Central limit theorems for sequential and random intermittent dynamical systems.

Matthew Nicol * Andrew Török † Sandro Vaienti ‡

December 15, 2015

Contents

1 Introduction 2
2 Cones and Martingales 7
3 Central Limit Theorem 15
4 Central Limit Theorem for nearby maps 22
5 Random compositions of intermittent maps 24
6 Appendices 27
   6.1 Gál-Koksma Theorem. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 27
   6.2 Proof of Lemma 3.4 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 28

Abstract

We establish self-norming central limit theorems for non-stationary time series arising as observations on sequential maps possessing an indifferent fixed point. These

*Department of Mathematics, University of Houston, Houston Texas, USA. e-mail: <nicol@math.uh.edu>.
†Department of Mathematics, University of Houston, Houston Texas, USA and Institute of Mathematics of the Romanian Academy, P.O. Box 1–764, RO-70700 Bucharest, Romania. e-mail: <torok@math.uh.edu>.
‡Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France and Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France. e-mail:<vaienti@cpt.univ-mrs.fr>.
transformations are obtained by perturbing the slope in the Pomeau-Manneville map. We also obtain quenched central limit theorems for random compositions of these maps.

1 Introduction

In a preceding series of two papers [13], [3], we considered a few statistical properties of non-stationary dynamical systems arising by the sequential composition of (possibly) different maps. The first article [13] dealt with the Almost Sure Invariance Principle (ASIP) for the non-stationary process given by the observation along the orbit obtained by concatenating maps chosen in a given set. We choose maps in one and more dimensions which were piecewise expanding, more precisely their transfer operator (Perron-Frobenius, ”PF”) with respect to the Lebesgue measure was quasi-compact structure on a suitable Banach space. The ASIP was then proved by applying a recent result by Cuny and Merlevede [7], whose first step was to approximate the original process with a reverse martingale difference plus an error. The latter was essentially bounded due to the presence of a spectral gap in the PF operator on a Banach space continuously injected in $L^\infty$ (from now on all the $L^p$ spaces will be with respect to the ambient Lebesgue measure $m$ and they will be denoted with $L^p$ or $L^p(m)$). Moreover, the same spectral property allowed us to show that for expanding maps chosen close enough, the variance $\sigma_n^2$ grows linearly, which permit to approximate the original process almost everywhere with a finite sum of i.i.d. Gaussian variables with the same variance.

The second paper [3] considered composition of Pomeau-Manneville like maps, obtained by perturbing the slope at the indifferent fixed point 0. We got polynomial decay of correlations for particular classes of centered observables, which could also be interpreted as the decay of the iterates of the PF operator on functions of zero (Lebesgue) average, and this fact is better known as loss of memory. In this situation the PF operator is not quasi-compact and although the process given by the observation along a sequential orbit can be decomposed again as the sum of a reverse martingale difference plus an error, apriori the latter turns out to be bounded only in $L^1$ and this was an obstacle to obtain an almost sure result like the ASIP by only looking at the almost sure convergence of the reverse martingale difference. Instead one could hope to get a (distributional) Central Limit Theorem (CLT); in this regard a general approach to CLT for sequential dynamical systems has been proposed and developed in [6]. It basically applies to systems with a quasi-compact PF.
operator and it is not immediately transposable to maps with do not admit a spectral gap. The main goal of our paper is to prove the CLT for the sequential composition of Pomeau-Manneville maps with varying slopes. A fundamental tool in obtaining such a result will be the polynomial loss of memory bound obtained in [3]; we are now going to recall it also because it will determine the regularity of the observables to which our CLT will apply; see Theorem 1.2.

We consider the family of Pomeau-Manneville maps

\[ T_\alpha(x) = \begin{cases} 
  x + 2^\alpha x^{1+\alpha}, & 0 \leq x \leq 1/2 \\
  2x - 1, & 1/2 \leq x \leq 1 
\end{cases} \quad 0 < \alpha < 1. \quad (1.1) \]

