Ergod. Th. & Dynam. Sys., (2024), **44**, 118–158
doi:10.1017/etds.2023.23

Large deviations, moment estimates and almost sure invariance principles for skew products with mixing base maps and expanding-on-average fibers

YEOR HAFOUTA

Department of Mathematics, The Ohio State University, Columbus, OH, USA
and
Department of Mathematics, University of Maryland, College Park, MD, USA
(e-mail: yeor.hafouta@mail.huji.ac.il)

(*Received* 27 *March* 2022 *and accepted in revised form* 21 *February* 2023)

Abstract. In this paper we show how to apply classical probabilistic tools for partial sums $\sum_{j=0}^{n-1} \varphi \circ \tau^j$ generated by a skew product τ, built over a sufficiently well-mixing base map and a random expanding dynamical system. Under certain regularity assumptions on the observable φ, we obtain a central limit theorem (CLT) with rates, a functional CLT, an almost sure invariance principle (ASIP), a moderate-deviations principle, several exponential concentration inequalities and Rosenthal-type moment estimates for skew products with α-, ϕ- or ψ-mixing base maps and expanding-on-average random fiber maps. All of the results are new even in the uniformly expanding case. The main novelty here (in contrast to [2]) is that the random maps are not independent, they do not preserve the same measure and the observable φ depends also on the base space. For stretched exponentially α-mixing base maps our proofs are based on multiple correlation estimates, which make the classical method of cumulants applicable. For ϕ- or ψ-mixing base maps, we obtain an ASIP and maximal and concentration inequalities by establishing an L^∞ convergence of the iterates $\mathcal{K}^n$ of a certain transfer operator $\mathcal{K}$ with respect to a certain sub-σ-algebra, which yields an appropriate (reverse) martingale-coboundary decomposition.

Key words: limit theorems, random dynamical systems, skew products
2020 Mathematics Subject Classification: 37H12 (Primary)

1. *Introduction and a preview of the main results*

1.1. *Quenched limit theorems for random dynamical systems.* Let $(X, \mathcal{B}, m)$ be a probability space and let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be an invertible ergodic probability-preserving system. Let $T_\omega : X \to X$, $\omega \in \Omega$, be a family of non-singular maps (that is, $m \circ T_\omega^{-1} \ll m$) so that the corresponding skew product τ given by $\tau(\omega, x) = (\sigma\omega, T_\omega x)$ is measurable. A random dynamical system is formed by the sequence of compositions

$$T_\omega^n x, n \geq 0 \quad \text{where } T_\omega^n = T_{\sigma^{n-1}\omega} \circ \cdots \circ T_{\sigma\omega} \circ T_\omega,$$

taken along the orbit of a 'random' point ω. The system $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is often referred to as the *driving system*, and the map σ is often referred to as the *base map*.

Let $\varphi : \Omega \times X \to \mathbb{R}$ be a measurable function (an 'observable') and let μ be a τ-invariant probability measure on $\Omega \times X$. Then μ can be decomposed as $\mu = \int \mu_\omega \, d\mathbb{P}(\omega)$, where μ_ω is a family of probability measures on X so that $(T_\omega)_*\mu_\omega = \mu_{\sigma\omega}$ for $\mathbb{P}$-almost every (a.e.) ω. Set $S_n\varphi = \sum_{j=0}^{n-1} \varphi \circ \tau^j$. Then

$$S_n\varphi(\omega, x) := S_n^\omega \varphi(x) = \sum_{j=0}^{n-1} \varphi_{\sigma^j\omega} \circ T_\omega^j,$$

where $\varphi_\omega(\cdot) = \varphi(\omega, \cdot)$. For $\mathbb{P}$-a.e. ω we can consider the sequence of functions $S_n^\omega\varphi(\cdot)$ on the probability space $(X, \mathcal{B}, \mu_\omega)$ as random variables. Limit theorems for such sequences are called *quenched limit theorems*. Among the first papers dealing with quenched limit theorems for random dynamical systems are [36, 37], where in [36] a quenched large-deviations principle was obtained, and in [37] a central limit theorem (CLT) and a law of the iterated logarithm were established. Since then quenched limit theorems for random dynamical systems have been extensively studied. For instance, in [16, 20–22] almost sure invariance principle (ASIP, an almost sure approximation by a sum of independent Gaussians) was established for random expanding or hyperbolic maps T_ω, in [19, 31] Berry–Esseen theorems (optimal rates in the CLT) were obtained for similar classes of maps and in [17, 18, 23, 31] local CLTs were achieved. In addition, in [27] several limit theorems were extended to random non-uniformly hyperbolic or expanding maps. We would also like to refer to [3] for related results concerning mixing rates for random non-uniformly hyperbolic maps and to [32] for related results concerning sequential dynamical systems, where an ASIP was obtained. We note that in many of the examples these results are obtained for the unique measure μ such that μ_ω is absolutely continuous with respect to m. However, some results hold true even for maps $T_\omega : \mathcal{E}_\omega \to \mathcal{E}_{\sigma\omega} \subset X$ which are defined on random subsets of X (see [40]), where in this case the most notable choices of μ_ω are the so-called random Gibbs measures (see [31, 44]).

1.2. *Limit theorem skew products.* Let us consider the sums $S_n\varphi = \sum_{j=0}^{n-1} \varphi \circ \tau^j$ as random variables on the probability space $(\Omega \times X, \mathcal{F} \times \mathcal{B}, \mu)$. In this paper will focus on limit theorems for such sequences of random variables. In order to demonstrate the difference between such limit theorems and the quenched ones, let us focus of the CLT. The quenched CLT means that for $\mathbb{P}$-a.e. ω, for all real t, we have

$$\lim_{n\to\infty} \mu_\omega(\{x : S_n^\omega\varphi(x) - \mu_\omega(S_n^\omega\varphi) \leq t\sqrt{n}\}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^t e^{-s^2/2\sigma^2} ds$$

where $\sigma \geq 0$ is the number that satisfies $\sigma^2 = \lim_{n\to\infty}(1/n)\mathrm{Var}_{\mu_\omega}(S_n^\omega\varphi)$ for $\mathbb{P}$-a.e. ω (assuming that this limit exists and does not depend on ω, refer to [37, Theorem 2.3] for sufficient conditions). On the other hand, the CLT for the skew product means that for all real t we have

$$\lim_{n\to\infty} \mu(\{(\omega, x) : S_n\varphi(\omega, x) - \mu(S_n\varphi) \le t\sqrt{n}\}) = \frac{1}{\sqrt{2\pi}\Sigma}\int_{-\infty}^{t} e^{-s^2/2\Sigma^2}\,ds,$$

where $\Sigma^2 = \lim_{n\to\infty}(1/n)\mathrm{Var}_\mu(S_n\varphi)$. Note that, in contrast to the quenched case, the summands $X_j = \varphi \circ \tau^j$ form a stationary sequence and, in applications, the existence of the limit Σ^2 follows from a sufficiently fast decay of $\mathrm{Cov}(X_0, X_n)$ as $n \to \infty$. We also remark that both CLT's above are formulated when σ and Σ are positive, and when one of them vanishes the convergence is towards the constant function 0.

When $\mu_\omega(\varphi_\omega)$ does not depend on ω, we have that $\mu_\omega(\varphi_\omega) = \mu(\varphi)$ and $\sigma^2 = \Sigma^2$. In this case the quenched CLT implies the CLT for $S_n\varphi$ by integrating $\mu_\omega(\{x : S_n^\omega\varphi(x) - \mu_\omega(S_n^\omega\varphi) \le t\sqrt{n}\})$ with respect to $\mathbb{P}$ (and similarly other distributive limit theorems for the skew product follow from the quenched ones). However, it is less likely to be true when $\mu_\omega(\varphi_\omega)$ depends on ω. Remark that even when $\mu_\omega(\varphi_\omega)$ does not depend on ω other finer results like the ASIP do not follow by integration. Indeed the ASIP concerns an almost sure approximation of the partial sums in question by a sum of independent Gaussian random variables, but the quenched ASIP provides a construction of such a Gaussian process which depends on the fiber ω.

1.2.1. *Annealed limit theorems: i.i.d. maps.* A particular well-studied case is when the maps $T_{\sigma^j\omega}$ are independent. That is, $\Omega = \mathcal{Y}^{\mathbb{Z}}$ is a product space, the coordinates ω_j of $\omega = (\omega_j)$ are independent (with σ being the left shift) and $T_\omega = T_{\omega_0}$ depends only on the zeroth coordinate. In this case the statistical behavior of the skew product τ can be investigated using the so-called annealed transfer operator, given by (see [8, 9, 35])

$$\mathcal{A}g(x) = \int \mathcal{L}_\omega g(x)\,d\mathbb{P}(\omega),$$

where $\mathcal{L}_\omega$ is the transfer operator corresponding to T_ω and the underlying reference measure m. In [2] it was shown that for several classes of random expanding maps, the operator $\mathcal{A}$ is quasicompact. Using that, a variety of limit theorems were obtained (such as a CLT, a Berry–Esseen theorem, a local CLT, a local large-deviations principle and an ASIP) for random variables of the form

$$S_n\varphi(\omega, x) = \sum_{j=0}^{n-1} \varphi(T_{\omega_{j-1}} \circ \cdots \circ T_{\omega_0}x),$$

where (ω, x) are distributed according to a τ-invariant measure μ of the form $\mathbb{P} \times (h\,dm)$ for some continuous function h, which satisfies $\mathcal{A}h = h$. The latter assumption means that the maps T_ω preserve the same measure $\nu = h\,dm$. The point is that once quasi-compactness is achieved the classical Nagaev–Guivarch method (see [33]) can be applied. This method was applied successfully to obtain limit theorems for deterministic dynamical systems (that is, when $T_\omega = T$ does not depend on ω), and in [2] (see also [7]) this method was applied to obtain annealed limit theorems. We note that since both the function φ and the measure $h\,dm$ do not depend on ω, and all the maps T_ω preserve the measure $h\,dm$, the fiberwise centering constant $\mu_\omega(S_n^\omega\varphi)$ and the usual centering constant $\mu(S_n\varphi)$ are both equal to $n \int \varphi(x)h(x)\, dm(x)$. Hence, as discussed in the previous section, in this setup some annealed results such as the CLT already follow from the quenched ones.

Independence here is crucial, since it yields that the iterates on the annealed transfer operator can be written as

$$\mathcal{A}^n g = \int \mathcal{L}_\omega^n g \, d\mathbb{P}(\omega), \tag{1.1}$$

where $\mathcal{L}_\omega^n = \mathcal{L}_{\sigma^{n-1}\omega} \circ \cdots \circ \mathcal{L}_{\sigma\omega} \circ \mathcal{L}_\omega$, which is the transfer operator of T_ω^n. Hence, the statistical behavior of the iterates τ^n of the skew product can be described by the iterates of $\mathcal{A}$. Note that in this independent and identically distributed (i.i.d.) setup this approach works only when $\varphi(\omega, x) = \varphi(x)$ does not depend on ω since it requires substituting φ (and appropriate functions of φ) into the annealed operator.

1.2.2. *The motivation behind the present paper: non-i.i.d. maps and random functions.* The starting point of this paper is the observation that when the coordinates (ω_j) are not independent (that is, that maps $T_{\sigma^j\omega}$ are not i.i.d.) there is no apparent relation between the iterates τ^n of τ and the iterates of the annealed operator $\mathcal{A}$ defined above. Thus, a natural question arising from [2, 7] is which limit theorems hold true for mixing base maps with non-independent coordinates, and functions φ which depend on ω. Moreover, the assumptions in [2] require all the maps T_ω to preserve the same absolutely continuous measure $\nu = h\, dm$, and it is also desirable to prove limit theorems without such assumptions. (We refer to [46] for a CLT and large deviations for random i.i.d. intermittent maps in the case where the T_ω do not preserve the same measure.) We note that without the above assumptions even the CLT was not obtained before for the skew products considered in this paper, which will be our first result.

The question described above was also one of the main motivations in [26], where a CLT, a local CLT and a renewal theorem were obtained for several classes of skew products with mixing base maps such as Markov shifts and non-uniform Young towers, together with uniformly expanding random maps. These results were obtained by a certain type of integration argument; however, the method of [26] does not involve the iterates of an annealed transfer operator, and instead we studied directly integrals of the form $\int \mathcal{L}_\omega^n g_\omega \, d\mathbb{P}(\omega)$, and their complex perturbations (relying on the fiberwise 'spectral' properties and a certain type of periodic point approach which was introduced in [31]). While [26] was the first paper to discuss limit theorem for skew products with non-independent fiber maps and random observables, all the results there were obtained for fiberwise centered observables φ (that is, $\mu_\omega(\varphi_\omega) = 0$). Moreover, the maps T_ω in [26] were uniformly expanding, the base map had a periodic point and the random transfer operator satisfied certain regularity assumptions as functions of ω around the periodic orbit. From this point of view, a second motivation for the present paper is to prove limit theorems for skew products with non-independent fiber maps $T_{\sigma^j\omega}$ without the fiberwise centralization assumption and without additional topological assumptions such as the behavior around a periodic orbit. We note that, apart from the CLT, we did not consider in [26] any of the limit theorems obtained in the present paper, and so almost all the results in the present paper are new even under the fiberwise centering assumption.

1.3. *Our new results and the method of the proofs.* As explained in the previous section, the goal of this paper is to obtain limit theorems with deterministic centering conditions for skew products τ built over mixing base maps and non-uniformly expanding maps T_ω. More precisely, we still consider a product space $\Omega = \mathcal{Y}^{\mathbb{Z}}$, but with 'weakly dependent' coordinates ω_j instead of independent ones. We consider a family of non-uniformly expanding maps $T_\omega = T_{\omega_0}$ and observables of the form $\varphi(\omega, x) = \varphi_{\omega_0}(x)$ and prove limit theorems for sequences of the form $Z_n = S_n\varphi - n\int \varphi\, d\mu$, where

$$S_n\varphi(\omega, x) = \sum_{j=0}^{n-1} \varphi_{\omega_j}(T_{\omega_{j-1}} \circ \cdots \circ T_{\omega_0} x) = \sum_{j=0}^{n-1} \varphi_{\sigma^j\omega}(T_\omega^j x)$$

considered as a random variables on the probability space $(\Omega \times X, \mathcal{F} \times \mathcal{B}, \mu)$, where $\mu = \int \mu_\omega\, d\mathbb{P}$ is the unique τ-invariant measure with μ_ω being absolutely continuous with respect to m (or when μ_ω is a random Gibbs measure). In this setup we have $(T_\omega)_*\mu_\omega = \mu_{\sigma\omega}$, and in general the maps T_ω do not preserve the same measure. These results are obtained for a certain type of observables φ so that $\varphi_\omega(\cdot)$ has bounded variation, uniformly in ω. When the maps T_ω are expanding on average we will also have a certain scaling assumption (that is, $\text{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$ for some tempered random variable K), which was shown in [22] to be necessary for quenched limit theorems, and which is similarly necessary for obtaining limit theorems for the skew product. In what follows we will always assume that $\int \varphi\, d\mu = 0$, which is not really a restriction since we can always replace φ with $\varphi - \int \varphi\, d\mu$.

We obtain our results using two different methods, as described below.

1.3.1. *Limit theorems for skew products: (functional) CLT, moment estimates, moderate-deviations and exponential concentration inequalities for α-mixing driving systems via the method of cumulants.* Recall that the α-mixing (dependence) coefficient between two sub-σ-algebras $\mathcal{G}, \mathcal{H}$ of $\mathcal{F}$ is given by

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathcal{G}, B \in \mathcal{H}\}.$$

Let $\mathcal{F}_{-\infty,k}$ be the σ-algebra generated by the coordinates ω_j at places $j \le k$ and $\mathcal{F}_{m,\infty}$ be the σ-algebra generated by the coordinates ω_j at places $j \ge m$. Then the α-dependence coefficients of the sequence of coordinates (ω_n) are defined by

$$\alpha_n = \sup_k \alpha(\mathcal{F}_{-\infty,k}, \mathcal{F}_{k+n,\infty}) = \alpha(\mathcal{F}_{-\infty,0}, \mathcal{F}_{n,\infty}) \tag{1.2}$$

where the last equality is due to stationarity of the process (ω_n).

We assume first that $\alpha_n = O(e^{-cn^\eta})$ for some $c, \eta > 0$ (that is, it is stretched exponential). The first step towards limit theorems is standard for stationary processes: we show that under the weaker condition $\sum_n n\alpha_n < \infty$, the limit

$$s^2 = \lim_{n\to\infty} \frac{1}{n}\text{Var}_\mu(S_n), \quad S_n = S_n\varphi,$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

exists and that it vanishes if and only if φ admits an appropriate coboundary representation. When $s^2 > 0$ we show that $n^{-1/2}S_n$ converges in distribution towards a centered normal random variable with variance s^2. More precisely, we obtain the convergence rate

$$\sup_{t\in\mathbb{R}} \left| \mu(S_n \le ts\sqrt{n}) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-(1/2)x^2}\,dx \right| \le Cn^{-1/(2+4\gamma)}, \quad \gamma = 1/\eta.$$

An annealed CLT (that is, for independent maps) was obtained in [7] for random toral automorphisms and in [2] for more general maps. When the base map is only mixing (and φ depends on ω) it was obtained in [26] for fiberwise centered potentials (that is, $\mu_\omega(\varphi_\omega) = 0$). One of the results in this paper is the CLT for stretched exponentially α-mixing base maps but without the fiberwise centering assumption (in fact, we will obtain a functional CLT; see Theorem 2.19 and the last paragraph of §1.3.1).

We also obtain a certain type of large-deviations results, often referred to as a moderate-deviations principle (see [14]). These results yield, for instance, that for every closed interval $[a, b]$ we have

$$\lim_{n\to\infty} \frac{1}{a_n^2} \ln \mu\left\{(\omega, x) : \frac{S_n(\omega, x)}{a_n s n^{1/2}} \in [a, b]\right\} = -\frac{1}{2} \inf_{x\in[a,b]} x^2,$$

where a_n is a sequence such that $a_n \to \infty$ and $a_n = o(n^{1/(2+4\gamma)})$. We also obtain several types of 'stretched' exponential concentration inequalities ((2.20), (2.21)) and Gaussian moment estimates of Rosenthal type (2.22). These result are obtained using the method of cumulants. More precisely, we first obtain a certain type of multiple correlation estimates (see Proposition 3.4), and then by applying a general theorem we conclude that the kth cumulant of the sum S_n is at most of order $n(k!)^{1+\gamma}(c_0)^{k-2}$ for $k \ge 3$, where c_0 is some constant (see Theorem 3.1). Then we can apply the method of cumulants [15, 49]. In the annealed setup, using the quasicompactness of the annealed transfer operator, large-deviations principles and exponential concentration inequalities were obtained in [2], and the above results show that there is a similar behavior when the maps are not independent and the function φ depends on ω (see also the results in the next section where better exponential concentration inequalities are described).

