Almost sure invariance principle for random dynamical systems via Gouëzel’s approach

D. Dragičević\(^1\)\(^*,\) and Y Hafouta\(^2\)

\(^1\) Department of Mathematics, University of Rijeka, Rijeka, Croatia
\(^2\) Department of Mathematics, The Ohio State University, Columbus, OH, United States of America

E-mail: ddragicevic@math.uniri.hr and yeor.hafouta@mail.huji.ac.il

Received 15 September 2020, revised 7 July 2021
Accepted for publication 14 July 2021
Published 20 August 2021

Abstract
We extend the spectral approach of S Gouëzel for the vector-valued almost sure invariance principle (ASIP) to certain classes of non-stationary sequences with a weaker control over the behaviour of the covariance matrices, assuming only linear growth. Then we apply this extension to obtain a quenched vector-valued ASIP for random perturbations of a fixed Anosov diffeomorphism as well as random perturbations of a billiard map associated to the periodic Lorentz gas. We also consider certain classes of random piecewise expanding maps.

Keywords: almost sure invariance principle, random dynamical systems, hyperbolic dynamical systems
Mathematics Subject Classification numbers: 37H99, 37C30, 60F17.

1. Introduction

1.1. Almost sure invariance principle (ASIP)

The almost sure invariance principle (ASIP) represents a powerful statistical tool. Given a sequence of vector-valued random variables \( A_0, A_1, A_1, \cdots \), it provides a coupling with an independent sequence of Gaussian random vectors \( Z_0, Z_1, Z_2, \cdots \) such that

\[
\left| \frac{1}{n} \sum_{j=0}^{n-1} A_j - \frac{1}{n} \sum_{j=0}^{n-1} Z_j \right| = o(s_n),
\]

\(^*\) Author to whom any correspondence should be addressed.
Recommended by Dr Vaughn Climenhaga.

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
where $s_n = \|S_n\|_2$, $S_n = \sum_{j=0}^{n-1} A_j$, and the $L^2$-norm of the sum $\sum_{j=0}^{n-1} Z_j$ has the form $s_n(1 + o(1))$. We stress that ASIP implies several other important limit theorems, such as the central limit theorem and the functional central limit theorem. We refer to [30] for a detailed discussion.

1.2. Spectral approach for ASIP

As will be discussed below, many proofs of the ASIP in the scalar case rely on an appropriate martingale approximation. In [19], Gouëzel developed a spectral method for proving the ASIP for classes of non-stationary random vectors which are bounded in $L^p$ for some $p > 4$, satisfying certain mixing assumptions of a ‘spectral type’ and having a certain control over the behaviour of the covariance matrices of $S_n - S_m$ for $n > m$, where $S_n := \sum_{k=0}^{n-1} A_k$. More recently, in [15] we have extended this approach to real-valued sequences $\{A_j\}$ with the property that the variance of $S_n - S_m$ grows linearly fast in $n - m$. We stress that the purpose for this extension was to handle certain induced random dynamical systems, and so apart from weakening the type of a covariance control, we have established the version of ASIP when $\|A_n\|_{L^p} = O(n^{1/p})$.

In this paper we present a second extension of Gouëzel’s approach (see theorem 2.1), proving the ASIP under the spectral mixing assumptions for sequences bounded in $L^p$ with the property that the covariance matrix of $S_n$ grows linearly fast in $n$. We emphasise that we do not impose any assumptions related to the behaviour of the covariance matrices of $S_n - S_m$. The purpose of this extension is to obtain the ASIP for several classes of random expanding or random hyperbolic dynamical systems (see theorem 4.18 and examples presented in section 5). We note that in [14] we have proved several other limit theorems (such as the central limit theorem, Berry–Esseen bounds, Edgeworth expansions, the local limit theorem and several versions of a large deviations principle) for vector-valued observables and essentially the same classes of random dynamical systems as studied in the present paper. However, ASIP remained out of reach of the techniques developed in [14]. Conversely, the results in the present paper imply only the versions of the central limit theorem discussed in [14], while all other results in [14] are completely independent of those developed in the present paper.

1.3. ASIP for deterministic dynamical systems

We emphasise that the ASIP has been widely studied for deterministic dynamical systems. We in particular mention the works of Field, Melbourne and Török [17] as well as Melbourne and Nicol [28, 29] (completed by Korepanov [25]), in which the authors obtained ASIP for wide classes of (nonuniformly) hyperbolic maps. In contrast to their approaches which relied on martingale techniques, Gouëzel’s [19] spectral approach works for vector-valued sequences, and was applied in [19] to certain classes of deterministic dynamical systems, with the property that the corresponding transfer operator has a spectral gap on an appropriate Banach space. In situations when this method is applicable, it was pointed out in [19] that it gives better error rates in ASIP from those obtained in [28, 29]. Finally, we mention the recent important papers by Cuny and Merlevède [9], Korepanov, Kosloff and Melbourne [27], Korepanov [26], as well as Cuny, Dedecker, Korepanov and Merlevède [7, 8] in which the authors further improved the error rates in ASIP for a wide class of (nonuniformly) hyperbolic deterministic dynamical systems.

3 Roughly speaking, it is assumed that $\text{Cov}(S_n - S_m) \asymp (n - m)\Sigma^2$ for some positive definite matrix $\Sigma^2$. 

6774
1.4. ASIP for random dynamical systems

A random dynamical system is generated by random compositions
\[ T^n_\omega := T^\sigma_{n-1}\omega \circ \cdots \circ T^\sigma_1\omega \circ T_\omega, \quad \omega \in \Omega, \quad n \in \mathbb{N}, \]

of maps \( T_\omega \) (acting on some space \( M \)) which are driven by an invertible, measure-preserving transformation \( \sigma \) on some probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \). We stress that random dynamical systems are key tools to model many natural phenomena, including the transport in complex environments such as in the ocean or the atmosphere [2].

To the best of our knowledge, the ASIP in the context of random dynamical systems was first discussed by Kifer [24]. Indeed, in [24] it was mentioned that the techniques developed there can be used to obtain the scalar-valued quenched ASIP for random expanding dynamics. More recently, the annealed ASIP was obtained for several classes of random dynamical systems [1, 31, 32]. In these works, the idea is to consider the transfer operator associated to the skew-product transformation (see (4.27)) and to show that it has a spectral gap on an appropriate space. Then, it remains to apply Gouëzel’s ASIP [19]. However, in order for this method to work, it is necessary to make strong (mixing) assumptions on the base space \( (\Omega, \mathcal{F}, \mathbb{P}) \) of the random dynamical system (in all mentioned papers, \( (\Omega, \mathcal{F}, \mathbb{P}) \) is a Bernoulli shift). On the other hand, in [21] the authors proved a scalar-valued ASIP for certain classes of sequential expanding or hyperbolic dynamical systems with some assumptions on the growth rates of the variances. The approach in [21] relied on an appropriate martingale approximation combined with the results from [9]. Moreover, in [11] (by building on the approach developed in [9, 21]) the authors have obtained the quenched scalar-valued ASIP for certain classes of random piecewise expanding dynamics without any mixing assumptions for the base space. Finally, we mention two recent papers [33, 34] by Su devoted to the ASIP for certain classes of random expanding maps and maps which admit a random tower extension. We note that while the latter two papers concern random dynamics with sub-exponential decay of correlations, the ASIP rates obtained there are not as good as the ones in [11] and the present paper.

1.5. Contributions of the present paper

As we have already mentioned in subsection 1.2, in the present paper we establish (see theorem 2.1) an appropriate modification of [19, theorem 1.3.], and use it to establish the quenched ASIP for several classes of random dynamical systems (see theorem 4.18 and examples given in section 5). More precisely, we consider the following three cases:

- maps \( T_\omega, \omega \in \Omega \) are Anosov diffeomorphisms on a compact Riemannian manifold \( M \) that belong to a sufficiently small neighborhood of a fixed Anosov diffeomorphism \( T \) on \( M \);
- maps \( T_\omega, \omega \in \Omega \) are suitable perturbations of a billiard map associated to the periodic Lorentz gas studied by Demers and Zhang [10];
- \( (T_\omega)_{\omega \in \Omega} \) is a family of piecewise expanding maps on the unit interval satisfying appropriate conditions as in [11, 12].

For a sufficiently regular random vector-valued observable \( g_\omega : X \rightarrow \mathbb{R}^d, \omega \in \Omega \), our quenched ASIP implies that for \( \mathbb{P}\text{-a.e. } \omega \in \Omega \), the random Birkhoff sums \( \sum_{j=0}^{n-1} g_{\sigma^j\omega} \circ T^j_\omega \) can be approximated in the strong sense by a sum of Gaussian independent random vectors \( \sum_{j=0}^{n-1} Z_j \) (depending on \( \omega \)), with the error being negligible compared to \( n^{\frac{1}{2}} \). In fact, we will show that for every \( \epsilon > 0 \), the error term is at most \( o(n^{1/4+\epsilon}) \). In particular, we extend the scalar ASIP for
random expanding maps from [11, 21] to vector-valued observables. Furthermore, we for the first time obtain the quenched ASIP for some classes of random hyperbolic dynamics.

Finally, let us explain why our modification of [19, theorem 1.3] is needed. We observe that [19, theorem 1.3] can be applied when the asymptotic covariance matrix

\[ \Sigma^2 = \lim_{n \to \infty} \frac{1}{n} \text{Cov}(S_n) \]

exists and

\[ |\text{Cov}(S_n - S_m) - (n - m)\Sigma^2| \leq C(m - n)^\alpha, \] (1.1)

for some \( 0 < \alpha < 1 \) small enough. In [19], it is explained that this condition is satisfied for wide classes of deterministic dynamical systems with some expansion or hyperbolicity. However, from a general non-stationary point of view the above condition is strong. From a probabilistic point of view, random dynamical systems are examples of non-stationary processes with an asymptotic covariance matrix

\[ \Sigma^2 := \lim_{n \to \infty} \frac{1}{n} \text{Cov}_\omega \left( \sum_{j=0}^{n-1} g_{\sigma^j \omega} \circ T_j^\omega \right). \]

However, the rate of the convergence in the above expression depends on \( \omega \), since the covariance matrix of random Birkhoff sums is controlled by certain ergodic averages. Thus, (1.1) may fail to hold in a random setup.

Our modification of [19, theorem 1.3] is precisely tailored to overcome this difficulty. We stress that theorem 2.1 is completely different from [14, theorem 1], and that we expect that a combination of those two results will lead to an extension of [19, theorem 1.3] not only in the absence of (1.1), but also when the underlying sequence \( \{A_n\} \) satisfies \( \|A_n\|_p = O(n^{bp}) \) for some constant \( b_p > 0 \) and \( p > 4 \).

1.6. Organisation of the paper

The paper is organised as follows: in section 2 we formulate the main result of the present paper (see theorem 2.1), which gives an abstract ASIP result for certain classes of non-stationary sequences. The proof of theorem 2.1 is presented in section 3. In section 4, we give sufficient conditions under which theorem 2.1 can be applied in the context of random dynamics. In particular, we establish theorem 4.18 which gives an abstract version of the ASIP for random dynamical systems. Finally, in section 5 we discuss concrete examples of random dynamics to which theorem 4.18 is applicable.