Actually in [3] we considered a slightly different family of this type, but pointed out that both versions could be worked out with the same techniques (see [1]), and lead to the same result; here we prefer to use the classical version (1.1). As in [18], we identify the unit interval \([0,1]\) with the circle \(S^1\), so that the maps become continuous. If \(0 < \beta_k < 1\) are given, denote by \(P_{\beta_k}\) or \(P_k\) the Perron-Frobenius operator associated with the map \(T_k = T_{\beta_k}\) w.r.t. the measure \(m\), where \(0 < \beta_k \leq \alpha\). For concatenations we use equivalently the notations

\[ P_{n-m+1} = P_{\beta_n} \circ P_{\beta_{n-1}} \circ \cdots \circ P_{\beta_m} = P_n \circ P_{n-1} \circ \cdots \circ P_m. \]

\[ T_{n-m+1} := T_{\beta_n} \circ T_{\beta_{n-1}} \circ \cdots \circ T_{\beta_m} = T_n \circ T_{n-1} \circ \cdots \circ T_m. \]

where the exponent denotes the number of maps in the concatenation. We use for simplicity \(T^\infty := \cdots \circ T_2 \circ T_1\) for a given sequence of transformations.

The Perron-Frobenius operator \(P_k\) associated to \(T_k\) satisfies the duality relation

\[ \int_M P_k f \, g \, dm = \int_M f \circ T_k \, g \, dm, \quad \text{for all } f \in L^1, \ g \in L^\infty \]

and this is preserved under concatenation.

We next consider [18, 3] the cone \(C_2\) of functions given by (here \(X(x) = x\) is the identity function):

\[ C_2 := \{ f \in C^0((0,1]) \cap L^1(m) \mid f \geq 0, \ f \text{ decreasing}, \ X^{\alpha+1} f \text{ increasing}, \ f(x) \leq ax^{-\alpha} m(f) \}^1 \]

\(^{1}\text{By "decreasing" we mean "nonincreasing".}\)
Remark 1.1 Some coefficients that appear later depend on the value $a$ that defines the cone $C_2$; however, we will not write explicitly this dependence.

Fix $0 < \alpha < 1$; as proven in [3], provided $a$ is large enough, the cone $C_2$ is preserved by all operators $P_\beta$, $0 < \beta \leq \alpha < 1$. The following polynomial decay result holds:

**Theorem 1.2 ([3])** Suppose $\psi, \varphi$ are in $C_2$ with equal expectation $\int \varphi dm = \int \psi dm$. Then for any $0 < \alpha < 1$ and for any sequence $T_{\beta_1}, \ldots, T_{\beta_n}$, $n \geq 1$, of maps of Pomeau-Manneville type (1.1) with $0 < \beta_k \leq \alpha < 1$, $k \in [1, n]$, we have

$$\int |P_{\beta_n} \circ \cdots \circ P_{\beta_1}(\varphi) - P_{\beta_n} \circ \cdots \circ P_{\beta_1}(\psi)| dm \leq C_\alpha(||\varphi||_1 + ||\psi||_1)n^{-\frac{1}{\alpha} + 1}(\log n)^{\frac{1}{\alpha}},$$

(1.2)

where the constant $C_\alpha$ depends only on the map $T_\alpha$, and $|| \cdot ||_1$ denotes the $L^1$ norm.

A similar rate of decay holds for observables $\varphi$ and $\psi$ that are $C^1$ on $[0, 1]$; in this case the rate of decay has an upper bound given by

$$C_\alpha F(||\varphi||_{C^1} + ||\psi||_{C^1})n^{-\frac{1}{\alpha} + 1}(\log n)^{\frac{1}{\alpha}}$$

where the function $F : \mathbb{R} \to \mathbb{R}$ is affine.

For the proof of the CLT Theorem 3.1 we need better decay than in $L^1$. In this paper we improve the above result to decay in $L^p$, provided $\alpha$ is small enough.

Note that $P^n \varphi \in C_2$ if $\varphi \in C_2$ and $m(P^n \varphi) = m(\varphi)$, so

$$||P^n(\varphi) - P^n(\psi)||_{X} \leq |P^n(\varphi)|_X + |P^n(\psi)|_X \leq am(\varphi)x^{-\alpha} + am(\psi)x^{-\alpha}$$

**Proposition 1.3** Under the assumptions on Theorem 1.2, if $1 \leq p < 1/\alpha$ then

$$||P_{\beta_n} \circ \cdots \circ P_{\beta_1}(\varphi) - P_{\beta_n} \circ \cdots \circ P_{\beta_1}(\psi)||_{L^p(m)} \leq C_{\alpha,p}(||\varphi||_1 + ||\psi||_1)n^{1 - \frac{1}{\alpha} - \frac{1}{p} - \frac{1}{\alpha p}}(\log n)^{\frac{1}{\alpha p - \alpha}}$$

(1.3)

where the constant $C_{\alpha,p}$ depends only on the map $T_\alpha$ and $p$.