The above multiple correlation estimates together with the method of cumulants and the Rosenthal-type moment estimates also yield a functional CLT. Let us consider the random function $\mathcal{S}_n(t) = n^{-1/2}S_{[nt]}$ on $[0, 1]$. Then we show that it converges in distribution in the Skorokhod space $D[0, 1]$ to sW, where W is a standard Brownian motion and $s^2 = \lim_{n\to\infty}(1/n)\mathrm{Var}_\mu(S_n)$.

1.3.2. *Limit theorems for skew products with ϕ- or ψ-mixing driving systems via martingale methods: almost sure invariance principle, concentration inequalities and maximal moment estimates.* One of the strongest methods to prove CLTs and related results in probability theory and dynamical systems is the so-called martingale-coboundary representation (Gordin's method). For a sufficiently chaotic dynamical system $(Y, \mathcal{G}, \mu, T)$ and an observable $\varphi : Y \to \mathbb{R}$ it means that φ can be represented as $\varphi = u + \chi - \chi \circ T$ for some sufficiently regular function χ, and $(u \circ T^n)$ forms a reverse martingale difference. Such results are well known for deterministic expanding (or hyperbolic) dynamical

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

systems, and we refer to [16, 22, 42] for quenched and sequential versions of such martingale methods.

Recall that the ϕ-mixing and ψ (dependence) coefficient between two sub-σ-algebras $\mathcal{G}, \mathcal{H}$ of $\mathcal{F}$ is given by

$$\phi(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbb{P}(B|A) - \mathbb{P}(B)| : A \in \mathcal{G}, B \in \mathcal{H}, \mathbb{P}(A) > 0\}$$

and

$$\psi(\mathcal{G}, \mathcal{H}) = \sup \left\{ \left| \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)\mathbb{P}(B)} - 1 \right| : A \in \mathcal{G}, B \in \mathcal{H}, \mathbb{P}(A)\mathbb{P}(B) > 0 \right\}.$$

The reverse ϕ-mixing coefficients of the sequence of coordinates (ω_n) are defined by

$$\phi_{n,R} = \sup_k \phi(\mathcal{F}_{k+n,\infty}, \mathcal{F}_{-\infty,k}) = \phi(\mathcal{F}_{n,\infty}, \mathcal{F}_{-\infty,0}), \tag{1.3}$$

while the ψ-mixing coefficients of (ξ_n) are defined by

$$\psi_n = \sup_k \psi(\mathcal{F}_{-\infty,k}, \mathcal{F}_{k+n,\infty}) = \psi(\mathcal{F}_{-\infty,0}, \mathcal{F}_{n,\infty}), \tag{1.4}$$

where $\mathcal{F}_{n,m}$ is as defined before (1.2). It is clear from the definitions of the mixing coefficients that

$$\alpha_n \leq \phi_{n,R} \leq \psi_n.$$

When the sequence (ω_n) is (sufficiently fast) ϕ- or ψ-mixing we obtain a certain type of L^∞ martingale-coboundary representation (that is, $\chi \in L^\infty$) for the underlying class of observables φ with respect to the skew product τ. This was already established in [2] in the annealed setup (that is, when (ω_n) is an i.i.d. sequence), and here, using different arguments, we obtain such a representation for skew products with mixing base maps.

Once an L^∞ martingale-coboundary decomposition is achieved, as usual, we can apply the Azuma–Hoeffding inequality together with Chernoff's bounding method and obtain exponential concentration inequalities of the form

$$\mathbb{P}(|S_n - \mathbb{E}[S_n]| \geq tn + c_1) \leq c_2 e^{-c_3 n t^2}, \quad t > 0,$$

where c_1, c_2, c_3 are positive constants. These concentration inequities are better than the ones we obtain using the method of cumulants, although they involve the stronger notions of ϕ- or ψ-mixing instead of α-mixing. (However, they only require summable ϕ- or ψ-mixing coefficients and not stretched exponential ones.) Another immediate consequence is moment estimates of the form

$$\left\| \max_{1 \leq k \leq n} |S_k - \mathbb{E}[S_k]| \right\|_{L_p} = O(n^{1/2})$$

which hold for every $1 \leq p < \infty$. Such results are known in the annealed case [2], and we extend them to the skew products considered in this paper.

The idea behind the martingale-coboundary representation is as follows. Consider the sub-σ-algebra $\mathcal{F}_0$ of $\Omega \times X$ generated by the projection $\pi_0(\omega, x) = ((\omega_j)_{j \geq 0}, x)$, where $\omega = (\omega_j)_{j \in \mathbb{Z}}$. Then τ preserves $\mathcal{F}_0$ since $T_\omega = T_{\omega_0}$ depends only on ω_0, and $\mathcal{F}_0$ can be viewed as a subsystem (or a factor) given by $(\Omega \times X, \mathcal{F}_0, \mu, \tau)$. Our main argument is

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

as follows. Let $\mathcal{K}$ be the transfer operator corresponding to the invariant σ-algebra $\mathcal{F}_0$, namely the one defined by the duality relation

$$\int (\mathcal{K}g) f \, d\mu = \int (\mathcal{K}g \circ \tau) f \circ \tau \, d\mu = \int \mathbb{E}[g|\tau^{-1}\mathcal{F}_0] \cdot f \circ \tau \, d\mu = \int g \cdot f \circ \tau \, d\mu$$

where $g \in L^1(\Omega \times X, \mathcal{F}_0, \mu)$ and $f \in L^\infty(\Omega \times X, \mathcal{F}_0, \mu)$. Then we show that, under quite mild ϕ- or ψ-mixing rates for the sequence of coordinates (ω_n), the iterates $\mathcal{K}^n\varphi$ of the transfer operator $\mathcal{K}$ corresponding to this system converge fast enough in $L^\infty(\mu)$ towards $\mu(\varphi)\mathbf{1}$, where $\mathbf{1}$ is the function taking the constant value 1, and φ is our given observable. This convergence can be established for every function φ so that $\|\varphi\|_{K,2} = \operatorname{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi(\omega, \cdot)\|_{BV}) < \infty$ for an appropriate tempered random variable $K(\omega)$, or for any observable with $\operatorname{esssup}_{\omega\in\Omega} \|\varphi(\omega, \cdot)\|_{BV} < \infty$ when the maps T_ω are uniformly expanding. We stress that in any case this is not a spectral result (even under exponential mixing), since the convergence of $\mathcal{K}^n$ is not in an operator norm, and, in general, it does not have an exponential rate. Indeed, we only prove that

$$\|\mathcal{K}^n\varphi - \mu(\varphi)\|_{L^\infty} \le C\|\varphi\|_{K,2} \cdot \gamma_n, \tag{1.5}$$

where $\gamma_n = \delta^n + \phi_R([n/2])$ or $\gamma_n = \delta^n + \psi([n/2])$, and $\delta \in (0, 1)$ and $\phi_R(\cdot)$ and $\psi(\cdot)$ are the reverse ϕ-mixing coefficients and ψ-mixing coefficients defined in (1.3) and (1.4), respectively.

Another consequence of the martingale-coboundary representation is the ASIP, which in our context concerns almost sure approximation of the Birkhoff sum by Gaussians. The ASIP for random (and sequential) dynamical systems has been studied by several authors in recent years (see, for instance, [16, 20–22, 32, 50, 51]), and in this paper we will focus on the ASIP for Birkhoff sums generated by the skew product.

In [13] the authors proved that, under certain assumptions, a reverse martingale M_n can be approximated almost surely by a sum of independent Gaussians. One consequence of the methods in [13] is for sums of the form $W_n = \sum_{j=0}^{n-1} \varphi \circ \tau^j$. For such sums, the conditions of [13, Theorem 3.2] show that there is a coupling with a sequence of i.i.d. centered normal random variables Z_j with variance $s^2 = \lim_{n\to\infty}(1/n)\mathrm{Var}(W_n)$ so that

$$\sup_{1\le k\le n} \left| W_k - \sum_{j=1}^{k} Z_j \right| = O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) \quad \text{almost surely.}$$

In our notation, the first and second conditions of [13, Theorem 3.2] about $\mathcal{K}$ can be verified using (1.5). In order to show that the third (and last condition) about $\mathcal{K}$ in [13, Theorem 3.2] is in force we will also need to provide more general estimates on expression of the form

$$\|\mathcal{K}^i(\bar{\varphi}\mathcal{K}^j\bar{\varphi}) - \mu(\mathcal{K}^i(\bar{\varphi}\mathcal{K}^j\bar{\varphi}))\|_{L^\infty}$$

for $1 \le i, j \le n$, where $\bar{\varphi} = \varphi - \mu(\varphi)$.

We note that in [2] the annealed ASIP was obtained using Gouëzel's approach [24] and not the martingale-coboundary approach. Gouëzel's approach was also used in [5] to obtain an ASIP for non-independent maps with mixing base maps, but as indicated in [5] the results are mostly applicable for Gordin–Denker maps.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Finally, we also prove a vector-valued ASIP for skew products with uniformly expanding random maps and exponentially fast α-mixing base maps via the method of Gouëzel [24]. As we have mentioned above, this method was applied in [2] in the annealed setting, while in [5] it was applied for Gordin–Denker systems. In a final section we also discuss a few extensions such as different types of mixing base maps such as Young towers or Gibbs–Markov maps, application of the method of cumulants for non-conventional sums of the form $S_n = \sum_{m=1}^{n} \prod_{j=1}^{\ell} \varphi_j \circ \tau^{q_j(m)}$, for polynomial $q_j(m)$, as well as extensions of the results for different classes of random expanding maps (the ones in [44]).

2. *Preliminaries and main results*

2.1. *The random maps.* We begin by recalling the setup from [12]. Let $(X, \mathcal{G})$ be a measurable space endowed with a probability measure m and a notion of a variation $\mathrm{v}\colon L^1(X, m) \to [0, \infty]$ which satisfies the following conditions:

(V1) $\mathrm{v}(th) = |t|\mathrm{v}(h)$;
(V2) $\mathrm{v}(g + h) \le \mathrm{v}(g) + \mathrm{v}(h)$;
(V3) $\|h\|_{L^\infty} \le C_{\mathrm{v}}(\|h\|_1 + \mathrm{v}(h))$ for some constant $1 \le C_{\mathrm{v}} < \infty$;
(V4) for any $C > 0$, the set $\{h\colon X \to \mathbb{R} : \|h\|_1 + \mathrm{v}(h) \le C\}$ is $L^1(m)$-compact;
(V5) $\mathrm{v}(\mathbf{1}) = 0$, where $\mathbf{1}$ denotes the function equal to 1 on X;
(V6) $\{h\colon X \to \mathbb{R}_+ : \|h\|_1 = 1$ and $\mathrm{v}(h) < \infty\}$ is $L^1(m)$-dense in $\{h\colon X \to \mathbb{R}_+ : \|h\|_1 = 1\}$;
(V7) for any $f \in L^1(X, m)$ such that $\operatorname{essinf} f > 0$, we have

$$\mathrm{v}(1/f) \le \frac{\mathrm{v}(f)}{(\operatorname{essinf} f)^2};$$

(V8) $\mathrm{v}(fg) \le \|f\|_{L^\infty} \cdot \mathrm{v}(g) + \|g\|_{L^\infty} \cdot \mathrm{v}(f)$;
(V9) for $M > 0$, $f : X \to [-M, M]$ measurable and every C^1 function $h\colon [-M, M] \to \mathbb{C}$, we have $\mathrm{v}(h \circ f) \le \|h'\|_{L^\infty} \cdot \mathrm{v}(f)$.

We define

$$BV = BV(X, m) = \{g \in L^1(X, m) : \mathrm{v}(g) < \infty\}.$$

Then BV is a Banach space with respect to the norm

$$\|g\|_{BV} = \|g\|_{L^1} + \mathrm{v}(g).$$

Remark 2.1. Observe that (V3) and (V8) imply that

$$\|fg\|_{BV} \le C_{\mathrm{v}}\|f\|_{BV} \cdot \|g\|_{BV} \quad \text{for } f, g \in BV. \tag{2.1}$$

Remark 2.2. We observe that in [12], assumption (V5) is replaced by the weaker $\mathrm{v}(\mathbb{1}) < +\infty$. However, for the examples we have in mind, our stronger version is satisfied. In particular, (V5) implies that $\|\mathbf{1}\|_{BV} = 1$.

The rest of our setup is almost identical to [22], with a single additional requirement that will be indicated in what follows. Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$, be a probability space and $\sigma : \Omega \to \Omega$ an invertible ergodic measure-preserving transformation. Let $T_\omega\colon X \to X$, $\omega \in \Omega$ be a collection of non-singular transformations (that is, $m \circ T_\omega^{-1} \ll m$ for each ω) acting on

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

X. Each transformation T_ω induces the corresponding transfer operator $\mathcal{L}_\omega$ acting on $L^1(X, m)$ and defined by the duality relation

$$\int_X (\mathcal{L}_\omega \phi)\varphi \, dm = \int_X \phi(\varphi \circ T_\omega) \, dm, \quad \phi \in L^1(X, m), \varphi \in L^\infty(X, m). \tag{2.2}$$

Thus, we obtain a cocycle of transfer operators $(\Omega, \mathcal{F}, \mathbb{P}, \sigma, L^1(X, m), \mathcal{L})$ that we denote by $\mathcal{L} = (\mathcal{L}_\omega)_{\omega\in\Omega}$. For $\omega \in \Omega$ and $n \in \mathbb{N}$, set

$$\mathcal{L}_\omega^n := \mathcal{L}_{\sigma^{n-1}\omega} \circ \cdots \circ \mathcal{L}_{\sigma\omega} \circ \mathcal{L}_\omega.$$

We recall the notion of a tempered random variable.

Definition 2.3. We say that a measurable map $K : \Omega \to (0, +\infty)$ is *tempered* if

$$\lim_{n\to\pm\infty} \frac{1}{n} \log K(\sigma^n \omega) = 0 \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

In this paper we will consider the following assumptions on the random transfer operators.

Definition 2.4. A cocycle $\mathcal{L} = (\mathcal{L}_\omega)_{\omega\in\Omega}$ of transfer operators is said to be *good* if the following conditions hold.

- Ω is a Borel subset of a separable, complete metric space and σ is a homeomorphism. Moreover, $\mathcal{L}$ is $\mathbb{P}$-continuous, that is, Ω can be written as a countable union of measurable sets such that $\omega \mapsto \mathcal{L}_\omega$ is continuous on each of those sets.
- There is a tempered random variable $N(\omega)$ such that

$$v(g \circ T_\omega) \le N(\omega) v(g) \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega \text{ and } g \in BV. \tag{2.3}$$

- There exists a random variable $C : \Omega \to (0, +\infty)$ such that $\log C \in L^1(\Omega, \mathbb{P})$ and

$$\|\mathcal{L}_\omega h\|_{BV} \le C(\omega)\|h\|_{BV} \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega \text{ and } h \in BV.$$

- There exist $N \in \mathbb{N}$ and random variables $\alpha, K : \Omega \to (0, +\infty)$ such that

$$\int_\Omega \log \alpha \, d\mathbb{P} < 0, \quad \log K \in L^1(\Omega, \mathbb{P})$$

  and, for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $h \in BV$,

$$\mathrm{v}(\mathcal{L}_\omega^N h) \le \alpha(\omega)\mathrm{v}(h) + K(\omega)\|h\|_1.$$

- For each $a > 0$ and $\mathbb{P}$-a.e. $\omega \in \Omega$, there exist random numbers $n_c(\omega) < +\infty$ and $\alpha_0(\omega), \alpha_1(\omega), \ldots$ such that for every $h \in \mathcal{C}_a$,

$$\operatorname{essinf}_x(\mathcal{L}_\omega^n h)(x) \ge \alpha_n \|h\|_1 \quad \text{for } n \ge n_c, \tag{2.4}$$

  where

$$\mathcal{C}_a := \{h \in L^\infty(X, m) : h \ge 0 \text{ and } \mathrm{v}(h) \le a\|h\|_1\}. \tag{2.5}$$

- $\log(\operatorname{essinf}_{x\in X}(\mathcal{L}_\omega \mathbb{1})(x)) \in L^1(\Omega, \mathbb{P})$.

Finally, we say that the cocycle $\mathcal{L}$ is *uniformly random* if the random variables C, α^N, K^N and n_c are constants and $\alpha_n(\omega)$ does not depend on n and ω.

Remark 2.5

- Definition 2.4 almost coincides with [22, Definition 3], the only difference being the addition of (2.3) (which was considered in [22, §3].)
- The log-integrability assumption specified at the end of Definition 2.4 may easily be checked on explicit examples (see, for example, the discussion in [6, Remark 2.12]).
- Furthermore, this assumption implies a certain version of the 'random covering' similar to (2.4); see [22, Remark 4].

Let us now give examples of systems satisfying our requirements. Our first example is essentially taken from [12].

Example 2.6. (Lasota–Yorke cocycles) Consider $X = [0, 1]$, endowed with Lebesgue measure m and the classical notion of variation v. We say that $T : X \to X$ is a piecewise monotonic non-singular (p.m.n.s.) map if the following conditions hold.

- T is piecewise monotonic, that is, there exists a subdivision $0 = a_0 < a_1 < \cdots < a_N = 1$ such that for each $i \in \{0, \ldots, N-1\}$, the restriction $T_i = T_{|(a_i, a_{i+1})}$ is monotonic (in particular, it is a homeomorphism on its image).
- T is non-singular, that is, there exists $|T'| : [0, 1] \to \mathbb{R}_+$ such that, for any measurable $E \subset (a_i, a_{i+1})$, $m(T(E)) = \int_E |T'| \, dm$.

The intervals $(a_i, a_{i+1})_{i \in \{0,\ldots,N-1\}}$ are called the intervals of T. We also set $N(T) := N$ and $\lambda(T) := \operatorname{essinf}_{[0,1]} |T'|$.

We consider a family $(T_\omega)_{\omega \in \Omega}$ of random p.m.n.s. as above, and such that $T : \Omega \times [0, 1] \to [0, 1]$, $(\omega, x) \mapsto T_\omega(x)$ is measurable. Denoting $N_\omega = N(T_\omega)$ and $\lambda_\omega = \lambda(T_\omega)$, we make the following assumptions.