2. Gouëzel’s ASIP in the nonstationary setup

The purpose of this section is to provide a certain modified version of [19, theorem 1.3].

Let \((A_1, A_2, \cdots)\) be an \( \mathbb{R}^d \)-valued process on some probability space \((\Omega, \mathcal{F}, \mathbb{P})\), where \( d \in \mathbb{N} \). We will denote the scalar product of two vectors \( t \) and \( v \) in \( \mathbb{R}^d \) by \( t \cdot v \), and when it is more convenient we will also abbreviate and just write \( tv \). We will also denote the Euclidean norm of a vector \( v \) by \( |v| \). We first recall the condition introduced in [19]: there exists \( \varepsilon_0 > 0 \) and \( C, c > 0 \) such that for any \( n, m \in \mathbb{N}, b_1 < b_2 < \cdots < b_{n+m+1}, k \in \mathbb{N} \) and \( t_1, \cdots, t_{n+m} \in \mathbb{R}^d \)
Theorem 2.1. Let \( \{A_n\} \) be a centered sequence of \( \mathbb{R}^d \)-valued random variables which is bounded in \( L^p \) for some \( p > 4 \), and satisfies property (2.1). Assume, in addition, that there exists a constant \( c_1 > 0 \) so that for every sufficiently large \( n \) and \( v \in \mathbb{R}^d \) we have that

\[
\text{Cov} \left( \sum_{j=1}^{n} A_j \right) v \cdot v \geq c_1 n |v|^2. \tag{2.2}
\]

Then, there exists a coupling between \( \{A_n\} \) and a sequence of independent and centered Gaussian \( d \)-dimensional random vectors \( Z_1, Z_2, \cdots \) such that for all \( \delta > 0 \) we have

\[
\left| \sum_{j=1}^{n} A_j - \sum_{j=1}^{n} Z_j \right| = o \left( n^{a_p+\delta} \right) \quad \mathbb{P} \text{-a.s.}, \tag{2.3}
\]

where \( a_p = \frac{p}{4(p-1)} = \frac{1}{4} + \frac{1}{4(p-1)} \). Moreover, there exists a constant \( C = C_\delta > 0 \) so that for every \( n \geq 1 \),

\[
\left\| \sum_{j=1}^{n} A_j - \sum_{j=1}^{n} Z_j \right\|_{L^2} \leq C n^{a_p+\delta}. \tag{2.4}
\]

Finally, if there are \( C_0 > 0 \) and \( r \in (0, 1) \) so that for every unit vector \( v \in \mathbb{R}^d \) and \( n, k \in \mathbb{N} \) we have

\[
|\text{Cov}(A_n \cdot v, A_{n+k} \cdot v)| \leq C_0 r^k, \tag{2.5}
\]

then there is a constant \( C = C_\delta > 0 \) so that for all \( n \geq 1 \) and a unit vector \( v \in \mathbb{R}^d \),

\[
\left\| \sum_{j=1}^{n} A_j \cdot v \right\|_{L^2}^2 - C_\delta n^{2a_p+\delta} \leq \left\| \sum_{j=1}^{n} Z_j \cdot v \right\|_{L^2}^2 \leq \left\| \sum_{j=1}^{n} A_j \cdot v \right\|_{L^2}^2 + C_\delta n^{2a_p+\delta}. \tag{2.6}
\]

Remark 2.2. In the scalar case \( d = 1 \), (2.4) and the linear growth of the variance of \( \sum_{j=1}^{n} A_j \) yields that the difference between the variances of \( \sum_{j=1}^{n} Z_j \) and \( \sum_{j=1}^{n} A_j \) is only \( O(n^{a_p+\delta/2}) \), which is close to \( O(n^{a_p+\delta}) \) when \( p \) is large. However, (2.6) yields that the difference between the variances is \( O(n^{2a_p+\delta}) \), which is close to \( O(n^{2\alpha_p+\delta}) \) when \( p \) is large. Moreover, using (2.6)
together with [23, theorem 3.2A], we conclude that under (2.5) in the scalar case, for any \( \delta > 0 \) there is a coupling of \( \{ A_n \} \) with a standard Brownian motion \( \{ W(t) : t \geq 0 \} \) such that

\[
\left| \sum_{j=1}^{n} A_n - W(V_n) \right| = o \left( n^{\frac{1}{2} + \frac{1}{2p} + \delta} \right) \quad \text{P-a.s.,}
\]

(2.7)

where \( V_n = \text{Var} \left( \sum_{j=1}^{n} A_j \right) \). Using only (2.4) and [23, theorem 3.2A], we could have only concluded that the left-hand side in (2.7) is of magnitude \( o(n^{d/2 + 1/2 + \delta}) \), which is close to \( n^{3/8} \) when \( p \) is large, and not to \( n^{1/2} \).

3. Proof of theorem 2.1

The proof is a modification of the proof of theorem 1.3 in [19].

We consider the so-called big and small blocks as introduced in [19, p 1659]. Fix \( \beta \in (0, 1) \) and \( \varepsilon \in (0, 1 - \beta) \). Furthermore, let \( f = f(n) = \lfloor \beta n \rfloor \). Then, Gouëzel decomposes \([2^n, 2^{n+1})\) into a union of \( F = 2^j \) intervals \((h_n, j)_0 \leq j < F\) of the same length, and \( F \) gaps \((J_n)_{0 \leq j < F}\) between them. In other words, we have

\[
[2^n, 2^{n+1}) = J_{n,0} \cup I_{n,0} \cup J_{n,1} \cup I_{n,1} \cup \cdots \cup J_{n,F-1} \cup I_{n,F-1}.
\]

Let us outline the construction of this decomposition. For \( 1 \leq j < F \), we write \( j \) in the form \( j = \sum_{k=0}^{r-1} \alpha_k(j)2^k \) with \( \alpha_k \in \{0, 1\} \). We then take the smallest \( r \) with the property that \( \alpha_r(j) \neq 0 \) and take \( 2^{[\alpha_r]}2^r \) to be the length of \( J_{n,j} \). In addition, the length of \( J_{n,0} \) is \( 2^{[\alpha]}2^f \). Finally, the length of each interval \( I_{n,j} \) is \( 2^{n-f} - (f + 2)2^{[\alpha]}-1 \).

In addition, we recall some notations from [19] which we will also use. We define a partial order on \( \{(n, j) : n \in \mathbb{N}, 0 \leq j < F(n)\} \) by writing \((n, j) \prec (n', j')\) if the interval \( I_{n,j} \) is to the left of \( I_{n',j'} \). Observe that a sequence \((n_k, j_k)\) tends to infinity for this order if and only if \( n_k \to \infty \). Moreover, let

\[
X_{n,j} := \sum_{t \in I_{n,j}} A_t
\]

and

\[
I := \bigcup_{n,j} I_{n,j} \quad \text{and} \quad J := \bigcup_{n,j} J_{n,j}.
\]

Finally, let us recall [19, proposition 5.1].

**Proposition 3.1.** There exists a coupling between \( (X_{n,j}) \) and a family of independent random vectors \( (Y_{n,j}) \) such that \( Y_{n,j} \) and \( X_{n,j} \) are equally distributed and almost surely, when \((n, j)\) tends to infinity,

\[
\left| \sum_{(n', j') \prec (n, j)} X_{n',j'} - Y_{n',j'} \right| = o(2^{j/2 + \varepsilon n/2}).
\]

(3.1)
The first modification (in comparison to [19]) that we need is a certain $L^2$-version of proposition 3.1.

**Proposition 3.2.** There exists a coupling between $(X_{n,j})$ and $(Y_{n,j})$ from proposition 3.1 such that (in addition to (3.1)), for some $C > 0$ and all $n \in \mathbb{N}$ we have that

$$\left\| \sum_{(i',j') \prec (n,j)} X_{n',j'} - Y_{n',j'} \right\|_{L^2} \leq C 2^{3n/2}. \tag{3.2}$$

**Proof.** We will show that we can couple $(X_{n,j})$ and $(Y_{n,j})$ so that

$$\left\| \sum_{j=0}^{F(n)-1} (X_{n,j} - Y_{n,j}) \right\|_{L^2} \leq C 2^{3n/2} \tag{3.3}$$

for every $n \in \mathbb{N}$, but it will be clear from the arguments in the proof that the same estimate holds true for the sum of $X_{n,j} - Y_{n,j}$, $j = 0, 1, \cdots, f$ where $f < F(n)$. Using this extended version of (3.2), it is clear that proposition 3.2 follows.

Let $\tilde{X}_{n,j} = X_{n,j} + V_{n,j}$ and $\tilde{Y}_{n,j} = Y_{n,j} + V_{n,j}'$ where the $V_{n,j}$'s and the $V_{n,j}'$'s are independent copies of the symmetric random vector $V$ constructed in [19, proposition 3.8], which are independent of everything else and from each other (enlarging our probability space if necessary).

We will first return to the arguments in step 1 of the proof of [19, theorem 1.3] (i.e. to the proof of proposition 3.1), and show that for every $s > 4$ there exists a constant $C_s > 0$ and a coupling between the $\tilde{X}_{n,j}$'s and the $\tilde{Y}_{n,j}$'s such that for all $0 \leq j < F(n)$ we have that

$$\mathbb{P}(\{ |\tilde{X}_{n,j} - \tilde{Y}_{n,j}| \geq C_s s^{-n} \}) \leq C_s s^{-n}. \tag{3.4}$$

Indeed, this was proved for $s = 4$ in [19, section 5] (see [19, p 1663]), but taking a careful look at the proofs of [19, lemma 5.2] and [19, lemma 5.4] one observes that $4^s$ can be replaced with $s^s$, for any $s > 4$, since the upper bounds there on the Prokhorov distances appearing in the proofs of these lemmas have the form $e^{-\delta_1 e^{-\delta_2 s}}$ for some positive $\delta_1$ and $\delta_2$.

