\sum_{1\le j\le[nt]-1}(\chi\otimes v)\circ T^j-\chi\otimes\sum_{1\le j\le[nt]-1}v\circ T^j.$$

Recall that $v\in L^\infty$, $\chi\in L^1$ and χ_+, $m\in L^2$. By the ergodic theorem, $\frac1nB_n(1)\to\int_\Lambda(\chi\otimes v-m\otimes(\chi_+\circ T))\,d\mu$a.e. and hence

$$\frac1n\sup_{t\in[0,K]}\left|B_n(t)-t\int_\Lambda\big(\chi\otimes v-m\otimes(\chi_+\circ T)\big)\,d\mu\right|\to0\quad\text{a.e.}$$

Also,

$$\frac1n\Big|\sup_{t\in[0,K]}A_n(t)\Big|_1\le\frac{1}{\sqrt n}\left|\max_{1\le\ell\le nK}\left|\sum_{i<\ell}m\circ T^i\right|\right|_2\frac{1}{\sqrt n}\left|\max_{1\le\ell\le nK}\chi_+\circ T^\ell\right|_2\le2\sqrt K|m|_2\frac{1}{\sqrt n}\left|\max_{1\le\ell\le nK}\chi_+\circ T^\ell\right|_2\to0$$

by Doob's inequality and Proposition 3.12. Hence

$$\sup_{t\in[0,K]}\left|\mathbb{W}_n(t)-\widehat{\mathbb{M}}_n(t)-t\int_\Lambda\big(\chi\otimes v-m\otimes(\chi_+\circ T)\big)\,d\mu\right|\to_p 0.$$

By this combined with Lemma 3.14 and Proposition 3.15, $(W_n,\mathbb{W}_n)\to_w(W,\mathbb{W})$ on (Λ,μ).

Finally, we consider the case where ν is a general probability measure absolutely continuous with respect to μ. As in the proof of Theorem 2.1, it suffices to establish (2.9) for all $\epsilon>0$. The estimates

$$\big|W_n(t)\circ T-W_n(t)\big|\le 2n^{-1/2}|v|_\infty,\qquad \big|\mathbb{W}_n(t)\circ T-\mathbb{W}_n(t)\big|\le 2n^{-1}|v|_\infty\max_{1\le k\le nK}|v_k|$$

hold as before for all $t\in[0,K]$. Hence it suffices to show that $|\max_{1\le k\le nK}|v_k||_2\ll n^{1/2}$.

Write $v=w+w'$ where $w=\mathbb{E}_0v$, $w'=v-\mathbb{E}_0v$. By Propositions 3.5 and 3.9,

$$\Big|\max_{1\le k\le nK}|w_k|\Big|_2\ll\left(n|v|_\infty\sum_{j\ge0}\big|\mathbb{E}_0\big(w\circ T^{-j}\big)\big|_1\right)^{1/2}=\left(n|v|_\infty\sum_{j\ge0}\big|\mathbb{E}_0\big(v\circ T^{-j}\big)\big|_1\right)^{1/2}$$

and

$$\Big|\max_{1\le k\le nK}\big|w_k'\big|\Big|_2\ll\left(n|v|_\infty\sum_{j\ge0}\big|\mathbb{E}_0\big(w'\circ T^j\big)-w'\circ T^j\big|_1\right)^{1/2}=\left(n|v|_\infty\sum_{j\ge0}\big|\mathbb{E}_0\big(v\circ T^j\big)-v\circ T^j\big|_1\right)^{1/2}.$$

Hence the required estimate for $\max_{1\le k\le nK}|v_k|$ follows from (3.1). □

4. Examples

In this section, we consider examples consisting of time-one maps of nonuniformly expanding semiflows and nonuniformly hyperbolic flows to which our theory applies and gives new results.

4.1. *Noninvertible setting*

We begin by revisiting nonuniformly expanding maps modelled by one-sided Young towers [51]. Optimal results for the iterated WIP were obtained by [23] and we recover their result. In particular, [51] proved results on decay of correlations; Theorem 2.1 applies whenever the decay of correlations is summable by Proposition 2.5. As described below, we are moreover able to treat time-one maps of nonuniformly expanding semiflows, significantly improving on existing results.