- The map $\omega \mapsto (\mathrm{v}(1/|T'_\omega|), N_\omega, \lambda_\omega, a_1, \ldots, a_{N_\omega - 1})$ is measurable.
- We have the following expanding-on-average property:

$$\lim_{K \to \infty} \int_\Omega \log \min(\lambda_\omega, K) \, d\mathbb{P}(\omega) > 0.$$

- The maps $\log(N_\omega)$ and $\log^+(N_\omega/\lambda_\omega)$ are integrable.
- The map $\log^+(\mathrm{v}(1/|T'_\omega|))$ is integrable.
- T_ω is covering, that is, for any interval $I \subset [0, 1]$, there exists a random number $n_c(\omega) > 0$ such that, for any $n \geq n_c$, one has

$$\operatorname{essinf}_{[0,1]} \mathcal{L}^n_\omega(\mathbb{1}_I) > 0. \tag{2.6}$$

- $\log(\operatorname{essinf}_{x \in X} (\mathcal{L}_\omega \mathbb{1})(x)) \in L^1(\Omega, \mathbb{P})$.

We will call a cocycle satisfying the previous assumptions an *expanding-on-average Lasota–Yorke cocycle*. For a countably valued measurable family $(T_\omega)_{\omega \in \Omega}$ of expanding-on-average Lasota–Yorke cocycles, the associated cocycle of transfer operators $(\mathcal{L}_\omega)_{\omega \in \Omega}$ is good (see [23]).

The following example can be fruitfully compared to a similar one by Kifer [38].

Example 2.7. We consider $X = \mathbb{S}^1$, endowed with the Lebesgue measure m and the notion of variation given by $\mathrm{v}(\phi) := \int_X |\phi'| \, dm = \|\phi'\|_{L^1}$. We consider a measurable map

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

$T : \Omega \times X \to X$ such that $T_\omega := T(\omega, \cdot)$ is C^r, $r \geq 2$. In addition, we make the following assumptions.

- There exists a tempered random variable $N(\omega)$ so that (2.3) holds true.
- The map $\omega \in \Omega \mapsto (\int_X (|T''_\omega|/(T'_\omega)^2)\, dm, \lambda_\omega)$ is measurable, where $\lambda_\omega = \inf_{[0,1]} |T'_\omega|$.
- The following expanding-on-average property holds:

$$\int_\Omega \log(\lambda_\omega)\, d\mathbb{P}(\omega) > 0. \tag{2.7}$$

- The map $\log(\int_X |T''_\omega|/(T'_\omega)^2\, dm)$ is $\mathbb{P}$-integrable.
- $\log(\operatorname{essinf}_{x\in X}(\mathcal{L}_\omega \mathbb{1})(x)) \in L^1(\Omega, \mathbb{P})$.

We call a family $(T_\omega)_{\omega\in\Omega}$ satisfying the previous assumptions a *smooth expanding-on-average cocycle*. For a family $(T_\omega)_{\omega\in\Omega}$, countably valued and measurable, of smooth expanding-on-average cocycles which satisfy (2.3), the associated cocycle of transfer operators $(\mathcal{L}_\omega)_{\omega\in\Omega}$ is good (see [23, Example 16]). We note that our expansion-on-average condition (2.7) implies that $\mathbb{P}$-almost surely, T_ω has non-vanishing derivative, hence is a local diffeomorphism and a monotonic map of the circle. As noted in [22, Example 6], smooth expanding-on-average cocycles satisfy a stronger version of the random covering property (which by [12, Remark 0.1] implies the one formulated in (2.6)): for each non-trivial interval $I \subset X$, for $\mathbb{P}$-a.e. $\omega \in \Omega$, there is an $n_c := n_c(\omega, I) < \infty$ such that, for all $n \geq n_c$,

$$T^n_\omega(I) = X.$$

2.2. *The one-dimensionality of the top Oseledets space: a summary of known results.* In this section we recall two results from [22] that will be in constant use in the course of the proofs of all of our results.

THEOREM 2.8. [22, Theorem 12] *Let $\mathcal{L} = (\mathcal{L}_\omega)_{\omega\in\Omega}$ be a good cocycle of transfer operators. Then the following assertions hold.*

- *There exists an essentially unique measurable family $(h_\omega)_{\omega\in\Omega} \subset BV$ such that $h_\omega \geq 0$, $\int_X h_\omega\, dm = 1$ and*

$$\mathcal{L}_\omega h_\omega = h_{\sigma\omega} \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega.$$

- *There is a random variable $\ell : \Omega \to (0, +\infty)$ such that, for $\mathbb{P}$-a.e. $\omega \in \Omega$,*

$$h_\omega \geq \ell(\omega), \quad m\text{-a.e.} \tag{2.8}$$

- *For $\mathbb{P}$-a.e. $\omega \in \Omega$,*

$$BV = \operatorname{span}\{h_\omega\} \oplus BV^0, \tag{2.9}$$

where

$$BV^0 = \left\{ h \in BV : \int_X h\, dm = 0 \right\}.$$

- *$\omega \mapsto \|h_\omega\|_{BV}$ is tempered.*

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

- *There exist $\lambda > 0$ and for each $\epsilon > 0$, a tempered random variable $D = D_\epsilon\colon \Omega \to (0, +\infty)$ such that, for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $n \in \mathbb{N}$,*

$$\|\mathcal{L}_\omega^n \Pi(\omega)\|_{BV} \le D(\omega)e^{-\lambda n} \tag{2.10}$$

and

$$\|\mathcal{L}_\omega^n (\mathrm{Id} - \Pi(\omega))\|_{BV} \le D(\omega)e^{\epsilon n}, \tag{2.11}$$

where $\Pi(\omega)\colon BV \to BV^0$ is a projection associated to the splitting (2.9).

Finally, for uniformly random cocycles the random variables $\ell(\omega)$ and $D(\omega)$ can be replaced with positive constants and $\omega \to \|h_\omega\|_{BV}$ is a bounded random variable.

COROLLARY 2.9. [22, Corollary 13] *Let $\mathcal{L} = (\mathcal{L}_\omega)_{\omega\in\Omega}$ be a good cocycle of transfer operators. Then the following assertions hold.*

- *If $(h_\omega)_{\omega\in\Omega} \subset BV$ is given by Theorem 2.8, then*

$$\omega \mapsto \|1/h_\omega\|_{BV} \text{ is tempered.} \tag{2.12}$$

- *For $\mathbb{P}$-a.e. $\omega \in \Omega$,*

$$BV = \mathrm{span}\{\mathbb{1}\} \oplus BV_\omega^0, \tag{2.13}$$

where

$$BV_\omega^0 = \left\{h \in BV : \int_X h\, d\mu_\omega = 0\right\},$$

in which $d\mu_\omega = h_\omega dm$, $\omega \in \Omega$;

- *there exist $\lambda' > 0$ and a tempered random variable $\tilde{D}\colon \Omega \to (0, +\infty)$ such that, for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $n \in \mathbb{N}$,*

$$\|L_\omega^n \tilde{\Pi}(\omega)\|_{BV} \le \tilde{D}(\omega)e^{-\lambda' n}, \tag{2.14}$$

$$\|L_\omega^n (\mathrm{Id} - \tilde{\Pi}(\omega))\|_{BV} \le \tilde{D}(\omega), \tag{2.15}$$

where $\tilde{\Pi}(\omega)\colon BV \to BV_\omega^0$ is a projection associated to the splitting (2.13), and

$$L_\omega^n g = \mathcal{L}_\omega^n(g h_\omega)/h_{\sigma^n\omega}, \quad g \in BV,\ n \in \mathbb{N}.$$

Finally, for uniformly random cocycles the random variable $\tilde{D}(\omega)$ can be replaced with a positive constant.

Since $\mathcal{L}_\omega h_\omega = h_{\sigma\omega}$ and $\mathcal{L}_\omega$ satisfy the duality relation (2.2), the measure μ_ω satisfies that for $\mathbb{P}$-a.e. ω we have $(T_\omega)_*\mu_\omega = \mu_{\sigma_\omega}$. Thus, if $\tau : \Omega \times X \to \Omega \times X$ is defined by $\tau(\omega, x) = (\sigma\omega, T_\omega x)$ then μ_ω gives rise to a τ-invariant probability measure μ on $\Omega \times X$ such that

$$\mu(A \times B) = \int_A \mu_\omega(B)\, d\mathbb{P}(\omega) = \int_{A\times B} h(\omega, x)\, d\mathbb{P}(\omega)dm(x)$$

for every measurable set A in Ω and B in X, where $h(\omega, x) = h_\omega(x)$.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

2.3. *Main results: limit theorems for mixing base maps*

2.4. *The observable.* Let us take a measurable $\varphi : \Omega \times X \to \mathbb{R}$ so that $\int \varphi \, d\mu = 0$. Let $\tilde{K}(\omega)$ be the tempered random variable defined by

$$\tilde{K}(\omega) = \max(D(\omega), \tilde{D}(\omega), N(\omega), \|1/h_\omega\|_{BV}),$$

where $D(\omega)$, $\tilde{D}(\omega)$ and $N(\omega)$ are specified in the definition of a good cocycles and in Theorem 2.8 and Corollary 2.9. In order to describe our assumptions on the observable φ, we will need the following classical result (see [4, Proposition 4.3.3.]).

PROPOSITION 2.10. *Let $\tilde{K} : \Omega \to (0, +\infty)$ be a tempered random variable. For each $\epsilon > 0$, there exists a tempered random variable $\tilde{K}_\epsilon : \Omega \to (1, +\infty)$ such that*

$$\frac{1}{\tilde{K}_\epsilon(\omega)} \le \tilde{K}(\omega) \le \tilde{K}_\epsilon(\omega) \quad \text{and} \quad \tilde{K}_\epsilon(\omega)e^{-\epsilon|n|} \le \tilde{K}_\epsilon(\sigma^n\omega) \le \tilde{K}_\epsilon(\omega)e^{\epsilon|n|},$$

for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $n \in \mathbb{Z}$.

Next, using the notation of Proposition 2.10, let $K(\omega) = \tilde{K}_\varepsilon(\omega)$ for some $\varepsilon < \lambda''/3$, where $\lambda'' = \min(\lambda, \lambda')$, and λ and λ' are specified in Theorem 2.8 and Corollary 2.9, respectively.

Remark 2.11. From now on we will replace both λ and λ' by their minimum, which for notational convenience will be denoted by λ.

In what follows we will consider an observable $\varphi : \Omega \times X \to \mathbb{R}$ satisfying the scaling condition

$$\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty \tag{2.16}$$

which was first introduced in [23]. In the uniformly random case $\tilde{K}(\omega)$ (and hence $K(\omega)$) can be replace by a positive constant, and so the scaling condition reads

$$\operatorname{esssup}_{\omega\in\Omega} \|\varphi_\omega\|_{BV} < \infty.$$

The main goal in this paper is to obtain limit theorems for the sequence of functions

$$S_n = S_n\varphi = \sum_{j=0}^{n-1} \varphi \circ \tau^j$$

under certain mixing assumptions on the driving system $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ and the above assumptions on the observable φ.

Remark 2.12. For expanding-on-average maps the scaling condition (2.16) is necessary for limit theorems (see [22, Appendix]). In any case, our results are also new in the uniformly random case, and readers who would prefer can just consider this case together with the assumption that $\operatorname{esssup}_{\omega\in\Omega} \|\varphi_\omega\|_{BV} < \infty$.

Let us also note that, in general, the random variable $K(\omega)$ comes from Oseledets theorem and it is not computable. In order to provide explicit conditions for quenched limit

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

theorems, in [28] several examples of non-uniformly expanding maps (which are stronger than expansion on average) were given with the property that

$$\|L_\omega^n - \mu_\omega\|_{BV} \leq B(\sigma^n\omega)\prod_{j=0}^{n-1}\rho(\sigma^j\omega). \tag{2.17}$$

Here the BV norm is with respect to the choice of variation $v(g) = v_\alpha(g)$, where v_α is the Hölder constant corresponding to some exponent α and $B(\omega)$ and $\rho(\omega) \in (0, 1)$ are random variables with explicit formulas, and they depend only on the zeroth coordinate ω_0. Moreover, for several of these examples we already have $B(\omega) \leq B$ for some constant B. In this case (similarly to [22, §5.2]) we have the following assertions. Let ε be smaller than $1 - \mathbb{E}_{\mathbb{P}}[\rho]$ and let $A = \{\omega : \rho(\omega) \leq 1 - \varepsilon\}$. Then $\mathbb{P}(A) > 0$. Let $R_n(\omega) = \sum_{j=0}^{n-1}\mathbb{I}(\sigma^j\omega \in A)$. Then $R_n/n \to r = \mathbb{P}(A)$ ($\mathbb{P}$-almost surely). Let

$$N(\omega) = \inf\{N : R_n(\omega) \geq \tfrac{1}{2}rn, \text{ for all } n \geq N\}.$$

Then, for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $n \in \mathbb{N}$,

$$\|L_\omega^n - \mu_\omega\|_{BV} \leq K(\omega)(1-\varepsilon)^n,$$

where $K(\omega) = B(1-\varepsilon)^{N(\omega)}$. Observe that for $k \geq 1$,

$$\{N(\omega) = k+1\} \subset \left\{\left|\frac{R_k(\omega)}{k} - r\right| > \frac{1}{2}r\right\}.$$

Thus, if the stationary process $(\mathbb{I}_A \circ \sigma^n)$ satisfies an appropriate concentration inequality (for example, under appropriate mixing assumptions on (ξ_n)), we can conclude that $N(\omega)$ is integrable. Hence, $\log K$ is integrable and consequently also tempered.

The above means that in this situation we can express the condition on φ by means of the more explicit random variable $K(\omega)$ defined above. Still, in the setup of [28], under appropriate integrability conditions on $B(\omega)$ the main results in this paper can be obtained under conditions such as $\varphi \in L^p(\mu)$ for p large enough (depending on the desired result). Since this approach requires several non-trivial modifications to the arguments in this paper such results will be considered elsewhere.

2.5. *Limit theorems.* Let us first introduce our assumptions on the base map. Let (ξ_n) be a two-sided stationary sequence taking values on some measurable space $\mathcal{Y}$. We assume here that $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is the corresponding shift system. Namely, $\Omega = \mathcal{Y}^{\mathbb{Z}}$, $(\sigma\omega)_j = (\omega_{j+1})_j$ is the left shift and if $\pi_0 : \Omega \to \mathcal{Y}$ denotes the zeroth coordinate projection, then (ξ_n) has the same distribution as $(\pi_0 \circ \sigma^n)$. We also assume that $T_\omega = T_{\omega_0}$ and $\varphi(\omega, \cdot) = \varphi(\omega_0, \cdot)$ depend only on zeroth coordinate ω_0 of ω.

2.5.1. *Limit theorems for stretched exponentially fast α-mixing driving processes.* Let $(\Omega_0, \mathscr{F}, \mathbf{P})$ be the probability space on which (ξ_n) is defined. We recall that the α-mixing (dependence) coefficient between two sub-σ-algebras $\mathcal{G}, \mathcal{H}$ of $\mathscr{F}$ is given by

$$\alpha(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)| : A \in \mathcal{G}, B \in \mathcal{H}\}.$$

The α-dependence coefficients of (ξ_n) are defined by

$$\alpha_n = \sup_k \alpha(\mathscr{F}_{-\infty,k}, \mathscr{F}_{k+n,\infty}) = \alpha(\mathscr{F}_{-\infty,0}, \mathscr{F}_{n,\infty}), \tag{2.18}$$

where $\mathscr{F}_{-\infty,k}$ is the σ-algebra generated by ξ_j, $j \le k$, and $\mathscr{F}_{k+n,\infty}$ is generated by ξ_j, $j \ge k+n$. The last equality holds true due to stationarity. Let us consider the following class of mixing assumptions on the base map.

Assumption 2.13. (Stretched exponential α mixing rates) There exist positive constants c_1, c_2 and η such that $\alpha_n \le c_1 e^{-c_2 n^\eta}$ for every n.

Our first result concerns the variance of S_n and the CLT (with rates).

THEOREM 2.14. *Suppose that the cocycle $\mathcal{L}$ is good. Let φ be an observable such that $\|\varphi\|_K := \mathrm{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$, where $\varphi_\omega = \varphi(\omega, \cdot)$. Suppose that $\sum_n n\alpha_n < \infty$. Then the limit*

$$s = \lim_{n\to\infty} n^{-1/2}\|S_n - \mathbb{E}[S_n]\|_{L^2(\mu)}$$

exists and vanishes if and only if $\varphi = r \circ \tau - r$ for some $r \in L^2(\mu)$. If in addition Assumption 2.13 is satisfied then $n^{-1/2}S_n$ converges in distribution to sZ, where Z is a standard normal random variable. Moreover, there is a constant $C > 0$ such that. for all $n \in \mathbb{N}$,

$$\sup_{t\in\mathbb{R}} |\mu(S_n - \mathbb{E}[S_n] \le ts\sqrt{n}) - \Phi(t)| \le Cn^{-1/2+4\gamma}, \tag{2.19}$$

where $\gamma = 1/\eta$ and Φ is the standard normal distribution function. The constant C depends only on c_1, c_2, η, $\|\varphi\|_K$ and the constant C_{v} (from the definition of the variation $\mathrm{v}(\cdot)$), and an explicit formula for C can be recovered from the proof.

The proof of Theorem 2.14 appears in §3.2.1. As discussed in §§1.2 and 1.3, when the quenched CLT holds true with a deterministic centering, then the CLT for the skew product follows by integration. This was the approach for the CLT in [**2**], but in the setup of this paper the function φ and the measure μ_ω depend on ω, and so the quenched CLT only holds with fiberwise centering. Thus, the novelty of Theorem 2.14 is that the CLT is obtained for the skew product beyond the annealed case considered in [**2**]. Moreover, Theorem 2.19 also strengthens the CLT in [**26**], since our maps T_ω are not uniformly expanding, and the observable φ is not fiberwise centered.

Next, let us discuss our results concerning moderate-deviations and exponential concentration inequalities.

THEOREM 2.15. *Suppose that the cocycle $\mathcal{L}$ is good, and let φ be an observable so that $\|\varphi\|_K = \mathrm{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$. Let Assumption 2.13 hold and set $\gamma = 1/\eta$. Then there exist constants $a_1, a_2 > 0$ such that, for every $x > 0$ and $n \in \mathbb{N}$,*

$$\mu(|S_n - \mathbb{E}[S_n]| \ge x) \le 2\exp\left(-\frac{x^2}{2(a_1 + a_2 x n^{-1/(2+4\gamma)})^{(1+2\gamma)/(1+\gamma)}}\right). \tag{2.20}$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

All the constants depend only on c_1, c_2, η, $\|\varphi\|_K$ *and* $C_{\rm v}$ *from the definition of the variation* $\mathrm{v}(\cdot)$, *and an explicit formula for them can be recovered from the proof.*

We will also prove the following theorem.