Next, set $\Gamma_j = \Gamma_{j,n} = \{ |\tilde{X}_{n,j} - \tilde{Y}_{n,j}| \geq C_s s^{-n} \}$, where $0 \leq j < F(n)$. Then, by applying the Cauchy–Schwartz inequality, we have that

$$\left\| \sum_{j=0}^{F(n)-1} (\tilde{X}_{n,j} - \tilde{Y}_{n,j}) \right\|_{L^2}^2 \leq 2 \mathbb{P}^2(\cap_{j} \Gamma_j)(C_s F(n)s^{-n})^2 + 2 \mathbb{E}\left[ \sum_{j=0}^{F(n)-1} (\tilde{X}_{n,j} - \tilde{Y}_{n,j}) \right]^2 \leq 2(C_s 2^{3n} s^{-n})^2 + 2 \mathbb{P}^2(\cup_{j} \Gamma_j) \sum_{j=0}^{F(n)-1} (\tilde{X}_{n,j} - \tilde{Y}_{n,j})^2 \leq I_1 + I_2,$$

where $I_\Gamma$ denotes the indicator function of a set $\Gamma$, $\Gamma^c$ denotes its complement in the underlying probability space and $\mathbb{P}(\Gamma)^c = (\mathbb{P}(\Gamma))^c$ for $r > 0$. Observe that when $s > 4$, we have that

$$I_1 \leq C 4^{-n},$$
for some $C > 0$. On the other hand, since the $L^4$-norms of $\tilde{X}_{n,j}$ and $\tilde{Y}_{n,j}$ are bounded by $c|I_{n,j}|$ (where $c$ is some constant), then

$$
\left\| \sum_{j=0}^{F(n)-1} (\tilde{X}_{n,j} - \tilde{Y}_{n,j}) \right\|_{L^4}^2 \leq \left( 2c \sum_{j=0}^{F(n)-1} |I_{n,j}| \right)^2 \leq C2^{2n},
$$

where $C > 0$ is some constant. Note that in the above inequality we have used that the sum of all the lengths of intervals $I_{n,j}, 0 \leq j < F(n)$ does not exceed $2^{n+1}$. The above inequality together with (3.3) implies that

$$
I_2 \leq (F(n)C_3s^{-n})^{1/2} \cdot C2^n.
$$

Thus, $I_2$ is bounded in $n$ when $s > 32$.

Finally, since $V$ is symmetric and the $V_{n,j}$’s are i.i.d. we have that

$$
\left\| \sum_{j=0}^{F(n)-1} V_{n,j} \right\|_{L^2}^2 = F(n)\| V \|_{L^2}^2 \leq 2^{3n}\| V \|_{L^2}^2,
$$

which together with the above estimates on $I_1$ and $I_2$ completes the proof of the proposition. □

We now derive the following corollary.

**Corollary 3.3.** There exists a constant $c > 0$ such that for every sufficiently large $n \in \mathbb{N}$ and all $v \in \mathbb{R}^d$ we have that

$$
\text{Cov} \left( \sum_{m=1}^{n} \sum_{j=0}^{F(m)-1} Y_{m,j} \right) v \cdot v \geq c2^n|v|^2. 
$$

(3.4)

**Proof.** Firstly, observe that for any two centered random vectors $X = (X_i)_{i=1}^d$ and $Z = (Z_i)_{i=1}^d$, which are defined on the same probability space, and all $1 \leq i, j \leq d$ we have that

$$
|\mathbb{E}[X_iX_j] - \mathbb{E}[Z_iZ_j]| \leq \|X_i\|_{L^2}\|X_j - Z_j\|_{L^2} + \|Z_i\|_{L^2}\|X_i - Z_i\|_{L^2}.
$$

It follows that

$$
|\text{Cov}(X) - \text{Cov}(Z)| \leq C_d\|X - Z\|_{L^2}(\|X\|_{L^2} + \|Z\|_{L^2}),
$$

where $C_d > 0$ is a constant which depends only on $d$. Let $X = \sum_{m=1}^{2^{n+1}} A_m$ and $Z = \sum_{m=1}^{n} \sum_{j=0}^{F(m)-1} X_{m,j}$. Then, by [19, proposition 4.1] we have that

$$
\|X - Z\|_{L^2} \leq C2^{(1-\gamma)/2},
$$

for some constants $C > 0$ and $\gamma \in (0, 1)$ which do not depend on $n$. Indeed, we observe that the number of $A_m$’s appearing in $X - Z$ is at most of order $2^{(1-\gamma)n}, 0 < \gamma < 1$. Furthermore, by [19, proposition 4.1], the $L^2$-norms of $X$ and $Z$ are of order at most $2^{n/2}$, and hence

$$
|\text{Cov}(X) - \text{Cov}(Z)| \leq C'2^{(1-\gamma)/2 + n/2},
$$

where $C' > 0$ does not depend on $n$. Using (2.2) we conclude that for every $v \in \mathbb{R}^d$ and a sufficiently large $n$, we have
Proposition 3.4. Let $Y_n \cdots, Y_{n-1}$ be independent centered $\mathbb{R}^d$-valued random vectors. Let $q \geq 2$ and set $M = \left( \sum_{i=0}^{b-1} \mathbb{E} |Y|^q \right)^{1/q}$. Assume that there exists a sequence $0 = m_0 < m_1 < \cdots < m_b = b$ such that with $\zeta_k = Y_{m_k} + \cdots + Y_{m_{k+1} - 1}$ and $B_k = \text{Cov}(\zeta_k)$, for every $v \in \mathbb{R}^d$ and $0 \leq k < s$ we have that

$$100M^2|v|^2 \leq B_kv \cdot v \leq 100CM^2|v|^2,$$

where $C \geq 1$ is some constant. Then, there exists a coupling between $(Y_0, \cdots, Y_{b-1})$ and a sequence of independent Gaussian random vectors $(S_0, \cdots, S_{b-1})$ such that $\text{Cov}(S_j) = \text{Cov}(Y_j)$ for each $j \in \mathbb{N}$ and

$$\mathbb{P}\left( \max_{0 \leq k \leq b-1} \sum_{j=0}^{i} (Y_j - S_j) \geq Mz \right) \leq C'z^{-q} + \exp(-C'z),$$

for all $z \geq C' \log s$. Here, $C'$ is a positive constant which depends only on $C$, $d$ and $q$.

Now we can describe our next modification of [19]. In the proof of [19, lemma 5.6], proposition 3.4 was applied with $(Y_n, \cdots, Y_{n-1})$ and $\text{Cov}(S_j) = \text{Cov}(Y_j)$ for each $j \in \mathbb{N}$ and

$$\sum_{n} \mathbb{P} \left( \max_{(k,i) \in (0,F(n-1)]} \sum_{m=1}^{k} \sum_{j=0}^{i} Y_{m,j} - S_{k,j} \right) \geq 2^{(1-\beta)/2 + 3/p + \varepsilon/2}) < \infty.$$  

**Proof.** Take $q \in (2, p)$ and set

$$M = \left( \sum_{m=1}^{n} \sum_{j=0}^{F(m)-1} \mathbb{E} |Y_{m,j}|^q \right)^{1/q}.$$ 

By proposition 4.1 in [19] we have

$$\|Y_{m,j}\|_{L^q} = \|X_{m,j}\|_{L^q} \leq C\|I_{m,j}\|_{L^q} \leq C'2^{1-\beta}m^2/2,$$
and therefore

\[ M \leq \sum_{m=1}^{n} \left( \sum_{j=0}^{F(m)-1} \mathbb{E}[Y_{m,j}] \right)^{1/2} \leq C \sum_{m=1}^{n} (F(m))^{1/2} 2^{(1-\beta)m/2} \]

\[ \leq C 2^{\beta m/2} . 2^{(1-\beta)m/2}. \quad (3.9) \]

If we take \( q \) sufficiently close to \( p \), then \( M^2 \) is much smaller than \( 2^n \). On the other hand, by corollary 3.3 for any \( v \in \mathbb{R}^d \) we have

\[ \text{Cov} \left( \sum_{m=1}^{n} \sum_{j=0}^{F(m)-1} Y_{m,j} \right) v \cdot v \geq c 2^n |v|^2. \]

Observe next that for all \( m, j \) and \( v \in \mathbb{R}^d \),

\[ |\text{Cov}(Y_{m,j})||v|^2 \leq \| Y_{m,j} \|_2^2 |v|^2 \leq \| Y_{m,j} \|_2^2 |v|^2 \leq M^2 |v|^2. \]

Therefore, we can regroup \( \{Y_{m,j}\}, (m, j) \prec (n, F(n-1)) \) so that \((3.6)\) holds true with some \( C' \) which does not depend on \( n \) and with some \( s \) (whose order in \( n \) does not exceed \( 2^n \)). Taking \( z \) of the form \( z = 2^n \) in \((3.7)\) we obtain \((3.8)\). In the last argument, we have used \((3.9)\), which insures that \( M 2^n \) is much smaller than \( 2((1-\beta)/2+\beta/2)q^n \), when \( q \) is close enough to \( p \) and \( \epsilon \) is sufficiently small.

**Remark 3.6.** If a nonnegative random variable \( X \) satisfies \( P(X \geq Mz) \leq C' z^{-q} + \exp(-C'z) \) for all \( z \geq C' \log s \), where \( C', M > 0, s \geq 2 \) and \( q > 2 \) then

\[ \mathbb{E}[X^2] = \int_0^\infty \mathbb{P}(X^2 \geq \nu) \, d\nu = 2M^2 \int_0^\infty \log(X \geq Mz) \, dz \leq 2M^2 \left( \int_0^{C' \log s} z \, dz + C' \int_{C' \log s}^\infty (z^{-q+1} + z \exp(-C'z)) \, dz \right) \leq C'' M^2 (\log^2 + 1) \]

and so \( \|X\|_2 \leq CM \log s \) for some \( C \) which depends only on \( C' \) and \( q \). Using this, it follows from the proof of lemma 3.5 that for every \( q \in (2, p) \) there is a constant \( C_q > 0 \) so that for all \( n \geq 2 \),

\[ \left\| \max_{(k,i)=(n,F(n)-1)} \sum_{j=0}^{k-1} Y_{k,j} - S_{k,i} \right\|_2 \leq C_q 2^{(2+2\beta)(1-\beta)} \log n. \quad (3.10) \]

**Completing of the proof of theorem 2.1.** The proof of theorem 1.3 in [19] is separated into six steps. All of these steps proceed exactly as in [19] except from lemmas 5.6 and 5.7 there. In lemma 3.5 we have proved a slightly weaker version of lemma 5.6 in [19] which is clearly enough in order to obtain the desired approximation by sums of independent Gaussian random vectors. The purpose of lemma 5.7 was to prescribe the covariances of the approximating Gaussians. In the first statement of theorem 2.1 we have not claimed anything about the variances of these Gaussians, and so we can skip in the corresponding part from [19] and complete the proof of (2.3) by taking \( \beta = \frac{p}{2(p-1)}. \)
Next, let us show that (2.4) holds true. Firstly, by applying proposition 3.2 with the finite sequence, we derive that
\[
\left\| \sum_{(n', f) \in (n J)} X_{n', f} - Y_{n', f} \right\|_2 \leq C 2^n \beta^{n/2}.
\]
Using also (3.10) with \(q\) sufficiently close to \(p\), and noting that \(\beta/p + (1 - \beta) = \beta/2\) we obtain that there is a coupling of \(\{X_{n', f}\}\) with \(\{S_{n', f}\}\) so that
\[
\left\| \sum_{(n', f) \in (n J)} X_{n', f} - \sum_{(n', f) \in (n J)} S_{n', f} \right\|_2 \leq C 2^n \beta^{n/2 + \delta}.
\] (3.11)

Of course, to obtain the above coupling we have also used the so-called Berkes–Philipp lemma (see [5, lemma A.1], [19, lemma 3.1]).

