BERRY ESSEEN THEOREMS FOR SEQUENCES OF EXPANDING MAPS

DMITRY DOLGOPYAT AND YEOR HAFOUTA

dimadolgopyat@gmail.com, yeor.hafouta@mail.huji.ac.il

Abstract. We prove Berry-Esseen theorems for sums $S_n = \sum_{j=0}^{n-1} f_j \circ T_{j-1} \circ \cdots \circ T_1 \circ T_0$ where f_j are functions with uniformly bounded "variation" and T_j is a sequence of expanding maps. Using symbolic representations similar result follow for maps T_j in a small C^1 neighborhood of an Axiom A map and Hölder continuous functions f_j. All of our results are already new for a single map $T_j = T$ and a sequence of different functions (f_j).

1. Introduction

1.1. **Non-autonomous dynamical systems.** A great discovery made in the last century is that deterministic systems could exhibit random behavior. One of the most notable results in this direction is the fact that ergodic averages of deterministic systems could satisfy the Central Limit Theorem (CLT). Since then statistical properties of autonomous hyperbolic dynamical systems have been studied extensively. However, many systems appearing in nature are time dependent due to an interaction with the outside world. In the context of dynamical systems this leads to the study of dynamics formed by a composition of different transformations rather than a single one. Such systems are often called sequential/time-dependent/non-autonomous dynamical systems.

Many powerful tools developed for studying autonomous systems are unavailable in non autonomous setting. In particular, the spectral approach developed by Nagaev [71] and extended to dynamical systems setting by Guivarch and Hardy [41], provides a powerful tool for obtaining asymptotics expansions in limit theorems for dynamical systems. It turns out that *complex Perron Frobenius Theorem* proven by Hafouta and Kifer in [46] (building on a previous work of Rugh [77] and Dubois [34]) provides a convenient tool for asymptotic computations of the characteristic functions in the setting of Markov chains and dynamical systems. This theorem has already found multiple applications to limit theorems [27, 33, 35, 46, 47, 48, 49, 51]. The goal of the present paper is to study the rate of convergence in the CLT (aka Berry-Esseen theorems) for non-autonomous dynamical systems without making any assumptions on the growth of variance of the underlying partial sum. In a forthcoming work [28] local limit theorems will be considered.

A particular case of a sequential dynamical systems are random dynamical systems. Ergodic theory of random dynamical systems has attracted a lot of attention in the past decades, see [58, 65, 22, 3, 25] and [62]. This includes, for instance, the theory of random invariant measures, entropy theory, thermodynamic formalism, multiplicative ergodic theory and many other classical topics in ergodic theory. Ergodic aspects of (non-random) sequential dynamical systems were studied for the first time in [12, 13], see also [4], [24, Sections 1-2] and [14]. We refer to [24, Section 3] and [48] for examples of expanding maps and to [8, 4, 74] for some examples of sequential hyperbolic sequences (see also Sections 4, 5 and Appendix C).

Next we discuss statistical properties of random or sequential dynamical systems. Decay of correlations for random dynamical systems were obtained in [19, 10, 7]. Large deviations were obtained in [59] and the CLT in [60]. Recently limit theorems for random dynamical systems have attracted a lot of attention. For hyperbolic maps in a $C^{3+\varepsilon}$ neighborhood of a $C^{3+\varepsilon}$ Anosov map, the CLT was obtained in [32] (see also [6, 23]). Berry-Esseen theorems (see (1.1)) were obtained in [46, 33, 49, 52]. Concerning statistical properties of non-random expanding sequential systems, the sharp CLT was obtained in [24] (see also [2, 72]). For hyperbolic systems the CLT is proven in [8] under the assumption that variance of S_n grows faster than $n^{2/3}$. Berry-Esseen theorems in the sequential setup were only obtained in [48] for some classes of maps, under the assumption that $\mathrm{Var}(S_n)$ grows linearly fast in n.

1.2. **Our results.** In this paper we obtain optimal CLT rates. A classical result due to Berry and Esseen [18, 37, 38] asserts that for partial sums $S_n = \sum_{j=1}^{n} X_j$ of zero mean of uniformly bounded iid random variables X_j the following optimal uniform CLT rates hold

$$\Delta_{0,n} := \sup_{t\in\mathbb{R}} |\mathbb{P}(S_n/\sigma_n \le t) - \Phi(t)| = O(\sigma_n^{-1}), \quad \sigma_n = \|S_n\|_{L^2} \tag{1.1}$$

where $\Phi(t)$ is the standard normal distribution function. By now the optimal convergence rate in the CLT was obtained for wide classes of stationary Markov chains [71, 55] and other weakly dependent random processes including chaotic dynamical systems [76, 41, 55, 44, 56, 57], random dynamical systems [46, 33, 49] uniformly bounded stationary sufficiently fast ϕ-mixing sequences [75], U-statistics [20, 39] and locally dependent random variables [11, 9, 21]. However, in all of these processes the variance of S_n is of linear order in the number of summands n. To the best of our knowledge, the only case where optimal rate was obtained without any growth rates on the variances is for additive functional of uniformly elliptic inhomogeneous Markov chains [27]. In the present paper we obtain optimal CLT rates (in various forms) for Birkhoff sums generated by a sequence of expanding maps and sufficiently regular functions. This was obtained in [48] for Hölder continuous functions when $\mathrm{Var}(S_n) \ge cn$ for some $c > 0$. Here we consider more general maps and more general functions (e.g. BV). Most importantly we will not assume any kind of growth rates for $\mathrm{Var}(S_n)$.

Our results are applicable to wide classes of expanding transformations including smooth expanding maps, piecewise expanding maps, Markov maps on the interval and sequential subshifts of finite type. The latter also has applications to additive functionals of finite state uniformly elliptic inhomogenuous Markov chains, where the main novelty here is that the functionals are allowed to depend on the entire path of the chain. Our main results also have applications to sequence of maps in a C^1 neighborhood of a given C^2-Axiom A map (see Appendix C). In [32] limit theorems were studied for random dynamical systems in $C^{3+\varepsilon}$ neighborhood of a $C^{3+\varepsilon}$ Anosov maps and Birkhoff sums formed by random $C^{2+\varepsilon}$ functions. Compared with [32] we can perturb more general maps in a weaker norm, consider functions which are only Hölder continuous and treat non-random systems. We would like to emphasize that when starting with sequential SRB measures the closeness of the maps is only needed here to overcome the difficulty that the variance of the underlying partial sum S_n does not have to grow linearly fast. We refer to Remark C.6 for a short discussion on this matter.

What we will actually prove is stronger than uniform rates. In Theorem 3.1(i) we will show that for every $p > 0$,

$$\Delta_{p,n} := \sup_{t\in\mathbb{R}} (1 + |t|^p) \, |\mathbb{P}(S_n/\sigma_n \le t) - \Phi(t)| = O(\sigma_n^{-1}) \tag{1.2}$$

The fact that we can take can positive p allows applications to optimal CLT rates in L^p (see Theorem 3.1(ii)), to Gaussian expectation estimates (Theorem 3.1(iii)) and to optimal CLT rates in the Wasserstein distances (Theorem 3.2).

The proof of (1.2) consists of several ingredients. First, using a general result from [52] it is enough to verify a logarithmic growth assumption introduced in [27]. This is done in Proposition 7.1. To prove this proposition we use ideas from [27] together with an appropriate martingale-coboundary decomposition. This decomposition is needed in order to obtain the moment estimates in Proposition 3.3, which are proven using the Burkholder inequality and ideas from [24]. Martingale-coboundary decomposition uses a version of real Ruelle-Perron-Frobenius (RPF), which is proven in Theorem 2.4. The second tool needed to prove Proposition 7.1 is an extension of Theorem 2.4 to complex operators (Theorem 7.3). This result generalizes the complex RPF theorem in [46, 48] and is proven using simpler arguments (see Appendix D) which result in conditions which are easier to verify.

Our setup includes a reference measure m_0 (e.g. Lebesgue or a time 0 Gibbs measure) and a notion of variation of functions (e.g. variation on $[0, 1]$ or Hölder continuity). Then our result hold true with respect to any initial measure μ_0 which is absolutely continuous with respect to the reference measure, with bounded density with finite variation. For autonomous systems the fact that the weak limit is preserved when changing densities is called Eagleson's theorem [36] (see also [45, 80]). Eagleson's theorem concerns the asymptotic behavior, and to show that the rates are persevered it is reasonable to require that the density is sufficiently regular. In [53] non-stationary versions of Eagleson's theorem were discussed. When applied to the setup of this paper, the results in [53, §3.3.1] show that optimal uniform CLT rates are preserved under a change of density with bounded variation if the variance of the underlying sum grows linearly fast in n (under one of the measures). Here we show that the non-uniform optimal rate holds for all the measures in the above class without any growth assumptions on the variance.

1.3. **The layout of the paper.** Section 2 describes the classes of maps we consider. Section 3 presents our main results. Sections 4 and 5 are devoted to examples. In the former section we present specific examples fitting into our abstract framework: expanding and piecewise expanding maps, subshifts of finite type, etc. In the next section we show that our abstract setting includes the setups of [24] and [48] as special cases. Section 6 discusses the moment estimates which play a key role in our analysis. Section 7 is devoted to the analysis of of the characteristic functions near the origin. This analysis plays a key role in the proof of our main results given in Section 7.3. In a forthcoming work, including [28], we combine the results of Section 7 with the estimates of the characteristic function away from the origin to get further asymptotic results related to limit theorems.

The examples in Sections 4 and the proof of the main results rely on many non-autonomous versions of known results for autonomous dynamical systems. Additionally, we find it more reader friendly to include a presentation of our results for hyperbolic maps in a separate section. For these reasons the paper has four appendixes. Appendix A is devoted to mixing properties of sequential Gibbs measures. It generalizes the fact that for a topologically mixing expanding system on a compact manifold there is a unique absolutely continuous invariant measure, which is mixing exponentially fast for Hölder or BV observables. Appendix B concerns extension of classical results for subshift of finite types to their non-autonomous counterparts, including the theory of sequential Gibbs measures. Appendix C is about non-autonomous hyperbolic systems formed by a sequence of maps (T_j) in a small C^1-neighborhood of a given Axiom A map. We show that such systems are sequentially Hölder conjugated to the Axiom A map, which will yield that all our results hold true for these systems. Lastly, Appendix D contains a short proof of the complex sequential Ruelle-Perron-Frobenius Theorem following the arguments of [55]. The assumptions of our Theorem D.2 are easier to verify in specific examples than the assumptions of the corresponding results in [46, 48].

2. Preliminaries

2.1. **Setup.** Our setup here will be a sequential (uniform) version of [19]. Let $(X_j, \mathcal{B}_j, m_j)_{j=0}^{\infty}$ be a sequence of probability spaces. For a measurable function $h : X_j \to \mathbb{C}$ denote $\|h\|_p = \|h\|_{p,j} = \|h\|_{L^p(X_j,\mathcal{B}_j,m_j)}$. We suppose that there are notions of variation $v_j : L^1(X_j, m_j) \to [0, \infty]$ satisfying

(V1) $v_j(th) = |t|v_j(h)$, for every $t \in \mathbb{C}$ and $h : X_j \to \mathbb{C}$;
(V2) $v_j(g + h) \le v_j(g) + v_j(h)$ for every $g, h : X_j \to \mathbb{C}$;
(V3) $\|h\|_\infty \le \|h\|_1 + C_{\text{var}} v_j(h)$ for some constant $C_{\text{var}} > 0$ (independent of j);
(V4) the functions $\mathbf{1}$ taking the constant value 1 have finite variation and $\sup\limits_j v_j(\mathbf{1})<\infty$;
(V5) there is a constant $C > 0$ such that for all j and $f, g : X_j \to \mathbb{C}$ we have $v_j(fg) \le C(\|f\|_\infty v_j(g) + \|g\|_\infty v_j(f))$;
(V6) if h is a positive function with bounded variation bounded below by a constant $c > 0$ then $v_j(1/h) \le C(c, v_j(h) + \|h\|_{1,j})$, where $C(x, y)$ is a function on $\mathbb{R}_+ \times \mathbb{R}_+$ which is increasing in both variables x and y;
(V7) the space of bounded functions with bounded variation is dense in $C(X_j)$ (the space of continuous bounded functions on X_j).

Example 2.1. (1) X_j are metric spaces such that $\sup\limits_j \operatorname{diam}(X_j) < \infty$ and v_j is the Hölder constant corresponding to some exponent $\alpha\in(0, 1]$ independent of j.

(2) X_j are Riemannian manifolds such that $\sup\limits_j \operatorname{diam}(X_j) < \infty$ and $v_j(g) = \sup |Dg|$.

(3) $X_j = [0, 1]$, m_j=Lebesgue and each $v_j(g)$ is the usual variation of a function g.

(4) $X_j = X$ is bounded subset of $\mathbb{R}^d$, $d > 1$ and all m_j coincide with the normalized Lebsegue measure on X. Moreover,

$$v_j(g) = \sup_{0<\varepsilon\le\varepsilon_0} \varepsilon^{-\alpha} \int_X \operatorname{Osc}(g, B_\varepsilon(x))dx$$

for some constants $\varepsilon_0 > 0$ and $\alpha \in (0, 1]$, where $B_\varepsilon(x)$ is the ball of radius ε around a point x and for a set A, $\operatorname{Osc}(g, A) = \sup\limits_{x_1,x_2\in A\cap X} |g(x_1) - g(x_2)|$.

Next, given a measurable function $h : X_j \to \mathbb{C}$, denote

$$\|h\|_{BV} = \|h\|_{BV,j} = \|h\|_{1,j} + v_j(h).$$

Then $\|\cdot\|_{BV}$ is a complete norm on the space of bounded functions B_j with finite variation. Note that $\|\cdot\|_{BV,j}$ are equivalent to the norms $\|g\|_{BV,\infty,j} = \|g\|_{\infty,j} + v_j(g)$, uniformly in j.

Let $T_j : X_j \to X_{j+1}$, $j \ge 0$ be a sequence of measurable maps, such that

$$\sup_j \sup_{h:v_{j+1}(h)\le 1} v_j(h \circ T_j) < \infty. \tag{2.1}$$

We also assume that the maps are absolutely continuous, that is, $(T_j)_* m_j \ll m_{j+1}$. Let $\mathcal{L}_j$ denote the transfer operator of T_j with respect to the measures m_j and m_{j+1}. Namely, if $\kappa_j = \rho_j dm_j$ for some $\rho_j \in L^1(X_j, m_j)$ then $\mathcal{L}_j \rho_j : X_{j+1} \to \mathbb{R}$ is the Radon–Nikodym derivative of the measure $(T_j)_*\kappa_j$. Then $\mathcal{L}_j$ is the unique linear operator satisfying the duality relation:

$$\int (f \circ T_j) g dm_j = \int f(\mathcal{L}_j g) dm_{j+1} \tag{2.2}$$

for all bounded measurable functions g on X_j and f on X_{j+1}. Denote

$$\mathcal{L}_j^n = \mathcal{L}_{j+n-1} \circ \cdots \circ \mathcal{L}_{j+1} \circ \mathcal{L}_j.$$

We make three assumptions which are sequential versions of the assumptions in [19].

(LY1) $\sup_j \|\mathcal{L}_j\|_{BV} < \infty$

(LY2) There are constants $\rho \in (0,1)$, $K \ge 1$ and $N \in \mathbb{N}$ such that for every j and a real function $h \in B_j$ we have

$$v_{j+N}(\mathcal{L}_j^N h) \le \rho v_j(h) + K\|h\|_1.$$

(SC) For $a > 0$ let $\mathcal{C}_{j,a} = \{h \in L^1(m_j) : h \ge 0,\ \ v_j(h) \le a m_j(h)\}$. Then for every a there are $n(a) \ge 1$ and $\alpha(a) > 0$ such that for all j, $n \ge n(a)$ and $h \in \mathcal{C}_{a,j}$

$$\text{ess-inf } \mathcal{L}_j^n h \ge \alpha(a) m_j(h).$$

Note that (LY) stands for "Lasota Yorke" and (SC) stands for "sequential covering".

In Section 4 we will verify the above assumptions for particular examples.

Remark 2.2. By iterating (LY2) it follows that there is a constant $K_0 > 0$ such that for all j and all $n \ge N$ and a real function $h \in B_j$ we have

$$v_{j+n}(\mathcal{L}_j^n h) \le \rho^{1+n-N} v_j(h) + K_0\|h\|_1. \tag{2.3}$$

Lemma 2.3. Under (LY1) and (LY2) we have

$$\sup_j \sup_n \|\mathcal{L}_j^n \mathbf{1}\|_\infty < \infty$$

where $\mathbf{1}$ denotes the function taking the constant value 1, regardless of its domain.

Proof. It follows from the remark and (LY1) that $\sup_j \sup_n v_{j+n}(\mathcal{L}_j^n \mathbf{1}) < \infty$. Next, by property (V3) we have $\|\mathcal{L}_j^n \mathbf{1}\|_\infty \le \|\mathcal{L}_j^n \mathbf{1}\|_1 + C_{var} v_{j+n}(\mathcal{L}_j^n \mathbf{1})$. To complete the proof of the lemma we note that by taking $f = \mathbf{1}$ in (2.2) with $\mathcal{L}_j^n$ instead of $\mathcal{L}_j$ and T_j^n instead of T_j and using the positivity of $\mathcal{L}_j^n$ we have $\|\mathcal{L}_j^n \mathbf{1}\|_1 = \|\mathbf{1}\|_1 = 1$. $\square$

Appendix A contains the following result.

Theorem 2.4. Suppose (LY1), (LY2) and (SC).

(i) There is a sequence of positive functions h_j which is uniformly bounded in B_j and uniformly bounded away from 0 with $m_j(h_j) = 1$, and constants $C > 0$, $\delta \in (0,1)$ such that for all $j \ge 0$ we have $\mathcal{L}_j h_j = h_{j+1}$ and for all $n \in \mathbb{N}$,

$$\left\| \mathcal{L}_j^n(\cdot) - m_j(\cdot) h_{j+n} \right\|_{BV} \le C\delta^n. \tag{2.4}$$

(ii) Let $\mu_j := h_j dm_j$. Then $(T_j)_*\mu_j = \mu_{j+1}$. Moreover if $\tilde{\mu}_j = g_j dm_j$ is another sequence satisfying

$$(T_j)_*\tilde{\mu}_j = \tilde{\mu}_{j+1} \tag{2.5}$$

then $\lim_{n\to\infty} \|h_n - g_n\|_1 = 0$. In fact, if $g_k \in BV$ for some k then $\lim_{n\to\infty} \|h_n - g_n\|_{BV} = 0$ exponentially fast.

Remark 2.5. In general, the measures μ_j are not unique. That is (2.5) does **not** imply that $\tilde{\mu}_j = \mu_j$ for all j. In fact, for every BV density $g_0 : X_0 \to \mathbb{R}$ the measures $\tilde{\mu}_j = (T_0^j)_*(g_0 dm_0)$ also satisfy (2.5) but in general they differ from μ_j (even when T_j does not depend on j). The point is that the density of $\tilde{\mu}_j$ is $g_j = \mathcal{L}_0^j g_0$. This density satisfies $\|g_n - h_n\|_{BV} = O(\delta^n)$ by Theorem 2.4(i). However g_n differs from h_n in general since h_n corresponds to some possibly other choice of g_0.