THEOREM 2.16. *Suppose that the cocycle* $\mathcal{L}$ *is good, and let* φ *be an observable such that* $\|\varphi\|_K = \mathrm{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$. *Let Assumption 2.13 hold and set* $\gamma = 1/\eta$. *Let us also assume that the asymptotic variance* s^2 *is positive.*

(i) *Set* $v_n = \sqrt{\mathrm{Var}(S_n)}$, *and when* $v_n > 0$ *also set* $Z_n = (S_n - \mathbb{E}[S_n])/v_n$. *Let* Φ *be the standard normal distribution function. Then there exist constants* $s_3, s_4, s_5 > 0$ *such that, that for every* $n \geq a_3$ *we have* $v_n > 0$, *and for every* $0 \leq x < a_4 n^{1/(2+4\gamma)}$,

$$\left|\ln \frac{\mu(Z_n \geq x)}{1-\Phi(x)}\right| \leq a_5(1+x^3)n^{-1/(2+4\gamma)} \quad \text{and} \\ \left|\ln \frac{\mu(Z_n \leq -x)}{\Phi(-x)}\right| \leq a_5(1+x^3)n^{-1/(2+4\gamma)}. \tag{2.21}$$

The constants a_4, a_5 *depend only on* c_1, c_2, η, $\|\varphi\|_K$ *and* $C_{\rm v}$, *and an explicit formula for them can be recovered from the proof.*

(ii) *Let* a_n, $n \geq 1$, *be a sequence of real numbers so that*

$$\lim_{n\to\infty} a_n = \infty \quad \text{and} \quad \lim_{n\to\infty} a_n n^{-1/(2+4\gamma)} = 0.$$

Then the sequence $W_n = (sn^{1/2}a_n)^{-1}S_n$, $n \geq 1$, *satisfies the moderate-deviations principle with speed* $s_n = a_n^2$ *and the rate function* $I(x) = x^2/2$. *Namely, for every Borel measurable set* $\Gamma \subset \mathbb{R}$,

$$-\inf_{x\in\Gamma^o} I(x) \leq \liminf_{n\to\infty} \frac{1}{a_n^2}\ln \mu(W_n \in \Gamma) \leq \limsup_{n\to\infty} \frac{1}{a_n^2}\ln \mu(W_n \in \Gamma) \leq -\inf_{x\in\overline{\Gamma}} I(x)$$

where Γ^o *is the interior of* Γ *and* $\overline{\Gamma}$ *is its closure.*

We also obtain the following Rosenthal-type moment estimates.

THEOREM 2.17. *Suppose that* $\mathcal{L}$ *is a good cocycle. If* $\|\varphi\|_K = \mathrm{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$, *then under Assumption 2.13 there exists a constant* c_0 *such that, with* $\gamma = 1/\eta$ *for every integer* $p \geq 1$, *we have*

$$|\mathbb{E}_\mu[(S_n - \mathbb{E}_\mu[S_n])^p] - (\mathrm{Var}_\mu(S_n))^{p/2}\mathbb{E}[Z^p]| \\ \leq (c_0)^p (p!)^{1+\gamma} \sum_{1\leq u\leq (p-1)/2} n^u \frac{p^u}{(u!)^2} = O(n^{[(p-1)/2]}), \tag{2.22}$$

where Z *is a standard normal random variable. In particular,* $\|S_n - \mathbb{E}_\mu[S_n]\|_{L^p} = O(\sqrt{n})$ *for every* p. *As in the previous theorems, the constant* c_0 *depends (explicitly) only on* c_1, c_2, η, $\|\varphi\|_K$ *and* $C_{\rm v}$.

We remark that Theorem 2.17 provides another proof of the CLT by the method of moments. Indeed, if $s^2 > 0$ then it follows that, for every integer $p \geq 1$, the pth moment of $(S_n - \mathbb{E}[S_n])n^{-1/2}s^{-1}$ converges to $\mathbb{E}[Z^p]$, where s^2 is the asymptotic variance. In fact, for even p we get the convergence rate $O(n^{-1/2})$, while for odd p we get the rate $O(n^{-1})$.

Remark 2.18. The proofs of Theorems 2.15–2.17 appear in §3.2.2.

Theorems 2.15–2.17 are well established for sufficiently fast mixing (in the probabilistic sense) sequences of random variables, where one of the most notable methods of proof is the so-called method of cumulants (see [**49**]). For random dynamical systems, a moderate-deviations principle was obtained in [**19**], using a random complex Perron–Frobenius theorem. In the setup of [**2**], annealed (local) large-deviations principles and exponential concentration inequalities were obtained for i.i.d. maps, and we expect that for independent maps the methods in [**2**] will yield results like Theorems 2.15–2.17 as well. The novelty in Theorems 2.15–2.17 is that we show how to apply the method of cumulants in the context of skew products with non-independent fiber maps, which results in concentration inequalities, moderate-deviations principles and Gaussian moment estimates beyond the annealed setup [**2**].

Finally, let us consider the random function $\mathcal{S}_n(t) = n^{-1/2}(S_{[nt]} - \mathbb{E}[S_{nt}])$ on $[0, 1]$. We also obtain a functional CLT.

THEOREM 2.19. *Let $\mathcal{L}$ be a good cocycle. Suppose that* $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$ *and that Assumption 2.13 holds true. Then the random function $\mathcal{S}_n$ converges in distribution towards the distribution of $\{sW_t\}$, where W is a standard Brownian motion (restricted to* $[0, 1]$*) and s^2 is the asymptotic variance.*

Remark 2.20. The proof of Theorem 2.19 appears in §3.3. In [**2**] an ASIP was obtained, which yields the functional CLT. In §2.5.2 below, using different mixing coefficients for the base map, we will obtain an ASIP for the more general skew products considered in this paper. However, Theorem 2.19 shows that the functional CLT already holds true for stretched exponential α-mixing base maps.

2.5.2. *An almost sure invariance principle and exponential concentration inequalities for ϕ- and ψ-mixing driving processes (via martingale methods).* Let $(\Omega_0, \mathscr{F}, \mathbf{P})$ be the probability space on which (ξ_n) is defined. We recall that the ϕ-mixing and ψ (dependence) coefficient between two sub-σ-algebras $\mathcal{G}, \mathcal{H}$ of $\mathscr{F}$ is given by

$$\phi(\mathcal{G}, \mathcal{H}) = \sup\{|\mathbf{P}(B|A) - \mathbf{P}(B)| : A \in \mathcal{G}, B \in \mathcal{H}, \mathbf{P}(A) > 0\}$$

and

$$\psi(\mathcal{G}, \mathcal{H}) = \sup\left\{\left|\frac{\mathbf{P}(A\cap B)}{\mathbf{P}(A)\mathbf{P}(B)} - 1\right| : A \in \mathcal{G}, B \in \mathcal{H}, \mathbf{P}(A)\mathbf{P}(B) > 0\right\}.$$

The reverse ϕ-mixing coefficients of (ξ_n) are defined by

$$\phi_{n,R} = \sup_k \phi(\mathscr{F}_{k+n,\infty}, \mathscr{F}_{-\infty,k}) = \phi(\mathscr{F}_{n,\infty}, \mathscr{F}_{-\infty,0}), \tag{2.23}$$

while the ψ-mixing coefficients of (ξ_n) are defined by

$$\psi_n = \sup_k \psi(\mathscr{F}_{-\infty,k}, \mathscr{F}_{k+n,\infty}) = \psi(\mathscr{F}_{-\infty,0}, \mathscr{F}_{n,\infty}), \tag{2.24}$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

where $\mathscr{F}_{-\infty,k}$ is the σ-algebra generated by ξ_j, $j \leq k$, and $\mathscr{F}_{k+n,\infty}$ is generated by ξ_j, $j \geq k+n$. It is clear from the definitions of the mixing coefficients that

$$\alpha_n \leq \phi_{n,R} \leq \psi_n.$$

THEOREM 2.21. (Exponential concentration and maximal inequalities) *Let $\mathcal{L}$ be a good cocycle. Suppose the observable satisfies* $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi_\omega\|_{BV}) < \infty$.

Let $\mathcal{F}_0$ be the σ algebra generated by the map $\pi(\omega, x) = ((\omega_j)_{j\geq 0}, x)$, namely the one generated by $\mathcal{B}$ and the coordinates with non-negative indexes in the ω direction. If either $\operatorname{essinf}\inf_x h_\omega(x) > 0$ *and $\sum_n \phi_{n,R} < \infty$ or $\sum_n \psi_n < \infty$ then there is an $\mathcal{F}_0$-measurable function $\chi \in L^\infty(\mu)$ such that if we set $u = \varphi + \chi\circ\tau - \chi$ then $(u\circ\tau^n)$ is a reverse martingale difference with respect to the reverse filtration $\{\tau^{-n}\mathcal{F}_0\}$. As a consequence, we have the following assertions.*

(i) *There are constants $a_1, a_2, a_3 > 0$ such that the following exponential concentration inequality holds true: for every $t > 0$, we have*

$$\mu(|S_n - \mathbb{E}_\mu[S_n]| \geq tn + a_1) \leq a_2 e^{-a_3 n t^2}. \tag{2.25}$$

The constants a_1, a_2, a_3 depend only on $\tilde{\Phi} = \sum_n \phi_{n,R} < \infty$ and c (or $\tilde{\Psi} = \sum_n \psi_n < \infty$), the constant C_{v} and $\|\varphi\|_{K,2} = \operatorname{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi_\omega\|_{BV})$, and an explicit formula for them can be recovered from the proof.

(ii) *For every $p \geq 2$, we have*

$$\Big\|\max_{1\leq k\leq n} |S_k - \mathbb{E}[S_k]|\Big\|_{L^p(\mu)} \leq C_p n^{1/2}, \tag{2.26}$$

where $C_p > 0$ is a constant (which can be recovered from the proof and depends only on p and the above constants).

We refer readers to [43] for some related moment bounds for random intermittent maps.

The proof of Theorem 2.21 appears in §4. Let us note that once the martingale-coboundary representation $\varphi = u + \chi - \chi\circ\tau$ is established, Theorem 2.21(i) follows from the Azuma–Hoeffding inequality together with Chernoff's bounding method, and Theorem 2.21(ii) follows from the so-called Rio inequality [48] (see [45, Proposition 7]).

To obtain the martingale-coboundary representation we show that if $\mathcal{K}$ is the transfer operator (namely, the one satisfying the duality relation

$$\int (\mathcal{K}g)\cdot f\, d\mu = \int g\cdot(f\circ\tau), \quad g \in L^1(\Omega\times X, \mathcal{F}_0, \mu),\ f \in L^\infty(\Omega\times X, \mathcal{F}_0, \mu)).$$

corresponding to the system $(\Omega\times X, \mathcal{F}_0, \mu, \tau)$ then there is a constant $C > 0$ such that

$$\|\mathcal{K}^n\varphi - \mu(\varphi)\|_{L^\infty} \leq C(\delta^n + \gamma_{[n/2]}), \tag{2.27}$$

where γ_n is either ψ_n or $\phi_{n,R}$, depending on the case, and $\delta \in (0, 1)$. Once this is established we can take

$$\chi = \sum_{n\geq 1} \mathcal{K}^n\varphi.$$

The proof of (2.27) is given in Proposition 4.3 (i).

Our next result is an ASIP.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

THEOREM 2.22. (ASIP) *Let $\mathcal{L}$ be a good cocycle, and suppose that the observable satisfies* $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi_\omega\|_{BV}) < \infty$.

When $\operatorname{essinf}\ \inf_x h_\omega(x) > 0$ *we set* $\gamma_n = \phi_{R,n}$*, while otherwise we set* $\gamma_n = \psi_n$*. In both cases, assume that*

$$\sum_{n\geq 2} n^{5/2}(\log n)^3\gamma_n^4 < \infty \quad \text{and} \quad \sum_{n\geq 2} n(\log n)^3\gamma_n^2 < \infty,$$

and

$$\sum_{n\geq 2}\frac{(\log n)^3}{n^2}\bigg(\sum_{k=0}^{n}(k+1)\gamma_k\bigg)^2 < \infty.$$

Then the limit

$$s^2 = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[(S_n - \mathbb{E}[S_n])^2]$$

exists and the following version of the ASIP holds true: there is a coupling of $(\varphi\circ\tau^n)$ *with a sequence of i.i.d. Gaussian random variables* Z_j *with zero mean and variance* s^2 *such that*

$$\sup_{1\leq k\leq n}\bigg|(S_k - \mathbb{E}[S_k]) - \sum_{j=1}^{k} Z_j\bigg| = O(n^{1/4}(\log n)^{1/2}(\log\log n)^{1/4}) \quad \textit{almost surely}.$$

Remark 2.23. The ASIP implies the functional CLT, see [**47**]. Thus, Theorem 2.22 yields better results than Theorem 2.19 for ϕ_R- or ψ-mixing driving sequences (which are not necessarily stretched exponentially mixing).

The proof of Theorem 2.22 appears in §4 and relies on an application of [**13**, Theorem 3.2]. In addition to (2.27), in order to apply [**13**, Theorem 3.2] we will show that for all $1 \leq i, j \leq n$ we have

$$\|\mathcal{K}^i(\bar\varphi\mathcal{K}^j\bar\varphi) - \mu(\mathcal{K}^i(\bar\varphi\mathcal{K}^j\bar\varphi))\|_{L^\infty} \leq C(\delta^n + \gamma_n), \tag{2.28}$$

where $\bar\varphi = \varphi - \mu(\varphi)$, C is a constant and δ and γ_n are as in (2.27). The proof of (2.28) is given in Proposition 4.3 (ii).

Remark 2.24. As discussed in §1.3.2, the martingale-coboundary decomposition in Theorem 2.21 (and its consequences) is comparable with the annealed case [**2**], and the main novelty is that we obtain it for more general skew products and functions φ which depend on ω. Moreover, we do not assume that all T_ω preserve the same absolutely continuous probability measure. The ASIPs we obtain are comparable to ASIPs in [**2**] (see the discussion in §1.3.2).

2.5.3. *A vector-valued almost sure invariance principle in the uniformly random case for exponentially fast α-mixing base maps.* Let us take a vector-valued measurable function $\varphi = (\varphi_1 \ldots \varphi_d) : \Omega\times X \to \mathbb{R}^d$ such that $\varphi_\omega = \varphi(\omega,\cdot)$ depend on ω only through ω_0 and $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_{\omega,i}\|_{BV}) < \infty$ for all $1 \leq i \leq d$. Let us also assume that $\mu(\varphi_i) = 0$ for every i. Set $S_n = \sum_{j=0}^{n-1}\varphi\circ\tau^j$.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

THEOREM 2.25. *Suppose that $\alpha_n = O(\alpha^n)$ for some $\alpha \in (0, 1)$. Then there is a positive semidefinite matrix Σ^2 such that*

$$\Sigma^2 = \lim_{n\to\infty} \frac{1}{n}\mathrm{Cov}(S_n).$$

Moreover, Σ^2 is positive definite if and only if $\varphi \cdot v \neq r - r \circ \tau$ for all unit vectors v and all $r \in L^2$.

Assume now that there are constants $C > 0$ and $\delta \in (0, 1)$ so that

$$\|\mathcal{L}_\omega^n \mathbf{1} - h_{\sigma^n\omega}\|_{BV} \leq C\delta^n, \tag{2.29}$$

namely, that $K(\omega)$ is a bounded random variable. Then there is a coupling of $(\varphi \circ \tau^n)$ with a sequence of independent Gaussian centered random vectors (Z_n) such that $\mathrm{Cov}(Z_n) = \Sigma^2$ and for every $\varepsilon > 0$,

$$\left|(S_n - \mathbb{E}[S_n]) - \sum_{j=1}^{n} Z_j\right| = o(n^{1/4+\varepsilon}) \quad \textit{almost surely.}$$

3. *Limit theorems via the method of cumulants for α-mixing driving processes*

We recall next that the kth cumulant of a random variable W with finite moments of all orders is given by

$$\Gamma_k(W) = \frac{1}{i^k}\frac{d^k}{dt^k}(\ln \mathbb{E}[e^{itW}])|_{t=0}.$$

Note that $\Gamma_1(W) = \mathbb{E}[W]$, $\Gamma_2(W) = \mathrm{Var}(W)$, and $\Gamma_k(aW) = a^k\Gamma_k(W)$ for any $a \in \mathbb{R}$ and $k \geq 1$.

From now on we will assume that $\mathbb{E}[S_n] = 0$ for all n, that is, we will replace φ by $\varphi - \mu(\varphi)$. The main result in this section is the following theorem.

THEOREM 3.1. *Let $\mathcal{L}$ be a good cocycle, and suppose that Assumption 2.13 holds true and that $\|\varphi\|_K = \mathrm{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$. Then, with $\gamma = 1/\eta$, there exists a constant c_0 which depends only on $\|\varphi\|_K$ and the constants from Assumption 2.13 such that, for any $k \geq 3$,*

$$|\Gamma_k(S_n)| \leq n(k!)^{1+\gamma}(c_0)^{k-2}.$$

We will prove Theorem 3.1 by applying the following Proposition 3.3, which appears in [25] as Corollary 3.2.

Let us start with a few preparations. Let V be a finite set and $\rho : V \times V \to [0, \infty)$ be such that $\rho(v, v) = 0$ and $\rho(u, v) = \rho(v, u)$ for all $u, v \in V$. For every $A, B \subset V$ set

$$\rho(A, B) = \min\{\rho(a, b) : a \in A, b \in B\}.$$

We assume here that there exist $c_0 \geq 1$ and $u_0 \geq 0$ such that

$$|\{u \in V : \rho(u, v) \leq s\}| \leq c_0 s^{u_0} \tag{3.1}$$

for all $v \in V$ and $s \geq 1$.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Next, let $X_v,\ v \in V$ be a collection of centered random variables with finite moments of all orders, and for each $v \in V$ and $t \in (0, \infty]$ let $\varrho_{v,t} \in (0, \infty]$ be such that $\|X_v\|_t \le \varrho_{v,t}$.

Assumption 3.2. For some $0 < \delta \le \infty$ and all $k \ge 1$, $b > 0$ and a finite collection A_j, $j \in \mathcal{J}$, of (non-empty) subsets of V such that $\min_{i \ne j} \rho(A_i, A_j) \ge b$ and $r := \sum_{j \in \mathcal{J}} |A_j| \le k$, we have

$$\bigg|\mathbb{E}\bigg[\prod_{j\in\mathcal{J}}\prod_{i\in A_j} X_i\bigg] - \prod_{i\in\mathcal{J}}\mathbb{E}\bigg[\prod_{j\in A_j} X_i\bigg]\bigg| \le (r-1)\bigg(\prod_{i\in\mathcal{J}}\prod_{i\in A_j}\varrho_{i,(1+\delta)k}\bigg)\gamma_\delta(b,k), \quad (3.2)$$

where $\gamma_\delta(b, r)$ is some non-negative number which depends only on δ, b and r, and $|\Delta|$ stands for the cardinality of a finite set Δ.