It is convenient to mention a specific family of dynamical systems. Prototypical examples of nonuniformly expanding map are given by intermittent maps of Pomeau-Manneville type [43]. For definiteness, we consider the example considered by [30], namely

$$f:[0,1]\to[0,1],\qquad f(x)=\begin{cases}x(1+2^\gamma x^\gamma) & x<\frac12\\ 2x-1 & x>\frac12.\end{cases}\tag{4.1}$$

Here $\gamma\in[0,1)$ is a parameter and there is a unique absolutely continuous invariant probability measure μ_0 for each γ.

Let $v:[0,1]\to\mathbb{R}^d$ be Hölder with $\int_0^1 v\,d\mu_0=0$. By [20,51], there is a constant $C>0$ such that $|\int_0^1 v\,w\circ T^m\,d\mu_0|\le Cn^{-(\gamma^{-1}-1)}|w|_\infty$ for all $w\in L^\infty([0,1],\mathbb{R})$. By Proposition 2.5, hypothesis (2.1) holds for $\gamma<\frac12$. Hence we obtain the iterated WIP, Theorem 2.1, for all $\gamma<\frac12$. This recovers a result of [23, Example 10.3] and it is sharp since even the CLT fails for $\gamma\in[\frac12,1)$ when $v(0)\ne0$ by [17].

Now we consider suspension semiflows and their time-one maps to obtain new examples where the iterated WIP holds. Again we consider the specific example (4.1) for definiteness, but f could be replaced by any nonuniformly expanding map modelled by a Young tower. Let $h:[0,1]\to(0,\infty)$ be a Hölder roof function and define the suspension semiflow $f_t:\Lambda\to\Lambda$ where

$$\Lambda=\big\{(x,u)\in[0,1]\times\mathbb{R}:0\le u\le h(x)\big\}/\sim,\quad \big(x,h(x)\big)\sim(fx,0),$$

and $f_t(x,u)=(x,u+t)$ computed modulo identifications. The probability measure $\mu=(\mu_0\times\mathrm{Leb})/\int_0^1 h\,d\mu_0$ is f_t-invariant and ergodic. At the level of the semiflow f_t, when $\gamma<\frac{1}{2}$ the iterated WIP holds for Hölder mean zero observables v by [23, Theorem 6.1].

Now consider the time-one map $T=f_1:\Lambda\to\Lambda$. In general, even the CLT is not known for such maps. By [33,35], typically (under a non-approximate eigenfunction condition due to [7]) T has decay of correlations at the same rate as f for sufficiently smooth observables. (We refer to [35, Section 3] for details regarding the class of observables v and [35, Section 5] for details regarding the word "typical".) A consequence [23, Theorem 4.3 and Proposition 4.4] is that the iterated WIP holds for $\gamma<\frac{1}{3}$. Previously, the range $\gamma\in[\frac{1}{3},\frac{1}{2})$ remained open. But Proposition 2.5 again implies that the L^1 Gordin criterion (2.1) holds for all $\gamma<\frac{1}{2}$. Hence, by Theorem 2.1, the iterated WIP holds in the optimal range $\gamma<\frac{1}{2}$.

4.2. *Invertible setting*

We begin by revisiting nonuniformly hyperbolic maps modelled by two-sided Young towers [50,51]. Optimal results in this setting were obtained by [40] (see also [23, Section 10.2]). (Unlike in the noninvertible setting, Theorem 3.13 does not recover this result since the iterated WIP is not known under the L^1 Gordin criterion (3.1).)

Examples of nonuniformly hyperbolic maps include intermittent solenoids [2, Section 5] and [40, Example 4.2]. These are invertible analogues of the intermittent maps in Section 4.1 and are obtained by adapting the classical Smale-Williams solenoid construction [45,49]. There is an invariant contracting stable foliation $\mathcal{W}^s$ as in Proposition 3.1 and the dynamics modulo the stable leaves is given by an intermittent map. In particular, condition (a) in Proposition 3.1 is satisfied for $p<\gamma^{-1}-1$ where γ is the parameter for the intermittent map. The examples in [2] and some of the examples in [40] have exponential contraction along stable leaves. For these examples and v Hölder, condition (b) in Proposition 3.1 is satisfied for all $p\in[1,\infty)$ and hence Theorem 3.13 applies for all $\gamma<\frac{1}{2}$. The remaining examples in [40] have contraction along stable leaves which is as slow as the expansion of the underlying intermittent map, and condition (b) in Proposition 3.1 is satisfied for $p<\gamma^{-1}-1$; hence Theorem 3.13 applies for $\gamma<\frac{1}{3}$.