Remark 2.6. Let $g \in B_j, f \in B_{j+n}$. Then for all j, n we have

$$m_j(g \cdot (f \circ T_j^n)) = m_{j+n}((\mathcal{L}_j^n g) f).$$

Plugging (2.4) into the RHS we get $|m_j(g \cdot f \circ T_j^n) - m_j(g)\mu_{j+n}(f)| \leq C_1 \|g\|_{BV} \|f\|_1 \delta^n$ for some constant $C_1 > 0$. Hence $m_{j+n}(f \circ T_j^n) \approx \mu_{j+n}(f)$. In particular if the operators T_j and measures m_j are defined for all $j \in \mathbb{Z}$ so that the assumpions (LY1), (LY2) and (SC) are valid for all $j \in \mathbb{Z}$ then $\mu_j(f) = \lim_{n\to\infty} m_{j-n}(f \circ T_{j-n}^n)$, and so in this case μ_j is the unique equivariant family of measures such that $\sup_j \left\| \frac{d\mu_j}{dm_j} \right\| < \infty$.

2.2. **Changing the reference measures.** It is important to note that once conditions (LY1), (LY2) and (SC) hold for a given sequence of measures m_j, they must hold with a uniformly equivalent sequence of measures (and hence, all the limit theorems stated in the next section are valid for initial measures having BV densities with respect to m_0). This is the content of the following result.

Proposition 2.7. Let μ_j be a sequence of probability measures such that $\mu_j = h_j dm_j$, with h_j uniformly bounded and bounded away from the origin and $\sup_j v_j(h_j) < \infty$.

Then (V1)-(V7) and (LY1), (LY2) and (SC) also hold when we repalce m_j by μ_j.

This proposition will allow us to prove limit theorems with respect to the Lebesgue measures (i.e. the case m_j=Lebesgue), as well as with respect to the unique absolutely continuous (sequentially) invariant measures μ_j or any other sequence of equivalent measures. This is one of the advantages of the setup presented in this section.

Proof. First, it is clear that there is a constant $C_1 > 0$ such that for every function $g : X_j \to \mathbb{C}$,

$$C_1 \|g\|_{L^1(m_j)} \leq \|g\|_{L^1(\mu_j)} \leq \|g\|_{L^1(m_j)}. \tag{2.6}$$

Consequently, the BV norms induced from both measures are equivalent. Moreover, since the measures m_j and μ_j are equivalent we have $\|\cdot\|_{L^\infty(m_j)} = \|\cdot\|_{L^\infty(\mu_j)}$. Therefore (V1)-(V7) remain true if we replace m_j with μ_j.

Next, the transfer operator L_j corresponding to the measures μ_j and μ_{j+1} is given by

$$L_j g = \frac{\mathcal{L}_j(g h_j)}{h_{j+1}}.$$

By (V6), $\sup_j v_j(1/h_j) < \infty$. So conditions (LY1) and (LY2) also hold true with L_j instead of $\mathcal{L}_j$ (possibly with different constants).

Finally, notice that the operator $\mathcal{L}_j$ is positive. Therefore, if c is a constant such that $h_j \geq c > 0$ then for every function $h \geq 0$,

$$\mathcal{L}_j^n(h h_j) \geq c \mathcal{L}_j^n(h), \;\; m_{j+1} \text{ a.s.}$$

Now, since $\|1/h_{j+1}\|_{L^\infty} \geq D$ for some constant $D < \infty$, we conclude that for every non-negative function h and all n we have $L_j^n h \geq D c \mathcal{L}_j^n h$, a.s. Consequently, the validity of condition (SC) for the operators L_j follows from the corresponding validity for $\mathcal{L}_j$, together with (2.6). □

3. Main results

Let $f_j : X_j \to \mathbb{R}$, $j \geq 0$ be a sequence of measurable functions such that $\sup_j \|f_j\|_{BV} < \infty$. Set

$S_n(x) = \sum_{j=0}^{n-1} f_j(T_0^j x)$ where $T_k^j = T_{k+j-1} \circ \cdots \circ T_{k+1} \circ T_k$.

We consider the sequence of functions $S_n : X_0 \to \mathbb{R}$ as random variables on the probability space $(X_0, \mathcal{B}_0, m_0)$. Our results will be limit theorems for such sequences.

Denote $\sigma_n = \sqrt{\mathrm{Var}_{m_0}(S_n)}$ and $\bar{S}_n = S_n - m_0(S_n)$. Let $F_n(t) = \mathbb{P}_{m_0}(\bar{S}_n/\sigma_n \le t)$ be the distribution function of $\bar{S}_n/\sigma_n$, and let

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx$$

be the standard normal distribution function.

Recall that the (self-normalized) central limit theorem (CLT) means that for every real t

$$\lim_{n\to\infty} F_n(t) = \Phi(t).$$

The CLT in our setup can be proven using martingale coboundary decomposition of §6.2 and applying an appropriate CLT for martingales (cf. [42, 24]). We refer to §6.2 for a characterization when $\sigma_n \to \infty$. In this paper we will not give a separate proof of the CLT since our main results give not only the CLT but also rate of convergence, in various metrics.

Theorem 3.1. Suppose $\sigma_n \to \infty$.

(i) (A non-uniform Berry-Esseen theorem). $\forall s \ge 0$ there is a constant C_s such that

$$\sup_{t\in\mathbb{R}} (1 + |t|^s) \, |F_n(t) - \Phi(t)| \le C_s \sigma_n^{-1}.$$

(ii) (A Berry-Esseen theorem in L^p). For all $p \ge 1$ we have $\|F_n - \Phi\|_{L^p(dx)} = O(\sigma_n^{-1})$.

(iii) For all $s \ge 1$ there is a constant C_s such that for every absolutely continuous function $h : \mathbb{R} \to \mathbb{R}$ such that $H_s(h) := \int \frac{|h'(x)|}{1+|x|^s} dx < \infty$ we have $\left| \mathbb{E}_{m_0}[h(\bar{S}_n/\sigma_n)] - \int h d\Phi \right| \le C_s H_s(h) \sigma_n^{-1}$.

Next, recall that the p-th Wasserstien distance between two probability measures μ, ν on $\mathbb{R}$ with finite absolute moments of order p is given by

$$W_p(\mu, \nu) = \inf_{(X,Y)\in\mathcal{C}(\mu,\nu)} \|X - Y\|_{L^p}$$

where $\mathcal{C}(\mu, \nu)$ is the class of all pairs of random variables (X, Y) on $\mathbb{R}^2$ such that X is distributed according to μ, and Y is distributed according to ν. Combining our estimates with the main results in [52] also yields the following result.

Theorem 3.2. [A Berry-Esseen theorem in W_p] For every p we have

$$W_p(dF_n, d\Phi) = O(\sigma_n^{-1})$$

where dG is the measure induced by a distribution function G.

A key ingredient in the proof of Theorems 3.1 and 3.2 is the following proposition, which we believe has its own interest.

Proposition 3.3. For every $p \ge 2$ there is a constant C such that for all $j \ge 0$ and $n \in \mathbb{N}$,

$$\|\bar{S}_{j,n} f\|_{L^p(m_0)} \le C(1 + \|\bar{S}_{j,n}\|_{L^2}).$$

Moreover, C depends only on p, $\sup_j \|f_j\|_{BV}$, and the constants from assumptions (V1)–(V7), (LY1), (LY2) and (SC).

4. Examples

Here we exhibit several classes of systems fitting in the abstract setup of Section 2.

4.1. **Piecewise expanding maps on the interval.** We take $X_j = I = [0,1]$ and m_j =Lebesgue for all j. Let $v_j = v$ be the usual variation of functions on $[0,1]$:

$$v(g) = \sup_n \sup_{t_0=0<t_1<...<t_n<t_{n+1}=1} \sum_{j=0}^{n} |g(t_{n+1}) - g(t_n))|.$$

We suppose that for each j we can write $[0,1] = \bigcup_{k=1}^{d_j} I_{j,k}$ where $I_{j,k}, 1 \le k \le d_j$ are intervals with disjoint interiors such that $\sup_j d_j < \infty$.

We also suppose that for all j and k the restriction $T_{j,k} := T_j | I_{j,k}$ is a C^2 expanding map so that $\sup_j \max_{1\le k\le d_j} \|T''_{j,k}\|_\infty < \infty$ and $\delta := \inf_j \min_{1\le k\le d_j} \inf |T'_{j,k}| > 1$. Moreover, we assume that

$$\inf_j \min_{1\le k\le d_j} |I_{j,k}| > 0.$$

Let $\mathcal{L}_j$ be the transfer operator of T_j, namely the operator given by

$$\mathcal{L}_j g(x) = \sum_{k:\, x\in T_{j,k}(I_{j,k})} \frac{g(T_{j,k}^{-1}x)}{T'_j(T_{j,k}^{-1}x)}.$$

Then (LY1) holds. A standard argument (see [63]) yields that if N satisfies $\delta^N > 2$ then there are constants $K \ge 1$ and $\rho \in (0,1)$ such that for all j we have $\mathrm{var}(\mathcal{L}_j g) \le \rho\, \mathrm{var}(g) + \|g\|_{L^1}$. Thus, (LY2) holds.

Next, in order to verify (SC) we can assume that for every interval $J \subset [0,1]$ there is $n(J) \in \mathbb{N}$ such that for every j we have

$$T_j^{n(J)} J = [0,1]. \tag{4.1}$$

Under the above condition the verification (SC) is carried out similarly to [19, §1.2].

For piecewise expanding maps satisfying the assumptions described above we get all the limit theorems described in the previous section for sums of the form $S_n = \sum_{j=0}^{n-1} f_j \circ T_0^j$ considered as random variables with respect to a measures which is absolutely continuous with respect to Lebesgue and its density is a BV function. Here, f_j must satisfy $\sup_j \|f_j\|_{BV} < \infty$.

4.2. **High dimensional piecewise expanding maps.** Let $X_j = X$ coincide with a single compact subset of $\mathbb{R}^k$ for some $k > 1$. Let m be the normalized Lebesgue measure on X and let v be the variation defined in Example 2.1(iv).

We suppose that the maps T_j have the following properties. There are constants $d \in \mathbb{N}$, $\gamma, C, \varepsilon > 0$ and $s \in (0,1)$ with the following properties. For each j there are disjoint sets $A_{j,i}, \tilde{A}_{j,i}, 1 \le i \le d_j \le d$ and maps $T_{j,i} : \tilde{A}_{j,i} \to X$ such that:

(i) The sets $A_{j,1}, ..., A_{j,d_j}$ are disjoint and $m(X \setminus \bigcup_{i=1}^{d_j} A_{j,i}) = 0$. Moreover, $A_{j,i} \subset \tilde{A}_{j,i}$;

(ii) Each $T_{j,i}$ is a $C^{1+\gamma}$ function;

(iii) $T_j|_{A_{j,i}} = T_{j,i}$ and for all j and i and $B_\varepsilon(T_{j,i}A_{j,i}) \subset T_{j,i}(\tilde{A}_{j,i})$, where $B_\varepsilon(A)$ is the ε-neighborhood of a set A;

(iv) For all j and i the function $J_{j,i} = \mathrm{Det}(DT_{j,i}^{-1})$ satisfies that for all $x, y \in T_{j,i}(A_{j,i})$,

$$\left|\frac{J_{j,i}(y)}{J_{j,i}(x)} - 1\right| \leq C\mathrm{dist}(x,y)^{\gamma};$$

(v) For every $x, y \in T_{j,i}(\tilde{A}_{j,i})$ with $\mathrm{dist}(x,y) \leq \varepsilon$ we have $\mathrm{dist}(T_{j,i}^{-1}x, T_{j,i}^{-1}y) \leq s \cdot \mathrm{dist}(x,y)$;

(vi) Each $\partial A_{j,i}$ is co-dimension one embedded compact C^1-submanifold and

$$s^{\gamma} + \frac{4s}{1-s} Z \frac{\Gamma_k}{\Gamma_{k-1}} < 1$$

where Γ_k is the volume of the unit ball in $\mathbb{R}^k$ and $Z = \sup_j \sup_x \sum_i \mathbb{I}(x \in \tilde{A}_{i,j})$.

Under the above assumptions (LY1) and (LY2) with N=1 are satisfied (see [78, Lemma 4.1]).

Next, we also assume that for any open set U there exists $n(U) \in \mathbb{N}$ such that for all j

$$T_j^{n(U)} U = X. \tag{4.2}$$

Under the above condition the verification (SC) is carried out similarly to [30, Lemma 3].

One example where (4.2) holds are *Markov* maps. That is, we assume that for each i, j the image $T_j A_{ij}$ is a union of some of the sets $A_{j+1,k}$. Moreover suppose that the system is uniformly mixing in the sense that $\exists \ell$ such that $T_j^{\ell} A_{ij} = X_{j+\ell} = X$ for each i, j. This condition can be verified as follows. Consider the adjacency matrix $\mathcal{A}_j$ such that $\mathcal{A}_j(i,k) = 1$ if $T_j A_{j,i} \supset A_{j+1,k}$. Then the uniform mixing assumption means that for each j all entries of $\mathcal{A}_{j+\ell-1} \cdots \mathcal{A}_{j+1}\mathcal{A}_j$ are positive. Now let U be an open set in $X_j = X$. Without loss of generality we may assume that U is an open ball with center x and radius r. Given k let $\mathcal{B}_{j,k}(x)$ denote the set of points $y \in X$ such that $T_j^m x$ and $T_j^m y$ belong to the same elements of our partition $(A_{j+m,q})_q$ for all $m < k$. By our assumptions $\mathrm{diam}(\mathcal{B}_{j,k}(x)) \leq C_0 s^k$ for some constant $C_0 > 0$, and so for sufficiently large $\bar{k}$ we have $\mathcal{B}_{j,\bar{k}}(x) \subset U$. Accordingly, $T_j^{\bar{k}} U$ contains one of elements of our Markov partition and then $T_j^{\bar{k}+\ell} U = X$.

4.3. **Covering maps and sequential SFT.** Suppose that each X_j is a metric space. Let d_j be the metric on X_j and suppose that $\mathrm{diam}(X_j) \leq 1$. Let $v_j = v_{j,\alpha}$ be the Hölder constant corresponding to some fixed exponent $\alpha \in (0,1]$.

Assumption 4.1. (Pairing). There are constants $\xi \leq 1$ and $\gamma > 1$ such that for every two points $x, x' \in X_{j+1}$ with $\mathsf{d}_{j+1}(x,x') \leq \xi$ we can write

$$T_j^{-1}\{x\} = \{y_i(x) : i \leq k\}, \quad T_j^{-1}(x') = \{y_i(x') : i \leq k\}$$

where $\mathsf{d}_j(y_i(x), y_i(x')) \leq \gamma^{-1}\mathsf{d}_{j+1}(x,x')$ for all i.
Moreover $\sup_j \deg(T_j) < \infty$, where $\deg(T)$ is the largest number of preimgaes that a point x can have under the map T.

Denote by $B_j(x,r)$ the open ball of radius r around a point $x \in X_j$.

Assumption 4.2. (Covering). There exists $n_0 \in \mathbb{N}$ such that for every j and $x \in X_j$ we have

$$T_j^{n_0}(B_j(x,\xi)) = X_{j+n_0}. \tag{4.3}$$

Fix some $\alpha \in (0,1]$ and a sequence of functions $\phi_j : X_j \to \mathbb{R}$ such that $\sup_j \|\phi_j\|_\alpha < \infty$. Here $\|\phi_j\|_\alpha = \sup|\phi_j| + v_j(\phi_j)$ and $v_j(\phi_j)$ is the Hölder constant of ϕ_j corresponding to the exponent α. Let L_j be the operator which maps a function $g : X_j \to \mathbb{R}$ to a function $L_j g : X_{j+1} \to \mathbb{R}$ given by $L_j g(x) = \sum_{T_j y = x} e^{\phi_j(y)} g(y)$. Then (see §5.2) there is a sequence of probability measures ν_j on X_j such that $(\mathcal{L}_j)^* \nu_{j+1} = \lambda_j \nu_j$, where $\lambda_j > 0$ is bounded and bounded away from 0. Then we can take any measure m_j of the form $m_j = u_j d\nu_j$ with $\sup_j \|u_j\|_\alpha < \infty$. This includes the unique sequence of measures μ_j which are absolutely continuous with respect to ν_j and $(T_j)_* \mu_j = \mu_{j+1}$ (see §5.2). This setup includes the following more concrete examples.

4.3.1. *Smooth expanding maps.* Let M be C^2 compact connected Riemannian manifold, and let $X_j = M$ for all j. Let $T_j : M \to M$ be C^2 endomorphisms of M such that

$$\sup_j \|DT_j\| < \infty \text{ and } \sup_j \|(DT_j)^{-1}\| < 1.$$

Then the arguments in [61, Section 4] (see [61, (4.6)]) yield that Assumption 4.1 is in force with some $\xi > 0$. Next, arguing like at the paragraph below [61, (4.19)]) we see that Assumption 4.2 holds if n_0 is large enough. Take $\phi_j = -\ln J_{T_j}$. Then by [48, Theorem 3.3 and Proposition 3.4] (see also [61, Theorem 2.2]) we see that the measures ν_j described after Assumptions 4.2 coincide with the normalized volume measure on M. Thus, we get all the limit theorems with respect to any absolutely continuous measure whose density is Hölder continuous.

4.3.2. *Subshifts of finite type.* Let $\mathcal{A}_j = \{0, 1, ..., d_j - 1\}$ with $\sup_j d_j < \infty$. Le $A^{(j)}$ be matrices of sizes $d_j \times d_{j+1}$ with 0-1 entries. We suppose that there exists an $M \in \mathbb{N}$ such that for every j the matrix $A^{(j)} \cdot A^{(j+1)} \cdots A^{(j+M)}$ has positive entries. Define

$$X_j = \left\{ (x_{j,k})_{k=0}^{\infty} : x_{j,k} \in \mathcal{A}_{j+k}, A^{(j+k)}_{x_{j,k}, x_{j,k+1}} = 1 \right\}. \tag{4.4}$$

Let $T_j : X_j \to X_{j+1}$ be the left shift. Consider a metric d_j on X_j given by

$$\mathsf{d}_j(x, y) = 2^{-\inf\{k : x_{j,k} \neq y_{j,k}\}}.$$

With this metric the maps T_j satisfy Assumptions 4.1 and 4.2. In order to introduce appropriate measures m_0 note that we can extend the dynamics for negative times by defining X_j for $j<0$ via appropriate extensions of the sequences (d_j) and $(A^{(j)})$. Note that such extensions are highly non-unique. Each one of these extensions gives raise to a unique time 0 Gibbs measure μ_0 (see Appendix B). Thus, all of our results hold true when starting with any measure which is absolutely continuous with Hölder continuous density with respect to $m_0 = \mu_0$.

4.3.3. *Two sided SFT.* Using the same notations like in the previous section but also considering negative integers j, we define

$$\tilde{X}_j = \left\{ (x_{j,k})_{k=-\infty}^{\infty} : x_{j,k} \in \mathcal{A}_{j+k}, \;\; A^{(j+k)}_{x_{j,k}, x_{j,k+1}} = 1 \right\}. \tag{4.5}$$

Let $\tilde{T}_j : \tilde{X}_j \to \tilde{X}_{j+1}$ be the left shift, and let $\tilde{T}_j^n = \tilde{T}_{j+n-1} \circ \cdots \circ \tilde{T}_{j+1} \circ \tilde{T}_j$. Consider the metric $\tilde{\mathsf{d}}_j$ on $\tilde{X}_j$ given by $\tilde{\mathsf{d}}_j(x, y) = 2^{-\inf\{|k| : x_{j,k} \neq y_{j,k}\}}$. Let $\pi_j : \tilde{X}_j \to X_j$ be the natural projection.