As in Theorem 1.2, a similar $L^p$-decay result also holds for observables $\varphi, \psi \in C^1([0,1])$.

**Proof** For functions in the cone $C_2$, Theorem 1.2 gives $L^1$-decay; then Lemma 2.7 together with the preceding discussion implies $L^p$-decay for $\alpha$ small enough. Note that we use this Lemma with $K = 2a(||\varphi||_1 + ||\psi||_1)$ and the $L^1$-bound given by the Theorem, and then the coefficient in the $L^p$-bound is proportional to $(||\varphi||_1 + ||\psi||_1)$ as well.
To prove the decay for $C^1$ observables, we use Lemma 2.4 (same approach as in the proof of Theorem 1.2).

Note that the convergence of the quantity (1.2) implies the decay of the non-stationary correlations with respect to $m$:

$$\left| \int \psi \varphi \circ T_{\beta_n} \circ \cdots \circ T_{\beta_1} dm - \int \psi dm \int \varphi \circ T_{\beta_n} \circ \cdots \circ T_{\beta_1} dm \right| \leq \| \varphi \|_\infty \left\| P_{\beta_n} \circ \cdots \circ P_{\beta_1} (\psi) - P_{\beta_n} \circ \cdots \circ P_{\beta_1} \left( 1 \left( \int \psi dm \right) \right) \right\|_1$$

provided $\varphi$ is essentially bounded and $(\int \psi dm)1$ is in the functional space where the convergence of (1.2) takes place. In particular, this holds for $C^1$ observables, by Theorem 1.2.

From now on we will take our observables as $C^1$ functions on the interval $[0,1]$ and for any $\varphi \in C^1$, we will consider the following observation along a sequential orbit:

$$\varphi_k = [\varphi]_k := \varphi - \int \varphi(T_k \circ \cdots \circ T_1 x) dm.$$ 

As it is suggested by the preceding loss of memory result, centering the observable is the good way to define the process when it is not stationary, in order to consider limit theorems.

Conze and Raugi [6] defined the sequence of transformations $T^\infty$ to be pointwise ergodic whenever the law of large numbers is satisfied, namely

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \left[ \varphi(T_k \circ \cdots \circ T_1 x) - \int \varphi(T_k \circ \cdots \circ T_1 x) dm \right] = 0 \text{ for Lebesgue-a.e. } x.$$ 

We will prove in Theorem 2.10 that such a law of large numbers holds for our observations provided $0 < \alpha < 1$. It is therefore natural to ask about a non-stationary Central Limit Theorem for the sums

$$S_n := \sum_{k=1}^{n} [\varphi]_k \circ T_k \circ \cdots \circ T_1 \quad \text{(1.4)}$$

for a given sequence $T^\infty := \cdots \circ T_n \circ \cdots \circ T_1$; this will be the content of the next sections.

To be more specific we will prove in Theorem 3.1 a non-stationary central limit theorem similar to that proved by Conze and Raugi [6] for (piecewise expanding) sequential systems:

$$\frac{S_n}{\sqrt{\text{Var}(S_n)}} \to^d N(0,1). \quad \text{(1.5)}$$

At this point, we would like to make a few comments about our result compared to that of Conze and Raugi. Theorem 5.1 in [6] shows that, when applied to the quantities defined above and for classes of maps enjoying a quasi-compact transfer operator:
(1) If the norms $||S_n||_2$ are bounded, then the sequence $S_n, n \geq 1$ is bounded.

(2) If $||S_n||_2 \to \infty$, then (1.5) holds.

We are not able to prove item (1) for the intermittent map following the same approach as in [6], since it uses the uniform boundedness of the sequence $H_n \circ T^k$, where the function $H_n$ is defined in (2.1) and is just the error in the martingale approximation as we discussed above. We can only prove that $H_n$ is bounded uniformly in $n$ on each set of the form $[a, 1), a > 0$, and do not expect it to be bounded near 0 (look at the stationary case).

Instead, our central limit theorem will satisfy item (2) under the assumption that the variance $||S_n||_2$ grows at a certain rate and for some limitation on the range of values of $\alpha$. It seems difficult to get such a result in full generality for the intermittent map considered here. Conze and Raugi proved the linear growth of the variance in their Theorem 5.3 under a certain number of assumptions, including the presence of a spectral gap for the transfer operator. We showed in our paper [13] that those assumptions apply to several classes of expanding maps even in higher dimensions.