Take \(n \in \mathbb{N}\), and let \(N_n\) be such that \(2^{N_n} \leq n < 2^{N_n + 1}\). Furthermore, let \(j_n\) be the largest index such that the left end point of \(I_{N_n, j_n}\) is smaller than \(n\). In the case when \(n \in I_{N_n, j_n}\) we have
\[
\sum_{i=1}^{n} A_i - \sum_{(n', f) \in (N_n, j_n)} X_{n', f} = \sum_{(n', f) \in (N_n, j_n)} \sum_{i \in J_{n, j_n}} A_i + \sum_{i \in J_{n, j_n}} A_i + \sum_{i \in J_{n, j_n}} A_i + \sum_{i=1}^{n} A_i := I_1 + I_2,
\]
where \(i_{d', f}\) denotes the left end point of \(I_{d', f}\). Recall next that by [19, (5.1)] the cardinality of \(J \cap [1, 2^{N_n + 1}]\) does not exceed \(C 2^{N_n + 1/2}\), which for our specific choice of \(N_n\) does not exceed \(C \beta^{3/2}\). Using [19, lemma 5.9] with a sufficiently small \(\alpha\) we derive that
\[
\|I_1\|_2 \leq C n^{\beta/2 + \varepsilon}.
\]
On the other hand, applying [19, proposition 4.1] we obtain that
\[
\|I_2\|_2 \leq C |I_{N_n, j_n}|^{1/2} \leq C 2^{N_n (1 - \beta)/2} \leq C n^{(1 - \beta)/2} \leq C n^{\beta/2},
\]
where we have used that for our specific choice of \(\beta\) we have \((1 - \beta)/2 = \beta/2 - \beta/p < \beta/2\). We conclude that there exists a constant \(C' > 0\) so that for every \(n \geq 1\),
\[
\left\| \sum_{j=1}^{n} A_j - \sum_{(n', f) \in (N_n, j_n)} X_{n', f} \right\|_2 \leq C n^{\beta/2 + \varepsilon + 1/p}.
\] (3.12)
The proof of (2.4) in the case when \(n \in I_{N_n, j_n}\) is completed now by (3.11). The case when \(n \notin I_{N_n, j_n}\) is treated similarly. We first write
\[
\sum_{j=1}^{n} A_j - \sum_{(n', f) \in (N_n, j_n)} X_{n', f} = \sum_{j \notin J_{n, j_n}} A_j + X_{N_n, j_n} := I_1 + I_2.
\]
Then the \(L^2\)-norms of the summands \(I_1\) and \(I_2\) are bounded exactly as in the case when \(n \in I_{N_n, j_n}\), and the proof of (2.4) in this case is also finalised by applying (3.11).
Finally, let us prove (2.6) under the exponential decay of correlations assumptions (2.5). By replacing $A_n$ with $A_n \cdot v$, where $v$ is a unit vector, it is enough to prove (2.6) in the scalar case $d = 1$. Recall first that the variances of $X_{\alpha',f}$ and $S_{\alpha',f}$ are equal. Therefore,

$$\left\| \sum_{(\alpha',f) \prec (n,j)} X_{\alpha',f} \right\|^2 = \left\| \sum_{(\alpha',f) \prec (n,j)} S_{\alpha',f} \right\|^2 + \frac{2}{d} \sum_{k \prec (\alpha',f) \not\prec (\alpha',f')} \text{Cov}(X_{\alpha',f}, X_{\alpha',f'}),$$

where the double sum ranges over the pairs $(\alpha',f') \neq (\alpha',f)$ so that $(\alpha',f'),(\alpha',f) \prec (n,j)$. Next, since the size of each $|I_{n,j}|$ is of order at most $2^n$ and the gaps between two different $I_{n,j}$’s is at least $2^{[\epsilon n]}$, using (2.5) we conclude that there is an $a > 0$ so that the above double sum is $O(2^{-an})$, and so

$$\left\| \sum_{(\alpha',f) \prec (n,j)} X_{\alpha',f} \right\|^2 = \left\| \sum_{(\alpha',f) \prec (n,j)} S_{\alpha',f} \right\|^2 + O(2^{-an}). \tag{3.13}$$

Next, set $S_n = \sum_{(\alpha',f) \prec (n,j)} (X_{\alpha',f} + \hat{X}_{\alpha,n,j})$, where $\hat{X}_{\alpha',f} = \sum_{k \prec (\alpha',f')} A_k$. Set also $G_n = \sum_{(\alpha',f) \prec (n,j)} X_{\alpha',f}$ and $\Delta_n = S_n - G_n$. Then

$$\text{Var}(S_n) = \text{Var}(G_n) + \text{Var}(\Delta_n) + 2 \text{Cov}(\Delta_n,G_n). \tag{3.14}$$

Now, by (3.12) we have $\text{Var}(\Delta_n) = O(n^{2+\epsilon})$. Moreover, as in the derivation of (3.13) we have

$$\text{Cov}(\Delta_n,G_n) = O(2^{-an}) + \sum_{(\alpha',f) \prec (n,j)} \text{Cov}(X_{\alpha',f}, \hat{X}_{\alpha',f}).$$

Because of (2.5) we have that $\text{Cov}(X_{\alpha',f}, \hat{X}_{\alpha',f})$ is uniformly bounded in $(\alpha',f')$. Therefore, $\text{Cov}(\Delta_n,G_n) = O(n^{\epsilon})$. We conclude that

$$\text{Var}(S_n) = \text{Var} \left( \sum_{(\alpha',f) \prec (n,j)} S_{\alpha',f} \right) + O(n^{2+\epsilon}).$$

Now, (2.5) implies that

$$\text{Var} \left( \sum_{j=1}^{n} A_j \right) = \text{Var}(S_n) + O(n^{2+\epsilon})$$

and therefore (2.6) follows by the above estimates, noting that $\beta = \frac{\epsilon}{2^{(d-1)}}$. \hfill $\square$

4. ASIP for random dynamical systems under abstract spectral conditions

4.1. Assumptions for a cocycle of transfer operators

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and assume that $\sigma : \Omega \to \Omega$ is an invertible measure-preserving transformation. Furthermore, we suppose that $\sigma$ is ergodic.

Let $M$ be a compact Riemannian manifold (possibly with a boundary) equipped with the Riemannian measure $m$ and assume that $B = (B, \| \cdot \|)$ is a Banach space whose elements are
distributions on $M$ of order at most $q$, $q \in \mathbb{N} \cup \{0\}$. More precisely, we require that there exists $C > 0$ such that for each $h \in \mathcal{B}$ and $\phi \in C^0(M, \mathbb{R}),$

$$|h(\phi)| \leq C |h| \cdot |\phi|_{C^0},$$

(4.1)

where $h(\phi)$ denotes the action of a distribution $h$ on a test function $\phi$. We say that $h \in \mathcal{B}$ is positive (and write $h \geq 0$) if $h(\phi) \geq 0$ for each test function $\phi \geq 0$.

We assume that we have a family $(T_\omega)_{\omega \in \Omega}$ of maps $T_\omega : M \to M$ and the associated family $(L_\omega)_{\omega \in \Omega}$ of bounded operators on $\mathcal{B}$ such that:

- $\omega \mapsto L_\omega$ is strongly measurable, i.e. $\omega \mapsto L_\omega h$ is measurable for each $h \in \mathcal{B}$;
- there is a norm $\|\cdot\|_w$ on $\mathcal{B}$ satisfying $\|\cdot\|_w \leq \|\cdot\|$ and constants $a \in (0,1)$ and $B_1, B_2 > 0$ such that

$$\|L_n^\omega h\| \leq B_1 a^n \|h\| + B_2 \|h\|_w \quad \text{for } P - \text{a.e. } \omega \in \Omega, h \in \mathcal{B} \text{ and } n \in \mathbb{N},$$

(4.2)

where

$$L_n^\omega = L_{\sigma_n - 1} \circ \cdots \circ L_{\sigma_\omega} \circ L_\omega;$$

(4.3)

- for $\omega \in \Omega$, $h \in \mathcal{B}$ and a test function $\phi$,

$$(L_\omega h)(\phi) = h(\phi \circ T_\omega),$$

(4.4)

where in the above formula it is implicitly assumed that the action $\phi \mapsto h(\phi)$ induces an action $\phi \mapsto h(\phi \circ T_\omega)$, for each $\omega$.

**Remark 4.1.** Observe that it follows from (4.2) (applied for $n = 1$) that

$$\operatorname{esssup}_{\omega \in \Omega} \|L_\omega\| < +\infty;$$

(4.5)

We consider $\xi \in \mathcal{B}^*$ given by

$$\langle \xi, h \rangle = h(1), \quad h \in \mathcal{B}.$$  

We suppose that there exist $D, \lambda > 0$ such that

$$\|L_n^\omega h\| \leq De^{-\lambda n} \|h\|,$$

(4.6)

for $P$-a.e. $\omega \in \Omega$, $n \in \mathbb{N}$ and $h \in B_0 := \{ h \in \mathcal{B} : \langle \xi, h \rangle = 0 \}$. By (4.4),

$$\langle \xi, L_\omega h \rangle = \langle \xi, h \rangle, \quad \text{for } h \in \mathcal{B} \text{ and } \omega \in \Omega.$$  

(4.7)

The following result is a minor modification of [16, proposition 7]. We include the proof for the sake of completeness.

**Proposition 4.2.** There exists a unique family $(h_\omega^0)_{\omega \in \Omega} \subset \mathcal{B}$ with the following properties:

- $\omega \mapsto h_\omega^0$ is measurable;
- $h_\omega^0 \geq 0$ and $\langle \xi, h_\omega^0 \rangle = 1$ for $P$-a.e. $\omega \in \Omega$;
- for $P$-a.e. $\omega \in \Omega$, $L_\omega h_\omega^0 = h_\omega^0$;
- we have that

$$\operatorname{esssup}_{\omega \in \Omega} \|h_\omega^0\| < +\infty.$$  

(4.8)
Proof. Let \(\mathcal{Y}\) denote the space of all measurable \(v : \Omega \to \mathcal{B}\) such that
\[
\|v\|_{\infty} := \text{esssup}_{\omega \in \Omega} \|v(\omega)\| < +\infty.
\]
Then, \((\mathcal{Y}, \| \cdot \|_{\infty})\) is a Banach space. By \(\mathcal{Z}\) we denote the subset of \(\mathcal{Y}\) consisting of all \(v \in \mathcal{Y}\) such that
\[
v(\omega) \geq 0 \quad \text{and} \quad \langle \xi, v(\omega) \rangle = 1, \quad \text{for } \mathbb{P} - \text{a.e. } \Omega \in \Omega.
\]
Using (4.1), it is easy to show that \(\mathcal{Z}\) is closed. We define \(L : \mathcal{Z} \to \mathcal{Z}\) by
\[
(Lv)(\omega) = L_{\sigma^{-1}\omega}v(\sigma^{-1}\omega), \quad \omega \in \Omega, \; v \in \mathcal{Z}.
\]
By ((4.4), (4.5) and (4.7)) and recalling that \(\omega \mapsto L_{\omega}\) is strongly measurable, we have that \(L\) is well-defined. Moreover, choosing \(N\) so that \(D e^{-\lambda N} < 1\), it follows from (4.6) that \(L_N\) is a contraction. Hence, \(L\) has a unique fixed point \(v \in \mathcal{Z}\). Set \(h_0(\omega) = v(\omega)\) for \(\omega \in \Omega\). Clearly, the family \((h_0(\sigma^{-1}\omega))_{\omega \in \Omega}\) satisfies the desired properties. Conversely, it is obvious that each family satisfying the properties as in the statement of the proposition gives rise to a fixed point of \(L\), which is unique. \(\square\)

From now on, we assume that \(h_0(\omega)\) is a probability measure on \(M\) for \(\mathbb{P} - \text{a.e. } \Omega \in \Omega\). (4.10)

Remark 4.4. In concrete applications, in order to fulfill the above requirement, it will be sufficient to show that \(h_0(\omega)\) is a Radon measure on \(M\). Then, the second assertion of proposition 4.2 will imply that \(h_0(\omega)\) is a probability measure.
for \( g_1, g_2 \in \mathcal{O} \). Moreover,
\[
\| g_1 g_2 \|_o \leq K_o \| g_1 \|_o \cdot \| g_2 \|_o,
\]
for some constant \( K_o > 0 \) which is independent on \( g_1, g_2 \).