Set $W = \sum_{v \in V} X_v$. In the course of the proof of Theorems 2.14–2.16 and 2.19 we will need the following general result.

PROPOSITION 3.3. [**25**, Corollary 3.2] *Suppose that inequality (3.1) and Assumption 3.2 are in force. Assume also that*

$$\tilde{\gamma}_\delta(m, k) := \max\{\gamma_\delta(m, r)/r : 1 \le r \le k\} \le d e^{-am^\eta}$$

for some $a, \eta > 0$, $d \ge 1$ and all $k, m \ge 1$. Then there exists a constant c which depends only on c_0, a, u_0 and η such that, for every $k \ge 2$,

$$|\Gamma_k(W)| \le d^k |V| c^k (k!)^{1+(u_0/\eta)} (M_k^k + M_{(1+\delta)k}^k) \quad (3.3)$$

where for all $q > 0$,

$$M_q = \max\{\varrho_{v,q} : v \in V\} \quad \text{and} \quad M_q^k = (M_q)^k.$$

When the X_v are bounded and (3.2) holds true with $\delta = \infty$ we can always take $\varrho_{v,t} = \varrho_{v,\infty}$, $t > 0$, and then, for any $k \ge 2$,

$$|\Gamma_k(W)| \le 2d^k |V| M_\infty^k c^k (k!)^{1+(u_0/\eta)}. \quad (3.4)$$

When $\delta < \infty$ and there exist $\theta \ge 0$ and $M > 0$ such that

$$(\varrho_{v,k})^k \le M^k (k!)^\theta \quad (3.5)$$

for any $v \in V$ and $k \ge 1$, we have that, for any $k \ge 2$,

$$|\Gamma_k(W)| \le 3C^{\theta/(1+\delta)} d^k |V| c^k (1+\delta)^k M^k (k!)^{1+(u_0/\eta)+\theta}, \quad (3.6)$$

where C is some absolute constant.

Theorem 3.1 will follow from the following result, which is proved in §3.1.

PROPOSITION 3.4. *For a good cocycle $\mathcal{L}$ and an observable φ satisfying (2.16) we have the following assertion. Fix some n and set $V = \{0, 1, \ldots, n-1\}$ and $X_v = \varphi \circ \tau^v$. Set also $\rho(x, y) = |x - y|$, and let $t = \delta = \infty$, $\gamma_\infty(b, k) = \gamma_b = e^{-(\lambda-\varepsilon)b/3} + \alpha_{[b/3]}$. Then condition (3.2) holds true with the above choices and with*

$$\varrho_{v,\infty} = A_0 \max(\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}), \|\varphi\|_{L^\infty}),$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

where A_0 is a constant which depends only on $\lambda - 3\varepsilon$ and on the constant C so that $\sup|g| \le C\|g\|_{BV}$ *for every function $g : X \to \mathbb{C}$ (and the dependence can be easily recovered from the proof).*

If, in addition, Assumption 2.13 holds then the conditions of Proposition 3.3 hold true with $u_0 = 1$, $c_0 = 2$ and $\gamma = 1/\eta$.

3.1. *Multiple correlation estimates: proof of Proposition 3.4.* Our goal is to show that (3.2) holds true with the desired upper bounds. We first need the following result.

LEMMA 3.5. *For every pair of measurable functions g, h on $\mathcal{Y}^{\mathbb{N}}$ with $g, h \in L^\infty$(with respect to the law of (ξ_n)) and all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, we have*

$$\begin{aligned}|\mathbb{E}[g(\ldots, \xi_{k-1}, \xi_k)h(\xi_{k+n}, \xi_{k+n+1}, \ldots)] \\ - \mathbb{E}[g(\ldots, \xi_{k-1}, \xi_k)] \cdot \mathbb{E}[h(\xi_{k+n}, \xi_{k+n+1}, \ldots)]| \\ \le \tfrac{1}{4}\|g(\ldots, \xi_{k-1}, \xi_k)\|_{L^\infty}\|h(\xi_{k+n}, \xi_{k+n+1}, \ldots)\|_{L^\infty}\alpha_n.\end{aligned} \tag{3.7}$$

Proof. By [11, Ch. 4], we have

$$\alpha(\mathcal{G}, \mathcal{H}) = \tfrac{1}{4}\sup\{\|\mathbb{E}[h|\mathcal{G}] - \mathbb{E}[h]\|_{L^1} : h \in L^\infty(\Omega, \mathcal{G}, \mathbf{P}), \|h\|_{L^\infty} \le 1\}.$$

Taking $g = g(\ldots, \xi_{k-1}, \xi_k)$ and $h = h(\xi_{k+n}, \xi_{k+n+1}, \ldots)$, $\mathcal{G} = \mathscr{F}_{-\infty,k}$ and $\mathcal{H} = \mathscr{F}_{k+n,\infty}$, we get

$$|\mathbb{E}[hg] - \mathbb{E}[g]\mathbb{E}[h]| = |\mathbb{E}[([h|\mathcal{G}] - \mathbb{E}[h])g]| \le \tfrac{1}{4}\alpha(\mathcal{G}, \mathcal{H})\|g\|_{L^\infty}\|h\|_{L^\infty}. \qquad \square$$

Next, is it clearly enough to prove Proposition 3.4 when $\|\varphi\|_{L^\infty}$ and $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV})$ do not exceed 1, for otherwise we can just divide φ by the maximum between the two. Recall also our assumption that $K(\omega)e^{-\varepsilon|m|} \le K(\sigma^m\omega) \le K(\omega)e^{\varepsilon|m|}$ for some $\varepsilon < \lambda/3$ (recall Remark 2.11).

The first step in the proof of Proposition 3.4 is the following result.

LEMMA 3.6. (Fiberwise multiple correlation estimates) *Let $B_1, B_2, \ldots, B_m$ be non-empty intervals in the non-negative integers so that B_i is to the left of B_{i+1} and B_1 contains* 0*. Let us denote by d_i the gap between B_i and B_{i+1} (namely, the distance). Let us fix some ω and let f_i be a family of functions such that $K(\sigma^i\omega)\|f_i\|_{BV} \le 1$ and $\|f_i\|_{L^\infty} \le 1$. Let us define $F_j = F_{B_j,\omega} = \prod_{i\in B_j} f_i \circ T^i_\omega$. Then*

$$\left|\left(\int\left(\prod_{j=1}^m F_j\right)d\mu_\omega\right) - \left(\prod_{j=1}^m\int F_j\,d\mu_\omega\right)\right| \le A\sum_{j=1}^{m-1} e^{-(\lambda-\varepsilon)d_j},$$

where $A = C^2\sup_{d\in\mathbb{N}} 2de^{-(\lambda-\varepsilon)d}$ and λ comes from (2.10) and (2.14) (recall Remark 2.11).

Proof. The proof will proceed by induction on m. Let us first prove the lemma in the case $m = 2$. We first note that for all functions $g_0, g_1, \ldots, g_q$, we have

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

$$\mathrm{v}\bigg(\prod_{k=0}^{q} g_k \circ T_\omega^k\bigg) \le \sum_{k=0}^{q} \bigg(\prod_{0\le s<k} \|g_s\|_\infty\bigg) \cdot (\mathrm{v}(g_k \circ T_\omega^k)) \cdot \bigg(\prod_{k<s\le q} \|g_s\|_\infty\bigg)$$

where $\|g_i\|_\infty = \sup \|g_i\|_{L^\infty}$, and hence

$$\bigg\|\prod_{k=0}^{q} g_k \circ T_\omega^k\bigg\|_{BV} \le \prod_{k=0}^{q} \|g_k\|_\infty + \sum_{k=0}^{q} \bigg(\prod_{0\le s<k} \|g_s\|_\infty\bigg) \times \bigg(\prod_{s=0}^{k-1} K(\sigma^s\omega)\mathrm{v}(g_k)\bigg)\bigg(\prod_{k<s\le q} \|g_s\|_\infty\bigg), \tag{3.8}$$

where we have used (2.3), that $N(\omega) \le K(\omega)$ and that

$$\bigg\|\prod_{k=0}^{q} g_k \circ T_\omega^k\bigg\|_{L^1} \le \bigg\|\prod_{k=0}^{q} g_k \circ T_\omega^k\bigg\|_{L^\infty} \le \prod_{k=0}^{q} \|g_k\|_\infty.$$

Let us write $B_1 = \{0, 1, \ldots, d\}$. Taking $g_k = f_k$ for $0 \le k \le d = q$ and noting that $K(\sigma^s\omega)\|g_s\|_\infty \le C$ for some constant C which depends (C is a constant which satisfies $\|g\|_\infty = \sup |g| \le C\|g\|_{BV}$ for every complex function on X) only the space X, we conclude that

$$\|F_1\|_{BV} \le C(d+1) \le 2Cd.$$

Now, if we write $B_2 = \{d+n, d+n+1, \ldots, d+n+L\}$ then

$$\mu_\omega(F_1F_2) = \mu_\omega(F_1 \cdot G_2 \circ T_\omega^{d+n}) = \mu_{\sigma^{n+d}\omega}(G_2 L_\omega^{n+d} F_1),$$

where

$$G_2 = \prod_{u\in B_2} f_u \circ T_{\sigma^u\omega}^{u-n-d}.$$

By (2.14) we have

$$\|L_\omega^{n+d} F_1 - \mu_\omega(F_1)\|_{BV} \le K(\omega)\|F_1\|_{BV} e^{-\lambda(d+n)} \le 2dCK(\omega)e^{-\lambda(d+n)}.$$

Therefore, using also that μ_ω is an equivariant family and that (since $n+d \in B_2$)

$$\|G_2\|_{L^\infty} \le \|f_{n+d}\|_{L^\infty} \le CK(\sigma^{n+d}\omega)^{-1},$$

we get that

$$\begin{aligned}
|\mu_\omega(F_1F_2) - \mu_\omega(F_1)\mu_\omega(F_2)| &= |\mu_{\sigma^{n+d}\omega}(G_2L_\omega^{n+d}F_1) - \mu_\omega(F_1)\mu_{\sigma^{d+n}\omega}(G_2)| \\
&= \bigg|\int (L_\omega^{d+n}F_1 - \mu_\omega(F_1))G_2\, d\mu_{\sigma^{d+n}\omega}\bigg| \\
&\le 2dCK(\omega)e^{-\lambda(d+n)}\|G_2\|_{L^\infty} \\
&\le 2dCK(\omega)e^{-\lambda(d+n)}K(\sigma^{n+d}\omega)^{-1} \\
&\le 2dC^2e^{-(\lambda-\varepsilon)(d+n)} \\
&= (2C^2de^{-(\lambda-\varepsilon)d})e^{-(\lambda-\varepsilon)n}.
\end{aligned}$$

This proves the lemma for $m = 2$.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Next, let us complete the induction step. Let d be the right end point of B_{m-1}. Then $d+d_m$ is the left end point of B_m and we can write

$$\mu_\omega\bigg(\prod_k F_k\bigg) = \mu_\omega\bigg(\prod_{k<m} F_k \cdot (G_m \circ T_\omega^{d+d_m})\bigg) = \mu_{\sigma^{d+d_m}\omega}\bigg(L_\omega^{d+d_m}\bigg(\prod_{k<m} F_k\bigg)\cdot G_m\bigg),$$

where G_m is some function. Now we observe that

$$\bigg\|\prod_{k<m} F_k\bigg\|_{BV} \le C(d+1) \le 2Cd,$$

which is proved exactly as in the previous case (even though there are gaps between the blocks B_j, we can set $g_i = 1$ when i does not belong to one of the B_j, and then $\mathrm{v}(g_i) = 0$). Thus, as in the case $m = 2$, we have

$$\bigg|\mu_\omega\bigg(\prod_k F_k\bigg) - \mu_\omega(F_m)\mu_\omega\bigg(\prod_{k<m} F_k\bigg)\bigg| \le (2C^2 de^{-(\lambda-\varepsilon)d})e^{-(\lambda-\varepsilon)d_m}.$$

The induction is completed by the above inequality, taking into account that $|\mu_\omega(F_m)| \le 1$. □

Integrating over ω yields the following corollary of Lemma 3.6.

COROLLARY 3.7. *Let τ be the skew product. Let B_j, $1 \le j \le m$, be blocks as in Lemma 3.6. Set $G_j = \prod_{i\in B_j} \varphi \circ \tau^i$. Let us denote by b_j the left end point of B_j. Then*

$$\bigg|\int \prod_{j=1}^m G_j \, d\mu - \int \bigg(\prod_{j=1}^m \int \bigg(\prod_{i\in B_j} \varphi_{\sigma^i\omega} \circ T_{\sigma^{b_j}\omega}^{i-b_j}\bigg) d\mu_{\sigma^{b_j}\omega}\bigg) d\mathbb{P}(\omega)\bigg| \le A \sum_{j=1}^d e^{-\lambda d_j}. \tag{3.9}$$

The next step of the proof is to estimate the second term inside the absolute value on the left-hand side of (3.9). To obtain appropriate estimates, we first need the following lemma.

LEMMA 3.8. *Let us fix some $k \in \mathbb{N}$ and set*

$$F_\omega = \prod_{j=0}^k \varphi_{\sigma^k\omega} \circ T_\omega^k.$$

Then, for every $n \in \mathbb{N}$ and for $\mathbb{P}$-a.e. ω, we have

$$|\mu_\omega(F_\omega) - m(F_\omega \mathcal{L}_{\sigma^{-n}\omega}^n \mathbf{1})| \le Ce^{-n(\lambda-\varepsilon)},$$

where C is such that $\|g\|_{L^\infty} \le C\|g\|_{BV}$ for every function g on X with bounded variation (recall that such a constant C exists by our assumption on the variation $\mathrm{v}(\cdot)$).

Proof. Using (2.10), that $K(\sigma^{-n}\omega) \le e^{\varepsilon n}K(\omega)$ and that $\|F_\omega\|_{L^\infty} \le \|\varphi_\omega\|_{L^\infty} \le C\|\varphi_\omega\|_{BV} \le CK(\omega)^{-1}$, we obtain that

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

$$
\begin{aligned}
|\mu_\omega(F_\omega) - m(F_\omega \mathcal{L}^n_{\sigma^{-n}\omega}\mathbf{1})| &= \bigg| \int (h_\omega - \mathcal{L}^n_{\sigma^{-n}\omega}\mathbf{1}) F_\omega \, dm \bigg| \\
&\le C K(\omega)^{-1} \int |h_\omega - \mathcal{L}^n_{\sigma^{-n}\omega}\mathbf{1}| \, dm \\
&\le K(\omega)^{-1} e^{-\lambda n} K(\sigma^{-n}\omega) \le C e^{-n(\lambda-\varepsilon)}. \qquad \square
\end{aligned}
$$

Taking into account that $|\mu_\omega(F_\omega)| \le 1$, that $|m(F_\omega \mathcal{L}^n_{\sigma^{-n}\omega}\mathbf{1})| = |m(F_\omega \circ T^n_{\sigma^{-n}\omega})| \le 1$ and that $|\prod_j \alpha_j - \prod_j \beta_j| \le \sum_j |\alpha_j - \beta_j|$ for all numbers α_j, β so that $|\alpha_j|, |\beta_j| \le 1$, we get the following result directly from Corollary 3.7 and Lemma 3.8.

COROLLARY 3.9. *Let b_j be the left end point of the block B_j. Let us also set $r_j = d_j/3$ and $r_0 = r_1$. Then there exists a constant $A_1 > 0$ which does not depend on ω or on the blocks so that in the notation of Corollary 3.7 and Lemma 3.8 we have*

$$
\bigg| \int \prod_{j=1}^{m} G_j \, d\mu - \int \bigg(\prod_{j=0}^{d} m(\varphi_{\omega,j} \mathcal{L}^{d_j}_{\sigma^{b_j-d_j}\omega} \mathbf{1}) \bigg) \, d\mathbb{P}(\omega) \bigg| \le A_1 \sum_{j=1}^{m-1} e^{-(\lambda-\varepsilon) r_j},
$$

where

$$
\varphi_{\omega,j} = \prod_{i \in B_j} \varphi_{\sigma^i\omega} \circ T^{i-b_j}_{\sigma^{b_j}\omega}.
$$

Now we observe that $m(\varphi_{\omega,j} \mathcal{L}^{d_j}_{\sigma^{n_j-d_j}\omega}\mathbf{1})$ is a function of $\xi_{b_j-r_j} \ldots \xi_{b_{j+1}-r_j}$ (that is, of the coordinates $\omega_{b_j-r_j} \ldots \omega_{b_{j+1}-r_j}$). Namely, in distribution it can be written as

$$
m(\varphi_{\omega,j} \mathcal{L}^{d_j}_{\sigma^{n_j-d_j}\omega}\mathbf{1}) = f_j(\xi_{b_j-r_j} \ldots \xi_{b_{j+1}-r_j})
$$

for some measurable function f_j. Since $m(\varphi_{\omega,j} \mathcal{L}^{d_j}_{\sigma^{n_j-d_j}\omega}\mathbf{1}) = m(\varphi_{\omega,j} \circ T^{d_j}_{\sigma^{n_j-d_j}\omega})$ and $|\varphi_{\omega,j}| \le 1$, we can ensure that $|f_j| \le 1$. Using [**25**, (2.20)] and Corollary 3.9 we conclude that the following result holds.

COROLLARY 3.10. *Let G_j, $1 \le j \le m$, be as in Corollary 3.7 (defined by some blocks B_j with gaps d_j). There are constants $A > 1$ and $\delta_0 \in (0, 1)$ which do not depend on the blocks so that*

$$
\bigg| \int \bigg(\prod_{j=1}^{m} G_j \bigg) d\mu - \bigg(\prod_{j=1}^{m} \int G_j \, d\mu \bigg) \bigg| \le A \sum_{j=1}^{m} (\delta_0^{r_j} + \alpha([r_j])).
$$

All that is left is to notice that Corollary 3.10 is a reformulation of Proposition 3.4, using the notation of this section.

3.2. *Limit theorems via the method of cumulants*

3.2.1. *The central limit theorem: proof of Theorem 2.14.* First, by Proposition 3.4 we have that (3.2) holds true with the numbers $\varrho_{i,(1+\delta)k}$ and $\gamma_\delta(b, k)$ specified in Proposition 3.4. By taking $r = 2$, $A_1 = \{0\}$ and $A_2 = \{n\}$ in (3.2) we see that

$$
|\mathbb{E}_\mu[\varphi \cdot \varphi^n]| = O(\delta^n + \alpha_{[n/3]})
$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

for some $\delta \in (0, 1)$. Hence, if $\sum n\alpha_n < \infty$ then $\sum_n n|\mathbb{E}_\mu[\varphi \cdot \varphi^n]| < \infty$ and the results concerning the asymptotic variance s^2 follow from the general theory of (weakly) stationary processes (see [34] and Lemma 3.11 below).