However, for concatenations given by the same intermittent map $T_\alpha$ with $\alpha < 1/2$, the variance is linear in $n$, provided the observable is not a coboundary for $T_\alpha$. In section 4 we prove that the linear growth of the variance still holds if we take maps $T_{\beta_n}$ with $\beta_n$ arbitrary but close to a fixed $\beta$, and an observable is not a coboundary for $T_\beta$; therefore, the CLT holds. See Theorem 4.1. Our proof of Theorem 4.1 uses an estimate of interesting related work of Leppänen and Stenlund [16], which we learnt about after a first version of this paper was completed. Their result allowed us to give another example where variance grows linearly for a sequential dynamical system of intermittent type maps, and hence the non-stationary CLT holds. The focus of [16] is however more on the strong law of large numbers and convergence in probability rather than the CLT. They also consider quasi static systems, introduced in [17].

In section 5 we show that the variance grows linearly for almost all sequences when we compose intermittent maps chosen from a finite set and we take them according to a fixed probability distribution. This means that for almost all sequences (with respect to the induced Bernoulli measure) of maps, the central limit theorem holds (a quenched-like CLT). See Theorem 5.2.

**Remark 1.4** For simplicity, in many of the following statements we will use as rate of
decay $n^{-\frac{1}{\alpha}+1}$, ignoring the log $n$-factor. This is correct if we take for $\alpha$ a slightly larger value (and is actually the correct rate of decay for the stationary case).

**Notation 1.5** For any sequences of numbers $\{a_n\}$ and $\{b_n\}$, we will write $a_n \approx b_n$ if $c_1 b_n \leq a_n \leq c_2 b_n$ for some constants $c_2 \geq c_1 > 0$.

## 2 Cones and Martingales

In order to get the right martingale representation, we begin by recalling a few formulas concerning the transfer operator; the conditional expectation is considered with respect to the measure $m$, and $\mathcal{B}$ denotes the Borel $\sigma$-algebra on $[0,1]$. We have:

$$
\mathbb{E}[\varphi \mid T^{-k}\mathcal{B}] = \frac{\mathcal{P}^k(\varphi)}{\mathcal{P}^k(1)} \circ T^k
$$

$$
P(\varphi \circ T \cdot \psi) = \varphi \cdot P(\psi)
$$

and therefore, for $0 \leq \ell \leq k$

$$
\mathbb{E}[\varphi \circ T^\ell \mid T^{-k}\mathcal{B}] = \frac{\mathcal{P}^{k-\ell}(\varphi \cdot \mathcal{P}^\ell(1))}{\mathcal{P}^k(1)} \circ T^k.
$$

Recall that for $L^2(m)$-functions these conditional expectations are the orthogonal projections in $L^2(m)$.

We denote as above: $\varphi - m(\varphi \circ T^j)$ by $\varphi_j$ or $[\varphi]_j$. However, to simplify notation, it is convenient to assume that $\varphi_0 = [\varphi]_0 = 0$. Therefore we have for the centered sum (1.4):

$$
S_n = \sum_{k=1}^n \varphi_k \circ T^k = \sum_{k=0}^n \varphi_k \circ T^k.
$$

Introduce

$$
\mathbf{H}_n \circ T^n := \mathbb{E}(S_{n-1} \mid T^{-n}\mathcal{B}).
$$

Hence $\mathbf{H}_1 = 0$, and the explicit formula for $\mathbf{H}_n$ is

$$
\mathbf{H}_n = \frac{1}{\mathcal{P}^n 1} \left[ P_n(\varphi_{n-1}P^{n-1} 1) + P_n P_{n-1}(\varphi_{n-2}P^{n-2} 1) + \cdots + P_n P_{n-1} \cdots P_1(\varphi_0 P^{0} 1) \right].
$$

(2.1)

It is not hard to check that setting

$$
S_n = M_n + \mathbf{H}_{n+1} \circ T^{n+1}
$$

7
the sequence \( \{M_n\} \) is a reverse martingale for the decreasing filtration \( \{\mathcal{B}_n := \mathcal{T}^{-n}\} \):

\[
E(M_n \mid \mathcal{B}_{n+1}) = 0.
\]