**Remark 4.5.** Without any loss of generality, we may assume (and from now on we will) that \( \| 1 \|_o = 1 \).

In addition, we suppose that there exists a bilinear operator \( \cdot : \mathcal{O} \times B \to B \) with the property that there exists \( K'_o > 0 \) (independent on \( g \) and \( h \)) such that:
\[
\| g \cdot h \| \leq K'_o \| g \|_o \cdot \| h \| 
\]
for every \( g \in \mathcal{O} \) and \( h \in B \), (4.12)

Moreover, we assume that the action of \( g \cdot h \) as a distribution is given by
\[
(g \cdot h)(\phi) = h(g\phi), \quad \text{for each test function } \phi.
\]

**Remark 4.6.** Clearly, we may take \( K_o = K'_o \).

Finally, we suppose that there is a continuous function \( D : [0, +\infty) \times [0, +\infty) \to (0, +\infty) \) with the following properties:
• \( D \) is increasing in the second variable and \( \lim_{\theta \to 0} D(\theta, x) = 0 \) for each \( x \geq 0 \);
• for each \( \theta \in \mathbb{C} \) and \( g \in \mathcal{O} \), \( e^{\theta g} \in \mathcal{O} \) and
\[
\| e^{\theta g} - 1 \|_o \leq D(|\theta|, \| g \|_o). \quad (4.14)
\]

**Remark 4.7.** In all of our examples, the last condition will be a simple consequence of the mean-value theorem.

### 4.3. Good observables and auxiliary results

Take now \( d \in \mathbb{N} \). We say that \( g : \Omega \times X \to \mathbb{R}^d, \ g = (g_1, \cdots, g_d) \) is a **good observable** if the following holds:
• \( g \) is measurable;
• \( g_\omega := g(\omega, \cdot) \in \mathcal{O} \) for \( \omega \in \Omega \) and \( 1 \leq i \leq d \);
•
\[
\| g \|_\infty := \max_{1 \leq i \leq d} \operatorname{esssup}_{\omega \in \Omega} \| g_\omega \|_o < +\infty.
\]

(4.15)

Take now a good observable \( g : \Omega \times X \to \mathbb{R}^d \). For \( \theta \in \mathbb{C}^d \), we define a linear operator \( \mathcal{L}_\theta^g : B \to B \) by
\[
\mathcal{L}_\theta^g h = \mathcal{L}_\omega(e^{\theta g} h), \quad h \in B.
\]

(4.16)

**Proposition 4.8.** There exists a continuous function \( L : \mathbb{C} \to (0, +\infty) \) such that
\[
\operatorname{esssup}_{\omega \in \Omega} \| \mathcal{L}_\omega^g \| \leq L(\theta),
\]
for every \( \theta \in \mathbb{C}^d \).

**Proof.** By (4.11) and (4.14), we have that
\[
\| e^{\theta g} \|_o = \prod_{i=1}^d \| e^{\theta g_i} \|_o \leq K_o \prod_{i=1}^d \| e^{\theta g_i} \|_0 \leq K_o \prod_{i=1}^d (D(|\theta_i|, \| g \|_\infty) + 1).
\]
Hence, (4.12) and (4.16) imply that (4.17) holds with $L : \mathbb{C} \to (0, +\infty)$ given by

$$\mathcal{L}(\theta) = \left(\operatorname{esssup}_{\omega \in \Omega} \| \mathcal{L}_\omega \| \right) K_{d+1} \prod_{i=1}^{d} \left( D(|\theta|, \|g\|_\infty) + 1 \right).$$

Clearly, $L$ is continuous. \hfill \Box

**Proposition 4.9.** There exists a continuous function $\tilde{L} : \mathbb{C} \to (0, +\infty)$ such that

$$\operatorname{esssup}_{\omega \in \Omega} \| \mathcal{L}_\omega - \tilde{L} \| \leq \tilde{L}(\theta),$$

for every $\theta \in \mathbb{C}^d$.

**Proof.** Observe that

$$(\mathcal{L}_\omega - \tilde{L})h = \mathcal{L}_\omega((e^{\theta \cdot g} - 1) \cdot h).$$

It follows from (4.11) and (4.14) and the triangle inequality that

$$\|e^{\theta \cdot g} - 1\|_\sigma = \left\| \prod_{i=1}^{d} e^{\theta_i g_i} - 1 \right\|_\sigma \leq \sum_{i=1}^{d} \left\| \prod_{j=1}^{i} e^{\theta_j g_j} - \prod_{j=1}^{i-1} e^{\theta_j g_j} \right\|_\sigma \leq K_o \sum_{i=1}^{d} \left\| e^{\theta_i g_i} - 1 \right\|_\sigma \leq \sum_{i=1}^{d} K_d D(|\theta|, \|g\|_\infty) \prod_{j=1}^{i-1} \left( D(|\theta|, \|g\|_\infty) + 1 \right),$$

and thus (4.18) holds with $\tilde{L} : \mathbb{C} \to (0, +\infty)$ given by

$$\mathcal{L}(\theta) = \left(\operatorname{esssup}_{\omega \in \Omega} \| \mathcal{L}_\omega \| \right) \sum_{i=1}^{d} K_{d+1} D(|\theta|, \|g\|_\infty) \prod_{j=1}^{i-1} \left( D(|\theta|, \|g\|_\infty) + 1 \right).$$

Obviously, $\tilde{L}$ is continuous and $\lim_{\theta \to 0} \tilde{L}(\theta) = 0$. \hfill \Box

As in (4.3), we set

$$\mathcal{L}^0_{\omega} = \mathcal{L}^0_{\omega_{n-1}} \circ \cdots \circ \mathcal{L}^0_{\omega}, \quad \text{for } \Omega \in \Omega, n \in \mathbb{N} \text{ and } \theta \in \mathbb{C}.$$

Choose $N \in \mathbb{N}$ such that $\gamma := B_2 a^N < 1$, where $B_1$ and $a$ are as in (4.2). The proof of the following result will be obtained by arguing as in the proof of [13, proposition 4.4].

**Proposition 4.10.** There exists $\tilde{\gamma} \in (0, 1)$ such that for all $\theta \in \mathbb{C}^d$ sufficiently close to 0,

$$\| \mathcal{L}_{\omega}^N h \| \leq \tilde{\gamma} \| h \| + B_2 \| h \|_\sigma, \quad \text{for } \mathbb{P} - \text{a.e. } \Omega \in \Omega \text{ and } h \in \mathcal{B}. \tag{4.19}$$

**Proof.** We have

$$\| \mathcal{L}_{\omega}^N h \| \leq \| \mathcal{L}_{\omega}^N h \| + \| \mathcal{L}_{\omega}^N - \mathcal{L}_{\omega}^N \| \cdot \| h \| \leq \gamma \| h \| + B_2 \| h \|_\sigma + \| \mathcal{L}_{\omega}^N - \mathcal{L}_{\omega}^N \| \cdot \| h \|. \tag{4.20}$$
On the other hand,
\[ L^{\theta N} - L^{N \omega} = \sum_{j=0}^{N-1} L^{\theta}_{j}^{\sigma^{N-1-j}} (L^{\theta}_{j}^{\sigma^{N-1-j}} - L^{\sigma^{N-1-j}}) L^{N-1-j}. \]  
(4.21)

Let \( L : \mathbb{C} \rightarrow (0, +\infty) \) be given by proposition 4.8. Since \( L \) is continuous, we have that
\[ M := \sup_{|\theta| \leq 1} L(\theta) < +\infty. \]  
(4.22)

Hence,
\[ \|L^{\theta}_{j}^{\sigma} L^{N-1-j}\| \leq M^{j} \quad \text{and} \quad \|L^{N-1-j}\| \leq M^{N-1-j}, \]
for \( P \)-a.e. \( \omega \in \Omega, |\theta| \leq 1 \) and \( 0 \leq j \leq N - 1 \). By (4.18) and (4.21), we have that
\[ \|L^{\theta N} - L^{N \omega}\| \leq NM^{N-1} \tilde{L}(\theta), \quad \text{for} \ P \ - \ a.e. \ \omega \in \Omega \ \text{and} \ |\theta| \leq 1. \]

The above estimate together with (4.20) readily implies the desired conclusion. \( \square \)

**Proposition 4.11.** Assume that there exists \( B_{3} > 0 \) such that
\[ \|L^{\theta N} h\|_{\omega} \leq B_{3} \|h\|_{\omega}, \quad \text{for} \ P \ - \ a.e. \ \omega \in \Omega, \ h \in \mathcal{B}, \ t \in \mathbb{R}^{d}, |t| \leq 1 \ \text{and} \ n \in \mathbb{N}. \]  
(4.23)

Then there exists \( \rho, C > 0 \) such that
\[ \|L^{\theta N}\| \leq C, \quad \text{for} \ P \ - \ a.e. \ \omega \in \Omega, n \in \mathbb{N} \ \text{and} \ t \in \mathbb{R}^{d}, |t| \leq \rho. \]  
(4.24)

**Proof.** Let \( \rho \in (0, 1) \) be such that (4.19) holds when \( |\theta| \leq \rho \). By iterating (4.19) and using (4.23), we find that there exists \( C_{1} > 0 \) such that
\[ \|L^{\theta N}\| \leq C_{1}, \quad \text{for} \ P \ - \ a.e. \ \omega \in \Omega, n \in \mathbb{N} \ \text{and} \ t \in \mathbb{R}^{d}, |t| \leq \rho. \]  
(4.25)

The desired conclusion follows readily from (4.17), (4.22) and (4.25). \( \square \)

**4.4. Almost sure invariance principle**

By \( E_{\omega}(\phi) \) we will denote the expectation of \( \phi \) with respect to \( h^{\omega}_{0} \) (see (4.10)).

Throughout this subsection, we take a good observable \( g \) such that (4.23) holds (for some \( B_{3} > 0 \)). Set
\[ T^{n}_{\omega} = T_{\sigma^{N-1} \omega} \circ \cdots \circ T_{\omega} \]
and
\[ S_{n} g(\omega, \cdot) = \sum_{i=0}^{n-1} g(\sigma^{i} \omega, T^{i}_{\omega}(\cdot)), \]
for \( \omega \in \Omega \) and \( n \in \mathbb{N} \).