As in Section 4.1, we consider intermittent solenoidal flows given by suspensions over intermittent solenoids. Optimal results on the iterated WIP for such flows follow by combining [40] and [23, Theorem 6.1]. Again we focus on time-one maps of intermittent solenoid flows, restricting to typical flows and sufficiently smooth observables. Previous results on the iterated WIP in this context apply only for $\gamma<\frac{1}{3}$; we considerably relax this restriction. By the arguments in [6], the conditions of Proposition 3.1 hold for the same values of p as in the case of intermittent solenoids described above. Hence for the examples with exponential contraction along stable leaves, we obtain the iterated WIP for all $\gamma<\frac{1}{2}$. Exponential contraction can be relaxed to moderately fast polynomial contraction as discussed in [1]. Explicit examples with $\gamma\in[\frac{1}{3},\frac{1}{2})$ and condition (b) of Proposition 3.1 holding for $p=2$ can be found in [12]. For such examples, the iterated WIP follows from Theorem 3.13; this is far beyond the scope of previous methods.

5. Application to homogenisation

Let $(\Lambda,\mathcal{F},\mu)$ be a probability space and $T:\Lambda\to\Lambda$ be an ergodic measure-preserving map. Consider the fast-slow system

$$x_{n+1}^{(\epsilon)}=x_n^{(\epsilon)}+\epsilon^2 a\big(x_n^{(\epsilon)}\big)+\epsilon b\big(x_n^{(\epsilon)}\big)v(y_n),\quad x_0^{(\epsilon)}=\xi\in\mathbb{R}^d,$$
$$y_{n+1}=Ty_n,$$

where $a:\mathbb{R}^d\to\mathbb{R}^d$, $b:\mathbb{R}^d\to\mathbb{R}^{d\times d}$ and $v\in L^\infty(\Lambda,\mathbb{R}^d)$ with $\int_\Lambda v\,d\mu=0$.

Define $\hat{x}_\epsilon(t)=x_{[t/\epsilon^2]}^{(\epsilon)}$ and

$$W_\epsilon(t)=\epsilon\sum_{0\le j\le[t/\epsilon^2]-1}v(y_j),\qquad \mathbb{W}_\epsilon(t)=\epsilon^2\sum_{0\le i<j\le[t/\epsilon^2]-1}v(y_i)\otimes v(y_j).$$

The aim is to prove homogenisation to a stochastic differential equation (SDE) of the type

$$dX=\tilde{a}(X)\,dt+b(X)\,dW,\quad X(0)=\xi,$$

where $\tilde{a}:\mathbb{R}^d\to\mathbb{R}^d$ is to be determined; i.e. to show that $\hat{x}_\epsilon\to_w X$ in $D[0,\infty),\mathbb{R}^d)$ as $\epsilon\to0$.

In the special case where a and b are Lipschitz, and b satisfies an exactness condition of the form $b=(dh)^{-1}$ for some $h:\mathbb{R}^d\to\mathbb{R}^d$, this problem was completely solved by [16]: It is necessary and sufficient that v satisfies the WIP. Hence the L^1 Gordin criterion suffices by Theorems 2.1 and 3.2.

When the exactness condition for b fails, [23,27] proved homogenisation for $a \in C^{1+}$ and $b \in C^{2+}$ under an L^4 Gordin criterion on v (see [6, Theorem 2.9]). Moreover, $\tilde{a}(X) = a(X) + \frac{1}{2}\sum_{\alpha,\beta,\gamma=1}^{d} E^{\gamma\beta}\partial_\alpha b^\beta(X) b^{\alpha\gamma}(X)$ where E is the matrix in the iterated WIP.[1]

In certain special cases, our results yield homogenisation theorems where the previous papers do not. One such example is the following:

Proposition 5.1. *Let $d = 2$ and write $x = (x^1, x^2)$. Let*

$$a(x) = \begin{pmatrix} 0 \\ g(x) \end{pmatrix}, \qquad b(x) = \begin{pmatrix} 1 & 0 \\ 0 & x^1 \end{pmatrix},$$

where $g : \mathbb{R}^2 \to \mathbb{R}$ is Lipschitz. Suppose that v satisfies either the L^1 Gordin criterion (2.1) in the noninvertible setting or the hybrid L^1–L^2 Gordin criterion (3.2) in the invertible setting. In particular, $\mathbb{W}_\epsilon^{12}(1) \to_w \mathbb{W}^{12}(1) + c$ for some $c \in \mathbb{R}$.