In particular,

\[
M_n - M_{n-1} = \psi_n \circ \mathcal{T}^n \quad \text{with} \quad \psi_n := \varphi_n + H_n - H_{n+1} \circ T_{n+1}. \tag{2.2}
\]

We recall three lemmas from [14], stated in the current context:

**Lemma 2.1 ([14, Lemma 2.6])**

\[
\sigma_n^2 := E[(\sum_{i=1}^{n} \varphi_i \circ T^i)^2] = \sum_{i=1}^{n} E[\psi_i^2 \circ T^i] - \int H_1^2 + \int H_{n+1}^2 \circ \mathcal{T}^{n+1}
\]

(and \( H_1 = 0 \)).

To prove this Lemma we replace our \( H_n \) with \( \omega_n \) in [14].

**Lemma 2.2 ([14, proof of Lemma 3.3])** Let \( H_j^\varepsilon = H_j 1_{|H_j| \leq \varepsilon \sigma_n} \), where for simplicity of notation we have left out the dependence on \( n \). Then

\[
\int \left( \sum_{j=1}^{n} \psi_j \circ T^j \cdot H_{j+1}^\varepsilon \circ \mathcal{T}^{j+1} \right)^2 = \sum_{j=1}^{n} \int (\psi_j \circ T^j \cdot H_{j+1}^\varepsilon \circ \mathcal{T}^{j+1})^2
\]

The last formula in the proof of [14, Lemma 2.6] equivalently gives:

**Lemma 2.3**

\[
\sigma_n^2 = \sum_{i=1}^{n} E[\varphi_i^2 \circ T^i] + 2 \sum_{i=1}^{n} E[(H_i \varphi_i) \circ T^i]
\]

The following Lemma plays a crucial role all along this paper. In a slightly different form it was introduced and used in [18, Sect. 4], without a proof, and subsequently in [3]. We now give a detailed proof in a more general setting.

**Lemma 2.4** Assume given a \( C^1 \)-function \( \varphi : [0, 1] \to \mathbb{R} \) and \( h \in C_2 \). where the cone \( C_2 \) is defined with \( a > 1 \).
Denote by $X$ the function $X(x) = x$. If

$$\lambda \leq -|\varphi'|_\infty$$

$$\nu \geq -|\varphi + \lambda X|_\infty$$

$$\delta \geq \frac{a}{\alpha + 1}(|\varphi'|_\infty + |\lambda|)m(h)$$

$$\delta \geq \frac{a}{a - 1}|\varphi + \lambda X + \nu|_\infty m(h)$$

then

$$(\varphi + \lambda X + \nu) h + \delta \in C_2.$$

**Remark 2.5** It follows immediately that if $\varphi \in C^1([0,1])$ and $h \in C_2$ then we can use Theorem 1.2 and Proposition 1.3 to obtain decay of $P^f(\varphi h - m(\varphi h))$: consider $\Phi := (\varphi + \lambda X + \nu) h + \delta$, $\Psi := (\lambda X + \nu) h + \delta + m(\varphi h)$, with constants chosen according to Lemma 2.4 so that $\Phi, \Psi \in C_2$ (by definition, $m(\Phi) = m(\Psi)$), and write

$$P^f(\varphi \cdot h - m(\varphi \cdot h)) = P^f(\Phi - \Psi).$$

**Corollary 2.6** In particular, for a sequence $\omega_k \in C^1([0,1])$ with $\|\omega_k\|_{C^1} \leq K$ and $h_k \in C_2$ with $m(h_k) \leq M$ (e.g., $h_k := P^k(1)$), one can choose constants $\lambda$, $\nu$ and $\delta$ so that

$$(\omega_k + \lambda X + \nu) h_k + \delta, (\lambda X + \nu) h_k + \delta + m(\omega_k h_k) \in C_2 \quad \text{for all } k \geq 1$$

and therefore

$$\|P^n(\omega_k h_k - m(\omega_k h_k))\|_1 \leq C_{\alpha,K,M} n^{-\frac{1}{\alpha} + 1}(\log n)^{\frac{1}{\alpha}} \quad \text{for all } n \geq 1, k \geq 1,$$

where the constant $C_{\alpha,K,M}$ has an explicit expression in terms of $\alpha$, $K$ and $M$. Decay in $L^p$ now follows from Lemma 2.7: if $1 \leq p < 1/\alpha$ then

$$\|P^n(\omega_k h_k - m(\omega_k h_k))\|_p \leq C_{\alpha,K,M,p} n^{-\frac{1}{p\alpha} + 1} \quad \text{for all } n \geq 1, k \geq 1$$

(ignoring the log-correction, see Remark 1.4) where the constant on the right hand side depends now upon $p$ too.