By Lemma B.2 (proven in Appendix B), given a sequence of functions $\mathbf{f}_j : \tilde{X}_j \to \mathbb{R}$ with $\sup_j \|\mathbf{f}_j\|_\alpha < \infty$ there is sequences of functions $f_j : X_j \to \mathbb{R}$ and $u_j : \tilde{X}_j \to \mathbb{R}$ such that $\sup_j \|f_j\|_{\alpha/2} < \infty$, $\mathfrak{u} = \sup_j \|u_j\|_{\alpha/2} < \infty$ and

$$\mathbf{f}_j = f_j \circ \pi_j + u_{j+1} \circ \tilde{T}_j - u_j.$$

Therefore, denoting $S_n \mathbf{f} = \sum_{j=0}^{n-1} \mathbf{f}_j \circ \tilde{T}_0^j$ we have $\sup_n \sup |S_n \mathbf{f} - S_n f| < \infty$.

Next, let γ_j denote a Gibbs measure of the two sided shift at time j (see Appendix B). Then $\mu_j = (\pi_j)_*\gamma_j$ are also Gibbs measures for one sided subshift and they satisfy the assumptions of Section 2 (see Appendix B for details). We view $S_n\mathbf{f}$ as random variables on $(\tilde{X}_0, \text{Borel}, \gamma_0)$. Then

$$A := \sup_n \|S_n\mathbf{f} - S_n f\|_{L^\infty(\gamma_0)} < \infty. \tag{4.6}$$

This is enough to do deduce all of our results for $S_n\mathbf{f}$, relying on the corresponding results for $S_n f$. Indeed, part (ii) of Theorem 3.1 is a direct consequence of part (i). Theorem 3.1(iii) also follows from Theorem 3.1(i). Indeed, for every random variable W with distribution function F and a function h satisfying $H_s(h) < \infty$ we have

$$E[h(W)] - h(\infty) = -E\left[\int_W^\infty h'(x)dx\right] = -\int_{-\infty}^\infty h'(x)P(W \le x)dx = -\int_{-\infty}^\infty h'(x)F(x)dx.$$

To show that Theorem 3.1(i) for $S_n f$ implies Theorem 3.1(i) for $S_n\mathbf{f}$, let

$$F_n(t) = \mathbb{P}\left(\frac{S_n\mathbf{f}}{\sigma_n} \le t\right), \quad G_n(t) = \mathbb{P}\left(\frac{S_n f}{\kappa_n} \le t\right) \text{ where } \kappa_n = \|S_n f\|_{L^2} \text{ and } \sigma_n = \|S_n\mathbf{f}\|_{L^2}.$$

By (4.6) and the triangle inequality, $|\sigma_n - \kappa_n| \le A$. To complete the proof fix $s \ge 0$. Then

$$\begin{aligned} F_n(t) &\le G_n(t\sigma_n/\kappa_n + A/\kappa_n) \le \Phi(t\sigma_n/\kappa_n + A/\kappa_n) + C_s\left(1 + |t\sigma_n/\kappa_n + A/\kappa_n|^s\right)^{-1} \\ &\le \Phi(t) + C(1+|t|)\sigma_n^{-1}e^{-ct^2} + \tilde{C}_s(1+|t|^s)^{-1} \end{aligned}$$

where in the penultimate inequality we have used that $|\Phi(x+\varepsilon) - \Phi(x)| \le C\varepsilon e^{-x^2/2}$ for every x and $\varepsilon > 0$ and that $|t\sigma_n/\kappa_n + A/\kappa_n - t| \le C(|t|+1)/\sigma_n$ for some constant $C > 0$. Similarly,

$$F_n(t) \ge G_n(t\sigma_n/\kappa_n - A/\kappa_n)) \ge \Phi(t) - C(1+|t|)\sigma_n^{-1}e^{-ct^2} - \tilde{C}_s(1+|t|^s)^{-1}.$$

Finally to deduce Theorem 3.2 for $S_n\mathbf{f}$ from the corresponding result for $S_n f$ let us fix some $p \ge 1$. Then by Theorem 3.2, we can couple $S_n f$ with a standard normal random variable Z so that $\|S_n f/\sigma_n - Z\|_{L^p} \le C\sigma_n^{-1}$. Now by Berkes–Philipp Lemma [17, Lemma A.1], we can also couple all three random variables $S_n f$, $S_n\mathbf{f}$ and Z so that (4.6) still holds under the new probability law.

5. Verification of (LY1), (LY2) and (SC) for some classes of maps

In this section we will show that the conditions in Section 3 are in force for the classes of expanding maps considered in both [24] and [48].

5.1. Verification of our assumptions in the setup of [24].

The following assumption is taken form [24].

Assumption 5.1. (i) $(X_j, \mathcal{B}_j, m_j)$ coincide with the same probability space $(X, \mathcal{B}, m)$, $v_j = v$ does not depend on j.

(ii) Conditions (LY1) and (LY2) hold with $m_j = m$ and $v_j = v$.

(iii) There is a constant $\delta_0 > 0$ such that

$$\text{ess-inf } \mathcal{L}_0^n\mathbf{1} \ge \delta_0 \tag{5.1}$$

(note that the condition (5.1) is denoted by (Min) in [24]). Here $\mathbf{1}$ denotes the function taking the constant value 1 on X.

(iv) There are constants $C_1 > 0$ and $\delta_1 \in (0,1)$ such that for every $h \in B$ with $m(h) = 0$

$$\|\mathcal{L}_j^n\|_{BV} \le C_1\delta_1^n. \tag{5.2}$$

Lemma 5.2. Assumption 5.1 implies the covering condition (SC).

Thus, Assumption 5.1 is less general than the combination of (LY1), (LY2) and (SC).

Proof. Since (LY1) and (LY2) hold, Lemma 2.3 implies that $C = \sup_k \|\mathcal{L}_0^k 1\|_\infty \in (0,\infty)$. Hence using the positivity of the operators $\mathcal{L}_k$ we see that m-a.s. we have

$$\delta_0 \le \mathcal{L}_0^{j+n}\mathbf{1} = \mathcal{L}_j^n(\mathcal{L}_0^j\mathbf{1}) \le C\mathcal{L}_j^n\mathbf{1}.$$

We thus conclude that $\inf_j \text{ess-inf}\, \mathcal{L}_j^n\mathbf{1} \ge \delta_2 = \delta_0/C$.

Next, fix some a and let $h \in \mathcal{C}_{a,j}$ where $\mathcal{C}_{a,j}$ comes from the condition (SC). Denote $\|\cdot\| = \|\cdot\|_{BV}$. Then by (5.2) we have

$$\|\mathcal{L}_j^n h - m(h)\mathcal{L}_j^n\mathbf{1}\| = \|\mathcal{L}_j^n(h - m(h)\mathbf{1})\| \le C_1\|h - m(h)\mathbf{1}\|\delta_1^n \le C_2\|h\|\delta_1^n \le C_3(a)m(h)\delta_1^n$$

for some constant $C_3(a)$ which depends only on a (we can take $C_3(a) = C_2(1+a)$ because $\|h\| = v(h) + m(h) \le (1+a)m(h)$). Using that $\mathcal{L}_j^n 1 \ge \delta_2 > 0$ we conclude that

$$\mathcal{L}_j^n h \ge \delta_2 m_j(h) - C_3(a)m(h)\delta_1^n = (\delta_2 - C_3(a)\delta_1^n)m(h) \quad (m\text{-a.s.}).$$

Let $n(a)$ be the smallest positive integer such that $C_3(a)\delta_1^{n(a)} \le \frac{1}{2}\delta_2$. Then for $n \ge n(a)$ we have $\mathcal{L}_j^n h \ge \dfrac{1}{2}\delta_2 m(h)$ (m-a.s.) and so (SC) holds with $\alpha(a) = \frac{1}{2}\delta_2$. □

5.2. **Verification of our assumptions in the setup of** [48]**.** Consider maps T_j from §4.3. We also suppose that we can extend the sequence $(T_j)_{j\ge0}$ to a two sided sequence $(T_j)_{j\in\mathbb{Z}}$ with the same properties. This is possible if there is a map $T_{-1} : X_0 \to X_0$ such that the sequence $(T_{-1}^j)_{j\ge0}$ has same paring property and covering assumption like the sequence $(T_j)_{j\ge0}$ (this is the case when $X_0 = X_1$). Indeed, in this case we can define $T_j = T_{-1}$ for $j < 0$. The need in a two sided sequences in this context arises from the lack of a given reference measure m_0, as will be elaborated in what follows.

Fix some Hölder exponent $\alpha \in (0,1]$ and let $\phi_j : X_j \to \mathbb{R}$, $j \in \mathbb{Z}$ be such that $\sup_j \|\phi_j\|_\alpha < \infty$. Let the operator $\mathbf{L}_j$ be given by

$$\mathbf{L}_j g(x) = \sum_{y: T_j y = x} e^{\phi_j(y)} g(y).$$

Then as proven[1] in [48], there are strictly positive functions $\bar{h}_j \in B_j$, probability measures ν_j on X_j and numbers $\lambda_j > 0$ such that $0 < \inf_j \inf \bar{h}_j \le \sup_j \|\bar{h}_j\|_\alpha < \infty$ and

$$\left\|\mathbf{L}_j^n/\lambda_{j,n} - \nu_j \otimes \bar{h}_{j+n}\right\|_\alpha \le C\delta^n, \quad \delta \in (0,1) \tag{5.3}$$

where $\lambda_{j,n} = \prod_{k=j}^{j+n-1} \lambda_k$ and $(\nu \otimes h)$ is the linear operator $g \to \nu(g)h$. Let

$$\mathcal{L}_j(g) = \mathbf{L}_j(g\bar{h}_j)/\lambda_j\bar{h}_{j+1}$$

and let $m_j = \mu_j$ be the sequential Gibbs measures corresponding to the potentials (ϕ_j), that is $\mu_j = \bar{h}_j\nu_j$. Then $\mathcal{L}_j$ is the dual of T_j with respect to μ_j. Moreover, $(T_j)_*\mu_j = \mu_{j+1}$ and $\mathcal{L}_j\mathbf{1} = \mathbf{1}$.

Proposition 5.3. The operators $\mathcal{L}_j$ obey conditions (LY1), (LY2) and (SC).

[1]This requires us to have a two sides sequence.

Proof. First, the uniform boundedness (LY1) of the operators $\mathcal{L}_j$ follows from the properties of the non-normalized RPF triplets (or from (5.4)). Second, the Lasota–Yorke inequality (LY2) was obtained in [46, Lemma 5.12.2]. Note that in [46, Lemma 5.12.2] the weak norm $\|\cdot\|_{L^1}$ is replaced with the (weak) norm $\|\cdot\|_\infty$. However, we have $\|g\|_\infty \leq \varepsilon_r v(g) + C_r\|g\|_{L^1}$ with $\varepsilon_r \to 0$ as $r \to 0$. Using this we have that the Lasota–Yorke inequality with respect to $(v_j(\cdot), \|\cdot\|_{L^1})$ is equivalent to the Lasota–Yorke inequality with respect to $(v_j(\cdot), \|\cdot\|_\infty)$.

Third, in [48] we proved that there are constants $C > 0$ and $\delta \in (0,1)$ such that for all j, n and a Hölder continuous function $h : X_j \to \mathbb{R}$,

$$\|\mathcal{L}_j^n h - m_j(h)\mathbf{1}\|_\alpha \leq C\delta^n \tag{5.4}$$

and so (5.2) holds. Since $\mathcal{L}_j\mathbf{1}=\mathbf{1}$ we have $\inf_{j,n} \inf \mathcal{L}_j^n\mathbf{1}=\delta_0=1$. Thus we can repeat the arguments from the previous section verbatim with time dependent v and m and obtain (SC). □

6. Moment estimates

Let $T_j^n = T_{j+n-1} \circ \cdots \circ T_{j+1} \circ T_j$. Consider a sequence of real valued functions $f_k \in B_k$, $k \geq 0$ such that $\sup_k \|f_k\|_{BV} < \infty$. Our goal is to obtain limit theorems for the sequence $S_n = \sum_{j=0}^{n-1} f_j \circ T_0^j$ considered as random variables on the probability space $(X, \mathcal{B}, m_0)$. It what follows it will be convenient to consider (f_j) as two sided sequence by setting $f_j = 0$ for $j < 0$.

6.1. **The pulled back measures and their transfer operators.** Consider the sequence of measures $\tilde{m}_j = (T_0^j)_* m_0$ on $(X_j, \mathcal{B}_j)$. Then $\tilde{m}_0 = m_0$ and for all $j > 0$ and each bounded measurable function $G : X_j \to \mathbb{C}$,

$$\tilde{m}_j(G) = m_0(G \circ T_0^j) = m_j(G \cdot \mathcal{L}_0^j\mathbf{1}).$$

Remark 6.1. Note that in the case when $m_j = \mu_j$ is an equivariant sequence (i.e. $(T_j)_*\mu_j = \mu_{j+1}$) then $\tilde{m}_j = \mu_j$.

Let the operator $\tilde{\mathcal{L}}_j$ be the dual of T_j with respect to the measure $\tilde{m}_j$ and $\tilde{m}_{j+1}$defined by the duality relation

$$\int f \cdot (\tilde{\mathcal{L}}_j g) d\tilde{m}_{j+1} = \int (f \circ T_j) \cdot g \cdot d\tilde{m}_j \tag{6.1}$$

for all functions such that both integrals are well defined. Then $\tilde{\mathcal{L}}_j\mathbf{1} = \mathbf{1}$ because $(T_j)_*\tilde{m}_j = \tilde{m}_{j+1}$. Let

$$\tilde{\mathcal{L}}_j^n = \tilde{\mathcal{L}}_{j+n-1} \circ \cdots \circ \tilde{\mathcal{L}}_{j+1} \circ \tilde{\mathcal{L}}_j.$$

Proposition 6.2. There are constants $C_0 > 0$, $J \in \mathbb{N}$ and $\delta_0 \in (0,1)$ such that for all $j \geq 0$ and $n \geq 1$ such that $j + n \geq J$ we have

$$\|\tilde{\mathcal{L}}_j^n(\cdot) - \tilde{m}_j(\cdot)\mathbf{1}\|_{BV} \leq C_0\delta_0^n.$$

Proof. Recall that the functions h_j in Theorem 2.4 satisfy $\eta := \inf_j \text{ess-inf } h_j{>}0$. Thus by (2.4) we see that $\text{ess-inf } \mathcal{L}_j^n\mathbf{1} \geq \eta - C\delta^n$. Consequently, if we take $\delta_0 = \frac{1}{2}\eta$ and $J \in \mathbb{N}$ such that $C\delta^J \leq \frac{1}{2}\eta$ we see that $\text{ess-inf } \mathcal{L}_0^m\mathbf{1} \geq \delta_0$ for all $m \geq J$. Next, a direct calculation shows that $\tilde{\mathcal{L}}_j^n g = \dfrac{\mathcal{L}_j^n(g \cdot \mathcal{L}_0^j\mathbf{1})}{\mathcal{L}_0^{j+n}\mathbf{1}}$. Now the proposition follows from Theorem 2.4, noting that $\|\mathcal{L}_0^m\mathbf{1} - h_m\|_{BV}$ decays exponentially fast and using that, due to (V6), $\sup_j \|1/h_j\|_{BV} < \infty$. □

6.2. **A martingale coboundary representation.**

Lemma 6.3. For every sequence of functions $f_j \in B_j$ such that $\sup_j \|f_j\|_{BV} < \infty$ there are functions $M_j = M_j(f) \in B_j$ and $u_j = u_j(f) \in B_j$ such that

$$\tilde{f}_j := f_j - m_0(f_j \circ T_0^j) = M_j + u_{j+1} \circ T_j - u_j, \;\; j \geq J. \tag{6.2}$$

Moreover, $\sup_j \|u_j\|_{BV} < \infty$ and $(M_j \circ T_0^j)_j$ is a reverse martingale difference with respect to the reverse filtration $\mathcal{A}_j = (T_0^j)^{-1}\mathcal{B}_j$ (on the probability space $(X_0, \mathcal{B}_0, m_0)$). Furthermore, $\sup_j \|u_j\|_{BV}$ is bounded above by a constant which depends only on the constants C and δ from Theorem 2.4 and on $\sup_j \|f_j\|_{BV}$.

Remark 6.4. Note that by (V8) we also get that $\sup_j \|M_j\|_{BV} < \infty$.

Proof of Lemma 6.3. Define $\tilde{\mathcal{L}}_j = 0$ for $j < 0$ and $\tilde{f}_j = 0$ for $j < J$. Set

$$u_j = \sum_{k=1}^{\infty} \tilde{\mathcal{L}}_{j-k}^k \tilde{f}_{j-k} = \sum_{k=1}^{j} \tilde{\mathcal{L}}_{j-k}^k \tilde{f}_{j-k}. \tag{6.3}$$

Since $\tilde{m}_s(\tilde{f}_s) = 0$ for all $s \geq 0$ by Proposition 6.2, $u_j \in B_j$ and $\sup_j \|u_j\|_{BV} < \infty$. Set $M_j = \tilde{f}_j + u_j - u_{j+1} \circ T_j$. It remains to show that $M_j \circ T_0^j$ is indeed a reverse martingale difference. To prove that we notice that $\mathbb{E}[M_j \circ T_0^j | \mathcal{A}_{j+1}] = \tilde{\mathcal{L}}_j(M_j) \circ T_0^{j+1}$. On the other hand, a direct calculation using (6.3) shows that $\tilde{\mathcal{L}}_j(M_j) = 0$. □

6.3. **On the divergence of the variance.** The first step in proving a central limit theorem is to show that the individual summands are negligible in comparison with the variance of the sum. In particular, we need to know when the variance is bounded. In this section we prove the following result.

Theorem 6.5. The following conditions are equivalent.

(1) $\liminf_{n\to\infty} \mathrm{Var}_{m_0}(S_n) < \infty$.

(2) $\sup_{n\in\mathbb{N}} \mathrm{Var}_{m_0}(S_n) < \infty$

(3) We can write $f_j = m_0(f_j \circ T_0^j) + M_j + u_{j+1} \circ T_j - u_j$ with $u_j, M_j \in B_j$ such that $M_j \circ T_0^j$ is a reverse martingale difference on $(X_0, \mathcal{B}_0, m_0)$ with respect to the reverse filtration $T_0^{-j}\mathcal{B}_j$, $\sup_j \|u_j\|_{BV} < \infty$, $\sup_j \|M_j\|_{BV} < \infty$, and $\sum_j \mathrm{Var}_{m_0}(M_j \circ T_0^j) < \infty$.

Proof. First, it is enough to prove the theorem when $m_0(f_j \circ T_0^j) = 0$ for all j. In this case, by Lemma 6.3 we can write

$$f_j = m_0(f_j \circ T_0^j) + M_j + u_{j+1} \circ T_j - u_j = M_j + u_{j+1} \circ T_j - u_j$$

with u_j and M_j like in (3), except that in general the sum of the variances of M_j might not converge. Notice now that

$$\|S_n - S_n M\|_{L^2(m_0)} \leq \|S_n - S_n M\|_{L^\infty(m_0)} \leq 2 \sup_j \|u_j\|_{L^\infty(m_j)} := U < \infty \tag{6.4}$$

where $S_n M = \sum_{j=0}^{n-1} M_j \circ T_0^j$.

Now assume (1), and let n_k be an increasing sequence such that $n_k \to \infty$ and $\sigma_{n_k} = \|S_{n_k}\|_{L^2} \leq C$ for some constant $C > 0$. Then by (6.4), $\|S_{n_k} M\|^2_{L^2(m_0)} \leq (C+U)^2 < \infty$. However, since $M_j \circ T_0^j$ is a reverse martingale, we have

$$\sum_{j=0}^{n_k - 1} \mathrm{Var}_{m_0}(M_j \circ T_0^j) = \|S_{n_k} M\|^2_{L^2(m_0)} \leq (C+U)^2.$$

Now, since $V_n := \|S_n M\|^2_{L^2} = \sum_{j=0}^{n-1} \mathrm{Var}_{m_0}(M_j \circ T_0^j)$ is increasing we conclude that the summability condition in (3) holds. This shows that (1) implies (3).