**Lemma 4.12.** For \( \omega \in \Omega, h \in \mathcal{B}, n \in \mathbb{N} \) and a test function \( \phi \), we have that
\[ (L^{\theta N} h)(\phi) = h(e^{\theta} S_{n} g(\omega, \cdot)(\phi \circ T^{n}_{\omega})). \]
**Proof.** The desired conclusion follows from (4.4) and (4.13) (by using induction on $n$). □

**Lemma 4.13.** Let $g : \Omega \times M \to \mathbb{R}^d$ be a good observable. For $\mathbb{P}$-a.e. $\omega \in \Omega$, there exist $C, c, \rho > 0$ such that for any $n, m > 0$, $b_1 < b_2 < \cdots < b_{n+m+1}$, $k > 0$ and $t_1, \cdots, t_{n+m} \in \mathbb{R}^d$ with $|t_j| \leq \rho$, we have that

$$
\left| \mathbb{E}_\omega \left( e^{\sum_{j=1}^m t_j \left( \sum_{\sigma \in \mathbb{B}_j} A_\sigma \right) + \sum_{j=1}^n (b_{j+1} + k) A_{t_j} \right) \right| 
= \mathbb{E}_\omega \left( e^{\sum_{j=1}^m t_j \left( \sum_{\sigma \in \mathbb{B}_j} A_\sigma \right) } \right) \cdot \mathbb{E}_\omega \left( e^{\sum_{j=1}^n (b_{j+1} + k) A_{t_j} \right)

\leq C(1 + \max |b_{j+1} - b_j|) e^{-ck},
$$

where

$$A_\ell := g_{\ell \omega} \circ T_\ell \quad \ell \in \mathbb{N} \cup \{0\}.$$

**Proof.** For $\omega \in \Omega$, let $Q_\omega$ be given by (4.9). By applying lemma 4.12, one can verify that

$$
\mathbb{E}_\omega \left( e^{\sum_{j=1}^m t_j \left( \sum_{\sigma \in \mathbb{B}_j} A_\sigma \right) + \sum_{j=1}^n (b_{j+1} + k) A_{t_j} \right)
= L^{b_{n+m} h_{n+m+1} - h_{n+m}} \cdots L^{b_{n+1} h_{n+1} - h_{n+1}} \cdot L_k^{b_{1} h_{1}}(1)
= L^{b_{n+m} h_{n+m+1} - h_{n+m}} \cdots L^{b_{k} h_{k}}(1)
+ L^{b_{n+m} h_{n+m+1} - h_{n+m}} \cdots L^{b_{1} h_{1}}(1)
.$$
In order to obtain our main result, we introduce an additional assumption. For \( h \in \mathcal{B} \), we assume that the action \( \phi \to h(\phi) \) induces an action \( \phi \to h(\phi) \) on \( \mathcal{O} \) and that there exists \( L > 0 \) such that
\[
|h(\phi)| \leq L \|h\| \cdot \|\phi\|_o, \quad \text{for } \phi \in \mathcal{O} \text{ and } h \in \mathcal{B}.
\] (4.26)

**Remark 4.14.** In principle, instead of assuming (4.26) we could have assumed from the start that \( \mathcal{B} \) is a space of distributions acting on \( \mathcal{O} \). In this case (4.1) and (4.26) coincide. For this to make sense, we would need to require that \( \mathcal{B} \) includes all finite measures, which holds when \( \|g\|_o \geq \sup |g| \), for all \( g \in \mathcal{O} \). For instance, in the setting of subsection 5.3, \( \mathcal{B} \) will be the space of functions \( h \) on \( M \) with bounded variation, identified as linear functionals given by \( h(\phi) = \int_M h \circ dm \). Note that the BV norm is larger than the supremum norm. In this case it will make no difference if we consider \( h \in \mathcal{B} \) as a distribution on \( C^0(M, \mathbb{R}) \) or on the space of BV functions.

The reason we have decided not to present our results in this way and to assume (4.26) is that it is less standard to consider an abstract class of test functions.

**Lemma 4.15.** There exists \( C > 0 \) such that
\[
|h^0_o(\phi(\psi \circ T^n_\omega))| \leq Ce^{-\lambda n}\|\phi\|_o \cdot \|\psi\|_o,
\]
for \( \mathbb{P}-a.e. \omega \in \Omega, n \in \mathbb{N} \) and \( \phi, \psi \in \mathcal{O} \) such that \( h^0_o(\phi) = 0 \).

**Proof.** We have that
\[
h^0_o(\phi(\psi \circ T^n_\omega)) = L^0_o(\phi \cdot h^0_o(\psi)).
\]
Since \( \phi \cdot h^0_o \in \mathcal{B}_0 \), the desired conclusion follows from (4.6), (4.8), (4.12) and (4.26). \( \square \)

We consider the skew-product transformation \( \tau : \Omega \times M \to \Omega \times M \) given by
\[
\tau(\omega, x) = (\sigma \omega, T_\omega(x)), \quad (\omega, x) \in \Omega \times M.
\] (4.27)

Furthermore, we consider the probability measure \( \mu \) on \( \Omega \times M \) given by
\[
\mu(A \times B) = \int_\Omega h^0_o(B) d\mathbb{P}(\omega), \quad \text{for } A \in \mathcal{F} \text{ and } B \subset M \text{ Borel}.
\]
It follows easily from the third assertion of proposition 4.2 that \( \mu \) is invariant for \( \tau \). Let us also assume that \( \mu \) is ergodic. Next, we need the following result.

**Proposition 4.16.** Let \( g : \Omega \times M \to \mathbb{R} \) be a good observable such that \( g \in L^2(\Omega \times M, \mu) \) and \( h^0_o(g_\omega) = 0 \) for \( \mathbb{P}-a.e. \omega \in \Omega \). Then, there exists \( \Sigma^2 \geq 0 \) such that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} g_{\omega \circ T^k_\omega} = \Sigma^2, \quad \text{for } \mathbb{P} - a.e. \Omega \in \Omega.
\]
Moreover,
\[
\Sigma^2 = \int_{\Omega \times M} g^2 d\mu + 2 \sum_{n=1}^{\infty} \int_{\Omega \times M} g(g \circ \tau^n) d\mu.
\] (4.28)

Finally, \( \Sigma^2 = 0 \) if and only if there exists \( q \in L^2(\Omega \times M, \mu) \) such that
\[
g = q - q \circ \tau, \quad \mu - a.s.
\] (4.29)
Proof. The proof of the following result can be obtained by arguing exactly as in the proofs of \cite[lemma 12]{11} and \cite[proposition 3]{11}. For reader’s convenience we provide here some of the details. Note first that
\[
\mathbb{E}_\omega \left( \sum_{k=0}^{n-1} g_{\sigma^k \omega} \circ T^k \right)^2 = \sum_{k=0}^{n-1} \mathbb{E}_\omega (g_{\sigma^k \omega}^2 \circ T^k) + 2 \sum_{0 \leq i < j \leq n-1} \mathbb{E}_\omega ((g_{\sigma^i \omega} \circ T^i)(g_{\sigma^j \omega} \circ T^j))
\]
\[
= \sum_{k=0}^{n-1} \mathbb{E}_\omega (g_{\sigma^k \omega}^2 \circ T^k) + 2 \sum_{j=0}^{n-1} \sum_{j=i+1}^{n-1} \mathbb{E}_\omega (g_{\sigma^i \omega} (g_{\sigma^j \omega} \circ T^{j-i})).\]

Set \( G(\omega) = \mathbb{E}_\omega (g_{\sigma^k \omega}^2), \omega \in \Omega \). By applying Birkhoff’s ergodic theorem for \( G \) over the ergodic measure-preserving system \((\Omega, \mathcal{F}, P, \sigma)\), we find that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}_\omega (g_{\sigma^k \omega}^2 \circ T^k) = \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} G(\sigma^k \omega) = \int_{\Omega} G(\omega) d\mathbb{P}(\omega)
\]
\[
= \int_{\Omega} \int_{M} g(\omega, x)^2 \, dh_0(x) \, d\mathbb{P}(\omega)
\]
\[
= \int_{\Omega \times M} g(\omega, x)^2 \, d\mu(\omega, x),
\]
for \( P \)-a.e. \( \omega \in \Omega \). Furthermore, set
\[
\Psi(\omega) = \sum_{n=1}^{\infty} \int_{M} g(\omega, x) g(\tau^n(\omega, x)) \, dh_0(x) = \sum_{n=1}^{\infty} \mathcal{L}^2_{\omega} (g_{\omega} \cdot h_0(\sigma^n \omega)).
\]

By (4.6), (4.8), (4.15) and (4.26), there is a constant \( c > 0 \) so that
\[
|\Psi(\omega)| \leq c, \quad \text{for } P - \text{a.e. } \Omega \in \Omega.
\]

In particular, \( \Psi \in L^1(\Omega) \) and thus it follows again from Birkhoff’s ergodic theorem that
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \Psi(\sigma^i \omega) = \int_{\Omega} \Psi(\omega) \, d\mathbb{P}(\omega) = \sum_{n=1}^{\infty} \int_{\Omega \times X} g(\omega, x) g(\tau^n(\omega, x)) \, d\mu(\omega, x), \quad (4.30)
\]
for \( P \)-a.e. \( \omega \in \Omega \). In order to complete the proof of the existence of \( \Sigma^2 \), it is enough to show that
\[
\lim_{n \to \infty} \frac{1}{n} \left( \sum_{i=0}^{n-1} \sum_{j=i+1}^{n-1} \mathbb{E}_\omega (g_{\sigma^i \omega} (g_{\sigma^j \omega} \circ T^{j-i})) - \sum_{i=0}^{n-1} \Psi(\sigma^i \omega) \right) = 0, \quad (4.31)
\]
for \( P \)-a.e. \( \omega \in \Omega \). However, (4.31) can be obtained by arguing exactly as in the proof of \cite[proposition 3]{11}. We conclude that the first assertion of the proposition holds.

To prove the characterisation for the positivity of \( \Sigma^2 \), let \( X_n = g \circ \tau^n \), considered as a random variable with respect to the measure \( \mu \). Then \( \{X_n\} \) is a stationary sequence satisfying \( \sum_n (n+1)|E_\mu[X_n X_0]| < \infty \) (using lemma 4.15). Thus, using classical results for stationary
sequences (see [22]) we conclude from (4.28) that
\[
\Sigma^2 = \lim_{n \to \infty} \frac{1}{n} \text{Var} \left( \sum_{j=0}^{n-1} g \circ \tau^j \right).
\]
Moreover, (4.29) is equivalent to \(\Sigma^2 = 0\).