**Proof of Lemma 2.4** Denote $\Phi := (\varphi + \lambda X + \nu) h + \delta$. There are three conditions for $\Phi$ to be in $C_2$. 
Φ nonnegative and decreasing. If λ ≤ −sup φ' and ν ≥ −inf(φ + λX) then φ + λX + ν is decreasing and nonnegative. Therefore Φ, is also decreasing (because h ∈ C2) and nonnegative provided δ ≥ 0.

ΦX^{1+α} increasing. For 0 < x < y ≤ 1, need

\[ [(φ(x) + λx + ν)h(x) + δ]^{x^{1+α}} ≤ [(φ(y) + λy + ν)h(y) + δ]^{y^{1+α}} \]

\[ \iff [φ(x) + λx + ν] ≤ [φ(y) + λy + ν] \frac{h(y)y^{α+1}}{h(x)x^{α+1}} + δ \left[ \frac{y^{α+1}}{x^{α+1}} - 1 \right] \frac{1}{h(x)} \]

Since hX^{α+1} ≥ 0 is increasing, 1 ≤ \frac{h(y)y^{α+1}}{h(x)x^{α+1}}, so it suffices to have

\[ φ(x) + λx + ν ≤ [φ(y) + λy + ν] + δ \left[ \frac{y^{α+1}}{x^{α+1}} - 1 \right] \frac{1}{h(x)} \]

\[ \iff δ ≥ -(φ(y) + λy + ν) - (φ(x) + λx + ν) \frac{h(x)}{\frac{y^{α+1}}{x^{α+1}} - 1}. \]

By the mean value theorem and using that α ≤ 1, y^{α+1} - x^{α+1} = (α + 1)ξ^α(y - x) ≥ (α + 1)x^α(y - x) ≥ (α + 1)x(y - x); therefore

\[ 0 ≤ \frac{h(x)}{\frac{y^{α+1}}{x^{α+1}} - 1} = \frac{h(x)x^{α+1}}{y^{α+1} - x^{α+1}} ≤ \frac{h(x)x^α}{(α + 1)(y - x)} ≤ \frac{am(h)}{(α + 1)(y - x)}. \]

Meanwhile,

\[ -(φ(y) + λy + ν) - (φ(x) + λx + ν) ≤ (|φ'|_∞ + |λ|)(y - x). \]

Using these in the above lower bound for δ, we conclude that it suffices to have

\[ δ ≥ \frac{a}{α + 1} (|φ'|_∞ + |λ|) m(h) \]

\[ ΦX^α ≤ am(Φ). Using that hX^α ≤ am(h), \]

\[ [(φ + λX + ν)h + δ]X^α ≤ (φ + λX + ν)hX^α + δ ≤ sup(φ + λX + ν)am(h) + δ. \]

On the other hand, am((φ + λX + ν)h + δ) ≥ a inf(φ + λX + ν)m(h) + aδ, so it suffices to have

\[ sup(φ + λX + ν)am(h) + δ ≤ a inf(φ + λX + ν)m(h) + aδ \]

\[ \iff δ ≥ a \frac{a}{a - 1} \left[ sup(φ + λX + ν) - inf(φ + λX + ν) \right] m(h). \]
Note that, since the transfer operators are monotone,
\[ |P_n \ldots P_{k+1}[\varphi P^k 1]|_x \leq P_n \ldots P_{k+1}[|\varphi|_\infty P^k 1]|_x = |\varphi|_\infty P_n \ldots P_{k+1}[P^k 1]|_x. \]

Since $|\varphi|_\infty P_n \ldots P_{k+1}[P^k 1]$ lies in the cone $C_2$ this implies that
\[ |P_n \ldots P_{k+1}[\varphi P^k 1]|_x \leq a|\varphi|_\infty x^{-\alpha}. \]

The following Lemma gives control over the $L^p$-norm of functions with such a bound.