Next, (2) clearly implies (1). Thus, to complete the proof it is enough to show that (3) implies (2), but this also follows from (6.4) since the latter yields $\|S_n\|^2_{L^2(m_0)} \leq (V_n + U)^2 < \infty$. □

6.4. **Quadratic variation and moment estimates.** Recall that the (unconditioned) quadratic variation difference of the reverse martingale difference $M_j \circ T_0^j$ is given by $q_j(M) := M_j^2 \circ T_0^j$. Henceforth we denote $Q_j = M_j^2$ and let

$$S_{j,n} = S_{j,n} f = \sum_{k=j}^{j+n-1} f_k \circ T_j^k, \quad \bar{S}_{j,n} = S_{j,n} - \tilde{m}_j(S_{j,n}) = S_{j,n} - m_0(S_{j,n} \circ T_0^j).$$

$S_{j,n} M$ and $S_{j,n} Q$ are defined similarly.

Proposition 6.6. There is a constant C which depends only on the constants from Theorem 2.4 and on $\sup_j \|f_j\|_{BV}$ such that for all $j \geq J$ (where J comes from Proposition 6.2) we have $\mathrm{Var}_{m_0}(S_{j,n} Q) \leq C(1 + \mathrm{Var}(S_{j,n} f))$.

Proof. First, to simplify the notation let us assume that $j = J = 0$. Denote

$$g_j = Q_j - \tilde{m}_j(Q_j) = Q_j - \tilde{m}_0(Q_j \circ T_0^j)$$

and $S_n g = \sum_{j=0}^{n-1} g_j \circ T_0^j$. The argument below is similar to the first part of the proof of [24, Theorem 4.1], but we provide the details to make our paper self contained. By Proposition 6.2 (using that $\tilde{m}_0 = m_0$) we have

$$\mathbb{E}_{m_0}[(S_n g)^2] \leq 2 \sum_{0 \leq \ell < n} \sum_{0 \leq k \leq \ell} \left| \tilde{m}_0\big((g_k \circ T_0^k) \cdot (g_\ell \circ T_0^\ell)\big) \right| = 2 \sum_{0 \leq \ell < n} \sum_{0 \leq k \leq \ell} \left| \tilde{m}_k (g_k \cdot \tilde{\mathcal{L}}_k^{\ell - k} g_\ell) \right|$$

$$\leq C_0 \sum_{0 \leq \ell < n} \sum_{0 \leq k \leq \ell} \tilde{m}_k(|g_k|) \|g_\ell\|_{BV} \delta_0^{\ell - k} = C_0 \sum_{0 \leq k \leq n} \tilde{m}_k(|g_k|) \left(\sum_{k \leq \ell < n} \|g_\ell\|_{BV} \delta_0^{\ell - k} \right)$$

$$\leq c_0 \sum_{0 \leq k \leq n} \tilde{m}_k(|g_k|) = c_0 \sum_{0 \leq k \leq n} m_0(|g_k \circ T_0^k|) \leq 2 c_0 \sum_{0 \leq k \leq n} m_0(Q_k \circ T_0^k)$$

for some constant c_0 (the first inequality of the last line uses that $\sup_j \|g_j\|_{BV} < \infty$). Here $\tilde{\mathcal{L}}_j$ are the transfer operators defined by (6.1). Observe that $m_0(Q_k \circ T_0^k) = m_0((M_k \circ T_0^k)^2)$ and, because of the orthogonality property, $\sum_{0\leq k<n} m_0((M_k\circ T_0^k)^2)=\mathrm{Var}_{m_0}(S_nM)$. Now the result follows from (6.4). □

Proof of Proposition 3.3. It is enough to prove the theorem for $j\geq J$, since to get the result for $0\leq j<J$ we can just take C large enough.

To simply the notation, we will only prove the theorem when $j=J=0$, the proof when $j\geq J>0$ is similar.

Notice that it is enough to prove the claim for $p=2^m$ for all m. Moreover, by replacing f_j with $f_j-\mu_0(f_j\circ T_0^j)$ we can and will assume that $\mu_0(S_n)=0$ for all n.

We use induction on m. For $m=1$ the result is trivial. Suppose that the statement is true for some $m\geq 1$. In order to estimate $\|S_nf\|_{2^{m+1}}$ we first use that[2]

$$\|S_nf\|_{2^{m+1}}\leq C+\|S_nM\|_{2^{m+1}}$$

for some constant C which depends only on the constants C and δ from Theorem 2.4 and on $\|f\|:=\sup_j\|f_j\|_{BV}$ since $\|S_nf-S_nM\|_{L^\infty}$ is bounded in n. So it suffices to show that

$$\|S_nM\|_{2^{m+1}}\leq C(1+\|S_nf\|_2) \tag{6.5}$$

for an appropriate constant C.

(6.5) follows from a version of Burkholder's inequality for martingales (see [69, Theorem 2.12]). Let $\mathfrak{d}_1,....,\mathfrak{d}_n$ be a martingale difference with respect to a filtration $(\mathcal{F}_j)_{j=1}^n$ on a probability space. Let $D_n=\mathfrak{d}_1+\mathfrak{d}_2+...+\mathfrak{d}_n$ and $E_n=\mathfrak{d}_1^2+\mathfrak{d}_2^2+...+\mathfrak{d}_n^2$. Then, for every $p\geq 2$ there are constants $c_p,C_p>0$ depending only on p such that

$$c_p\|E_n\|_{p/2}^{1/2}\leq\|D_n\|_p\leq C_p\|E_n\|_{p/2}^{1/2}. \tag{6.6}$$

Applying (6.6) with the (reverse) martingale difference $M_j\circ T_0^j$ we see that

$$\|S_nM\|_{2^{m+1}}\leq a_m\|S_nQ\|_{2^m}^{1/2} \tag{6.7}$$

where $S_nQ=S_{0,n}Q$ and a_m depends only on m. Applying the induction hypothesis with the sequence of functions $\tilde Q_j=Q_j-m_0(Q_j\circ T_0^j)$ we see that there is a constant $R_m>0$ depending only of m and the constants in the formulation of Proposition 3.3 such that

$$\|S_n\tilde Q\|_{2^m}\leq R_m(1+\|S_n\tilde Q\|_2).$$

Since $\mathbb{E}[S_nQ]=\mathrm{Var}(S_nM)$, Proposition 6.6 gives

$$\|S_nQ\|_{2^m}\leq\|S_n\tilde Q\|_{2^m}+\mathbb{E}[S_nQ]\leq R_m\left(1+C(1+\mathrm{Var}(S_nf))\right)+\mathrm{Var}(S_nM)$$

$$\leq R'_m(1+\mathrm{Var}(S_nf))+\mathrm{Var}(S_nM)$$

for some other constant R'_m. Using that $\sup_n\|S_nf-S_nM\|_{L^\infty}<\infty$ we see that there is a constant $C>0$ such that $\mathrm{Var}(S_nM)\leq C(1+\mathrm{Var}(S_nf))$. Thus, there is a constant $R''_m>0$ such that

$$\|S_n\tilde Q\|_{2^m}\leq R''_m(1+\mathrm{Var}(S_nf)).$$

Now (6.5) follows from (6.7), completing the proof of the proposition. □

[2]Where $S_nM=\sum_{j=0}^{n-1}M_j\circ T_0^j$, $M_j=M_j(f)$.

7. Logarithmic growth conditions

In this section we prove the following result.

Proposition 7.1. Let $f_j : X_j \to \mathbb{R}$ be a sequence of functions such that $\sup_j \|f_j\|_{BV} < \infty$ and $\tilde{m}_j(f_j) = 0$. Suppose $\sigma_n^2 = \mathrm{Var}_{m_0}(S_n) \to \infty$. Let $\Lambda_n(t) = \ln m_0(e^{itS_n f/\sigma_n})$. Then for every $k \geq 3$ there are constants $\delta_k > 0$ and $C_k > 0$ such that

$$\sup_{t \in [-\delta_k \sigma_n, \delta_k \sigma_n]} |\Lambda_n^{(k)}(t)| \leq C_k \sigma_n^{-(k-2)}$$

where $\Lambda_n^{(k)}$ is the k-th derivative of the function Λ_n.

The proof uses the complex Ruelle-Perron-Frobinuous theorem, see Theorem 7.3 below.

7.1. **Sequential complex RPF theorem.** Take a sequence of real valued functions $f_j \in B_j$ and $\|f\| = \sup_j \|f_j\|_{BV} < \infty$. Consider the operators $\mathcal{L}_{j,z}(h) = \mathcal{L}_j(he^{zf_j})$ (where $z \in \mathbb{C}$). We need the following result.

Lemma 7.2. (i) $\sup_j \|\mathcal{L}_{j,z}\|_{BV} \leq C(z)$ for some continuous function $C(z)$ (of exponential order in $|z|$).
(ii) The map $z \to \mathcal{L}_{j,z}$ is analytic and its k-order derivatives $\mathcal{L}_{j,t,k}$ satisfy $\sup_j \|\mathcal{L}_{j,t,k}\|_{BV} \leq C_k(z)$ for some continuous function $C_k(\cdot)$ (of exponential order in $|z|$).

Proof. Since $\|e^{zf_j}\|_{BV} \leq \sum_k \frac{|z|^k \|f_j^k\|_{BV}}{k!}$ and by (V5), $\|f_j^k\|_{BV} \leq C^k \|f_j\|_{BV}^k$ for some constant C, we get $\sup_j \|e^{zf_j}\|_{BV} \leq e^{c|z|}$ for some constant $c > 0$.

The analyticity of the operators follows from the analytiticty of the map $z \to e^{zf_j}$, and the estimate on the norms of the derivatives is obtained similarly to part (i). □

The following result is a consequence of the more general perturbation theorem (Theorem D.2 in Appendix D).

Theorem 7.3. There is a constant $r_0 > 0$ such that for every complex number with $|z| \leq r_0$ there are sequences $\lambda(z) = (\lambda_j(z)), \lambda_j(z) \in \mathbb{C} \setminus \{0\}$, $h^{(z)} = (h_j^{(z)}), h_j^{(z)} \in B_j$, $m^{(z)} = (m_j^{(z)}), m_j^{(z)} \in B_j^*$ with the following properties.

(i) The maps $z \to \lambda_j(z)$, $z \to h_j^{(z)}$ and $z \to m_j^{(z)}$ are analytic in z and their norms[3] are bounded uniformly in j.

(ii) We have $\lambda_j(0) = 1$, $m_j^{(0)} = m_j$, $h_j^{(0)} = h_j$, $m_j^{(z)}(\mathbf{1}) = m_j^{(z)}(h_j^{(z)}) = 1$ and

$$\mathcal{L}_{j,z} h_j^{(z)} = \lambda_j(z) h_{j+1}^{(z)}, \quad (\mathcal{L}_{j,z})^* m_{j+1}^{(z)} = \lambda_j(z) m_j^{(z)}.$$

(iii) There are $\delta_3 \in (0,1)$ and $C_3 > 0$ such that

$$\left\| \mathcal{L}_{j,z}^n - \lambda_{j,n}(z) m_j^{(z)}(\cdot) h_{j+n}^{(z)} \right\| \leq C_3 \delta_3^n$$

where $\mathcal{L}_{j,z}^n = \mathcal{L}_{j+n-1,z} \circ \cdots \circ \mathcal{L}_{j+1,z} \circ \mathcal{L}_{j,z}$ and $\lambda_{j,n}(z) = \lambda_{j+n-1}(z) \cdots \lambda_{j+1}(z) \lambda_j(z)$.

[3]Here if γ is a map from $\{|z| \leq r_0\}$ to a Banach space we take $\|\gamma\| = \sup_{|z| \leq r_0} \|\gamma(z)\|$.

7.2. **Proof of Proposition 7.1.** Let $\lambda_j(z)$ be like in Theorem 7.3, and let $\Pi_j(z)$ be an analytic branch of $\ln \lambda_j(z)$ so that $\Pi_j(0) = 1$.

Lemma 7.4. There are $\varepsilon_0 > 0$ and $A_0 > 0$ such that for every complex number z with $|z| \leq \varepsilon_0$ and all j, n with $j + n \geq J$ we have

$$\left| \ln \mathbb{E}_{m_0}[e^{z(S_{n+j}f - S_jf)}] - \sum_{k=j}^{j+n-1} \Pi_j(z) \right| \leq A_0.$$

Proof. Since $m_0 = \tilde{m}_0$ we have $\mathbb{E}_{m_0}[e^{z(S_{n+j}f - S_jf)}] = \tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,z}^n \mathbf{1})$ where

$$\tilde{\mathcal{L}}_{j,z}^n(g) = \tilde{\mathcal{L}}_j^n(e^{zS_{j,n}f}g) = \frac{\mathcal{L}_{j,z}^n(g\mathcal{L}_0^j \mathbf{1})}{\mathcal{L}_0^{j+n}\mathbf{1}}. \tag{7.1}$$

Using Theorem 7.3(iii), and the facts that $m_j(\mathcal{L}_0^j\mathbf{1})$=1, $\text{essinf}(\mathcal{L}_0^{j+n}\mathbf{1}) \geq c_0 > 0$ and $\sup_k \|\mathcal{L}_0^k\|_{BV} < \infty$ we get

$$\tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,z}^n \mathbf{1}) = \lambda_{j,n}(z)U_{j,n}(z) + O(\delta_3^n).$$

where $\delta_3 \in (0,1)$ and $U_{j,n}(z) = m_j^{(z)}(\mathcal{L}_0^j\mathbf{1})\tilde{m}_{j+n}(H_{j+n,z})$, $H_{j+n,z} = \dfrac{h_{j+n}^{(z)}}{\mathcal{L}_0^{j+n}\mathbf{1}}$. Next the function $G(z) = \tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,z}^n\mathbf{1}) - \lambda_{j,n}(z)U_{j,n}(z)$ is analytic. Since $\tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,0}^n\mathbf{1}) = \lambda_{j,n}(0)U_{j,n}(0) = 1$ we have $G(0)$=0. Applying the maximal modulus principle to the function $\frac{G(z)}{z}$ we get

$$\tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,z}^n \mathbf{1}) = \lambda_{j,n}(z)U_{j,n}(z) + zO(\delta_3^n)$$

namely, the error term is $O(z\delta_3^n)$ and not only $O(\delta_3^n)$. Next

$$\tilde{m}_{j+n}(H_{j+n,z}) = m_{j+n}(H_{j+n,z}\mathcal{L}_0^{j+n}\mathbf{1}) = m_{j+n}(h_{j+n}^{(z)}).$$

Since $\lambda_j(0) = 1$ and $\lambda_j(z)$ are uniformly bounded and analytic in z, we can write the above in the following form

$$\tilde{m}_{j+n}(\tilde{\mathcal{L}}_{j,z}^n \mathbf{1}) = \lambda_{j,n}(z)\left(U_{j,n}(z) + zO(\delta_4^n)\right), \quad \delta_4 \in (0,1).$$

The functions $U_{j,n}(z)$ are analytic, uniformly bounded in z and they take the value one at $z = 0$. Thus, we can take the logarithm of both sides to conclude that, if $|z|$ is small enough then

$$\ln \mathbb{E}_{m_0}[e^{z(S_{n+j}f - S_jf)}] = \ln \lambda_{j,n}(z) + V_{j,n}(z)$$

where $V_{j,n}(z) = \ln U_{j,n}(z) = \ln(1 + O(z) + zO(\delta_4^n)) = O(1)$. □

Using the Cauchy integral formula we derive the following result.

Corollary 7.5. Let $\tilde{\Lambda}_{j,n}(z) = \ln \mathbb{E}_{m_0}[e^{z(S_{j+n}f - S_jf)}]$. Then there exists $\varepsilon > 0$ such that for every $s \geq 1$ there is a constant $C_s > 0$ such that for every complex z with $|z| \leq \varepsilon$ and all j, n with $j + n \geq J$ we have

$$\left| \tilde{\Lambda}_{j,n}^{(s)}(z) - \sum_{k=j}^{j+n-1} \Pi_k^{(s)}(z) \right| \leq C_s \tag{7.2}$$

where $g^{(s)}$ denotes the s-th derivative of a function g.

Remark 7.6. Clearly, (7.2) also holds when $j + n < J$ since then the number of summands is uniformly bounded.

Set

$$\Pi_{j,n}(z) = \sum_{k=j}^{j+n-1} \Pi_k(z). \tag{7.3}$$

Lemma 7.7. Let B be a constant and let $k \geq 2$. Then if B is sufficiently large there are constants D and r_0 depending only on B and k so that for every $t \in [-r_0, r_0]$ and each j, n such that $B \leq \mathrm{Var}(S_{j+n}f - S_j f) \leq 2B$ we have

$$|\Pi_{j,n}^{(k)}(it)| \leq D.$$

Proof. Applying [27, Lemma 43] with $S = S_{j+n}f - S_j f$ we see that there is $r = r(B)$ such that if $t \in [-r, r]$ then

$$|\tilde{\Lambda}_{j,n}^{(k)}(it)| \leq D_k \mathbb{E}_{m_0}[|S|^k]$$

for some constant D_k which depends only on k. Using also Corollary 7.5 we derive that

$$|\Pi_{j,n}^{(k)}(it)| \leq D_k \mathbb{E}_{m_0}[|S|^k].$$

Next, by Proposition 3.3 we have $\mathbb{E}_{m_0}[|S|^k] \leq C(k, B)$. These estimates prove the lemma. □

Proof of Proposition 7.1. Fix some $k \geq 3$. Since $\sigma_n = \|S_n\|_{L^2(m_0)} \to \infty$, using the martingale coboundary representation from Lemma 6.3, given $B > 0$ large enough we can decompose $\{0, ..., n-1\}$ into a disjoint union of intervals $I_1, ..., I_{k_n}$ in $\mathbb{Z}$ so that I_j is to the left of I_{j+1} and

$$B \leq \mathrm{Var}_{m_0}(S_{I_j}) \leq 2B \tag{7.4}$$

where $S_I = \sum_{j \in I} f_j \circ T_0^j$ for every interval I. Now, by Lemma 6.3 there is a constant[4] $A > 0$ independent of B such that $\left| \|S_n f\|_{L^2} - \left(\sum_{k=J}^{n-1} \mathrm{var}_{m_0}(M_k \circ T_0^k) \right)^{1/2} \right| \leq A$ and for each $j > J$ we have $\left| \|S_{I_j}\|_{L^2} - \left(\sum_{k \in I_j} \mathrm{var}_{m_0}(M_k \circ T_0^k) \right)^{1/2} \right| \leq A.$

Hence, if we also assume that $B > (4A)^2$ then it follows that

$$C_1 \leq k_n / \sigma_n^2 \leq C_2 \tag{7.5}$$

for some constants $C_1, C_2 > 0$ which depend only on B. Next, let $\Pi_I(z) = \sum_{k \in I} \Pi_k(z)$. Then by Lemma 7.7 there are constants $r_k > 0$ and D_k such that $\sup_j \sup_{t \in [-r_k, r_k]} \left| \Pi_{I_j}^{(k)}(it) \right| \leq D_k$. Hence,

$$\sup_{t \in [-r_k, r_k]} \left| \Pi_{0,n}^{(k)}(it) \right| \leq D_k k_n \leq C_2 D_k \sigma_n^2. \tag{7.6}$$

Combining this with Corollary 7.5 and taking into account that $\sigma_n \to \infty$ we see that

$$\sup_{t \in [-r_k, r_k]} \left| \tilde{\Lambda}_n^{(k)}(it) \right| \leq \tilde{D}_k \sigma_n^2$$

for some constant $\tilde{D}_k$, and the proof of the proposition is complete. □

[4] We can take $A = 2 \sup_j \|u_j\|_{L^\infty}$, where u_j are the coboundaries from (6.2) .