Then $\hat{x}_\epsilon \to_w X$ in $D([0,\infty), \mathbb{R}^2)$ as $\epsilon \to 0$ where X is the solution to the SDE

$$dX = \tilde{a}(X)\,dt + b(X)\,dW, \quad X(0) = \xi,$$

with $\tilde{a}(X) = \begin{pmatrix} 0 \\ g(X)+c \end{pmatrix}$.

Proof. We have $\hat{x}_\epsilon^1(t) = \xi^1 + W_\epsilon^1(t)$ and

$$\begin{aligned}\hat{x}_\epsilon^2(t) &= \xi^2 + \epsilon^2 \sum_{j=0}^{[t\epsilon^{-2}]-1} g\big(x_j^{(\epsilon)}\big) + \epsilon \sum_{j=0}^{[t\epsilon^{-2}]-1} \big(x_j^{(\epsilon)}\big)^1 v^2(y_j) = \xi^2 + \int_0^t g\big(x_\epsilon(s)\big)\,dt + \int_0^t \hat{x}_\epsilon^1(s)\,dW_\epsilon^2(s) + A_\epsilon(t) \\ &= \xi^2 + \int_0^t g\big(x_\epsilon(s)\big)\,ds + \int_0^t \big(\xi^1 + W_\epsilon^1(s)\big)\,dW_\epsilon^2(s) + A_\epsilon(t) \\ &= \xi^2 + \int_0^t g\big(x_\epsilon(s)\big)\,ds + \xi^1 W_\epsilon^2(t) + \mathbb{W}_\epsilon^{12}(t) + A_\epsilon(t)\end{aligned}$$

where

$$\big|A_\epsilon(t)\big| \le \epsilon^2|g|_\infty + \epsilon|v|_\infty \max_{0\le j\le \epsilon^{-2}t} \big|\big(x_j^{(\epsilon)}\big)^1\big| \le \epsilon^2|g|_\infty + \epsilon|v|_\infty\xi^1 + \epsilon|v|_\infty \max_{0\le s\le t}\big|W_\epsilon^1(s)\big|.$$

In other words,

$$\hat{x}_\epsilon(t) = \xi + \int_0^t a\big(x_\epsilon(s)\big)\,ds + U_\epsilon(t)$$

where

$$U_\epsilon = \begin{pmatrix} W_\epsilon^1 \\ \xi^1 W_\epsilon^2 + \mathbb{W}_\epsilon^{12} + A_\epsilon \end{pmatrix}.$$

The resulting solution map $\hat{x}_\epsilon = \mathcal{G}(U_\epsilon)$ is continuous on $C([0,K],\mathbb{R}^2)$ for all $K > 0$ since a is Lipschitz. By the iterated WIP, $U_\epsilon \to_w U$ where

$$U = \begin{pmatrix} W^1 \\ \xi^1 W^2 + \mathbb{W}^{12} \end{pmatrix}, \quad \mathbb{W}^{12}(t) = \int_0^t W^1\,dW^2 + tE^{12}, \quad E^{12} = c,$$

so the continuous mapping theorem shows that $\hat{x}_\epsilon \to_w \mathcal{G}(U) = X$ where

$$dX = a(X)\,dt + dU, \quad X(0) = \xi.$$

But $U^1 = W^1$ and

$$U^2(t) = \int_0^t \big(\xi^1 + W^1\big)\,dW^2 + tE^{12} = \int_0^t X^1\,dW^2 + tE^{12} = \int_0^t X^1\,dW^2 + tc,$$

[1] There is a typo in [23] and subsequent papers; the matrix entry $E^{\beta\gamma}$ should be replaced by $E^{\gamma\beta}$ as written here.

yielding $\tilde{a}(X) = \binom{0}{g(X)+c}$ as required. □

Acknowledgements

The authors are grateful to the referee for helpful comments and suggestions.

References

[1] J. F. Alves and D. Azevedo. Statistical properties of diffeomorphisms with weak invariant manifolds. *Discrete Contin. Dyn. Syst.* **36** (2016) 1–41. MR3369212 https://doi.org/10.3934/dcds.2016.36.1

[2] J. F. Alves and V. Pinheiro. Slow rates of mixing for dynamical systems with hyperbolic structures. *J. Stat. Phys.* **131** (2008) 505–534. MR2386574 https://doi.org/10.1007/s10955-008-9482-6

[3] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang In *Multiscale Systems, Homogenization, and Rough Paths. Probability and Analysis in Interacting Physical Systems: In Honor of S.R.S. Varadhan* 17–48. P. Friz et al (Eds). *Berlin, August, 2016". Springer Proceedings in Mathematics & Statistics* **283**. Springer, Berlin, 2019. MR3968507 https://doi.org/10.1007/978-3-030-15338-0_2