Now suppose that $s^2 = \lim_{n\to\infty}(1/n)\mathrm{Var}_\mu(S_n) > 0$, where $S_n = S_n\varphi$. To prove the CLT and the convergence rate (2.19), by applying [49, Corollary 2.1], taking into account Theorem 3.1, we get the CLT and the rate (2.19) for $S_n/\sqrt{\mathrm{Var}(S_n)}$. To get the same rate for $S_n/\sqrt{n}$ we need the following general fact from the theory of stationary real-valued sequences, which for the sake of convenience is stated as a lemma.

LEMMA 3.11. *Let Y_n be a centered weakly stationary sequence of square integrable random variables. Set $b_n = \mathbb{E}[Y_0 Y_n]$ and $S_n = \sum_{j=1}^n Y_j$. Suppose that $\sum_k k|b_k| < \infty$. Then*

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[S_n^2] = b_0 + 2\sum_{n\geq 1} b_n := s^2$$

and

$$\left|\frac{1}{n}\mathbb{E}[S_n^2] - s^2\right| \leq 2n^{-1}\sum_{k=1}^{\infty} k|b_k|.$$

Let us give a reminder of the short proof. We have $(1/n)\mathbb{E}[S_n^2] = \sum_{k=1}^{n-1}(1-k/n)b_k + b_0$ and so

$$\begin{aligned}\left|\frac{1}{n}\mathbb{E}[S_n^2] - s^2\right| &= \left|2\sum_{k=n}^{\infty} b_k + 2n^{-1}\sum_{k=1}^{n-1} kb_k\right| \\ &\leq 2n^{-1}\left(\sum_{k=n}^{\infty} k|b_k| + \sum_{k=1}^{n-1} k|b_k|\right) \leq 2n^{-1}\sum_{k\geq 1} k|b_k|.\end{aligned}$$

Using this lemma together with [29, Lemma 3.3] with $a = 2$ and that

$$\left\|\frac{S_n}{\sqrt{\mathrm{Var}(S_n)}} - \frac{S_n}{s\sqrt{n}}\right\|_{L^2} = \|S_n\|_{L^2}\left|\frac{1}{\sqrt{\mathrm{Var}(S_n)}} - \frac{1}{s\sqrt{n}}\right| = O(n^{1/2})\cdot O(n^{-3/2}) = O(n^{-1}),$$

we obtain (2.19).

3.2.2. *A moderate-deviations principle, stretched exponential concentration inequalities and Rosenthal-type estimates: proof of Theorems 2.15–2.17.* First, Theorem 2.15 follows from Theorem 3.1 and [49, Lemma 2.3]. The estimates (2.21) stated in Theorem 2.16 follow from Theorem 3.1 and [15, Lemma 2.3] (which is a consequence of [49, Lemma 2.3]). The moderate-deviations principle stated in Theorem 2.16 follows from Theorem 3.1 and [15, Theorem 1.1]. We note that the conditions of [49, Lemma 2.3], [15, Lemma 2.3] and [15, Theorem 1.1] are certain estimates on the growth rates (in k) of the cumulants $\Gamma_k(S_n)$, and the role of Theorem 3.1 is that it shows that the conditions of all of these results are in force in the setup of this paper.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

3.3. *A functional central limit theorem via the method of cumulants: proof of Theorem 2.19.* Let us first show that the sequence $\mathcal{S}_n$ is tight. By Theorem 2.17 we have that

$$\|S_n\|_4 = O(\sqrt{n})$$

where $\|\cdot\|_4 = \|\cdot\|_{L^4}$, and therefore, using also stationarity and the Hölder inequality, we get that for all $t_1 < t_2 \le r_1 < r_2$,

$$\begin{aligned}\mathbb{E}[(\mathcal{S}_n(r_2) - \mathcal{S}_n(r_1))^2(\mathcal{S}_n(t_2) - \mathcal{S}_n(t_1))^2] &\le \|\mathcal{S}_n(r_2) - \mathcal{S}_n(r_1)\|_4^2\|\mathcal{S}_n(t_2) - \mathcal{S}_n(t_1)\|_4^2 \\ &\le C\bigg(\frac{[r_2 n] - [t_1 n]}{n}\bigg)^2.\end{aligned}$$

Thus, by [10, Ch. 15], $\mathcal{S}_n(\cdot)$ is a tight sequence in the Skorokhod space $D[0, 1]$.

Now let us show that the finite-dimensional distributions converge. Let us fix some $t_1 < t_2 < \cdots < t_d$. Set $X_k = \varphi \circ \tau^k$. Next, let us recall the following general fact. Given a vector-valued sequence of random variables $Y_n = (Y_{1,n}, \ldots, Y_{d,n})$, by the multidimensional version of Levi's theorem, in order to show that Y_n converges in distribution as $n \to \infty$ towards a given random variable $\mathcal{Z}$, it is enough to show that for every $a \in \mathbb{R}^d$ we have

$$\lim_{n\to\infty} \mathbb{E}[e^{i(a\cdot Y_n)}] = \mathbb{E}[e^{i(a\cdot\mathcal{Z})}].$$

Therefore, it is enough to show that any linear combination of $Y_{j,n}$, $j = 1, 2, \ldots, d$, converges in distribution towards the corresponding linear combination of the coordinates of $\mathcal{Z}$. Returning to our problem, to obtain the appropriate convergence of the distribution of $(\mathcal{S}_n(it_j))_{j=1}^d$ it is enough to show that any linear combination of $\mathcal{S}_n(t_j)$ converges towards a centered normal random variable with an appropriate variance. More precisely, let $a_1, \ldots, a_d \in \mathbb{R}$. Then we need to show that $\sum_{j=1}^d a_j\mathcal{S}_n(t_j)$ converges in distribution towards a centered normal random variable with variance

$$s^2\sum_{j=1}^d (a_j + \cdots + a_d)^2(t_j - t_{j-1}),$$

where $t_0 = 0$ and $s^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[S_n^2]$. We first notice that

$$\sum_{j=1}^d a_j\mathcal{S}_n(t_j) = n^{-1/2}\sum_{j=1}^d (a_j + \cdots + a_d)(S_{[nt_j]} - S_{[nt_{j-1}]}),$$

where we set $t_0 = 0$ and $S_0 = 0$. Thus, using stationarity, we have

$$\begin{aligned}\mathbb{E}\bigg[\bigg(\sum_{j=1}^d a_j\mathcal{S}_n(t_j)\bigg)^2\bigg] &= n^{-1}\sum_{j=1}^d (a_j + \cdots + a_d)^2\mathbb{E}[S^2_{[nt_j]-[nt_{j-1}]}] \\ &\quad + 2n^{-1}\sum_{1\le j_1<j_2\le d}(a_{j_1} + \cdots + a_{j_d})(a_{j_2} + \cdots + a_{j_d}) \\ &\quad \times \mathbb{E}[(S_{[nt_{j_2}]} - S_{[nt_{j_2-1}]})(S_{[nt_{j_1}]} - S_{[nt_{j_1-1}]})].\end{aligned}$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Now the first summand on the right-hand side above converges to

$$s^2 \sum_{j=1}^{d} (a_j + \cdots + a_d)^2 (t_j - t_{j-1}),$$

while the second summand (the double sum) converges to 0 because $|\mathbb{E}[\varphi \cdot \varphi \circ \tau^n]|$ converges to 0 stretched exponentially fast. Therefore, the asymptotic variance of $\sum_{j=1}^{d} a_j \mathcal{S}_n(t_j)$ has the desired form. Now, let us consider the following array of random variables. Set

$$Y_k = Y_k^{(n,a_1,\ldots,a_d,t_1,\ldots,t_d)} = (a_1 + \cdots + a_j)\varphi \circ \tau^k \ \text{ if } \ [nt_{j-1}] \le k < [nt_j].$$

Then

$$\sum_{j=1}^{d} (a_j + \cdots + a_d)(S_{[nt_j]} - S_{[nt_{j-1}]}) = \sum_{j=1}^{d} (a_j + \cdots + a_d) \sum_{s=[nt_{j-1}]}^{[nt_j]-1} \varphi \circ \tau^s$$

$$= \sum_{j=1}^{d} (a_j + \cdots + a_d) \sum_{s=0}^{[nt_d]-1} \mathbb{I}([nt_{j-1}] \le s < [nt_j])\varphi \circ \tau^s = \sum_{s=0}^{[nt_d]-1} Y_s.$$

On the other hand, arguing as in the proof of Theorem 3.1 (replacing each appearance of $\varphi \circ \tau^k$ by Y_k), we get the same kind of estimates on the cumulants of

$$\tilde{S}_n := \sum_{s=0}^{[nt_d]-1} Y_s,$$

that is, there exists a constant c_0 which might depend on t_j and a_j such that for every k we have

$$|\Gamma_k(\tilde{S}_n)| \le n(k!\,)^{1+\gamma}(c_0)^{k-2}.$$

Thus, by applying [49, Corollary 2.1] we get that

$$\sum_{s=0}^{[nt_d]-1} Y_s^{(n,a_1\ldots a_d)}/w_n$$

converges towards the standard normal distribution, where w_n is the standard deviation of the numerator. Note that, as we have shown, $w_n^2/n \to s^2 \sum_{j=1}^{d} (a_1 + \cdots + a_d)^2 (t_j - t_{j-1})$, which is positive unless either $s = 0$ or $a_1 = \cdots = a_d = 0$, which are both trivial cases. Thus, in any case we obtain the desired convergence of the linear combination $\sum_{j=1}^{d} a_j \mathcal{S}_n(t_j)$ and the proof of Theorem 2.19 is complete.

4. *Limit theorems via martingale approximation for ϕ- and ψ-mixing driving processes*

4.1. *Some expectation estimates using mixing coefficients.*

In the course of the proof of Theorem 2.22 we will need the following two relatively simple lemmas.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

LEMMA 4.1. *Let $\mathcal{G}, \mathcal{H}$ be two sub-σ-algebras of a given σ-algebra on some space measure space. Let g be a real-valued bounded $\mathcal{G}$-measurable function and h be an $\mathcal{H}$-measurable real-valued integrable function. Then*

$$|\mathbb{E}[hg] - \mathbb{E}[h]\mathbb{E}[g]| \le \tfrac{1}{2}\|h\|_{L^\infty}\|g\|_{L^1}\phi(\mathcal{G}, \mathcal{H}).$$

Proof. By [11, Ch. 4] we have

$$\|\mathbb{E}[h|\mathcal{G}] - \mathbb{E}[h]\|_{L^\infty} \le \tfrac{1}{2}\|h\|_{L^\infty}\phi(\mathcal{G}, \mathcal{H}),$$

which clearly implies the lemma. $\square$

LEMMA 4.2. *Let $\mathcal{G}, \mathcal{H}$ be two sub-σ-algebras of a given σ-algebra on some measure space. Let g a real-valued bounded $\mathcal{G}$-measurable function and h be an $\mathcal{H}$-measurable real-valued integrable function. Suppose also that $\psi = \psi(\mathcal{G}, \mathcal{H}) < 1$. Then*

$$|\mathbb{E}[hg] - \mathbb{E}[h]\mathbb{E}[g]| \le 4\|hg\|_{L^1}C_\psi\psi,$$

where $C_\psi = (1-\psi)^{-1}$.

Proof. By [11, Ch. 4] we have

$$\|\mathbb{E}[h|\mathcal{G}] - \mathbb{E}[h]\|_{L^\infty} \le \|h\|_{L^1}\psi(\mathcal{G}, \mathcal{H}).$$

Hence

$$|\mathbb{E}[hg] - \mathbb{E}[h]\mathbb{E}[g]| \le \|h\|_{L^1}\|g\|_{L^1}\psi.$$

Taking $h, g \ge 0$, we get that

$$|\mathbb{E}[hg] - \mathbb{E}[h]\mathbb{E}[g]| \le \mathbb{E}[h]\mathbb{E}[g]\psi.$$

Thus,

$$\mathbb{E}[h]\mathbb{E}[g] \le (1-\psi)^{-1}\mathbb{E}[hg] = C_\psi\mathbb{E}[hg].$$

Therefore, for non-negative functions we have

$$|\mathbb{E}[hg] - \mathbb{E}[h]\mathbb{E}[g]| \le C_\psi\psi\mathbb{E}[hg].$$

Now the general result follows by writing $h = h^+ - h^-$ and $g = g^+ - g^-$, where $h^\pm$ and $g^\pm$ are non-negative functions such that $h^+ + h^- = |h|$ and $g^+ + g^- = |g|$, and using that both $(g, h) \to \mathbb{E}[g]\mathbb{E}[h]$ and $(g, h) \to \mathbb{E}[hg]$ are bilinear in (g, h). $\square$

4.2. *Convergence of the iterates of the transfer operator with respect to a sub-σ-algebra.* Let $\mathcal{F}_0$ be the σ-algebra generated by the map $\pi(\omega, x) = ((\omega_j)_{j\ge 0}, x)$, namely, the one generated by $\mathcal{B}$ and the coordinates with non-negative indexes in the ω direction. Then $(\tau^{-k}\mathcal{F}_0)_{k\ge 0}$ is a decreasing sequence of σ-algebras and $\tau^{-k}\mathcal{F}_0$ is generated by τ^k and the coordinates ω_j for $j \ge k$. In particular, τ preserves $\mathcal{F}_0$.

Next, let us define a transfer operator with respect to $\mathcal{F}_0$. For each function $g \in L^1(\mu)$ there is a unique $\mathcal{F}_0$-measurable function G such that

$$\mathbb{E}[g|\tau^{-1}\mathcal{F}_0] = G \circ \tau.$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Let us define $\mathcal{K}g = G$, where we formally set G to be 0 outside the image of τ (if τ is not onto). Then

$$\mathbb{E}[g|\tau^{-1}\mathcal{F}_0] = \mathcal{K}g \circ \tau.$$

Notice that for $g \in L^1(\Omega \times X, \mathcal{F}_0, \mu)$, $f \in L^\infty(\Omega \times X, \mathcal{F}_0, \mu)$ we have

$$\int (\mathcal{K}g) f \, d\mu = \int (\mathcal{K}g \circ \tau) f \circ \tau \, d\mu = \int \mathbb{E}[g|\tau^{-1}\mathcal{F}_0] \cdot f \circ \tau \, d\mu = \int g \cdot f \circ \tau \, d\mu,$$

and therefore $\mathcal{K}$ can also be defined using the usual duality relation with respect to the above σ-algebra. That is, it is the transfer operator of τ with respect to $(\Omega \times X, \mathcal{F}_0, \mu)$.

The proof of Theorems 2.21 and 2.22 is based on the following result.

PROPOSITION 4.3. *Under the assumptions of Theorems 2.21 and 2.22, and when $\mu(\varphi) = 0$, we have the following assertions.*

(i) *We have*

$$\|\mathcal{K}^n \varphi\|_{L^\infty} \le C(e^{-(\lambda - 2\varepsilon)n/2} + \psi_{[n/2]}) := C\gamma_{2,n}. \tag{4.1}$$

Moreover, if $h_\omega \ge c^{-1} > 0$ for some constant $c > 1$ then

$$\|\mathcal{K}^n \varphi\|_{L^\infty} \le Cc(e^{-(\lambda - 2\varepsilon)n/2} + \phi_{[n/2],R}) := C\gamma_{1,n}. \tag{4.2}$$

Here $C = C_\varphi$ is a constant having the form $C_\varphi = AC_{\mathrm{v}}\, \mathrm{esssup}_{\omega \in \Omega}(K(\omega)^2 \|\varphi_\omega\|_{BV})$, where A is an absolute constant and C_0 is any constant satisfying $\|g\|_{L^\infty} \le C_0 \|g\|_{BV}$ and $\|fg\|_{BV} \le C_0 \|g\|_{BV} \|f\|_{BV}$ for all functions $g, f : X \to \mathbb{C}$.

(ii) *We have*

$$\|\mathcal{K}^i(\varphi \mathcal{K}^j \varphi) - \mu(\mathcal{K}^i(\varphi \mathcal{K}^j \varphi))\|_{L^\infty} \le C\gamma_{2,\max(i,j)}.$$

If $h_\omega \ge c^{-1} > 0$ for some constant $c > 1$ then

$$\|\mathcal{K}^i(\varphi \mathcal{K}^j \varphi) - \mu(\mathcal{K}^i(\varphi \mathcal{K}^j \varphi))\|_{L^\infty} \le Cc\gamma_{1,\max(i,j)}.$$

Proof of Theorems 2.21 and 2.22 based on Proposition 4.3. First, Theorem 2.21(i) follows since if we set $\chi = \sum_{n=1}^\infty K^n \varphi$ and $u = \varphi + \chi \circ \tau - \chi$, then $\|\chi\|_{L^\infty} < \infty$ and $(u \circ \tau^n)$ is a reverse martingale difference with respect to the reverse filtration $\{\tau^{-n}\mathcal{F}_0\}$. Moreover, the differences $u \circ \tau^n$ are uniformly bounded (as χ and φ are in L^∞). Thus, by the Azuma–Hoeffding inequality, for every $\beta > 0$, we have

$$\mathbb{E}_\mu[e^{\lambda \sum_{j=0}^{n-1} u \circ \tau^j}] \le e^{\beta^2 n \|u\|_{L^\infty}^2}.$$

Now the proof proceeds by using the Chernoff bounding method. By the Markov inequality for all $t > 0$ we have

$$\mu\left\{\sum_{j=0}^{n-1} u \circ \tau^j \ge tn\right\} \le e^{-\beta t n} e^{\beta^2 n \|u\|_{L^\infty}^2}.$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Taking $\beta = \beta_t = t/2\|u\|_{L^\infty}$ and replacing u with $-u$, we get that

$$\mu\bigg\{\pm\sum_{j=0}^{n-1} u\circ\tau^j \geq tn\bigg\} \leq e^{-nt^2/4\|u\|_{L^\infty}}.$$

The proof of Theorem 2.21(i) is completed now by noticing that

$$\bigg\|S_n\varphi - \sum_{j=0}^{n-1} u\circ\tau^j\bigg\|_{L^\infty} = \|\chi - \chi\circ\tau^n\|_{L^\infty} \leq 2\|\chi\|_{L^\infty}. \tag{4.3}$$

Next, the proof of Theorem 2.21(ii) is completed by applying [**45**, Proposition 7] with the reverse martingale $(u\circ\tau^n)$ and using (4.3).