7.3. **Proof of Theorems 3.1 and 3.2.** Let $\Lambda_n(t) = m_0(e^{itS_n/\sigma_n})$. Applying Theorems 5 and 9, and Corollary 11 from [52] to $W_n = S_n/\sigma_n$, we see that in order to prove Theorems 3.1 and 3.2 it is enough to show that for every $u \geq 3$ there are constants $\delta_u, C_u > 0$ such that the function $\Lambda_n(t)$ is well defined, differentiable u times on $[-\delta_u\sigma_n, \delta_u\sigma_n]$ and $\sup_{|t|\leq\sigma_n\delta_u} |\Lambda_n^{(u)}(t)| \leq C_u\sigma_n^{-(u-2)}$. However, this is exactly Proposition 7.1.

Appendix A. Proof of Theorem 2.4

A.1. **Proof of Theorem 2.4(i).** Let $\varepsilon_0 > 0$ be such that $\eta := \rho + \varepsilon_0 < 1$ and let $a > 0$ be a constant such that $a\varepsilon_0 > K_0$, where ρ and K_0 come from (2.3). Then it follows from (2.3) that for every $n \geq N$ and all $j \geq 0$ we have

$$\mathcal{L}_j^n \mathcal{C}_{j,a} \subset \mathcal{C}_{j+n,a\rho^{1+N-n}+K_0} \subset \mathcal{C}_{j+n,\eta a}.$$

Given $C > 0$ consider the cone

$$\mathcal{K}_j(C) = \{g \in B_j, \ \ g > 0, \quad \text{esssup } g \leq C \text{ essinf } g\}$$

where the essential supremum and infimum are with respect to m_j. Denote by d_j the projective Hilbert metric associated with the cone $\mathcal{C}_{j,a}$ (see [16, 66]). Then by [19, (1.1)] the projective diameter of $\mathcal{K}_j(C) \cap \mathcal{C}_{j,\eta a}$ inside $\mathcal{C}_{j,a}$ does not exceed a constant $R = R(a, \eta, V, C)$ which depends only on a, η, C and $V := \sup_j v_j(\mathbf{1})$. Next, combining (V3), (SC), (2.3) and the relation $(\mathcal{L}_j^n)^* m_{j+n} = m_j$, we see that there are constants $M \geq N$ and C (depending on a but not on j) such that for every j and $n \geq M$ we have

$$\mathcal{L}_j^n \mathcal{C}_{j,a} \subset \mathcal{K}_{j+n}(C).$$

We conclude that for every j and $n \geq M$ the projective diameter of $\mathcal{L}_j^n \mathcal{C}_{j,a}$ inside $\mathcal{C}_{j+n,a}$ does not exceed R.

The next step of the proof is the following result.

Lemma A.1. (a) There is a constant $A > 0$ such that for every a, j and $h \in \mathcal{C}_{j,a}$,

$$\|h\| \leq A m_j(h)$$

(b) If $a > V$ then there is a constant $r_0 > 0$ such that every real valued $g \in B_j$ can be written in the form $g = g_1 - g_2$ where $g_i \in \mathcal{C}_{j,a}$ and $\|g_1\| + \|g_2\| \leq r_0\|g\|$.

Proof. To prove (a) we note that if $h \in \mathcal{C}_{j,a}$ then $\|h\| = m_j(h) + v_j(h) \leq (1+a)m_j(h)$ so we can take $A = 1 + a$.

In order to prove (b), given $g \in B_j$ we take $C_0 = C_0(g) = \|g\|_\infty + \frac{1+a}{a-V}\|g\|_{BV}$. It is immediate to check that $g + C_0 \in \mathcal{C}_{j,a}$. Since $C_0 \leq C\|g\|_{BV}$ for some constant C which does not depend on j we have $g = (g + C_0) - C_0$ and, with $\|\cdot\| = \|\cdot\|_{BV}$,

$$\|g + C_0\| + \|C_0\| \leq \|g\| + 2\|C_0\| \leq \|g\|(1 + 2C_0 + 2C_0V) \leq r\|g\|$$

for some constant $r = r(a, V, C)$. Thus we can take $g_1 = (g + C_0)$ and $g_2 = C_0$ (the constant function). Note that $g_2 \in \mathcal{C}_{j,a}$ since $a > V$. $\square$

Based on Lemma A.1, the proof of proof of Theorem 2.4(ii) is completed like in [46, Chapters 4 and 5]. In fact, here the situation is much simpler, and for readers' convenience we provide the details. First, by taking $f = 1$ in (2.2) and letting g vary we see that $(\mathcal{L}_j^n)^* m_{j+n} = m_j$ for all j and n. Next, fix some a large enough and take two sequences of functions $g_n, f_n \in \mathcal{C}_{n,a}$. Denote $G_{j,n} = \mathcal{L}_j^n g_j$ and $F_{j,n} = \mathcal{L}_j^n f_j$. Then, since the projective diameter of the image of $\mathcal{L}_k^s \mathcal{C}_{a,k}$ inside

$\mathcal{C}_{a,k+s}$ does not exceed R if $s \geq M$, we see by [66, Theorem 1.1] that for every j and $n \geq M$ we have

$$\mathsf{d}_j(F_{j,n}, G_{j,n}) \leq (c(R))^{[(n-1)/M]}$$

where $c(R) = \tanh(R/4) \in (0,1)$.

For a positive function g, define the normalized iterates by

$$\bar{\mathcal{L}}_j^n g = \frac{\mathcal{L}_j^n g}{m_{j+n}(\mathcal{L}_j^n g)} = \frac{\mathcal{L}_j^n g}{m_j(g)}.$$

Then by [46, Theorem A.2.3] and Lemma A.1(1) we have

$$\left\|\bar{\mathcal{L}}_j^n f_j - \bar{\mathcal{L}}_j^n g_j\right\|_{BV} \leq \frac{1}{2} A(c(R))^{[(n-1)/M]}. \tag{A.1}$$

Next, fix some function $h_0 \in \mathcal{C}_{a,0}$ such that $m_0(h_0) = 1$ and for $j > 0$ define $h_j = \mathcal{L}_0^j h_j$. Then $h_{j+1} = \mathcal{L}_j h_j$ and $m_j(h_j) = 1$ for all $j \geq 0$. Therefore,

$$\bar{\mathcal{L}}_j^n h_j = h_{j+n} \tag{A.2}$$

for all $j \geq 0$ and $n > 0$. Moreover, $h_j \in \mathcal{C}_{a,j}$ (due to the equivariance of the cones). By taking $g_j = h_j$ in (A.1) and using (A.2) we see that

$$\left\|\bar{\mathcal{L}}_j^n f_j - h_{j+n}\right\|_{BV} \leq \frac{1}{2} A(c(R))^{[(n-1)/M]}.$$

Multiplying by $m_j(f_j)$ we get that, with $\delta = (c(R))^{1/M}$, for all $f \in \mathcal{C}_{a,j}$ we have

$$\|\mathcal{L}_j^n(f) - m_j(f) h_{j+n}\| \leq m_j(f) B \delta^n$$

where B is some constant. Using Lemma A.1(b) we obtain (2.4) with $C = 2r_0 B$. $\square$

A.2. **Proof of Theorem 2.4(ii).** To see why $(T_j)_*\mu_j = \mu_{j+1}$, take a measurable function $g : X_j \to \mathbb{R}$ and write

$$\left[(T_j)_*\mu_j\right](g) = \mu_j(g \circ T_j) = m_j(h_j g \circ T_j) = m_{j+1}((\mathcal{L}_j h_j) \cdot g) = m_{j+1}(h_{j+1} g) = \mu_{j+1}(g).$$

To prove asymptotic uniqueness of μ_j, let $\tilde{\mu}_j = g_j dm_j$ be another sequence of probability measures such that $(T_j)_*\tilde{\mu}_j = \tilde{\mu}_{j+1}$. We show that $\lim_{j\to\infty} \|h_j - g_j\|_1 = 0$. We claim that

$$\mathcal{L}_j g_j = g_{j+1}, \quad m_{j+1} - \text{a.e.} \tag{A.3}$$

Indeed given a function $f : X_{j+1} \to \mathbb{R}$ we have

$$m_{j+1}((\mathcal{L}_j g_j) f) = m_j(g_j(f \circ T_j)) = \tilde{\mu}_j(f \circ T_j) = \tilde{\mu}_{j+1}(f) = m_{j+1}(g_{j+1} f)$$

proving (A.3). Iterating (A.3) we see that $g_n = \mathcal{L}_0^n g_0$, m_n-a.e. Next, let $\varepsilon > 0$ and let $q_0 : X_0 \to \mathbb{R}$ be a BV function such that $\|q_0 - g_0\|_1 < \varepsilon$. Normalizing q_0 if needed we can always assume that $m_0(q_0) = 1$. By Theorem 2.4(i) we see that

$$\|\mathcal{L}_0^n q_0 - h_n\|_{BV} \leq C \|q_0\|_{BV} \delta^n \to 0, \text{ as } n \to \infty.$$

On the other hand,

$$\|\mathcal{L}_0^n q_0 - g_n\|_1 = \|\mathcal{L}_0^n q_0 - \mathcal{L}_0^n g_0\|_1 \leq \|\mathcal{L}_0^n(|q_0 - g_0|)\|_1 = \|g_0 - d_0\| < \varepsilon.$$

Therefore, by taking n large enough we get that $\|h_n - g_n\|_1 < \varepsilon$, and the proof of Theorem 2.4(ii) is complete. $\square$

Appendix B. Gibbs measures for SFT.

In this section we denote by $T_j : X_j \to X_{j+1}$, a one sided topologically mixing sequential subshift of finite type and by $\tilde{T}_j : \tilde{X}_j \to \tilde{X}_{j+1}$, $j \in \mathbb{Z}$ the corresponding two sided shift. For a point $(x_j, x_{j+1}, ...) \in X_j$ and $m \in \mathbb{N}$ let

$$[x_j, x_{j+1}, ..., x_{j+m-1}] = \left\{(x'_{j+k})_{k\geq 0} \in X_j : x'_{j+k} = x_{j+k}, \forall k < m\right\}$$

be the corresponding cylinder of length m. Cylinders in $\tilde{X}_j$ are denoted similarly.

Remark B.1. The reason we consider here negative indexes j is to provide a complete theory which includes uniqueness of Gibbs measures (see Section B.2) and relations with two sided shift via a non-stationary version of Sinai's lemma (see Lemma B.2). If we begin with a one sided shift then there are many Gibbs measures corresponding to a given sequence of potentials, each of which corresponds to an extension of the shift to negative indexes and an extension of the sequence of potentials to a two sided sequence. Considering negative indexes j also allows us to uniquely define two sided shifts which have applications to non-autonomous hyperbolic systems (see Appendix C).

B.1. **A sequential Sinai's Lemma.**

Let $\pi_j : \tilde{X}_j \to X_j$ be given by

$$\pi_j((x_{j+k})_{k\in\mathbb{Z}}) = (x_{j+k})_{k\geq 0}.$$

Lemma B.2. Fix some $\alpha \in (0, 1]$ and let $\psi_j : \tilde{X}_j \to \mathbb{R}$ be uniformly Hölder continuous with exponent α. Then there are uniformly Hölder continuous functions $u_j : \tilde{X}_j \to \mathbb{R}$ with exponent $\alpha/2$ and $\phi_j : X_j \to \mathbb{R}$ such that

$$\psi_j = u_j - u_{j+1} \circ \tilde{T}_j + \phi_j \circ \pi_j.$$

Moreover if $\|\psi_j\|_\alpha \to 0$ then $\|u_j\|_{\alpha/2} \to 0$.

Proof. The proof is a modification of the proof of [15, Lemma 1.6]. For each j and t take a point $a^{(j,t)} = (a^{(j,t)}_{j+k})_k \in \tilde{X}_j$ such that $a^{(j,t)}_j = t$. For $y = (y_{j+k})_k \in \tilde{X}_j$ define

$$r_j(y) = (y^*_{j+k})_{k\in\mathbb{Z}} \in \tilde{X}_j$$

where $y^*_{j+k} = y_{j+k}$ if $k \geq 0$ and $y^*_{j+k} = a^{(j,y_j)}_{j+k}$ if $k \leq 0$. Let

$$u_j(y) = \sum_{k=0}^{\infty} \Big(\psi_{j+k}(\tilde{T}^k_j y) - \psi_{j+k}(\tilde{T}^k_j r_j(y))\Big). \tag{B.1}$$

To see that the RHS of (B.1) converges note that since y and $r_j(y)$ have the same values at the coordinates indexed by $j+k$ for $k\geq 0$ (i.e. with indexes to the right of j) we have

$$|\psi_{j+k}(\tilde{T}^k_j y) - \psi_{j+k}(\tilde{T}^k_j r_j(y))| \leq v_\alpha(\psi_{j+k})2^{-\alpha k} \tag{B.2}$$

and so $\sup|u_j| \leq \sum_{k=0}^{\infty} v_\alpha(\psi_{j+k})2^{-\alpha k}$. Thus, $\sup_j \sup_{\tilde{X}_j} |u_j| < \infty$. Moreover, if $v_\alpha(\psi_j) \to 0$ then $\sup_{\tilde{X}_j} |u_j| \to 0$.

Next, we claim that the function $\psi_j - u_j + u_{j+1} \circ \tilde{T}_j$ depends only on the coordinates indexes by $j + k$ for $k \geq 0$ (so it has the form $\phi_j \circ \pi_j$). Indeed,

$$u_j - u_{j+1} \circ \tilde{T}_j = \psi_j + \sum_{k=0}^{\infty} \Big(\tilde{T}_{j+k} \circ \tilde{T}^k_j \circ r_j - \psi_{j+k+1} \circ \tilde{T}^k_{j+1} \circ r_{j+1} \circ \tilde{T}_j\Big).$$

Note the the sum on the above right hand side depends only on the coordinates indexed by $j+k$ for $k\geq 0$, and the claim follows.

In view of the above, in order to complete the proof of the lemma it is enough to obtain appropriate bounds on the Hölder constants of the functions u_j (corresponding to the exponent $\alpha/2$). For that end, let y and y' in $\tilde X_j$ be such that $y_{j+k}=y'_{j+k}$ for every $|k|\leq n$ for some $n>0$. Using (B.2) we have

$$|u_j(y)-u_j(y')|\leq \sum_{k=0}^{[n/2]}\left|\psi_{j+k}(\tilde T_j^k y)-\psi_{j+k}(\tilde T_j^k y')\right|+\sum_{k=0}^{[n/2]}\left|\psi_{j+k}(\tilde T_j^k r_j(y))-\psi_{j+k}(\tilde T_j^k r_j(y'))\right|$$

$$+2\sum_{k>[n/2]}v_\alpha(\psi_{j+k})2^{-\alpha k}:=I_1+I_2+I_3.$$

To show that u_j is Hölder continuous with exponent $\alpha/2$ (uniformly in j) we use that

$$\left|\psi_{j+k}(\tilde T_j^k y)-\psi_{j+k}(\tilde T_j^k y')\right|\leq \sup_s v_\alpha(\psi_s)2^{-(n-k)\alpha}$$

and similarly with $r_j(y)$ and $r_j(y')$ instead of y and y', respectively. So $I_1+I_2\leq C2^{-n\alpha/2}$. Moreover, we note that $I_3\leq C2^{-n\alpha/2}$.

Next, if $\|\psi_j\|_\alpha\to 0$ then $I_1+I_2\leq \sum_{k=0}^{[n/2]}v_\alpha(\psi_{j+k})2^{-(n-k)\alpha}\leq \varepsilon_j 2^{-n\alpha/2}$ with $\varepsilon_j\to 0$. Similarly, $I_3\leq \varepsilon_j 2^{n\alpha/2}$ for ε_j with the same properties. □

B.2. **Sequential Gibbs measures for one sided shifts.**

Definition B.3. Let $\phi_j:X_j\to\mathbb{R}$ be a sequences of functions such that $\sup_j\|\phi_j\|_\alpha<\infty$ for some $\alpha\in(0,1]$. We say that a sequence of probability measures μ_j on X_j is a *sequential Gibbs family for* (ϕ_j) if:

(i) For all j we have $(T_j)_*\mu_j=\mu_{j+1}$

(ii) There is a constant $C>1$ and a sequence of positive numbers (λ_j) such that for all j, every point $(x_{j+k})_k$ in X_j and every $r>0$ we have

$$C^{-1}e^{S_{j,r}\phi(x)}/\lambda_{j,r}\leq \gamma_j([x_j,...,x_{j+r-1}])\leq Ce^{S_{j,r}\phi(x)}/\lambda_{j,r}$$

where $S_{j,r}\phi(x)=\sum_{s=0}^{r-1}\phi_{j+s}(T_j^s x)$ and $\lambda_{j,r}=\prod_{k=j}^{j+r-1}\lambda_k$.

We say that two sequences (α_j) and (β_j) of positive numbers are *equivalent* if there is a sequence (ζ_j) of positive numbers which is bounded and bounded away from 0 such that for all j we have $\alpha_j=\zeta_j\beta_j/\zeta_{j+1}$.

We need the following simple result.

Lemma B.4. Two positive sequences (α_j) and (β_j) are equivalent if and only if there is a constant $C>0$ such that for all j and n we have $C\leq \alpha_{j,n}/\beta_{j,n}\leq C^{-1}$.

Proof. It is clear that there exists such a constant C if (α_j) and (β_j) are equivalent. On other hand, suppose that such a constant exists. Then, for $j\geq 0$ define $\zeta_j=\beta_{0,j}/\alpha_{0,j}$. Clearly, $\alpha_j/\beta_j=\zeta_j/\zeta_{j+1}$. For $j<0$ we define similarly $\zeta_j=\alpha_{j+1,|j|}/\beta_{j+1,|j|}$. □

Theorem B.5. For every sequence of functions $\phi_j : X_j \to \mathbb{R}$ such that $\sup_j \|\phi_j\|_\alpha < \infty$ for some $\alpha \in (0, 1]$ there exists unique Gibbs measures μ_j. Moreover the sequence (λ_j) is unique up to equivalence.

We will see below that the measures μ_j can be expressed in terms of the strong limits of the normalized operators $\mathcal{L}_j^n$ associated with the sequence (ϕ_j) defined in (B.4). The rate of convergence is exponential, see (B.5) below.

Proof. Existence of μ_j was proven in [48]. Let us recall the main arguments. First, we define the operator L_j which maps a function $g : X_j \to \mathbb{R}$ to a function $L_j g : X_{j+1} \to \mathbb{R}$ given by

$$L_j g(x) = \sum_{T_j y = x} e^{\phi_j(y)} g(y).$$

Then, in [48] it was shown that there is a positive sequence (λ_j) which is bounded and bounded away from the origin, a sequence of positive functions h_j with uniformly bounded Hölder norms (corresponding to the same exponent α of ϕ_j), which are also uniformly bounded below by a positive constant, and a sequence of probability measures ν_j on X_j such that $\nu_j(h_j)$=1 and for all j, n and a Hölder continuous function g on X_j,

$$\|(\lambda_{j,n})^{-1} L_j^n g - \nu_j(g) h_{j+n}\|_\alpha \leq C_0 \|g\|_\alpha \delta^n \tag{B.3}$$

for some constants $C > 0$ and $\delta \in (0, 1)$ which do not depend on j, n and g. Here

$$L_j^n = L_{j+n-1} \circ \cdots \circ L_{j+1} \circ L_j.$$

Then the Gibbs measures constructed in [48] are given by $\mu_j = h_j d\nu_j$. Moreover, the operator $\mathcal{L}_j$ given by

$$\mathcal{L}_j g(x) = \frac{L_j(g h_j)}{\lambda_j h_{j+1}} \tag{B.4}$$

is the dual of the Koopman operator corresponding to T_j with respect to μ_j and μ_{j+1} and (because of (B.3)),

$$\|\mathcal{L}_j^n g - \mu_j(g)\mathbf{1}\|_\alpha \leq C \|g\|_\alpha \delta^n \tag{B.5}$$

for some $C > 0$. In the derivation of (B.5) we used that $\sup_j \|1/h_j\|_\alpha < \infty$.