[4] I. Chevyrev, P. K. Friz, A. Korepanov, I. Melbourne and H. Zhang Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 2. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear, 2021. MR3968507 https://doi.org/10.1007/978-3-030-15338-0_2

[5] J. Dedecker and E. Rio. On the functional central limit theorem for stationary processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** (2000) 1–34. MR1743095 https://doi.org/10.1016/S0246-0203(00)00111-4

[6] M. Demers, I. Melbourne and M. Nicol. Martingale approximations and anisotropic Banach spaces with an application to the time-one map of a Lorentz gas. *Nonlinearity* **33** (2020) 4095–4113. MR4169659 https://doi.org/10.1088/1361-6544/ab7d22

[7] D. Dolgopyat. Prevalence of rapid mixing in hyperbolic flows. *Ergodic Theory Dynam. Systems* **18** (1998) 1097–1114. MR1653299 https://doi.org/10.1017/S0143385798117431

[8] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689. MR2034323 https://doi.org/10.1090/S0002-9947-03-03335-X

[9] D. Dolgopyat. Averaging and invariant measures. *Mosc. Math. J.* **5** (2005) 537–576, 742. MR2241812 https://doi.org/10.17323/1609-4514-2005-5-3-537-576

[10] P. K. Friz and M. Hairer. *A Course on Rough Paths: With an Introduction to Regularity Structures*. *Universitext*. Springer, Cham, 2014. MR3289027 https://doi.org/10.1007/978-3-319-08332-2

[11] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge University Press, Cambridge, 2010. MR2604669 https://doi.org/10.1017/CBO9780511845079

[12] M. Galton. Limit theorems for slowly mixing dynamical systems. Ph. D. Thesis, University of Warwick, 2020.

[13] M. I. Gordin. The central limit theorem for stationary processes. *Sov. Math., Dokl.* **10** (1969) 1174–1176. MR0251785

[14] M. I. Gordin. Central limit theorems for stationary processes without the assumption of finite variance. In *Abstracts of Communications, T.1:A-K. Pages 173–174, International Conference on Probability Theory and Mathematical Statistics, June 25–30, 1973, Vilnius*, 1973.

[15] M. I. Gordin and M. Peligrad. On the functional central limit theorem via martingale approximation. *Bernoulli* **17** (2011) 424–440. MR2797997 https://doi.org/10.3150/10-BEJ276

[16] G. A. Gottwald and I. Melbourne. Homogenization for deterministic maps and multiplicative noise. *Proc. R. Soc. Lond. Ser. A, Math. Phys. Sci.* **469** (2013), 20130201. MR3064670 https://doi.org/10.1098/rspa.2013.0201

[17] S. Gouëzel. Central limit theorem and stable laws for intermittent maps. *Probab. Theory Related Fields* **128** (2004) 82–122. MR2027296 https://doi.org/10.1007/s00440-003-0300-4

[18] C. C. Heyde. On the central limit theorem and iterated logarithm law for stationary processes. *Bull. Aust. Math. Soc.* **12** (1975) 1–8. MR0372954 https://doi.org/10.1017/S0004972700023583

[19] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* **180** (1982) 119–140. MR0656227 https://doi.org/10.1007/BF01215004

[20] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory Dynam. Systems* **24** (2004) 495–524. MR2054191 https://doi.org/10.1017/S0143385703000671

[21] A. Jakubowski, J. Mémin and G. Pagès. Convergence en loi des suites d'intégrales stochastiques sur l'espace $\mathbf{D}^1$ de Skorokhod. *Probab. Theory Related Fields* **81** (1989) 111–137. MR0981569 https://doi.org/10.1007/BF00343739

[22] G. Keller. Un théorème de la limite centrale pour une classe de transformations monotones per morceaux. *C. R. Math. Acad. Sci. Paris* **291** (1980) 155–158. MR0605005

[23] D. Kelly and I. Melbourne. Smooth approximation of stochastic differential equations. *Ann. Probab.* **44** (2016) 479–520. MR3456344 https://doi.org/10.1214/14-AOP979

[24] D. Kelly and I. Melbourne. Homogenization for deterministic fast-slow systems with multidimensional multiplicative noise. *J. Funct. Anal.* **272** (2017) 4063–4102. MR3626033 https://doi.org/10.1016/j.jfa.2017.01.015

[25] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104** (1986) 1–19. MR0834478

[26] 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. MR3795019 https://doi.org/10.1016/j.anihpc.2017.08.005

[27] A. Korepanov, Z. Kosloff and I. Melbourne Deterministic homogenization under optimal moment assumptions for fast-slow systems. Part 1. *Ann. Inst. Henri Poincaré Probab. Stat.* To appear, 2021.