**Proposition 4.17.** Let \(g : \Omega \times M \to \mathbb{R}^d\) be a good observable such that \(g \in L^2(\Omega \times M, \mu)\) and \(h_0^d(g_\omega) = 0\) for \(\mathbb{P}\)-a.e. \(\omega \in \Omega\). Then, there exists a positive semi-definite \(d \times d\) matrix \(\Sigma^2\) such that for \(\mathbb{P}\)-a.e. \(\omega \in \Omega\) we have that
\[
\lim_{n \to \infty} \frac{1}{n} \text{Var} (S_n g(\omega, \cdot))^2 = \Sigma^2.
\]
Moreover, \(\Sigma^2\) is not positive definite if and only if there exist \(v \in \mathbb{R}^d \setminus \{0\}\) and an \(\mathbb{R}\)-valued function \(r \in L^2(\Omega \times M, \mu)\) such that
\[
v \cdot g = r - r \circ \tau, \quad \mu - \text{a.e.}\ldots
\]
(4.32)

**Proof.** Let \(v \in \mathbb{R}^d\) and consider the real valued function \(g_v = v \cdot g\). By applying proposition 4.16, we obtain that there exists \(\Sigma^2_v \geq 0\) such that
\[
\lim_{n \to \infty} \frac{1}{n} \text{Var} (S_n g_v(\omega, \cdot))^2 = \Sigma^2_v, \quad \text{for } \mathbb{P} - \text{a.e. } \omega \in \Omega.
\]
(4.33)

Moreover, \(\Sigma^2_v = 0\) if and only if there exists \(r \in L^2(\Omega \times M, \mu)\) such that
\[
g_v = r - r \circ \tau.
\]
(4.34)

For \(1 \leq i, j \leq d\), we claim that there exists a real number \(\Sigma^2_{i,j}\) so that \(\mathbb{P}\)-a.e. we have that
\[
\lim_{n \to \infty} \frac{1}{n} \text{Var} (S_n g_i(\omega, \cdot) \cdot S_n g_j(\omega, \cdot))^2 = \Sigma^2_{i,j}.
\]
(4.35)

Clearly, it follows from (4.33) that \(\Sigma^2_{i,j} = \left( \Sigma^2_{e_i + e_j} - \Sigma^2_{e_i} - \Sigma^2_{e_j} \right) / 2\) satisfies (4.35), where \(e_i\) denotes the standard \(i\)th unit vector. The resulting matrix \(\Sigma^2 = (\Sigma^2_{i,j})\) is positive semi-definite. Indeed, it is easy to verify that \(\Sigma^2 v \cdot v = \Sigma^2_v\) for \(v \in \mathbb{R}^d\). From this we also see that \(\Sigma^2\) is not positive definite if and only if there exists \(v \in \mathbb{R}^d, v \neq 0\) such that \(\Sigma^2_v = 0\). However, it follows from the previous paragraph that this happens if and only if \(v \cdot g = g_v\) can be written in the form (4.32). The proof of the proposition is completed.

**Theorem 4.18.** Let \(g : \Omega \times M \to \mathbb{R}^d\) be a good observable such that \(g \in L^\infty(\Omega \times M, \mu)\) and \(h_0^d(g_\omega) = 0\) for \(\mathbb{P}\)-a.e. \(\omega \in \Omega\). Furthermore, suppose that \(\Sigma^2\) given by proposition 4.17 is positive definite. Then, for \(\mathbb{P}\)-a.e. \(\omega \in \Omega\) and every \(\delta > 0\), there exists a coupling between \(\{g_\sigma \circ T^n : n \geq 0\}\), considered as a sequence of random variables on \((M, \mathcal{M})\), and a sequence \((Z_k)\) of independent centered (i.e. of zero mean) Gaussian random vectors such that
\[
\left| \sum_{i=0}^{n-1} g_\sigma \circ T^n_i - \sum_{i=1}^n Z_i \right| = O(n^{1/4+\delta}), \quad \text{almost surely}.
\]
Moreover, there exists a constant $C = C_\delta(\omega) > 0$ so that for every $n \geq 1$,
\[
\left\| \sum_{i=0}^{n-1} g^{\sigma_n} \circ T^i_\omega - \sum_{i=1}^{n} Z_i \right\|_{L^2} \leq C n^{1/4+\delta}.
\]

Furthermore, there exists a constant $C' = C'_\delta(\omega) > 0$ so that for every unit vector $v \in \mathbb{R}^d$,
\[
\left\| \sum_{i=0}^{n-1} g^{\sigma_n} \circ T^i_\omega \cdot v \right\|_{L^2} \leq \left\| \sum_{i=1}^{n} Z_i \cdot v \right\|_{L^2} + C' n^{1/2+\delta}.
\]

**Proof.** The conclusion of the theorem follows from proposition 4.17, lemma 4.13 and theorem 2.1 (by noting that the sequence $(g^{\sigma_n} \circ T^i_\omega)_{n \in \mathbb{N}}$ is bounded in $L^p(M, h_0^\omega)$ for each $p \geq 1$).

4.5. A general scheme for verifying condition (4.6)

We will now discuss a general framework under which (4.6) holds true. In order to do so, we introduce some additional assumptions:

- $\mathcal{B} = (\mathcal{B}, \| \cdot \|)$ can be compactly embedded in $\mathcal{B}_w = (\mathcal{B}_w, \| \cdot \|_w)$ and $\| \cdot \|_w \leq \| \cdot \|$ on $\mathcal{B}$;
- there is a bounded linear operator $\mathcal{L}$ on $\mathcal{B}$ (that admits an extension to a bounded operator on $\mathcal{B}_w$) with the property that there exist $K, \rho > 0$ such that
  \[
  \|\mathcal{L}^n h\| \leq Ke^{-\rho n}\|h\|, \quad \text{for } h \in \mathcal{B}, h(1) = 0 \text{ and } n \in \mathbb{N}.
  \] (4.36)
- there exist $C_i > 0, i \in \{1, 2, 3\}$ and $b \in (0, 1)$ such that
  \[
  \|\mathcal{L}_n \cdots \mathcal{L}_1\|_w \leq C_1,
  \] (4.37)
  and
  \[
  \|\mathcal{L}_n \cdots \mathcal{L}_1 h\| \leq C_2 b^n\|h\| + C_3\|h\|_w,
  \] (4.38)
  for each $n \in \mathbb{N}, h \in \mathcal{B}$ and $\mathcal{L}_1, \cdots, \mathcal{L}_n \in \mathcal{P}$, where $\mathcal{P}$ is a family of bounded operators on $\mathcal{B}$, $\mathcal{L} \in \mathcal{P}$ such that each element of $\mathcal{P}$ admits an extension to a bounded operator on $\mathcal{B}_w$.

Moreover, $\{h \in \mathcal{B} : h(1) = 0\}$ is invariant for each operator in $\mathcal{P}$.

For $\epsilon_0 > 0$, set
\[
\mathcal{P}(\mathcal{L}, \epsilon_0) := \left\{ \mathcal{L}' \in \mathcal{P} : \|\mathcal{L}' - \mathcal{L}\|_w = \sup_{\|h\| \leq 1} \|(\mathcal{L}' - \mathcal{L}) h\|_w \leq \epsilon_0 \right\}.
\]

Provided that $\epsilon_0 > 0$ is sufficiently small, it follows from [6, proposition 2.10] that there exist $D, \lambda > 0$ such that
\[
\|\mathcal{L}_n \cdots \mathcal{L}_1 h\| \leq D e^{-\lambda n}\|h\|, \quad \text{for } h \in \mathcal{B}, h(1) = 0 \text{ and } n \in \mathbb{N}.
\]

Hence, if we build our cocycle $(\mathcal{L}_\omega)_{\omega \in \Omega}$ so that $\mathcal{L}_\omega \in \mathcal{P}(\mathcal{L}, \epsilon_0)$ for each $\omega \in \Omega$, we have that (4.6) holds.
5. Examples

We now discuss various classes of random dynamical systems to which theorem 4.18 is applicable.

5.1. Random hyperbolic dynamics

Let $M$ be a $C^\infty$ compact connected Riemannian manifold and let $T$ be a topologically transitive Anosov map of class $C^r$, where $r > 2$. For $\epsilon > 0$, set

\[ M_\epsilon(T) := \{ S : M \to M : S \text{ is an Anosov map of class } C^{r+1} \text{ and } d_{C^{r+1}}(S, T) < \epsilon \} \]

Let $\omega \mapsto T_\omega \in M_\epsilon(T)$ be a measurable map. By $L^1_\omega$, we denote the transfer operator associated to $T_\omega$. Provided that $\epsilon > 0$ is sufficiently small, it is proved in [13, section 3] (by using arguments in subsection 4.5) that there exists a Banach space $B = (B, \| \cdot \|)$ and two norms $\| \cdot \|$ and $\| \cdot \|_w$ on $B$ such that (4.2) and (4.6) hold. Moreover, elements of $B$ are distributions of order at most 1. The strong measurability of the map $\omega \mapsto L_\omega$ is verified in [13, subsection 3.1], while (4.10) is proved in [13, proposition 3.3].

Set $O = C^r(M, C)$. It follows from [18, lemma 3.2] that (4.12) holds. Finally, (4.23) can be established as in [13]. Indeed, (4.23) was proved in [13, p 653–654] in the case when $d = 1$, i.e. when $g$ is a real-valued observable. As for the case when $d > 1$, it is sufficient to note that one only needs to justify that the version of [13, (58)] holds true (with a constant independent on $\omega$, $n$ and $t$ with $|t| \leq 1$). However, for this we only need to apply [13, (58)] for $g := t \cdot g$ instead of $g$, which can be done since

\[ \sup_{|t| \leq 1} \| g(\omega, \cdot) \|_{C^r} \leq C_d \| g(\omega, \cdot) \|_{C^r}, \]

where $C_d > 0$ is some constant which depends only on $d$. We conclude that theorem 4.18 can be applied in this setting.

**Remark 5.1.** We would like to explain the reason why we took $O = C^r(M, C)$. Namely, this was done in order to satisfy assumptions introduced in subsection 4.2, which in particular require that for $g \in O$ and $h \in B$, we can define $g \cdot h \in B$. Moreover, for a fixed $g \in O$, $h \mapsto g \cdot h$ needs to be a bounded operator on $B$. These properties were crucial for the twisted transfer operators $L^1_\omega$ to be well defined and for our approach to work. Since our $B$ belongs to the class of anisotropic Banach spaces introduced in [18] (which are precisely defined as closures of $C^r(M, C)$ with respect to certain norms), it was necessary to consider observables in $C^r(M, C)$. We also note that the version of the central limit theorem case in the deterministic case in [18] is also stated for $C^r$-observables (see [18, remark 2.10]).