Now, let us suppose that there is another Gibbs measure $\tilde{\mu}_j$ associated with a sequence $\tilde{\lambda}_j$. Namely, $(T_j)_* \tilde{\mu}_j = \tilde{\mu}_{j+1}$ and for every point $x \in X_j$,

$$(\tilde{C})^{-1} e^{S_{j,r}\phi(x)} / \tilde{\lambda}_{j,r} \leq \tilde{\mu}_j([x_j, x_{j+1}, ..., x_{j+r-1}]) \leq \tilde{C} e^{S_{j,r}\phi(x)} / \tilde{\lambda}_{j,r}$$

for some constant $\tilde{C} > 0$. Let us first show that the sequences (λ_j) and $(\tilde{\lambda}_j)$ are equivalent. Indeed, for all j and n, by the Gibbs property of both μ_j and $\tilde{\mu}_j$,

$$C^{-1} \sum_{[y_j,...,y_{j+n-1}]} e^{S_{j,n}\phi(x_y)} / \lambda_{j,n} \leq \sum_{[y_j,...,y_{j+n-1}]} \mu_j([y_j, ..., y_{j+n-1}]) \leq C \sum_{[y_j,...,y_{j+n-1}]} e^{S_{j,n}\phi(x_y)} / \lambda_{j,n}$$

and

$$(\tilde{C})^{-1} \sum_{[y_j,...,y_{j+n-1}]} e^{S_{j,n}\phi(x_y)} / \tilde{\lambda}_{j,n} \leq \sum_{[y_j,...,y_{j+n-1}]} \tilde{\mu}_j([y_j, ..., y_{j+n-1}]) \leq \tilde{C} \sum_{[y_j,...,y_{j+n-1}]} e^{S_{j,n}\phi(x_y)} / \tilde{\lambda}_{j,n}$$

where the point x_y is an arbitrary point inside the cylinder $[y_j, ..., y_{j+n-1}]$. On the other hand,

$$1 = \sum_{[y_j,...,y_{j+n-1}]} \mu_j([y_j, ..., y_{j+n-1}]) = \sum_{[y_j,...,y_{j+n-1}]} \tilde{\mu}_j([y_j, ..., y_{j+n-1}]).$$

Hence,

$$(C\tilde{C})^{-1} \le \frac{\lambda_{j,n}}{\tilde{\lambda}_{j,n}} \le C\tilde{C}$$

and the equivalence of the sequences (λ_j) and $(\tilde{\lambda}_j)$ follows from Lemma B.4.

Next, we show that $\tilde{\mu}_j = \mu_j$ for all j. We first need the following result. Let $H_{j,\alpha}$ denote the space of Hölder continuous functions on X_j with the Hölder exponent α, equipped with the usual α-Hölder norm $\|\cdot\|_\alpha$. Let $H^*_{j,\alpha}$ denote its dual, and denote by $\|\cdot\|_\alpha$ the (operator) norm on the dual, as well. Let $\mathcal{L}_j^* : H^*_{j+1,\alpha} \to H^*_{j,\alpha}$ be the dual operators of $\mathcal{L}_j$. Then it follows from (B.5) that for every j,n and $\kappa_{j+n} \in H^*_{j+n,\alpha}$,

$$\|(\mathcal{L}_j^n)^*\kappa_{j+n} - \kappa_{j+n}(\mathbf{1}_{j+n})\mu_j\|_\alpha \le C\|\kappa_{j+n}\|_\alpha \delta^n \tag{B.6}$$

where $\mathbf{1}_k$ is the function on X_k which takes constant value 1. In particular, if we take $\kappa_{j+n} = \tilde{\mu}_{j+n}$ then, since $\tilde{\mu}_{j+n}\mathbf{1}_{j+n} = 1$, we see that

$$(\mathcal{L}_j^n)^*\tilde{\mu}_{j+n} \to \mu_j \text{ as } n \to \infty. \tag{B.7}$$

Next, we claim that

$$(\mathcal{L}_j^n)^*\tilde{\mu}_{j+n} = \tilde{\mu}_j. \tag{B.8}$$

Once the claim is proven, combining it with (B.7) we obtain $\tilde{\mu}_j = \mu_j$, as required.

Now, because of the Gibbs properties of both μ_j and $\tilde{\mu}_j$ and since (λ_j) and $(\tilde{\lambda}_j)$ are equivalent we see that $\tilde{\mu}_j \ll \mu_j$. Let $p_j = \frac{d\tilde{\mu}_j}{d\mu_j}$ denote the corresponding Radon Nikodym derivative. In order to prove (B.8), we will show that for every j,

$$p_{j+1} \circ T_j = p_j, \quad \mu_j - \text{a.s.} \tag{B.9}$$

(B.9) implies that $p_{j+n} \circ T_j^n = p_j$. Therefore, for every bounded measurable function $g : X_j \to \mathbb{R}$,

$$(\mathcal{L}_j^n)^*\tilde{\mu}_{j+n}(g) = \tilde{\mu}_{j+n}(\mathcal{L}_j^n g) = \mu_{j+n}(p_{j+n}\mathcal{L}_j^n g) = \mu_j(g \cdot (p_{j+n} \circ T_j^n)) = \mu_j(g \cdot p_j) = \tilde{\mu}_j(g)$$

and (B.8) follows (note that in the third equality above we have used that $\mathcal{L}_j^n$ is the transfer operator of T_j^n with respect to μ_j and μ_{j+n}).

To complete the proof of the theorem we need to prove (B.9). By identifying both measures μ_j and $\tilde{\mu}_j$ as measure on the two sided shift $\tilde{X}_j$ (see next section), we can assume that μ_j and $\tilde{\mu}_j$ are mapped by $\tilde{T}_j$ to μ_{j+1} and $\tilde{\mu}_{j+1}$, respectively. In what follows we will abuse the notation and denote the lifted measures by μ_j and $\tilde{\mu}_j$. The identification of the function p_j to $\tilde{X}_j$ will also be denoted by p_j, namely, we write $p_j(..., x_{j-1}, x_j, x_{j+1}, ...) = p_j(x_j, x_{j+1}, ...)$. Now, since $(\tilde{T}_j)_*\mu_j = \mu_{j+1}$ and $(\tilde{T}_j)_*\tilde{\mu}_j = \tilde{\mu}_{j+1}$ for every bounded measurable function $f : \tilde{X}_{j+1} \to \mathbb{R}$ we have

$$\mu_j(p_j(f \circ \tilde{T}_j)) = \tilde{\mu}_j(f \circ \tilde{T}_j) = \tilde{\mu}_{j+1}(f) = \mu_{j+1}(p_{j+1}f) = \mu_j((p_{j+1} \circ \tilde{T}_j) \cdot (f \circ \tilde{T}_j)).$$

Since T_j is invertible we can replace $f \circ T_j$ with a general bounded measurable function g on X_j and (B.9) follows. □

B.3. Gibbs measure for two sided shifts.

Definition B.6. Let $\psi_j : \tilde{X}_j \to \mathbb{R}$ be a sequences of functions such that $\sup_j \|\psi_j\|_\alpha < \infty$ for some $\alpha \in (0,1]$. A sequence of probability measures γ_j on $\tilde{X}_j$ is a Gibbs family for (ψ_j) if:

(i) For all j we have $(\tilde{T}_j)_*\gamma_j = \gamma_{j+1}$.

(ii) There is a constant $C>1$ and a sequence of positive numbers (λ_j) so that for every point $(y_{j+k})_k$ in $\tilde{X}_j$ we have

$$C^{-1}e^{S_{j,r}\psi(y)}\lambda_{j,r}\leq\gamma_j([y_j,y_{j+1},...,y_{j+r-1}]])\leq Ce^{S_{j,r}\psi(y)}\lambda_{j,r}$$

where $S_{j,r}\psi(y)=\sum_{s=0}^{r-1}\phi_{j+s}(\tilde{T}_j^s y)$ and $\lambda_{j,r}=\prod_{k=j}^{j+r-1}\lambda_k$.

Proposition B.7. Every uniformly Hölder continuous sequence ψ_j (with respect to some exponent $\alpha\in(0,1]$) admits a unique sequence of Gibbs measures γ_j.

Proof. Let ϕ_j be the functions obtained in Lemma B.2. Let μ_j be the unique sequential Gibbs measure on the one sided shift space X_j corresponding to the functions (ϕ_j). Since $(T_{j-k})_*\mu_{j-k}=\mu_j$ using Kolmogorov's extension theorem we can extend μ_j to the space of two sided sequences $\tilde{X}_j$. Let us denote this measure by γ_j. It is clear that $(\tilde{T}_j)_*\gamma_j=\gamma_{j+1}$. To prove the second condition in the definition of a sequential Gibbs measure we notice that for every point $x\in\tilde{X}_j$,

$$\gamma_j([x_j,...,x_{j+r-1}])=\mu_j([x_j,...,x_{j+r-1}])\asymp\lambda_{j,n}e^{S_{j,r}\phi\circ\pi_j(x)}$$

where (λ_j) is the sequence associated with the Gibbs measure μ_j. Now, by Lemma B.2 we have that

$$e^{S_{j,r}\phi\circ\pi_j(x)}\asymp e^{S_{j,r}\psi(x)}$$

and hence γ_j is a sequential Gibbs measure corresponding to the functions ψ_j, and (λ_j) is the corresponding associated sequence.

To prove uniqueness of γ_j constructed above, we note that if $\tilde{\gamma}_j$ is another Gibbs measure then the measure $\tilde{\mu}_j$ defined by the restriction of γ_j to the σ-algebra generated by the coordinates with indexes $j+k, k\geq 0$ projects[5] to a Gibbs measure on X_j. Hence $\tilde{\mu}_j=\mu_j$, where μ_j is the unique sequential Gibbs measure corresponding to the functions ϕ_j. On the other hand, since $(\tilde{T}_j)_*\tilde{\gamma}_j=\tilde{\gamma}_{j+1}$, for every $r>0$ the restriction of $\tilde{\gamma}_j$ to the σ-algebra generated by the coordinates indexed by $j+k$ for $k\geq -r$ coincides with the restriction of $\tilde{\gamma}_{j-r}$ to the σ-algebra generated by the coordinates indexed by $j-r+k$ for $k\geq 0$, which, as explained above, projects on X_{j-r} to μ_{j-r}. Hence, for every r the measures γ_j and $\tilde{\gamma}_j$ agree on the σ-algebra generated by the coordinates indexed by $j+k$ for $k\geq -r$ (as both coincide with μ_{j-r}). By taking $r\to-\infty$ we conclude that $\tilde{\gamma}_j=\gamma_j$. ☐

Appendix C. Small perturbations of hyperbolic maps

C.1. **Hyperbolic sets.** Let M be a compact C^2 Riemannian manifold equipped with its Borel σ-algebra $\mathcal{B}$. Denote by $\mathsf{d}(\cdot,\cdot)$ the induced metric. Let $T:M\to M$ be a C^2 diffeomorphism.

Definition C.1. A compact T invariant subset $\Lambda\subset M$ is called a *hyperbolic set* for T if there exists an open set V with compact closure, constants $\lambda\in(0,1)$ and $\alpha_0,A_0,B_0>0$ and subbudnles Γ^s and Γ^u of the tangent bundle $T\Lambda$ such that:

(i) The set $\{x\in M:\text{dist}(x,\Lambda)<\alpha_0\}$ is contained in a open subset U of V such that $TU\subset V$ and $T|_U$ is a diffeomorphism with $\sup_{x\in U}\max(\|D_xT\|,\|D_xT^{-1}\|)\leq A_0$;

(ii) $T\Lambda=\Gamma^s\oplus\Gamma^u$, $DT(\Gamma^s)=\Gamma^s$, $DT(\Gamma^u)=\Gamma^u$ and the minimal angle between Γ^s and Γ^u is bounded below by α;

(iii) For all $n\in\mathbb{N}$ we have

$$\|D_xT^nv\|\leq B_0\lambda^n\|v\|\ \ \forall\, v\in\Gamma^s_x\quad\text{and}\quad\|D_xT^{-n}v\|\leq B_0\lambda^n\|v\|\ \ \forall\, v\in\Gamma^u_x.$$

[5]By "projects" we mean that if the restriction is denoted by γ_j^* then $(\pi_j)_*\gamma_j^*=\mu_j$.

Definition C.2. A hyperbolic set is called

(i) *locally maximal* if the set U above could be chosen so that $\Lambda = \bigcap_{n\in\mathbb{Z}} T^n U$ (that is, Λ is the largest hyperbolic set contained in U);

(ii) hyperbolic attractor, if in addition, U could be chosen so that $TU \subset U$ (in the case when $M = \Lambda$ T is called *Anosov*).

We say that Λ is the *basic hyperbolic set* if it is infinite locally maximal hyperbolic set such that $T|_\Lambda$ is topologically transitive.

Henceforth, we assume that Λ is topologically mixing[6] basic hyperbolic set.

A powerful tool for studying hyperbolic maps is given by a symbolic representations. Namely, every topologically mixing basic set Λ admits a Markov partition (see [79, Chapter 10]) which gives raise to a semiconjugacy $\pi : \Sigma \to M$ where Σ is a topologically mixing subshift of a finite type.

C.2. **Structural stability.** Now, consider a sequence of maps $\mathcal{T} = (T_j : M \to M)_{j\in\mathbb{Z}}$. Denote by $\mathsf{d}_1(f, g)$ the C^1 distance between f and g. We have the following result.

Theorem C.3. If $\delta_1(\mathcal{T}) := \sup_j \mathsf{d}_1(T, T_j)$ is small enough then there is a sequence of sets $\Lambda_j \subset M$ and homeomorphisms $h_j : \Lambda \to \Lambda_j$ (that we think of as a "sequential conjugacy") such that h_j and h_j^{-1} are uniformly Hölder continuous,

$$T_j\Lambda_j = \Lambda_{j+1} \text{ and } T_j \circ h_j = h_{j+1} \circ T. \tag{C.1}$$

Moreover $\sup_j \|h_j - \mathrm{Id}\|_{C^0} \to 0$ as $\delta_1(\mathcal{T}) \to 0$. [7]

The sets (Λ_j) are sequentially hyperbolic for the sequence $\mathcal{T}$ in the following sense. They are compact, satisfy $T_j\Lambda_j = \Lambda_{j+1}$ and there exist constants $\lambda' \in (0, 1)$ and $\alpha_1, A_1, B_1 > 0$ and sequences of subbudnles $\Gamma_j^s=\{\Gamma_{j,x}^s : x \in \Lambda_j\}$ and $\Gamma_j^u=\{\Gamma_{j,x}^u : x \in \Lambda_j\}$ of the tangent bundle $T\Lambda_j$ such that, for each j:

(i) The set $\{x \in M : \mathsf{d}(x, \Lambda_j) < \alpha_1\}$ is contained in an open subset U_j of V such that $T_jU_j \subset V$ and $T_j|_{U_j}$ is a diffeomorphism satisfying

$$\sup_j \sup_{x\in U_j} \max(\|D_xT_j\|, \|D_xT_j^{-1}\|) \le A_1;$$

(ii) $T\Lambda_j = \Gamma_j^s \oplus \Gamma_j^u$, $DT_j(\Gamma_j^s) = \Gamma_{j+1}^s$, $DT_j(\Gamma_j^u) = \Gamma_{j+1}^u$ and the minimal angle between Γ_j^s and Γ_j^u is bounded below by α_1;

(iii) For every $n \in \mathbb{N}$ and all j we have $\|D_xT_j^n v\| \le B_1 \cdot (\lambda')^n\|v\|$ for every $v \in \Gamma_{j,x}^s$ and $\|D_xT_j^{-n}v\| \le B_1 \cdot (\lambda')^n\|v\|$ for every $v \in \Gamma_{j,x}^u$ where $T_j^{-n} = (T_{j-n}^n)^{-1}$.

(iv) $T_jU_j \subset U_{j+1}$ and $\bigcap_{n=0}^{\infty} T_{j-n}^n U_{j-n} = \Lambda_j$.

[6]Topological mixing assumption can be made without a loss of generality. Indeed (see e.g. [79, Chapter 8]) an arbitrary basic set Λ can be decomposed as $\Lambda = \bigcup_{j=1}^{p} \Lambda_j$ so that $T\Lambda_j = \Lambda_{j+1 \mod p}$ where Λ_j are topologically mixing basic hyperbolic sets for T^p. Then we could apply the results discussed below to (T^p, Λ_j).

[7]Note that in Theorem C.3 we can also consider one sided sequences $(T_j)_{j\ge0}$ because they can be extended to two sided ones. The reason we consider two sided sequences is because the definition of hyperbolicity requires considering negative times to define the unstable subspaces.

Theorem C.3 is proven in [62, Example 4.2.3], [40, Example 2.5] and [64, Theorem 1.1]) except the Hölder continuity of h_j and h_j^{-1} was not discussed there. The proof of Hölder continuity is quite standard but for completeness it is included in §C.4. We note that in [62, 40, 64] the smallness of $\|T_j - T\|_{C^1}$ was not uniform in j, which led to non-uniform in j hyperbolicity. However, when $\delta_1(\mathcal{T})$ is uniformly small then the arguments in the proofs yield the uniform hyperbolicity.

C.3. **Limit Theorems.** Let $\pi_j = h_j \circ \pi$. Then π_j provides a semiconjugacy between the sequence $\mathcal{T}$ and the subshift Σ describing the symbolic dynamics of T. Given a sequence of Hölder functions ϕ_j on Λ_j let $\psi_j = \phi_j \circ \pi_j$. Note that ψ_j are Hölder continuous due to Hölder continuity of π_j. Let ν_j be the Gibbs measures for $\{\psi_j\}$ which exist due to the results of Appendix B. Define measures μ_j on Λ_j by $\mu_j(A) = \nu_j(\pi_j^{-1}A)$. Given a sequence of Holder functions f_j on Λ_j let $g_j = f_j \circ \pi_j$. Thus $S_n f = S_n g \circ \pi_0^{-1}$ and so $(S_n f)(x)$ when x is distributed according to μ_0 has the the distribution as $(S_n g)(\omega)$ when ω is distributed according to ν_0. We thus obtain

Corollary C.4. Theorems 3.1 and 3.2 are valid for $(S_n f)(x)$ where x is distributed according to μ_0.

We also have the following result (see [40, Theorem 4.3]).

Theorem C.5 (Sequential SRB measures)**.** Suppose that Λ is a hyperbolic attractor. Then there is a sequence of probability measures μ_j on Λ such that $(T_j)_*\mu_j = \mu_{j+1}$ and

$$\mu_j = \lim_{n\to\infty} (T_{j-n}^n)_*(\rho_{j-n} d\mathrm{Vol})$$

for every uniformly bounded sequence of probability densities ρ_n on Γ_n. The measures (μ_j), $j \in \mathbb{Z}$ are the unique family of equivariant measures such that the conditional measures of μ_j on the unstable manifolds (see §C.4) at time j are absolutely continuous with respect to the Riemannian volume on these submanifolds.