[28] T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** (1991) 1035–1070. MR1112406

[29] C. Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems* 56–75. F. Ledrappier, J. Lewowicz and S. Newhouse (Eds) *Pitman Research Notes in Math.* **362**. Longman Group Ltd, Harlow, 1996. MR1460797

[30] C. Liverani, B. Saussol and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** (1999) 671–685. MR1695915 https://doi.org/10.1017/S0143385799133856
[31] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** (1998) 215–310. MR1654527 https://doi.org/10.4171/RMI/240
[32] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** (2000) 713–724. MR1782272 https://doi.org/10.1214/aop/1019160258
[33] I. Melbourne. Decay of correlations for slowly mixing flows. *Proc. Lond. Math. Soc.* **98** (2009) 163–190. MR2472164 https://doi.org/10.1112/plms/pdn028
[34] I. Melbourne. Large and moderate deviations for slowly mixing dynamical systems. *Proc. Amer. Math. Soc.* **137** (2009) 1735–1741. MR2470832 https://doi.org/10.1090/S0002-9939-08-09751-7
[35] I. Melbourne. Superpolynomial and polynomial mixing for semiflows and flows. *Nonlinearity* **31** (2018) R268–R316. MR3851312 https://doi.org/10.1088/1361-6544/aad309
[36] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Comm. Math. Phys.* **260** (2005) 131–146. MR2175992 https://doi.org/10.1007/s00220-005-1407-5
[37] I. Melbourne and M. Nicol. Large deviations for nonuniformly hyperbolic systems. *Trans. Amer. Math. Soc.* **360** (2008) 6661–6676. MR2434305 https://doi.org/10.1090/S0002-9947-08-04520-0
[38] I. Melbourne and A. Stuart. A note on diffusion limits of chaotic skew product flows. *Nonlinearity* **24** (2011) 1361–1367. MR2776125 https://doi.org/10.1088/0951-7715/24/4/018
[39] I. Melbourne and A. Török. Central limit theorems and invariance principles for time-one maps of hyperbolic flows. *Comm. Math. Phys.* **229** (2002) 57–71. MR1917674 https://doi.org/10.1007/s00220-002-0676-5
[40] 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. MR3471365 https://doi.org/10.1142/S0219493716600121
[41] F. Merlevède, M. Peligrad and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** (2006) 1–36. MR2206313 https://doi.org/10.1214/154957806100000202
[42] M. Peligrad and S. Utev. A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33** (2005) 798–815. MR2123210 https://doi.org/10.1214/009117904000001035
[43] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74** (1980) 189–197. MR0576270
[44] E. Rio. *Théorie asymptotique des processus aléatoires faiblement dépendants*. *Mathématiques & Applications (Berlin) [Mathematics & Applications]* **31**. Springer-Verlag, Berlin, 2000. MR2117923
[45] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967) 747–817. MR0228014 https://doi.org/10.1090/S0002-9904-1967-11798-1
[46] M. Tyran-Kamińska. An invariance principle for maps with polynomial decay of correlations. *Comm. Math. Phys.* **260** (2005) 1–15. MR2175987 https://doi.org/10.1007/s00220-005-1400-z
[47] D. Volný. Approximating martingales and the central limit theorem for strictly stationary processes. *Stochastic Process. Appl.* **44** (1993) 41–74. MR1198662 https://doi.org/10.1016/0304-4149(93)90037-5
[48] W. Whitt. Proofs of the martingale FCLT. *Probab. Surv.* **4** (2007) 268–302. MR2368952 https://doi.org/10.1214/07-PS122
[49] R. F. Williams. One-dimensional non-wandering sets. *Topology* **6** (1967) 473–487. MR0217808 https://doi.org/10.1016/0040-9383(67)90005-5
[50] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650. MR1637655 https://doi.org/10.2307/120960
[51] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188. MR1750438 https://doi.org/10.1007/BF02808180
[52] R. Zweimüller. Mixing limit theorems for ergodic transformations. *J. Theoret. Probab.* **20** (2007) 1059–1071. MR2359068 https://doi.org/10.1007/s10959-007-0085-y