As pointed to us by M Demers, the class of anisotropic Banach spaces introduced in [3] have the property that the multiplication by a Hölder continuous observable (with a suitably chosen Hölder constant) acts as a bounded linear operator. We note that the same holds true for the class of spaces introduced Demers and Zhang [10], which are build to handle also the presence of singularities in the two-dimensional setting. Consequently, using spaces introduced in [3], it seems reasonable that one can extend theorem 4.18 by dealing with observables $g$ which have the property that $g_\omega$ is Hölder continuous. We differ from pursuing this direction because it would require a lot of additional technical work. We feel that this would somehow hide the main contribution of this paper, which is the adaptation of the spectral approach for ASIP given in [19] to the case of quenched random dynamics, and its applications.
5.2. Random perturbations of the Lorentz gas

Let us consider the two dimensional torus $\mathbb{T}^2$ on which we place finitely many (disjoint) scatterers $\Gamma_i$, $i = 1, \cdots, d$ which have $C^3$ boundaries with strictly positive curvature. We stress that in what follows these scatterers will be allowed to move but their number and the arclengths of their boundaries will not change. Set

$$M := \bigcup_{i=1}^d I_i \times [-\pi/2, \pi/2],$$

where $I_i$ is an interval with endpoints identified such that $|I_i| = |\partial\Gamma_i|$. Furthermore, let $m$ be the normalised Lebesgue measure on $M$, i.e. $dm = \frac{1}{\pi} dr d\varphi$ where $L = \sum_{i=1}^d |I_i|$. We consider the class $\mathcal{F}$ of maps on $M$ introduced in [10, section 3]. We stress that $\mathcal{F}$ contains various perturbations of the billiard map associated to periodic Lorentz gas (see [10, section 2.4] for details).

Let now $F'$ consist of all those $T \in F$ that preserve measure $\mu$ given by $d\mu = \frac{1}{\pi} \cos \varphi dm$. Hence, for $T \in F'$ we have that [10, (H5)] holds with $\eta = 1$.

Let $\| \cdot \|_w$ and $\| \cdot \|_B$ be norms on $C^1(M, \mathbb{C})$ introduced in [10, section 3.2]. Moreover, let $B = (B, \| \cdot \|)$ be the completion of $C^1(M, \mathbb{C})$ with respect to $\| \cdot \|_B$. It follows from [10, lemma 3.4] that elements of $B$ are distributions of order at most 1. Finally, let $d_F$ be the distance on $\mathcal{F}$ introduced in [10, section 3.4].

Let us now fix $T \in F'$ such that $(T, \mu)$ is mixing. Thus, (4.36) holds (with some $K, \rho > 0$), where $L = L_T$. For $\epsilon > 0$, set

$$\mathcal{M}_\epsilon(T) := \{ T' \in F' : d_F(T, T') < \epsilon \}.$$ 

Let $\omega \mapsto T_\omega \in \mathcal{M}_\epsilon(T)$ be a measurable map. By [10, theorem 2.3], there exist $C, \beta > 0$ such that

$$\| L_{T_1} - L_{T_2} \|_w = \sup_{|h| \leq 1} \| (L_{T_1} - L_{T_2}) h \|_w \leq Ce^{\beta/2}, \quad \text{for } T_1, T_2 \in F'.$$

Moreover, it follows from the proof of [10, proposition 5.6] that (4.37) and (4.38) hold (with some $b \in (0, 1), C_0 > 0, i \in \{1, 2, 3\}$), where $L_i = L_{T_i}, T_i \in \mathcal{M}_\epsilon(T)$. Provided that $\epsilon > 0$ is sufficiently small, it follows from the discussion in subsection 4.5 that (4.6) holds. Moreover, the strong measurability of $\omega \mapsto L_\omega$ can be obtained by arguing as in [13, subsection 3.1]. In this setting $h_0^\omega = \mu$ and therefore (4.10) is trivially satisfied.

Let $\gamma \in (0, 1)$ be as in the statement of [10, lemma 5.3.] and let $\mathcal{O} = C^1(M)$ be the space of all Hölder continuous functions $\varphi : M \to \mathbb{C}$ with Hölder exponent $\gamma$. Hence, [10, lemma 5.3.] implies that (4.12) holds. Finally, one verifies (4.23) by arguing as in the previous subsection. We conclude that theorem 4.18 can be applied in this setting.

5.3. Random Lasota–Yorke maps

Our results hold true for the random expanding maps considered by Buzzi [4], for which the spectral method was developed in [12]. The results also hold true for the expanding maps considered in [20, ch 5]. In order not to overload the paper we will only present a more concrete and relatively simple example.

Let $M = [0, 1]$ equipped with the Lebesgue measure $m$. For a piecewise $C^2$ map $T : M \to M$, set $\delta(T) = \text{essinf}_{x \in [0,1]} |T'|$ and let $b(T)$ denote the number of intervals of monotonicity (branches) of $T$. Consider now a measurable map $\omega \mapsto T_\omega$, $\omega \in \Omega$ of piecewise $C^2$ maps on $[0, 1]$ such that

$$b := \text{esssup}_{\omega \in \Omega} b(T_\omega) < \infty, \quad \delta := \text{essinf}_{\omega \in \Omega} \delta(T_\omega) > 1, \quad D := \text{esssup}_{\omega \in \Omega} \| T'_\omega \|_w < \infty.$$
In addition, we impose the following condition:

for every subinterval $J \subset M$, $\exists k = k(J)$ such that for $\mathbb{P}$ – a.e. $\omega \in \Omega$, $T^k_\omega(J) = M$.

Finally, let $B = (B, \| \cdot \|) = (BV, \| \cdot \|_{BV})$ and $\| \cdot \|_w = \| \cdot \|_{L^1(m)}$. By identifying each $v \in BV$ with the functional $\phi \mapsto \int_M \phi v dm$ on $C^0(M, \mathbb{C})$, we have that $B$ consists of distributions of order 0 (i.e. measures). It is proved in [12, subsection 2.3.1] that (4.2) and (4.6) hold. Moreover, proposition 4.2 (see remark 4.4) implies that (4.10) holds.

Set $O = B = BV$. Then, (4.12) holds. Finally, (4.23) holds with $B_3 = 1$. We conclude that theorem 4.18 can be applied in this setting.

Acknowledgments

We would like to express our gratitude to anonymous referees and the handling editor for many useful comments and suggestions that helped us to improve our paper. In addition, we thank Mark Demers for his comments in relation to remark 5.1. DD was supported in part by Croatian Science Foundation under the project IP-2019-04-1239 and by the University of Rijeka under the projects uniri-prirod-18-9 and uniri-prprirod-19-16.

ORCID iDs

D. Dragičević © https://orcid.org/0000-0002-1979-4344

References

[1] Aimino R, Nicol M and Vaienti S 2015 Annealed and quenched limit theorems for random expanding dynamical systems Probab. Theor. Relat. Fields 162 233–74
[2] Arnold L 1998 Random Dynamical Systems (Springer Monographs in Mathematics) (Berlin: Springer)
[3] Baladi V and Tsujii M 2007 Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms Ann. Inst. Fourier 57 127–54
[4] Buzzi J 1999 Exponential decay of correlations for random Lasota–Yorke maps Commun. Math. Phys. 208 25–54
[5] Berkes I and Philipp W 1979 Approximation theorems for independent and weakly dependent random vectors Ann. Probab. 7 29–54
[6] Conze J-P and Raugi A 2007 Limit theorems for sequential expanding dynamical systems on [0, 1] Ergodic Theory and Related Fields (Contemporary Mathematics) vol 430 (Providence, RI: American Mathematical Society) pp 89–121
[7] Cuny C, Dedecker J, Korepanov A and Merlevède F 2020 Rates in almost sure invariance principle for slowly mixing dynamical systems Ergod. Theor. Dyn. Syst. 40 2317–48
[8] Cuny C, Dedecker J, Korepanov A and Merlevède F 2020 Rates in almost sure invariance principle for quickly mixing dynamical systems Stoch. Dyn. 20 2050002
[9] Cuny C and Merlevède F 2015 Strong invariance principles with rate for ‘reverse’ martingale differences and applications J. Theor. Probab. 28 137–83
[10] Demers M F and Zhang H-K 2013 A functional analytic approach to perturbations of the Lorentz gas Commun. Math. Phys. 324 767–830
[11] Dragičević D, Froyland G, Gonzalez-Tokman C and Vaienti S 2018 Almost sure invariance principle for random piecewise expanding maps Nonlinearity 31 2252–80
[12] Dragičević D, Froyland G, Gonzalez-Tokman C and Vaienti S 2018 A spectral approach for quenched limit theorems for random expanding dynamical systems Comm. Math. Phys. 360 1121–87
[13] Dragičević D, Froyland G, Gonzalez-Tokman C and Vaienti S 2020 A spectral approach for quenched limit theorems for random hyperbolic dynamical systems Trans. Am. Math. Soc. 373 629–64
[14] Dragičević D and Hafouta Y 2020 Limit theorems for random expanding or Anosov dynamical systems and vector-valued observables Ann. Henri Poincaré 21 3869–917
[15] Dragičević D and Hafouta Y 2021 Almost sure invariance principle for random distance expanding maps with a nonuniform decay of correlations Thermodynamic Formalism (CIRM Jean-Morlet Chair Subseries) ed M Pollicott and S Vaienti (Berlin: Springer)
[16] Dragičević D and Sedro J 2020 Statistical stability and linear response for random hyperbolic dynamics (arXiv:2007.06088)
[17] Field M, Melbourne I and Török A 2003 Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions Ergod. Theo. Dyn. Syst. 23 87–110
[18] Gouëzel S and Liverani C 2006 Banach spaces adapted to Anosov systems Ergod. Theo. Dyn. Syst. 26 123–51
[19] Gouëzel S 2010 Almost sure invariance principle for dynamical systems by spectral methods Ann. Probab. 38 1639–71
[20] Hafouta Y and Kifer Y 2018 Nonconventional Limit Theorems and Random Dynamics (Singapore: World Scientific)
[21] Haydn N, Nicol M, Török A and Vaienti S 2017 Almost sure invariance principle for sequential and non-stationary dynamical systems Ann. Math. Soc. 369 5293–316
[22] Ibragimov I A and Linnik Y V 1971 Independent and Stationary Sequences of Random Variables (Groningen: Wolters-Noordhoff)
[23] Kifer Y 1998 Limit theorems for random transformations and processes in random environments Trans. Am. Math. Soc. 350 1481–518
[24] Korepanov A 2018 Equidistribution for nonuniformly expanding dynamical systems, and application to the almost sure invariance principle Commun. Math. Phys. 359 1123–38
[25] Korepanov A 2018 Rates in almost sure invariance principle for dynamical systems with some hyperbolicity Commun. Math. Phys. 363 173–90
[26] Korepanov A, Kosloff Z and Melbourne I 2018 Martingale-coboundary decomposition for families of dynamical systems Ann. Inst. Henri Poincaré C. Anal. Non Linéaire 35 859–85
[27] Melbourne I and Nicol M 2005 Almost sure invariance principle for nonuniformly hyperbolic systems Commun. Math. Phys. 260 131–46
[28] Melbourne I and Nicol M 2009 A vector-valued almost sure invariance principle for hyperbolic dynamical systems Ann. Probab. 37 478–505
[29] Philipp W and Stout W F 1975 Almost sure invariance principles for partial sums of weakly dependent random variables Mem. Amer. Math. Soc. 161
[30] Stenlund M 2014 A vector-valued almost sure invariance principle for Sinai billiards with random scatterers Commun. Math. Phys. 325 879–916
[31] Stenlund M and Su Y 2019 Random young towers and quenched limit laws preprint (arXiv:1907.12199)
[32] Su Y 2019 Vector-valued almost sure invariance principle for non-stationary dynamical systems (arXiv:1903.09763)