Moreover μ_j can be obtained by the construction described above corresponding to the sequence of functions $\phi_j = -\ln J(T_j|_{\Gamma_j^u})$, where $J(\cdot)$ stands for the Jacobian matrix.

Therefore $W_n = \sum_{j=0}^{n-1} f_j(T_0^j x)$ satisfies all the limit theorems in Section 3 when x is distributed according to a measure having a Hölder density with respect to μ_0.

Remark C.6. When starting with a time zero sequential SRB measure Theorems 3.1 and 3.2 hold for uniformly hyperbolic sequences (T_j) without the assumption that the maps T_j are close, if $\liminf_{n\to\infty} \frac{1}{n}\mathrm{Var}(S_n) > 0$. Indeed, in this setup we can use the functional analytic approach of Bakhtin [8] together with the perturbative approach in Section D. The reason this works well only when $\liminf_{n\to\infty} \frac{1}{n}\mathrm{Var}(S_n) > 0$ is that when the variance grows sublinearly fast our proofs rely on a block decomposition technique. This involves a sufficiently smooth martingale coboundary representations, a tool which is not available in the setup of [8] since the transfer operators acts on spaces of distributions and not functions. See also [43] for a similar functional approach, which has applications to sufficiently close non-autonomous hyperbolic sequences (see [43, Section 7]).

C.4. **Hölder continuity of the conjugacies.** To prove the Hölder continuity of h_j and h_j^{-1} we need some background.

Local stable and unstable manifolds. For ε small enough, and $x \in \Lambda_j$ define $W_j^s(x,\varepsilon)$ to be the set of all points $y \in \Lambda_j$ such that $\mathsf{d}(T_j^n x, T_j^n y) \le \varepsilon$ for all n and $\mathsf{d}(T_j^n x, T_j^n y) \to 0$. Similarly, we define $W_j^u(x,\varepsilon)$ to be the set of all points $y \in \Lambda_j$ such that

$$\mathsf{d}((T_{j-n}^n)^{-1}x, (T_{j-n}^n)^{-1}y) \le \varepsilon$$

for all n and $\mathsf{d}((T_{j-n}^n)^{-1}x, (T_{j-n}^n)^{-1}y) \to 0$. Then (see [40, 62] and [70]) $W_j^s(x,\varepsilon)$ and $W_j^u(x,\varepsilon)$ are manifolds and tangent space of $W_j^s(x,\varepsilon)$ at x is $\Gamma_{x,j}^s$ while the tangent space of $W_j^u(x,\varepsilon)$ at x is $\Gamma_{x,j}^u$. Moreover, there are constants $C > 0$ and $\delta \in (0,1)$ such that for every j,

$$\mathsf{d}(T_j^n x, T_j^n y) \le C\delta^n \text{ for all } y \in W_j^s(x,\varepsilon) \tag{C.2}$$

and

$$\mathsf{d}((T_{j-n}^n)^{-1}x, (T_{j-n}^n)^{-1}y) \le C\delta^n \text{ for all } y \in W_j^u(x,\varepsilon) \tag{C.3}$$

Remark C.7. Like in the autonomous case, we have the following (see [70, Proposition 3.5]). Define $\overline{W}_j^s(x,\beta)$ to be the set of all points $y \in \Lambda_j$ such that $\mathsf{d}(T_j^n x, T_j^n y) \le \beta$ for all $n \ge 0$. Similarly, let $\overline{W}_j^u(x,\beta)$ be the set of all points $y \in \Lambda_j$ such that $\mathsf{d}((T_{j-n}^n)^{-1}x, (T_{j-n}^n)^{-1}y) \le \beta$ for all $n \ge 0$. Then for every $\beta < \frac{1}{2}\varepsilon$,

$$\overline{W}_j^s(x,\beta) \subset W_j^s(x,\varepsilon) \quad \text{and} \quad \overline{W}_j^u(x,\beta) \subset W_j^u(x,\varepsilon).$$

This essentially means that, like in the autonomous case, up to replacing ε with $\frac{1}{2}\varepsilon$, the local stable/unstable manifolds can be defined using only the condition about the ε-closeness of the forward/backward orbits.

Let us also denote by $W^s(x,\varepsilon)$ and $W^u(x,\varepsilon)$ the local stable and unstable manifolds of x with respect to T (then all the above properties hold true).

Proof of Hölder continuity. The proof is a minor modification of the proof of [54, Proposition 19.1.2], but for readers' convenience we provide the details. We only prove that h_j is Hölder continuous, the proof that h_j^{-1} is Hölder continuous is analogous, see below. Fix some ε small enough (in a way that will be determined later). We say that $x, y \in \Lambda$ are s-equivalent if $y \in W^s(x,\varepsilon)$. Similarly, we say that they are u-equivalent if $y \in W^u(x,\varepsilon)$. Then, using local coordinates and that the angles between the stable and unstable directions are uniformly bounded below we see that if ε is small enough then there is a constant K_1 such that

$$\mathsf{d}(x,y)^2 + \mathsf{d}(y,z)^2 \le K_1 \mathsf{d}(x,z)^2 \tag{C.4}$$

for every $x, y, z \in \Lambda$ such that x is s-equivalent to y and y is u-equivalent to z, and $\mathsf{d}(x,z) \le \varepsilon$. In view of (C.4) in order to show that h_j is Hölder continuous it suffices to show that restrictions of h_j to both unstable and stable manifolds are uniformly Hölder (see e.g. [54, Proposition 19.1.1]). We will consider $h_j|W^u(x,\varepsilon)$, the result for $h_j|W^s(x,\varepsilon)$ follows by replacing T by T^{-1}.

Let us prove that $h_j|W^u(x,\varepsilon)$ are uniformly Hölder continuous. Fix some $\varepsilon_0 > 0$. By uniform continuity of h_j proven in [64], there exists $0 < \delta_0 < \varepsilon$ such that for every j and every $x, y \in \Lambda$ with $\mathsf{d}(x,y) < \delta_0$ we have $\mathsf{d}(h_j(x), h_j(y)) < \varepsilon_0$. Let $x, y \in \Lambda$ be such that $y \in W^u(x,\varepsilon)$. Denote $\rho = \mathsf{d}(x,y)$ and let $K > 1$ be a Lipschitz constant for T. Assuming that $\rho \le K^{-2}\delta_0$ there is a unique natural number n such that $K^n\rho < \delta_0 \le K^{n+1}\rho$. Then $\mathsf{d}(T^k x, T^k y) \le K^n \rho < \delta_0$ for all $k \le n$ and so $\mathsf{d}(h_{j+n}T^n x, h_{j+n}T^n y) < \varepsilon_0$. Since $y \in W^u(x,\varepsilon)$ and $\mathsf{d}(T^k x, T^k y) < \delta_0 < \varepsilon$ for $0 \le k \le n$ we have $T^n y \in W^u(x,\varepsilon)$. Using Remark C.7 and the equicontinuity of (h_k), we see that $h_{j+n}T^n y \in W_{j+n}^u(h_{j+n}T^n x, \varepsilon')$ where $\varepsilon' = \varepsilon'(\varepsilon) \to 0$ as $\varepsilon \to 0$ (we can take $\varepsilon' = \sup_k(\sup\{\mathsf{d}(h_k(a), h_k(b)) : \mathsf{d}(a,b) \le \varepsilon\})$. Thus, by fixing ε small enough and using (C.3) (with the above ε' instead of ε) we see that there are constants $C > 0$ and $\delta \in (0,1)$ such that

$$\mathsf{d}(h_j(x), h_j(y)) = \mathsf{d}((T_j^n)^{-1} h_{j+n} T^n x, (T_j^n)^{-1} h_{j+n} T^n y) \le C\delta^n \varepsilon_0$$

$$= C\varepsilon_0 K^{-cn} = C\varepsilon_0 K^c (K^{-n-1})^c \le (C\varepsilon_0 K^c \delta_0^{-c})\rho^c = (C\varepsilon_0 K^c \delta_0^{-c})\mathsf{d}(x,y)^c$$

where $c = \frac{|\ln \delta|}{\ln K} > 0$.

This shows Hölder continuity of h_j. The proof that h_j^{-1} is Hölder continuous is similar except we use the equivalence relations induced by $W_j^u(x,\varepsilon)$ and by $W_j^s(x,\varepsilon)$. □

Appendix D. Perturbation theorem

D.1. **The statement.** Let $(B_j, \|\cdot\|_j)_{j\ge 0}$ be a sequence of Banach spaces and $A_j : B_j \to B_{j+1}$ be a sequence of bounded linear operators. We assume here that there are $(\lambda_j, h_j, \nu_j) \in (\mathbb{C} \setminus \{0\}) \times B_j \times B_j^*$ where B_j^* denotes the dual space of B_j so that

$$A_j h_j = \lambda_j h_{j+1}, \quad (A_j)^* \nu_{j+1} = \lambda_j \nu_j. \tag{D.1}$$

We will also assume that $\dim(B_j) > 1$ and that $\nu_j(h_j) = 1$.

Let $\mathbf{C}$ be a complex Banach space. Let $\varepsilon > 0$. Denote by $B_{\mathbf{C}}(0,\varepsilon)$ the open ball in $\mathbf{C}$ around 0 with radius ε. For each $t \in B_{\mathbf{C}}(0,\varepsilon)$ let $A_j^{(t)} : B_j \to B_{j+1}$ be a sequence of bounded operators such that $A_j^{(0)} = A_j$. Set $A_j^{t,n} = A_{j+n-1}^{(t)} \circ \cdots \circ A_j^{(t)}$. In the sequel, for any sequence of operators $R_j : B_j \to B_{j+1}$ we will denote $R_j^n = R_{j+n-1} \circ \cdots \circ R_{j+1} \circ R_j$. Write $\mathbb{C}' = \mathbb{C} \setminus \{0\}$.

Assumption D.1. (i) $\inf_j |\lambda_j| > 0$.

(ii) There are constants $C_0, \delta_0{>}0$ such that for all j and $g \in B_j$ so that $\nu_j(g){=}0$ we have

$$\|A_j^n g\| \le C_0 \|g\| \delta_0^n. \tag{D.2}$$

Moreover, $\liminf_{n\to\infty} \inf_j |\lambda_{j,n}|^{1/n} > \delta_0$, where $\lambda_{j,n} = \prod_{k=j}^{j+n-1} \lambda_k$.

(iii) The maps $t \to A_j^{(t)}$ are analytic in some neighborhood $\mathcal{U}$ of the origin (which does not depend on j) and $\sup_{t\in\mathcal{U}} \|A_j^{(t)}\|$ are uniformly bounded in j.

The main result in this section is as follows.

Theorem D.2 (A sequential perturbation theorem)**.** Under Assumption D.1, for every $\delta_1 > \delta_0$ there exist $r_1, C_1 > 0$ such that for every $j \in \mathbb{N}$ and $t \in B_{\mathbf{B}}(0, r_1)$ we have the following.
(i) There are triplets $(\lambda_j(t), h_j^{(t)}, \nu_j^{(t)}) \in \mathbb{C}' \times B_j \times B_j^*$, which are uniformly bounded in j, t so that

$$A_j^{(t)} h_j^{(t)} = \lambda_j(t) h_{j+1}^{(t,y)}, \quad (A_j^{(t)})^* \nu_{j+1}^{(t)} = \lambda_j(t) \nu_{k,j}^{(t,y)}.$$

Moreover, $\lambda_j(0) = \lambda_j$, $h_j^{(0)} = h_j$, $\nu_j^{(0)} = \nu_j$, and the triplets are analytic functions of t and their norm is uniformly bounded. Furthermore, $\nu_j(h_j^{(t)}) = \nu_j^{(t)}(h_{\ell,j}) = \nu_j^{(t)}(h_j^{(t)}) = 1$. If instead of analyticity of $t \mapsto A_j^{(t)}$ we assume that $\mathbf{C}$ is a real Banach space and $t \mapsto A_j^{(t)}$ is C^k for some $k \in \mathbb{N}$ with uniform bounds on the C^k norms, then the above triplets are C^k functions of t with uniformly bounded C^k norms.
(ii) Consider the operators $P_j^{(t)} : B_j \to B_{j+1}$ given by $P_j^{(t)} g = \lambda_j \nu_j^{(t)}(g) h_{j+1}^{(t)}$. Then

$$P_j^{t,n} = \lambda_{j,n}(t) \nu_j^{(t)}(\cdot) h_{j+n}^{(t)}$$

and denoting $E_j^{(t)} = A_j^{(t)} - P_j^{(t)}$ we have $A_j^{t,n} = P_j^{t,n} + E_j^{t,n}$ and

$$\|E_j^{t,n}\| \le C_1 \delta_1^n. \tag{D.3}$$

D.2. **Preparations for the proof of Theorem D.2.** Our first result shows that we can always consider two sided sequences instead of one sided ones.

Lemma D.3. We can extend both sequences $(B_j, \|\cdot\|_j)_{j\ge 0}$ and $(A_j)_{j\ge 0}$ to two sided sequences such that (D.1) and Assumption D.1 still hold.

Proof. For $j < 0$ we take $B_j = B_0$. Then we take a nonzero $\nu_{-1} \in B_0^*$ and set $\nu_j = \nu_{-1}$ for $j < -1$. We also define $\lambda_j = 1+\delta_0$ for $j < 0$. Finally, we take a nonzero $h_{-1} \in B_0$ such that $\nu_{-1}(h_{-1}) = 1$ and define $h_j = h_{-1}$ for all $j < 0$. For $j < 0$, define $A_j(g) = (1+\delta_0)\nu_{-1}(g)h_{j+1} = \lambda_j \nu_j(g) h_{j+1}$. It is clear that (D.1) holds with the new two sided sequences. Assumption D.1 is in force since for $j < 0$, if $\nu_j(g) = 0$ then $A_j(g) = 0$ and so $A_j^n g = 0$. □

Henceforth we assume that we have a two sided sequence satisfying (D.1) and Assumption D.1.

Remark D.4. If the operators A_j are defined for all $j \in \mathbb{Z}$ then the proof of Theorem D.2 (which proceeds by applying the Implicit Function Theorem) shows that the projections P_j constructed in the theorem are unique. The uniqueness does not hold if the operators are only defined for $j \in \mathbb{N}$ (cf. the discussion after Theorem 2.4).

Lemma D.5. (i) Let the operator $Q_j : B_j \to B_j$ be given by $Q_j(g) = \nu_j(g)h_j$. Then for every j and $g \in B_j$ we have $\nu_j(Q_j g) = \nu_j(g)$.

(ii) Consider the operator $P_j : B_j \to B_{j+1}$ given by

$$P_j g = \lambda_j \nu_j(g) h_{j+1}.$$

Then P_j coincides with A_j on $\eta_j := \mathrm{span}\{h_j\}$ and

$$A_{j+1}P_j = P_{j+1}A_j = P_{j+1}P_j, \text{ and } P_j^n = \lambda_{j,n}\nu_j(\cdot)h_{j+n}.$$

Therefore, with $E_j = A_j - P_j$ we have

$$E_j^n = A_j^n - P_j^n = A_j^n - \lambda_{j,n}\nu_j(\cdot)h_{j+n}.$$

In particular, $A_j^n = P_j^n + E_j^n$.

Proof. (i) $\nu_j(\nu_j(g)h_j) = \nu_j(g)\nu_j(h_j) = \nu_j(g)$.

(ii) We have $P_j h_j = \lambda_j \nu_j(h_j) h_j = \lambda_j h_{j+1}$. So A_j and P_j coincide on η_j. Next, the identity $P_j^n = \lambda_{j,n}\nu_j(\cdot)h_{j+n}$ follows by induction on n. For $n = 1$ is it just the definition of P_j, and to prove the induction step, if $P_j^n = \lambda_{j,n}\nu_j(\cdot)h_{j+n}$ then

$$P_j^{n+1}(g) = P_{j+n}(P_j^n g) = \lambda_{j,n}\nu_j(g)P_{j+n}(h_{j+n}) = \lambda_{j,n}\nu_j(g)\lambda_{j+n}h_{j+n+1} = \lambda_{j,n+1}\nu_j(g)h_{j+n+1}.$$

Since $(A_j)^*\nu_{j+1} = \lambda_j \nu_j$ we have $A_{j+1}P_j = P_{j+1}A_j = P_{j+1}P_j$, because all three expression above coincide with the operator

$$g \to \lambda_j \lambda_{j+1} \nu_j(g) h_{j+2} = P_j^2(g).$$

Using this the proof that $E_j^n = A_j^n - P_j^n$ proceeds by induction on n. □

Lemma D.6. Suppose that Assumption D.1 holds and let C be a constant so that $\sup_j \|Q_j\| \le C$. Then for all $g \in B_j$,

$$\|A_j^n g - \lambda_{j,n}\nu_j(g)h_{j+n}\| = \|A_j^n - P_j^n\| \le (C+1)C_0\|g\|\delta_0^n. \tag{D.4}$$

Proof. Set $R_j g = g - Q_j g$. By Lemma D.5, $\nu_j(R_j g) = 0$. So by (D.2) applied to $R_j g$

$$\|A_j^n g - \lambda_{j,n}\nu_j(g)h_{j+n}\| = \|A_j^n(R_j g)\| \le (C+1)C_0\|g\|\delta_0^n.$$ □

Lemma D.7. Let $R_j : B_j \to B_{j+1}$ be a sequence of operators. Suppose that there are constants $\delta_0, C_0 > 0$ such that for all j, n we have

$$\|R_j^n\| \leq C_0 \delta_0^n.$$

Then for every $\delta_1 > \delta_0$ there exist constants $\varepsilon_0 > 0$ and $C_1 > 0$ with the following property. For any other family $S_j : B_j \to B_{j+1}$ of linear operators satisfying $\sup_j \|R_j - S_j\| < \varepsilon_0$ we have

$$\|S_j^n\| \leq C_1 \delta_1^n.$$

Proof. Let $\delta_1 > \delta_0$ and let n_0 be so that $C_0 \delta_0^{n_0} < \frac{1}{2}\delta_1^{n_0}$. Let $\varepsilon_0 > 0$ be so small such that

$$\sup_j \|R_j - S_j\| < \varepsilon_0$$

implies $\sup_j \|S_j^{n_0}\| < \delta_1^{n_0}$. It is indeed possible to find such ε_0 since

$$\|S_j^{n_0} - R_j^{n_0}\| \leq C(C_0, \delta_0, n_0)\varepsilon_0$$

for some constant $C(C_0, \delta_0, n_0)$ which depends only on C_0, δ and n_0.

Now, given $n \in \mathbb{N}$ let us write $n = kn_0 + \ell$ for some $0 \leq \ell < n_0$. Then

$$\|S_j^n\| \leq \prod_{u=0}^{k-1} \|S_{j+un_0}^{n_0}\| \cdot \|S_{j+kn_0}^{\ell}\| \leq (\varepsilon_0 + C_0\delta_0)^\ell \delta_1^{-\ell} \delta_1^n. \qquad \square$$

D.3. **Proof of Theorem D.2.**

Proof. We first prove the existence of triplets. Denote

$$\mathbf{H}_j = \{g \in B_j : \nu_j(g) = 0\}$$

Since $A_j^*\nu_{j+1}=\lambda_j\nu_j$ we have $A_j\mathbf{H}_j \subset \mathbf{H}_{j+1}$. Also, for every $g_j \in \mathbf{H}_j$ we have $\nu_j(h_j + g_j)=1$ and

$$\nu_{j+1}\big(A_j(h_j + g_j)\big) = \lambda_j(A_j^*\nu_{j+1})(h_j + g_j) = \lambda_j\nu_j(h_j + g_j) = \lambda_j.$$

Consider the space $\mathbf{H}$ of sequences $\bar{g} = (g_j)$ so that $g_j \in \mathbf{H}_j$ for all j, equipped with the norm $\|\bar{g}\| := \sup_j \|g_j\| < \infty$. Define a function $F : \mathbf{C} \times \mathbf{H} \to \mathbf{H}$ by

$$\left(F(t, \bar{g})\right)_j = A_{j-1}^{(t)}(h_{j-1} + g_{j-1}) - \nu_j\big(A_{j-1}^{(t)}(h_{j-1} + g_{j-1})\big)\left(h_j + g_j\right).$$

Then $F_k(0,0) = 0$, and F is analytic in t (in the case C^k dependence on t this map is differentiable k times).