In order to prove Theorem 2.22, we apply [**13**, Theorem 3.2] with the bounded function φ and the probability-preserving system $(\Omega\times X, \mathcal{F}_0, \mu, \tau)$, whose transfer operator is $\mathcal{K}$. Now, since we have assumed that $\mu(\varphi) = 0$, in order for the conditions of [**13**, Theorem 3.2] to be in force we need the estimates

$$\sum_{n\geq 2} n^{5/2}(\log n)^3\|\mathcal{K}\varphi\|^4_{L^4(\mu)} < \infty,$$
$$\sum_{n\geq 2} n(\log n)^3\|\mathcal{K}\varphi\|^2_{L^2(\mu)} < \infty,$$
$$\sum_{n\geq 2}\frac{(\log n)^3}{n^2}\bigg(\sum_{i=1}^{n}\sum_{j=0}^{n-i}\|\mathcal{K}^i(\varphi\mathcal{K}^j(\varphi)) - \mu(\varphi\mathcal{K}^j(\varphi))\|_{L^2(\mu)}\bigg)^2 < \infty$$

to hold. These three conditions are verified by Proposition 4.3 and the mixing rates specified in the formulation of Theorem 2.22, and the proof of Theorem 2.22 is complete. □

Proof of Proposition 4.3. (i) Since $L^\infty(\mu)$ is the dual of $L^1(\mu)$, and φ and $\mathcal{K}^n\varphi$ are $\mathcal{F}_0$-measurable, it is enough to show that, for every $g\in L^1(\Omega\times X, \mathcal{F}_0, \mu)$ such that $\|g\|_{L^1}\leq 1$, we have

$$\bigg|\int g\cdot(\mathcal{K}^n\varphi)\,d\mu\bigg| \leq \gamma_n\|g\|_{L^1(\mu)},$$

where γ_n is one of the desired upper bounds. To achieve that let us first note that $\mathcal{K}^n$ is the dual of the restriction of the Koopman operator $f\to f\circ\tau^n$ acting on $\mathcal{F}_0$-measurable functions. Thus,

$$\int g\cdot(\mathcal{K}^n\varphi)\,d\mu = \int\varphi\cdot(g\circ\tau^n)\,d\mu = \int\bigg(\int\varphi_\omega\cdot(g_{\sigma^n\omega}\circ T^n_\omega)\,d\mu_\omega\bigg)\,d\mathbb{P}(\omega) \tag{4.4}$$
$$= \int\bigg(\int(L^n_\omega\varphi_\omega)\cdot g_{\sigma^n\omega}\,d\mu_{\sigma^n\omega}\bigg)\,d\mathbb{P}(\omega).$$

Now, using (2.14) and that $\|\varphi\|_K = \operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV}) < \infty$, we get that

$$\|L^n_\omega\varphi_\omega - \mu_\omega(\varphi_\omega)\|_{L^\infty} \leq C_0\|\varphi\|_K e^{-\lambda n}.$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Hence, using also the σ-invariance of $\mathbb{P}$,

$$\int g \cdot (\mathcal{K}^n \varphi)\, d\mu = \int \mu_\omega(\varphi_\omega)\mu_{\sigma^n\omega}(g_{\sigma^n\omega})\, d\mathbb{P}(\omega) + I,$$

where $|I| \le Ce^{-\lambda n}\|g\|_{L^1(\mu)}$. Next, let us write

$$\mu_{\sigma^n\omega}(g_{\sigma^n\omega}) = m(g_{\sigma^n\omega}h_{\sigma^n\omega}).$$

By (2.10) we have

$$\|h_{\sigma^n\omega} - \mathcal{L}^{n-[n/2]}_{\sigma^{[n/2]}\omega}\mathbf{1}\|_{L^\infty} \le C_0 K(\sigma^{[n/2]}\omega)e^{-\lambda n/2} \le C_0 K(\omega)e^{-(\lambda-\varepsilon)n/2}.$$

Observe next that since $\|1/h_\omega\|_{BV} \le K(\omega)$ we have

$$m(|g|) = \mu_{\sigma^n\omega}(|g|/h_{\sigma^n\omega}) \le C_0 K(\sigma^n\omega)\mu_\omega(|g|)$$

for every function g, and recall that $K(\sigma^n\omega) \le K(\omega)e^{\varepsilon n}$. Combining this with the previous estimates, we get that

$$\begin{aligned} |m(g_{\sigma^n\omega}h_{\sigma^n\omega}) - m(g_{\sigma^n\omega}\mathcal{L}^{n-[n/2]}_{\sigma^{[n/2]}\omega}\mathbf{1})|C_0 &\le K(\omega)e^{-(\lambda-\varepsilon)n/2}m(|g_{\sigma^n\omega}|) \\ &\le CK(\omega)^2\mu_{\sigma^n\omega}(|g_{\sigma^n\omega}|)e^{-(\lambda-3\varepsilon)n/2}. \end{aligned} \tag{4.5}$$

Therefore,

$$\int g \cdot (\mathcal{K}^n\varphi)\, d\mu = \int \mu_\omega(\varphi_\omega)m(g_{\sigma^n\omega}\mathcal{L}^{n-[n/2]}_{\sigma^{[n/2]}\omega}\mathbf{1})\, d\mathbb{P}(\omega) + I + J, \tag{4.6}$$

where $|I| \le Ce^{-\lambda n}\|g\|_{L^1(\mu)}$ and $|J| \le C'e^{-(\lambda-3\varepsilon)n/2}\|g\|_{L^1(\mu)}$ and we have used that $K(\omega)^2\|\varphi_\omega\|_{BV}$ is bounded.

Next, using (2.10) and that $K(\omega)$ is tempered, we have $h_\omega = \lim_{n\to\infty}\mathcal{L}^n_{\sigma^{-n}\omega}\mathbf{1}$, and therefore h_ω depends only on the coordinates ω_j for $j \le 0$. Thus,

$$\mu_\omega(\varphi_\omega) = F(\omega_j;\, j \le 0)$$

for some measurable function F so that $|F| \le \|\varphi\|_{L^1(\mu)}$. Observe also that the random variable

$$G_n(\omega) = m(g_{\sigma^n\omega}\mathcal{L}^{n-[n/2]}_{\sigma^{[n/2]}\omega}\mathbf{1})$$

depends only on ω_j, $j \ge [n/2]$, since $g_\omega(x)$ is a function of x and ω_j, $j \ge 0$ (that is, it factors through π_0). In the case where $h_\omega \ge c^{-1} > 0$ for some constant $c > 0$ we have

$$|G_n(\omega)| = |\mu_{\sigma^n\omega}(g_{\sigma^n\omega}L^{n-[n/2]}_{\sigma^{[n/2]}\omega}(1/h_{\sigma^{[n/2]}\omega}))| \le c\mu_{\sigma^n\omega}(|g_{\sigma^n\omega}|).$$

Thus, using also Lemma 4.1, we see that there is a constant $C > 0$ so that

$$\begin{aligned} &\left|\int \mu_\omega(\varphi_\omega)m(g_{\sigma^n\omega}\mathcal{L}^{n-[n/2]}_{\sigma^{[n/2]}\omega}\mathbf{1})\, d\mathbb{P}(\omega)\right| \\ &\quad \le C\phi_{[n/2],R}\int |G_n(\omega)|\, d\mathbb{P}(\omega) \le cC\phi_{[n/2],R}\|g\|_{L^1(\mu)}, \end{aligned}$$

where we have taken into account that $\int \mu_\omega(\varphi_\omega)\, d\mathbb{P}(\omega) = \mu(\varphi) = 0$. This, together with (4.6) and the previous estimates on I and J, proves (4.2).

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

To prove (4.1), we first use (4.5) in order to obtain that

$$|G_n(\omega)| \leq C\mu_{\sigma^n\omega}(|g_{\sigma^n\omega}|)(1 + CK^2(\omega)e^{-(\lambda-3\varepsilon)n/2}) \leq C'\mu_{\sigma^n\omega}(|g_{\sigma^n\omega}|)K(\omega)^2. \quad (4.7)$$

Taking into account that

$$\mathrm{esssup}_{\omega\in\Omega}(\|\varphi_\omega\|_{L^\infty}K(\omega)^2) \leq C\mathrm{esssup}_{\omega\in\Omega}(\|\varphi_\omega\|_{BV}K(\omega)^2) < \infty,$$

we conclude that $G_n(\omega)\mu_\omega(\varphi_\omega)$ is integrable. We would now like to apply Lemma 4.2, but the problem is that G_n is not bounded. To overcome that, for each $M > 0$ set $G_n^{(M)}(\omega) = G_n(\omega)\mathbb{I}(|G_n(\omega)| \leq M)$. Then, since $G_n(\omega)\mu_\omega(\varphi_\omega)$ is integrable, by the dominated convergence theorem we have

$$\int \mu_\omega(\varphi_\omega)G_n(\omega)\, d\mathbb{P}(\omega) = \lim_{M\to\infty}\int \mu_\omega(\varphi_\omega)G_n^{(M)}(\omega)\, d\mathbb{P}(\omega).$$

Now, taking n so that $\psi_{[n/2]} \leq 1/2$ and using that $\mu(\varphi) = 0$, we get from Lemma 4.2 that

$$\begin{aligned}\bigg|\int \mu_\omega(\varphi_\omega)G_n^{(M)}(\omega)\, d\mathbb{P}(\omega)\bigg| &\leq 2\bigg(\int |G_n^{(M)}(\omega)\mu_\omega(\varphi_\omega)|\, d\mathbb{P}(\omega)\bigg)\psi_{[n/2]}\\ &\leq 2\bigg(\int |G_n(\omega)\mu_\omega(\varphi_\omega)|\, d\mathbb{P}(\omega)\bigg)\psi_{[n/2]}.\end{aligned}$$

Using also (4.7) and that $\mathrm{esssup}_{\omega\in\Omega}(\|\varphi_\omega\|_{BV}K(\omega)^2) < \infty$, we conclude that

$$\bigg|\int \mu_\omega(\varphi_\omega)G_n(\omega)\, d\mathbb{P}(\omega)\bigg| \leq 2(\mathrm{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi_\omega\|_{BV}))C'\|g\|_{L^1}\psi_{[n/2]}$$

and (4.1) follows (using also (4.6)).

(ii) First, since $\mathcal{K}$ weakly contracts the L^∞ norm (being defined through conditional expectations) and φ is bounded we have

$$\|\mathcal{K}^i(\varphi\mathcal{K}^j\varphi) - \mu(\mathcal{K}^i(\varphi\mathcal{K}^j\varphi))\|_{L^\infty} \leq 2\|\varphi\|_{L^\infty}\|\mathcal{K}^j\varphi\|_{L^\infty}.$$

This, together with Proposition 4.3(i), provides the desired estimate when $j \geq i$. The estimate in the case $i > j$ is found similarly to the proof of (i). Let $g \in L^1(\Omega \times X, \mu, \mathcal{F}_0)$. Let us first show that

$$\int \mathcal{K}^i(\varphi\mathcal{K}^j\varphi)g\, d\mu = \int \mu_\omega(\varphi_\omega \cdot (\varphi_{\sigma^j\omega} \circ T_\omega^j))\mu_{\sigma^{i+j}\omega}(g_{\sigma^{i+j}\omega})\, d\mathbb{P}(\omega) + I \quad (4.8)$$

where $|I| \leq C_2e^{-\lambda i}$, in which C_2 is some constant.

In order to prove (4.8), using that $\mathcal{K}$ satisfies the duality relation and the disintegration $\mu = \int \mu_\omega\, d\mathbb{P}(\omega)$, we first have that

$$\begin{aligned}\int \mathcal{K}^i(\varphi\mathcal{K}^j\varphi)g\, d\mu &= \int (\varphi\mathcal{K}^j\varphi)\cdot g\circ\tau^i\, d\mu = \int \mathcal{K}^j\varphi\cdot(\varphi\cdot(g\circ\tau^i))\, d\mu\\ &= \int (\varphi\cdot(\varphi\circ\tau^j))\cdot g\circ\tau^{i+j}\, d\mu\\ &= \int\bigg(\int \varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j)\cdot(g_{\sigma^{i+j}\omega}\circ T_\omega^{i+j})\, d\mu_\omega\bigg)\, d\mathbb{P}(\omega) \quad (4.9)\\ &= \int\bigg(\int L_\omega^{i+j}(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j))g_{\sigma^{i+j}\omega}\, d\mu_{\sigma^{i+j}\omega}\bigg)\, d\mathbb{P}(\omega).\end{aligned}$$

Next, since $L_\omega^n(f\circ T_\omega^n)=f$ for every function f and n, we have

$$L_\omega^{i+j}(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j))=L^i_{\sigma^j\omega}(\varphi_{\sigma^j\omega}L_\omega^j\varphi_\omega).$$

By (2.14) we have

$$\|L_\omega^j\varphi_\omega-\mu_\omega(\varphi_\omega)\|_{BV}\leq K(\omega)\|\varphi_\omega\|_{BV}e^{-\lambda j}.$$

In particular,

$$\|L_\omega^j\varphi_\omega\|_{BV}\leq CK(\omega)\|\varphi_\omega\|_{BV}$$

for some constant C. Since $\|uv\|_{BV}\leq C_0\|u\|_{BV}\|v\|_{BV}$ for every pair of functions u, v, we have

$$\|\varphi_{\sigma^j\omega}L_\omega^j\varphi_\omega\|_{BV}\leq C_0CK(\omega)\|\varphi_\omega\|_{BV}\|\varphi_{\sigma^j\omega}\|_{BV}.$$

Thus by (2.14),

$$\begin{aligned}&\|L^i_{\sigma^j\omega}(\varphi_{\sigma^j\omega}L_\omega^j\varphi_\omega)-\mu_{\sigma^j\omega}(\varphi_{\sigma^j\omega}L_\omega^j\varphi_\omega)\|_{BV}\\&\quad\leq C_0CK(\omega)K(\sigma^j\omega)\|\varphi_\omega\|_{BV}\|\varphi_{\sigma^j\omega}\|_{BV}e^{-\lambda i}\leq C_0C\|\varphi\|_K^2e^{-\lambda i},\end{aligned}$$

where $\|\varphi\|_K=\text{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV})$. Observe next that

$$\mu_{\sigma^j\omega}(\varphi_{\sigma^j\omega}L_\omega^j\varphi_\omega)=\mu_\omega(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j)).$$

The desired inequality (4.8) follows from the above estimates.

Observe that the function $\mu_\omega(\varphi_\omega\cdot\varphi_{\sigma^j\omega}\circ T_\omega^j)$ depends only on ω_k for $k\leq j$ and that it is bounded by $CK^{-2}(\omega)$ for some constant $C>0$ (since $\text{esssup}_{\omega\in\Omega}(K(\omega)^2\|\varphi_\omega\|_{BV})<\infty$). Therefore, the same arguments in the proof of (i) yield that

$$\begin{aligned}&\int\mu_\omega(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j))\mu_{\sigma^{i+j}\omega}(g_{\sigma^{i+j}\omega})\,d\mathbb{P}(\omega)\\&\quad=\int\mu_\omega(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j))\,d\mathbb{P}(\omega)\cdot\int\mu_{\sigma^{i+j}\omega}(g_{\sigma^{i+j}\omega})\,d\mathbb{P}(\omega)+J\end{aligned}$$

where $|J|\leq\gamma_i\|g\|_{L^1}$ and γ_i is one of the right-hand sides on the upper bounds in (i) (depending on the case) with n replaced by i. Notice next that

$$\int\mu_\omega(\varphi_\omega\cdot(\varphi_{\sigma^j\omega}\circ T_\omega^j))\,d\mathbb{P}(\omega)=\int\mathcal{K}^i(\varphi\mathcal{K}^j\varphi)\,d\mu$$

(this can be seen by taking $g=1$ in (4.9)). Hence,

$$\left|\int(\mathcal{K}^i(\varphi\mathcal{K}^j\varphi)-\mu(\mathcal{K}^i(\varphi\mathcal{K}^j\varphi))g\,d\mu\right|\leq C(e^{-\lambda i}+\gamma_i)\|g\|_{L^1}$$

and the desired estimate follows again since L^∞ is the dual of L^1. $\square$

5. *A vector-valued almost sure invariance principle for skew products with uniformly expanding fiber maps and exponentially fast α-mixing base maps*

Let us first explain why the matrix Σ^2 exists. For a fixed vector v the limit $s_v^2=\lim_{n\to\infty}(1/n)\mathbb{E}[(S_n\cdot v)^2]$ exists, by considering the real-valued observable $\varphi\cdot v$. Then the matrix Σ^2 from Theorem 2.25 is given by $(\Sigma^2)_{i,j}=\frac{1}{2}(s^2_{e_i+e_j}-s^2_{e_i}-s^2_{e_j})$. This

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

matrix satisfies $\Sigma^2 v \cdot v = s_v^2$ and so it is not positive definite if and only if $\varphi \cdot v$ is a coboundary for some unit vector v. Note that this part does not require T_ω to be uniformly expanding.

We assume next that there exist constants $C > 0$ and $\delta \in (0, 1)$ such that, for $\mathbb{P}$-a.e. ω, we have

$$\|\mathcal{L}_\omega^n \mathbf{1} - h_{\sigma^n \omega}\|_{BV} \leq C\delta^n \tag{5.1}$$

(this is the uniform expansion assumption).

The proof of Theorem 2.25 relies on an application of [**24**, Theorem 1.2]. The main condition of [**24**, Theorem 1.2] is the content of the following lemma. Once the lemma is proven Theorem 2.25 follows from [**24**, Theorem 1.2] applied with an arbitrary large p.

LEMMA 5.1. *There exist $\varepsilon_0 > 0$, $c, C > 0$ such that for any $n, m > 0$, $b_1 < b_2 < \cdots < b_{n+m+1}$, $k > 0$ and $t_1, \ldots, t_{n+m} \in \mathbb{R}^d$ with $|t_j| \leq \varepsilon_0$ we have*

$$\begin{aligned}\Big|&\mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j \cdot (\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell) + i\sum_{j=n+1}^{n+m} t_j \cdot (\sum_{\ell=b_j+k}^{b_{j+1}+k-1} B_\ell)}\big) \\ &\quad - \mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j \cdot (\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell)}\big) \cdot \mathbb{E}_\mu\big(e^{i\sum_{j=n+1}^{n+m} t_j \cdot (\sum_{\ell=b_j+k}^{b_{j+1}+k-1} B_\ell)}\big)\Big| \\ &\leq C^{n+m} e^{-ck},\end{aligned} \tag{5.2}$$

where $B_\ell = \varphi \circ \tau^\ell$.