We claim that the partial derivative $\Phi := (\partial F/\partial \mathbf{H})(0,0)$ is an isomorphism. Indeed, a direct computation shows that $\left(\Phi\bar{g}\right)_j = A_{j-1}g_{j-1} - \lambda_j g_j$. Assume first that there is $\bar{g} \in \mathbf{H}$ so that

$$A_{j-1}g_{j-1} = \lambda_j g_j. \tag{D.5}$$

When the restrictions $A_j : \mathbf{H}_j \to \mathbf{H}_{j+1}$ satisfy

$$\|A_j^n g_j\| \leq C_0 \|g_j\| \delta_0^n \tag{D.6}$$

since on $\mathbf{H}_j$ the functionals ν_j vanish. Iterating (D.5) we see that $A_j^n g_j = \lambda_{j,n} g_{j+n}$. Plugging in $j - n$ instead of j in (D.6) we get that $g_j = \frac{A_{j-n}^n g_{j-n}}{\lambda_{j-n,n}}$ and so

$$\|g_j\| \leq C_0 \|\bar{g}\| \frac{\delta_0^n}{|\lambda_{j,n}|} \leq C'\varepsilon^n \to 0$$

for some $\varepsilon \in (0,1)$ and a constant C', where in the second inequality uses Assumption D.1(ii). Hence $\bar{g} = 0$.

Now we show that Φ is surjective. Take $\bar{g}\in\mathbf{H}$. Set $h_j=-\sum_{n=0}^{\infty}\frac{A_{j-n}^n g_{j-n}}{\lambda_{j-n,n}}$. This series converges uniformly in j since $\left\|\frac{A_{j-n}^n g_j}{\lambda_{j-n,n}}\right\|\leq C'\|\bar{g}\|\varepsilon^n$ where ε and C' are like in the previous estimates. It is immediate that

$$A_{j-1}h_{j-1}-\lambda_j h_j=\lambda_j g_j.$$

Applying the implicit function theorem in Banach spaces (e.g. [5, Theorem 3.2]) we get that there is $\varepsilon_0>0$ so that for every $t\in B_{\mathbf{C}}(0,\varepsilon_0)$ there is $\bar{g}^{(t)}=(g_j^{(t)})_j\in\mathbf{H}$ so that

$$F(t,\bar{g}^{(t)})=0.$$

Moreover, the function $t\to\bar{g}^{(t)}$ is analytic and so by possibly reducing ε_0 we can assume that it is bounded (in the case of C^k dependence on t we get that this map is of class C^k).

Now we construct the families $\nu_j^{(t)}$. First, we obtain from (D.4) that

$$\|(A_j^n)^*-\lambda_{j,n}h_{j+n}^*\nu_j\|\leq C_0\delta_0^n \tag{D.7}$$

where $h_{j+n}^*(\mu)=\mu(h_{j+n})$. Let $\mathbf{E}_j\subset B_j^*$ be the space of all functionals μ_j such that $\mu_j(h_j)=0$. Then $(A_j)^*\mathbf{E}_{j+1}\to\mathbf{E}_j$ since

$$(A_j)^*\mu_{j+1}(h_j)=\mu_{j+1}(A_jh_j)=\lambda_j\mu_{j+1}(h_{j+1})=0.$$

Let $\mathbf{E}$ be the Banach space of sequences $\bar{\mu}=(\mu_j)$ such that $\mu_j\in\mathbf{E}_j$ for all j and $\|\bar{\mu}\|:=\sup_j\|\mu_j\|<\infty$. Define

$$\big(G(t,\bar{\mu})\big)_j=(A_j^{(t)})^*(\nu_{j+1}+\mu_{j+1})-\lambda_j(t)(\nu_j+\mu_j)$$

where $\lambda_j(t)=\nu_j(h_j^t)$. Then $G(0,0)=0$ and G is a well defined function on $\{t:\|t\|<\varepsilon_0\}$ for some $\varepsilon_0>0$. Moreover, G is analytic in t. We consider G as a function from a neighborhood of $(0,0)\in\mathbf{C}\times\mathbf{E}$ to $\mathbf{E}$.

Now, a direct computation shows that the derivative of G at $(0,0)$ in the direction $\mathbf{E}$ is the operator given by

$$\Psi:=(\partial G/d\mathbf{E})(0,0)\bar{\mu})_j=(A_j)^*\mu_{j+1}-\lambda_j\mu_j.$$

Let us show that Ψ is injective. If $\bar{\mu}$ belongs to its kernel then $(A_j)^*\mu_{j+1}=\lambda_j\mu_j$, and so $(A_j^n)^*\mu_{j+n}=\lambda_{j,n}\mu_j$. Using (D.7) we get that $\|\mu_j\|\leq C\varepsilon^n$ for some constants $\varepsilon\in(0,1)$ and $C>0$. Therefore $\bar{\mu}=0$.

Next, Ψ is also surjective since for any $\bar{\mu}\in\mathbf{E}$ we have $(A_j)^*\kappa_{j+1}-\lambda_j\kappa_j=\lambda_j\mu_j$ where $\kappa_j=-\sum_{n=0}^{\infty}\frac{(A_{j+1}^n)^*\mu_{j+n+1}}{\lambda_{j,n}}$ (the convergence of this series follows from (D.7)).

Hence, by the implicit function theorem there is an analytic function $t\mapsto\bar{\kappa}^{(t)}=(\kappa_j^{(t)})_j$ with values in $\mathbf{E}$ so that

$$(A_j^{(t)})^*(\nu_{kj+1}+\kappa_{j+1}^{(t)})-\lambda_j(t)(\nu_j+\kappa_j^{(t)})=0.$$

Take $\nu_j^{(t)}=\nu_j+\kappa_j^{(t)}$ and notice that $\nu_j^{(t)}(h_j^{(t)})$ does not depend on j. Indeed, we have

$$\lambda_j(t)\nu_{j+1}^{(t)}(h_{j+1}^{(t)})=\nu_{j+1}^{(t)}(\lambda_j(t)h_{j+1}^{(t)})=\nu_{j+1}^{(t)}(A_j^{(t)}h_j^{(t)})=(A_j^{(t)})^*\nu_{j+1}^{(t)}(h_j^{(t)})=\lambda_j(t)\nu_j^{(t)}(h_j^{(t)}).$$

We conclude that there is a function $c(t)$ such that $\nu_j^{(t)}(h_j^{(t)})=c(t)$ for every $j\geq0$. The functions $c(t)$ are bounded and analytic and $c(0)=1$. Thus, by decreasing $\|t\|$ if necessary we can assume that $|c(t)|\geq\frac{1}{2}$. Next, by replacing $h_j^{(t)}$ with $h_j^{(t)}/c(t)$ (or $\nu_j^{(t)}$ by $\nu_j^{(t)}/c(t)$) we can just assume that $c(t) = 1$. Like in the construction of $g^{(t)}$, in the case $t \to (A_j^{(t)})$ is of class C^k all the above functions of t are differentiable k times.

Finally, the exponential convergence follows by Lemma D.7 applied with $R_j = A_j - P_j$ and $S_j = A_j^{(t)} - P_j^{(t)}$ (when t is close enough to 0) where $P_j^{(t)} = \lambda_j(t)\nu_j^{(t)}(\cdot)h_{j+1}^{(t)}$ taking into account that by Lemma D.5, $S_j^n = A_j^{t,n} - P_j^{t,n} = A_j^{t,n} - \lambda_{j,n}(t)\nu_j^{(t)}(\cdot)h_{j+n}^{(t)}$. □

References

[1] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps*, Stoch. Dyn. **1** (2001) 193–237.
[2] R. Aimino, M. Nicol, S. Vaienti *Annealed and quenched limit theorems for random expanding dynamical systems,* Prob. Th. Rel. Fields **162** (2015) 233–274.
[3] L. Arnold, *Random Dynamical Systems*, Springer-Verlag, New York, Berlin (1998).
[4] P. Arnoux, A.M. Fisher *Anosov families, renormalization and non-stationary subshift*, Ergod. Th. Dynam. Sys. **25** (2005) 661–709.
[5] A. Avez *Differential Calculus*, Wiley, Chichester (1986).
[6] A. Ayyer, C. Liverani, M. Stenlund *Quenched CLT for random toral automorphism,* DCDS **24** (2009) 331–348.
[7] A. Ayyer, M. Stenlund *Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms,* Chaos **17** (2007) paper 043116.
[8] V. Bakhtin *Random processes generated by a hyperbolic sequence of mappings-I,* Izvestiya Math. **44** (1995) 247–279.
[9] A.D. Barbour and L.H.Y. Chen, *An Introduction to Stein's Method*, Singapore University Press, Singapore, 2005.
[10] V. Baladi, M. Benedicks, V. Maume–Deschamps *Almost sure rates of mixing for i.i.d. unimodal maps* Ann. ENS **35** (2002) 77–126.
[11] P. Baldi, Y. Rinott *On normal approximations of distributions in terms of dependency graphs,* Ann. Prob. **17** (1989) 1646–1650.
[12] D. Berend, V. Bergelson *Ergodic and mixing sequences of transformations*, ETDS **4** (1984) 353–366.
[13] D. Berend, V. Bergelson, *Mixing sequences in Hilbert spaces*, Proc. AMS **98** (1986) 239–246.
[14] A. Berger, L. Bunimovich, T. Hill *One-dimensional dynamical systems and Benfords law*, Trans. AMS **357** (2005) 197–219.
[15] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, 2d revised edition, Springer Lecture Notes in Math. **470** (2008).
[16] G. Birkhoff, *Extension of Jentzsch's theorem*, Trans. AMS **85** (1957) 219–227.
[17] I. Berkes, W. Philipp *Approximation theorems for independent and weakly dependent random vectors*. Ann. Probab. **7** (1979) 29–54.
[18] W. Berry, *The accuracy of the Gaussian approximation to the sum of independent variates*, Trans. AMS **49** (1941) 122–136.
[19] J. Buzzi, *Exponential Decay of Correlations for Random Lasota-Yorke Maps*, Comm. Math. Phys. **208** (1999) 25–54.
[20] H. Callaert, P. Janssen, *The Berry-Esseen theorem for U statistics*, Ann. Stat. **6** (1978) 417–421.
[21] L.H.Y Chen, Q.M. Shao, *Normal approximation under local dependence*, Ann. Probab. **32** (2004) 1985–2028.
[22] N.-D. Cong, *Topological Dynamics of Random Dynamical Systems*, Oxford Univ. Press, Oxford (1997).
[23] J.–P. Conze, S. Le Borgne, M. Roger *Central limit theorem for stationary products of toral automorphisms,* DCDS–S **32** (2012) 1597–1626.
[24] J.–P. Conze, A. Raugi *Limit theorems for sequential expanding dynamical systems on [0,1],* Contemp. Math. **430** (2007) 89–121.
[25] H. Crauel *Random Probability Measures on Polish Spaces,* Taylor & Francis, London (2002)
[26] C. Cuny, F. Merlevede, *Strong invariance principles with rate for reverse martingale differences and applications*, J. Theor. Prob. **28** (2015) 137–183.
[27] D. Dolgopyat, Y. Hafouta *A Berry-Esseen theorem and Edgeworth expansions for uniformly elliptic inhomogeneous Markov chains*, Prob. Th. Rel. Fields **186** (2023) 439–476.
[28] D. Dolgopyat, Y. Hafouta *A local limit theorem for expanding maps,* in preparation.
[29] D. Dolgopyat, O. Sarig *Local limit theorems for inhomogeneous Markov chains*, Springer Lect. Notes Math. **2331** (2023).

[30] D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti, *Almost sure invariance principle for random piecewise expanding maps*, Nonlinearity **31** (2018) 2252–2280.

[31] D. Dragičević, G. Froyland, C. González-Tokman, S. Vaienti. A Spectral Approach for Quenched Limit Theorems for Random Expanding Dynamical Systems, *Comm. Math. Phys.* **360** (2018) 1121–1187.

[32] D. Dragičević, G. Froyland, C. González-Tokman, S. Vaienti. *A Spectral Approach for Quenched Limit Theorems for Random Expanding Dynamical Systems*, Tran. AMS **360** (2018) 1121–1187.

[33] D. Dragicevic, Y. Hafouta *Limit Theorems for Random Expanding or Anosov Dynamical Systems and Vector-Valued Observables*, Ann. Hen. Poinc. **21** (2020) 3869-3917.

[34] L. Dubois *Projective metrics and contraction principles for complex cones,* J. Lond. Math. Soc. **79** (2009) 719–737.

[35] L. Dubois *An explicit Berry-Esseen bound for uniformly expanding maps on the interval,* Israel J. Math. **186** (2011) 221–250.

[36] G. K. Eagleson, *Some simple conditions for limit theorems to be mixing*, Teor. Verojatnost. i Primenen. **21** (1976) 653–660.

[37] C.-G. Esseen, *On the Liapounoff limit of error in the theory of probability.* Ark. Math. **28** (1942) 1–19.

[38] C.-G. Esseen *Fourier analysis of distribution functions. A mathematical study of the Laplace-Gaussian law,* Acta Math. **77** (1945) 1–125.

[39] W.F. Grams, R.J. Serfling, *Convergence rates for U-statistics and related statistics*, Ann. Stat. **1** (1973) 153–160.

[40] V.M. Gundlach, Y. Kifer, *Random hyperbolic systems*, Stochastic dynamics, Springer, 1999.

[41] Y. Guivarch J. Hardy, *Theoremes limites pour une classe de chaines de Markov et applications aux diffeomorphismes d'Anosov*, AIHP **24** (1988) 73–98.

[42] M. I. Gordin, B. A. Livsic *Central limit theorem for stationary Markov processes,* Dokl. Akad. Nauk SSSR **239** (1978) 766–767.

[43] S. Gouëzel, C. Liverani, *Banach spaces adapted to Anosov systems*, Ergodic. Theory. Dynam. Systems 26, 189–217.

[44] S. Gouëzel. *Berry–Esseen theorem and local limit theorem for non uniformly expanding maps*, Annales IHP Prob. & Stat. **41** (2005) 997–1024.

[45] S. Gouëzel, *Variations around Eagleson's theorem on mixing limit theorems for dynamical systems*, Ergodic. Theory. Dynam. Systems **40** (2020) 3368–3374.

[46] Y. Hafouta, Yu. Kifer *Nonconventional limit theorems and random dynamics,* World Scientific, Hackensack, NJ (2018).

[47] Y. Hafouta, *On the Asymptotic moments and Edgeworth expansions for some processes in random dynamical environment*, J. Stat. Phys. **179** (2020) 945–971.

[48] Y. Hafouta, *Limit theorems for some time dependent dynamical systems*, Nonlinearity **33** (2020) 6421–6460.

[49] Y. Hafouta, *Limit theorems for random non-uniformly expanding or hyperbolic maps with exponential tails*, Ann. Hen. Poinc. **23** (2022), 293-332.

[50] Y. Hafouta, *An almost sure invariance principle for some classes of non-stationary mixing sequences*, Stat. Prob. Let. **193** (2023) paper 109728.

[51] Y.Hafouta, Explicit conditions for the CLT and related results for non-uniformly partially expanding random dynamical systems via effective RPF rates, Adv. Math. **426** (2023) paper 109109.

[52] Y. Hafouta, *Non-uniform Berry-Esseen theorem and Edgeworth expansions with applications to transport distances for weakly dependent random variables,*, preprint, https://arxiv.org/abs/2210.07204

[53] Y. Hafouta, *On Eagleson's theorem in the non-stationary setup*, Bull. LMS **53** (2021) 991–1008.

[54] B. Hasselblatt, A. Katok *Introduction to the modern theory of dynamical systems*, Cambridge, 1995.

[55] H. Hennion, H. Herve, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness* Springer Lect. Notes Math. **1766** (2001).

[56] M. Jirak *Berry-Esseen theorems under weak dependence*, Ann. Prob. **44** (2016) 2024–2063.

[57] M. Jirak *A Berry–Esseen bound with (almost) sharp dependence conditions,* Bernoulli **29** (2005) 1219–1245.

[58] Yu. Kifer, *Ergodic Theory of Random Transformations*, Birkhäuser, Boston (1986).

[59] Yu. Kifer, *Perron-Frobenius theorem, large deviations, and random perturbations in random environments*, Math. Z. **222** (1996) 677–698.

[60] Yu. Kifer, *Limit theorems for random transformations and processes in random environments*, Trans. AMS **350** (1998) 1481–1518.

[61] Yu. Kifer, *Thermodynamic formalism for random transformations revisited*, Stoc. Dyn. **8** (2008) 77–102.

[62] Yu. Kifer, P.-D Liu, *Random Dynamics*, Handbook of dynamical systems, Cambridge, 1995.

[63] A. Lasota and J.A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. AMS **186** (1973) 481–488.

[64] P.-D. Liu, *Random perturbations of Axiom A sets.* Jour. Stat. Phys. **90** (1998) 467–490.
[65] P.-D. Liu, M. Qian, *Smooth Ergodic Theory of Random Dynamical Systems*, Lecture Notes Math. **1606** (1995) Springer, Berlin.
[66] C. Liverani, *Decay of correlations*, Ann. Math. **142** (1995) 239–301.
[67] I. Melbourne, M. Nicol, *Almost sure invariance principle for nonuniformly hyperbolic systems*, Commun. Math. Phys. **260** (2005) 131–146.
[68] I. Melbourne, M. Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Prob. **37** (2009) 478-505.
[69] F. Merlevede, M. Peligrad, S. Utev *Functional Gaussian approximation for dependent structures,* Oxford Studies Prob. **6** (2019).
[70] J. Muentes, R. Ribeiro *Expansiveness, Shadowing and Markov Partition for Anosov Families*, preprint, https://arxiv.org/abs/2007.07424
[71] S.V. Nagaev *Some limit theorems for stationary Markov chains*, Th. Prob. Appl. **2** (1957) 378–406.
[72] P. Nandori, D. Szasz, T. Varju *A central limit theorem for time-dependent dynamical systems,* J. Stat. Phys **146** (2012) 1213–1220.
[73] W. Philipp, W. F. Stout, *Almost sure invariance principles for sums of weakly dependent random variables*, Memoirs AMS **161** (1975).
[74] L. Polterovich, Z. Rudnick *Stable mixing for cat maps and quasi-morphisms of the modular group*, ETDS **24** (2004) 609–619.
[75] E. Rio, *Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes*, Probab. Th. Relat. Fields **104** (1996) 255–282.
[76] J. Rousseau-Egele. *Un théoreme de la limite locale pour une classe de transformations dilatantes et monotones par morceaux,* Ann. Prob. **11** (1983) 772–788.
[77] H.H. Rugh, *Cones and gauges in complex spaces: Spectral gaps and complex Perron-Frobenius theory*, Ann. Math. **171** (2010) 1707–1752.
[78] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps*, Israel J. Math. **116** (2000) 223-248.
[79] M. Shub *Global stability of dynamical systems,* Springer, New York, 1987, xii+150 pp.
[80] R. Zweimüller *Mixing limit theorems for ergodic transformations*, J. Theor. Prob. **20** (2007) 1059–1071.