Proof. First, denoting by $\mathbb{E}_\omega$ the expectation with respect to μ_ω, by [**22**, Lemma 24] there exist $\varepsilon_0 > 0$, $c, C > 0$ with the property that for every $n, m > 0$, $b_1 < b_2 < \cdots < b_{n+m+1}$, $k > 0$ and $t_1, \ldots, t_{n+m} \in \mathbb{R}^d$ such that $|t_j| \leq \varepsilon_0$,

$$\begin{aligned}\Big|&\mathbb{E}_\omega\big(e^{i\sum_{j=1}^n t_j \cdot (\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell) + i\sum_{j=n+1}^{n+m} t_j \cdot (\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell)}\big) \\ &\quad - \mathbb{E}_\omega\big(e^{i\sum_{j=1}^n t_j \cdot (\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell)}\big) \cdot \mathbb{E}_\omega\big(e^{i\sum_{j=n+1}^{n+m} t_j \cdot (\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell)}\big)\Big| \\ &\leq C^{n+m} e^{-ck},\end{aligned} \tag{5.3}$$

where $\mathbb{E}_\omega(g) = \int g h_\omega \, dm$ and

$$A_\ell := \varphi_{\sigma^\ell \omega} \circ T_\omega^\ell, \quad \ell \in \mathbb{N}.$$

Let

$$G(\omega) = \mathbb{E}_\omega\big(e^{i\sum_{j=1}^n t_j \cdot (\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell)}\big)$$

and

$$F(\omega) = \mathbb{E}_\omega\big(e^{i\sum_{j=n+1}^{n+m} t_j \cdot (\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell)}\big).$$

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

Then with $B_\ell = \varphi \circ \tau^\ell$ we have

$$\Big|\mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j\cdot(\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell)+i\sum_{j=n+1}^{n+m} t_j\cdot(\sum_{\ell=b_j+k}^{B_{j+1}+k-1} B_\ell)}\big)$$
$$-\mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j\cdot(\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell)}\big)\cdot\mathbb{E}_\mu\big(e^{i\sum_{j=n+1}^{n+m} t_j\cdot(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} B_\ell)}\big)\Big|$$
$$\le C^{n+m}e^{-ck}+|\mathrm{Cov}_{\mathbb{P}}(G,F)|. \tag{5.4}$$

Using (5.1) and that $(T_\omega)_*\mu_\omega = \mu_{\sigma\omega}$, we get that there are $k_0 \in \mathbb{Z}$ and functions G_1 and F_1 such that

$$\|G(\omega) - G_1(\ldots, \omega_{k_0-1}, \omega_{k_0+[k/4]})\|_{L^\infty} \le C'\delta^{k/4}$$

and

$$\|G(\omega) - G_1(\omega_{k_0+k-[k/4]}, \omega_{k_0+k-[k/4]+1}, \ldots)\|_{L^\infty} \le C'\delta^{k/4}.$$

Thus,

$$|\mathrm{Cov}_{\mathbb{P}}(G,F)| \le |\mathrm{Cov}_{\mathbb{P}}(G_1,F_1)| + C''\delta^{k/4},$$

where we have used that G_1, G_2, G and F are uniformly bounded (so the above constants C', C'' do not depend on the choice of b_j, t_j, etc.). On the other hand, by (3.7),

$$|\mathrm{Cov}_{\mathbb{P}}(G_1,F_1)| \le C'''\alpha^{k/2}.$$

Thus,

$$\Big|\mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j\cdot(\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell)+i\sum_{j=n+1}^{n+m} t_j\cdot(\sum_{\ell=b_j+k}^{B_{j+1}+k-1} B_\ell)}\big)$$
$$-\mathbb{E}_\mu\big(e^{i\sum_{j=1}^n t_j\cdot(\sum_{\ell=b_j}^{b_{j+1}-1} B_\ell)}\big)\cdot\mathbb{E}_\mu\big(e^{i\sum_{j=n+1}^{n+m} t_j\cdot(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} B_\ell)}\big)\Big|$$
$$\le C^{n+m}e^{-ck}+C''\delta^{\delta k/4}+C'''\alpha^{k/2}. \tag{5.5}$$

$\square$

6. *Extensions, generalizations, additional results and a short discussion*

In this section we will describe a few additional results which can also be obtained using the methods of the current paper. In order not to overload the paper the section is presented in a form of a discussion rather than explicit formulations of theorems.

6.1. *More general mixing base maps for continuous in ω transfer operators.*

Let $(\xi_n)_{n\in\mathbb{Z}}$ be a stationary process taking values on a metric space $(\mathcal{Y}, d)$ satisfying the following approximation and mixing conditions.

There are sub-σ-algebras $\mathcal{G}_{n,m}$ on the underlying probability space such that $\mathcal{G}_{n,m} \subset \mathcal{G}_{n_1,m_1}$ if $[n,m] \subset [n_1,m_1]$ and for each r and n there is an $\mathcal{G}_{n-r,n+r}$ measurable random variable $\xi_{n,r}$ so that the following assertions hold.

(1) *Approximation*. $\|d(\xi_n, \xi_{n,r})\|_{L^\infty} \le A_1\beta^r$, $\beta \in (0,1)$.

(2) *Mixing*. The sequences $(\xi_{2nr,r})_{n\in\mathbb{Z}}$ are α-mixing (or ϕ_R- or ψ-mixing) uniformly in r.

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

We note that the above uniform approximation by α-mixing sequences applies to the case where ξ_j has the form $\xi_j = S^j\xi_0$, in which S is an invertible Young tower. In this case we take

$$\mathcal{G}_{n,m} = \bigwedge_{j=n}^{m} S^{-j}\mathcal{A},$$

where $\mathcal{A}$ is the partition that defines the tower. Then $\alpha_n = O(n^{-(p-2)})$ if the tails of the tower are $O(n^{-p})$ for some $p \geq 3$. We can also consider several classes of smooth maps S on the interval or Gibbs–Markov maps [1] for which such an approximation holds with $\psi_n = O(\delta^n)$ for some $\delta \in (0, 1)$.

Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be the shift system constructed as before. Then all the results stated in the paper hold true when $\omega \to \mathcal{L}_\omega$ and $\omega \to \varphi_\omega$ are Hölder continuous in ω (on a set with probability 1). The main point is that Lemma 3.8 and the similar approximations used in the construction of the martingale (that is, in the proof of Proposition 4.3) can be obtained by first approximating (taking $r = r_n = \varepsilon_0 n$ for some small ε_0) and using the mixing conditions on the approximating sequences. The main reason we did not include such results in the body of the paper is that it would make the notation more complicated, and that the additional essentially global regularity assumptions on the transfer operators are somehow less natural.

6.2. *Extension to random Gibbs measures.* Let us now consider the random expanding maps T_ω as in [44]. Let $\mu_\omega = h_\omega \nu_\omega$ be a random Gibbs measure corresponding to a given random logarithmically α-Hölder continuous potential, and let λ_ω be the exponent of the random pressure. Namely, if $\mathcal{L}_\omega$ is the transfer operator corresponding to the random potential, then

$$\mathcal{L}_\omega h_\omega = \lambda_\omega h_{\sigma\omega}, \; (\mathcal{L}_\omega)^* \nu_{\sigma\omega} = \lambda_\omega \nu_\omega.$$

Next, for the sake of simplicity let us consider here random expanding maps as in [31, Ch. 5]. Then there is a constant $K > 0$ such that with $\tilde{\mathcal{L}}_\omega = \mathcal{L}_\omega/\lambda_\omega$ we have

$$\|\tilde{\mathcal{L}}_\omega^n - \nu_\omega \otimes h_{\sigma^n\omega}\|_{\text{Holder}} \leq K e^{-\lambda n},$$

where $\|\cdot\|_{\text{Holder}}$ is the usual Hölder norm corresponding to the exponent α and $\nu \otimes h(g) = \nu(g)h$. Plugging in $g = \mathbf{1}$ we get similar estimates to those we had in (2.10):

$$\|\tilde{\mathcal{L}}_\omega^n - h_{\sigma^n\omega}\|_{\text{Holder}} \leq K e^{-\lambda n}.$$

Remark also that $h_\omega \geq c > 0$ for some constant $c > 0$ (see [31]).

The main additional difficulty here is to estimate expressions of the form $\mu_\omega(F_\omega)$ (as in Lemma 3.8) by functions of the coordinates in places j for $|j| \leq n$. Once this is achieved, we can use the approximation argument (similarly to Lemma 3.8) which was essential in the proofs of all of the results stated in the body of the paper. The main difference in comparison to the case where $\nu_\omega = m$ does not depend on ω is that now we need to approximate ν_ω by functions of the first n coordinates (exponentially fast in n). For uniformly expanding maps, this follows from the construction of ν_ω as a certain uniform limit (see [31, Chs. 4–5]).

https://doi.org/10.1017/etds.2023.23 Published online by Cambridge University Press

6.3. *Extension to non-conventional sums (multiple recurrences).* Let us consider partial 'non-conventional' sums of the form

$$S_n\varphi = \sum_{m=1}^{n}\prod_{j=1}^{\ell}\varphi\circ\tau^{q_j(m)},$$

where ℓ is an integer and $q_j(n)$ are positive integer-valued sequences. The statistical properties of such sums have been studied for several classes of expanding or hyperbolic maps (in particular); see [25, 39, 41] and references therein. When all the q_j are polynomials, we believe that all the results obtained using the method of cumulants (that is, Theorems 2.14–2.17 and an appropriate version of Theorem 2.19) can be obtained for such sums exactly as in [25], relying on a version of Proposition 3.4 applied with $\rho(n,m) = \max_{1\leq i,j\leq \ell}|q_i(m)-q_j(n)|$. The main idea is that by induction on the number of blocks we can show that the conditions of Proposition 3.3 with $\rho=\rho_\ell$ hold true for

$$X_m = \prod_{j=1}^{\ell}\varphi\circ\tau^{q_j(m)}.$$

That is, by an inductive argument similar to that in [31, Corollary 1.3.11], we can prove the following result.

LEMMA 6.1. *Let $r\in\mathbb{N}$ and let $B_1, B_2, \ldots, B_k$ be finite subsets of $\mathbb{N}$ so that the distance between B_j and B_{j+1} is d_j. Set $r_j=[d_j/3]$. Let $\mathcal{C}=\{\mathcal{C}_j : 1\leq j\leq s\}$ be a partition of $\{1,2,\ldots,k\}$ and set $Y_j=\prod_{k\in\mathcal{C}_j}\prod_{u\in B_k}\varphi\circ\tau^u$. Then, assuming that $\|\varphi\|_{L^\infty}\leq 1$ and that $\operatorname{esssup}_{\omega\in\Omega}(K(\omega)\|\varphi_\omega\|_{BV})\leq 1$, there is an absolute constant $A>1$ such that*

$$\bigg|\mathbb{E}_\mu\bigg[\prod_{j=1}^{s}Y_j\bigg]-\prod_{j=1}^{s}\mathbb{E}_\mu[Y_j]\bigg|\leq A^m\sum_{j=1}^{m}(\delta^{r_j}+\alpha([r_j]))$$

where $\delta=e^{-(\lambda-3\varepsilon)/2}\in(0,1)$.

We note that in order to prove a version of the functional CLT for the sums above we first need to use the arguments in [30, 41] to compute the variance of the limiting Gaussian, which for general polynomials might differ from a Brownian motion, and this can also be done by using the above lemma.

REFERENCES

[1] J. Aaronson and M. Denker. Local limit theorems for Gibbs–Markov maps. *Stoch. Dyn.* **1** (2001), 193–237.
[2] R. Aimino, M. Nicol and S. Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields* **162** (2015), 233–274.
[3] J. F. Alves, W. Bahsoun and R. Ruziboev. Almost sure rates of mixing for partially hyperbolic attractors. *J. Differential Equations* **311** (2022), 98–157.
[4] L. Arnold. *Random Dynamical Systems (Springer Monographs in Mathematics)*. Springer, Berlin, 1998.
[5] J. Atnip. An almost sure invariance principle for several classes of random dynamical systems. *Preprint*, 2018, arXiv:1702.07691.
[6] J. Atnip, G. Froyland, C. González-Tokman and S. Vaienti. Thermodynamic formalism for random weighted covering systems. *Comm. Math. Phys.* **386** (2021), 819–902.
[7] A. Ayyer, C. Liverani and M. Stenlund. Quenched CLT for random toral automorphism. *Discrete Contin. Dyn. Syst.* **24** (2009), 331–348.

[8] V. Baladi. Correlation spectrum of quenched and annealed equilibrium states for random expanding maps. *Comm. Math. Phys.* **186** (1997), 671–700.
[9] V. Baladi and L.-S. Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.* **156** (1993), 355–385.
[10] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.
[11] R. C. Bradley. *Introduction to Strong Mixing Conditions*, Vol. 1. Kendrick Press, Heber City, UT, 2007.
[12] J. Buzzi. Exponential decay of correlations for random Lasota–Yorke maps. *Comm. Math. Phys.* **208** (1999), 25–54.
[13] C. Cuny and F. Merlevede. Strong invariance principles with rate for 'reverse' martingale differences and applications. *J. Theoret. Probab.* **28** (2015), 137–183.
[14] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications (Applications of Mathematics, 38)*, 2nd edn. Springer, New York, 1998.
[15] H. Döring and P. Eichelsbacher. Moderate deviations via cumulants. *J. Theoret. Probab.* **26** (2013), 360–385.
[16] D. Dragičević, G. Froyland, C. González-Tokman and S. Vaienti. Almost sure invariance principle for random piecewise expanding maps. *Nonlinearity* **31** (2018), 2252–2280.
[17] D. Dragičević, G. Froyland, C. González-Tokman and S. Vaienti. A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys.* **360** (2018), 1121–1187.
[18] D. Dragičević, G. Froyland, C. González-Tokman and S. Vaienti. A spectral approach for quenched limit theorems for random hyperbolic dynamical systems. *Trans. Amer. Math. Soc.* **373** (2020), 629–664.
[19] D. Dragičević and Y. Hafouta. Limit theorems for random expanding or Anosov dynamical systems and vector-valued observables. *Ann. Henri Poincaré* **21** (2020), 3869–3917.
[20] D. Dragičević and Y. Hafouta. Almost sure invariance principle for random distance expanding maps with a nonuniform decay of correlations. *Thermodynamic Formalism: CIRM Jean-Morlet Chair*. Eds. M. Pollicott and S. Vaienti. Springer-Verlag, Cham, 2021.
[21] D. Dragičević and Y. Hafouta. Almost sure invariance principle for random dynamical systems via Gouëzel's approach. *Nonlinearity* **34** (2021), 6773.
[22] D. Dragičević, Y. Hafouta and J. Sedro. A vector-valued almost sure invariance principle for random expanding on average cocycles. *J. Stat. Phys.* **190** (2023), 54.
[23] D. Dragičević and J. Sedro. Quenched limit theorems for expanding on average cocycles. *Preprint*, 2021, arXiv:2105.00548.
[24] S. Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.* **38** (2010), 1639–1671.
[25] Y. Hafouta. Nonconventional moderate deviations and exponential concentration inequalities. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**(1) (2020), 428–448.
[26] Y. Hafouta. Limit theorems for some skew products with mixing base maps. *Ergod. Th. & Dynam. Sys.* **41**(1) (2021), 241–271.
[27] Y. Hafouta. Limit theorems for random non-uniformly expanding or hyperbolic maps. *Ann. Henri Poincaré* **23** (2022), 293–332.
[28] Y. Hafouta. Explicit conditions for the CLT and related results for non-uniformly partially expanding random dynamical systems via effective RPF rates. *Preprint*, 2022, arXiv:2208.00518.
[29] Y. Hafouta and Y. Kifer. Berry–Esseen type estimates for nonconventional sums. *Stoch. Process. Appl.* **126** (2016), 2430–2464.
[30] Y. Hafouta and Y. Kifer. Nonconventional polynomial CLT. *Stochastics* **89** (2017), 550–591.
[31] Y. Hafouta and Y. Kifer. *Nonconventional Limit Theorems and Random Dynamics*. World Scientific, Singapore, 2018.
[32] N. Haydn, M. Nicol, A. Török and S. Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Trans. Amer. Math. Soc.* **369** (2017), 5293–5316.
[33] H. Hennion and L. Hervé, *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness (Lecture Notes in Mathematics, 1766)*. Springer, Berlin, 2001.
[34] I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen, 1971.
[35] H. Ishitani. Central limit theorems for the random iterations of 1-dimensional transformations (dynamics of complex systems). *RIMS Kokyuroku, Kyoto Univ.* **1404** (2004), 21–31.
[36] Y. Kifer. Perron–Frobenius theorem, large deviations, and random perturbations in random environments. *Math. Z.* **222**(4) (1996), 677–698.
[37] Y. Kifer. Limit theorems for random transformations and processes in random environments. *Trans. Amer. Math. Soc.* **350** (1998), 1481–1518.
[38] Y. Kifer. Thermodynamic formalism for random transformations revisited. *Stoch. Dyn.* **8** (2008), 77–102.

[39] Y. Kifer. Nonconventional limit theorems. *Probab. Theory Related Fields* **148** (2010), 71–106.
[40] Y. Kifer and P. D. Lui. Random dynamics. *Handbook of Dynamical Systems 1B*. Eds. B. Hasselblatt and A. Katok. Elsevier, Amsterdam, 2006, pp. 379–499.
[41] Y. Kifer and S. R. S. Varadhan. Nonconventional limit theorems in discrete and continuous time via martingales. *Ann. Probab.* **42** (2014), 649–688.
[42] 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.
[43] A. Korepanov and J. Leppanen. Loss of memory and moment bounds for nonstationary intermittent dynamical systems. *Comm. Math. Phys.* **385**(2) (2021), 905–935.
[44] V. Mayer, B. Skorulski and M. Urbański. *Distance Expanding Random Mappings, Thermodynamical Formalism, Gibbs Measures and Fractal Geometry (Lecture Notes in Mathematics, 2036).* Springer, Berlin, 2011.
[45] F. Merlevéde, M. Peligrad and S. Utev. Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3** (2006), 1–36.
[46] M. Nicol, F. P. Pereira and A. Török. Large deviations and central limit theorems for sequential and random systems of intermittent maps. *Ergod. Th. & Dynam. Sys.* **41**(9) (2021), 2805–2832.
[47] W. Philipp and W. F. Stout. *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables (Memoirs of the American Mathematical Society, 161).* American Mathematical Society, Providence, RI, 1975.
[48] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants (Mathématiques & Applications (Berlin) [Mathematics & Applications], 31)*. Springer Verlag, Berlin, 2000.
[49] L. Saulis and V. A. Statulevicius. *Limit Theorems for Large Deviations*. Kluwer Academic, Dordrecht, 1991.
[50] Y. Su. Almost surely invariance principle for non-stationary and random intermittent dynamical systems. *Discrete Contin. Dyn. Syst.* **39**(11) (2019), 6585–6597.
[51] Y. Su. Vector-valued almost sure invariance principles for (non)stationary and random dynamical systems. *Trans. Amer. Math. Soc.* **375**(7) (2022), 4809–4848.