**Lemma 2.7** Suppose that $f \in L^1(m)$ and $|f(x)| \leq Kx^{-\alpha}$. Then, provided $p \geq 1$ and $\alpha p < 1$,
\[ \|f\|_p \leq C_{\alpha,p}\|f\|_{\frac{1-\alpha p}{1-\alpha}} K^{\frac{p-1}{p-\alpha}}. \]

In particular, if $|f(x)| \leq Kx^{-\alpha}$ and $\|f\|_1 \leq Mn^{1-\frac{1}{\alpha}}$, then
\[ \|f\|_p \leq C_{K,M,\alpha,p}n^{1-\frac{1}{\alpha}} \text{ for } 1 \leq p < 1/\alpha. \]

Therefore, for $1 \leq p < 1/(2\alpha)$, there is $\delta > 0$ such that $\|f\|_p \leq C_{K,M,\alpha,p}n^{1-\delta}$.

**Proof** The case $p = 1$ is obviously true, so we assume from now on that $p > 1$. Denote $C_1 := \|f\|_1$. Compute, for $0 < x_* \leq 1$, and $\alpha p < 1$:
\[ \int_0^{x_*} |f|^p dx \leq \sup\{|f(x)|^{p-1} | x_\ast \leq x \leq 1\} \int_0^{x_*} |f|^p dx \leq K^{p-1}x_*^{-\alpha(p-1)}C_1, \text{ and } \int_0^{x_*} |f|^p dx \leq K^p \int_0^{x_*} x^{-\alpha p}dx = \frac{K^p x_*^{1-\alpha p}}{1-\alpha p}. \]

We want to minimize over $x_*$ the quantity
\[ G(x_*) := K^p C_1 x_*^{-\alpha(p-1)} + K^p \frac{1}{1-\alpha p} x_*^{1-\alpha p} = Ax_*^{-\alpha(p-1)} + Bx_*^{1-\alpha p}. \]

It reaches its minimum value for $x_*^{\alpha-1} = \frac{B(1-\alpha p)}{A\alpha(p-1)}$, which gives for the minimum of $G^{1/p}$ the value
\[ C_{\alpha,p} C_1^{1-\alpha p} K^{\frac{p-1}{1-\alpha}}. \]

For the last statement notice that $\frac{1-p\alpha}{p\alpha} > 1 \iff 0 < \alpha p < 1/2$. 

**Corollary 2.8** We have:

1. $\|H_n\|_q$ is uniformly bounded in $n$ for $1 \leq q < \frac{1}{\alpha}$. 

11
2. \( ||H_n \circ T^n||_r \) is uniformly bounded in \( n \) for \( 1 \leq r < \frac{1}{2\alpha} - \frac{1}{2} \).

**Proof** Recall that \( H_n \) is given in (2.1). By [3, Remark 1.3], \( P^n(1) \geq D_\alpha > 0 \) on (0,1]. We now apply Minkowski’s inequality in the sum defining \( H_n \). Thanks to Lemma 2.7 each term of the form \( P_{n}P_{n-1} \ldots P_{n-\ell}(\varphi_{n-\ell-1}P^{n-\ell-1}) \), \( \ell \in [0,n-1] \) will be bounded in \( L^p \) by \( \frac{2}{2^\alpha} C_{\alpha,K,p} \ell^{1-\frac{1}{p}} \), where \( K \) is the \( C^1 \) norm of \( \varphi \). The role of \( h_k \) in Lemma 2.6 is now played by \( P^{n-\ell-1} \) and therefore \( M = 1 \). By summing over \( \ell \) from 1 to infinity, we get a convergent series whenever \( p\alpha < 1/2 \). We now write \( \int |H_n \circ T^n|^r \, dx = \int |H_n|^r P^n 1 \, dx \). Since \( P^n 1 \) belongs to \( L^p(m) \) for \( 1 \leq p < \frac{1}{\alpha} \) by the definition of \( C_2 \) and its invariance property, the function \( |H_n|^r P^n 1 \) must be uniformly in \( L^1(m) \) and therefore, by the previous item, \( r \frac{L}{p-1} < \frac{1}{2\alpha} \). Thus we need \( 1 \leq r < \frac{1}{2\alpha} - \frac{1}{2} \).

As we said in the Introduction, we will also have a pointwise bound on the \( H_n \)'s.

**Lemma 2.9** For \( 0 < \alpha < 1/2 \), there is a constant \( C \) depending on \( \alpha \) and \( K = ||\varphi||_{C^1} \), such that

\[
|H_n(x)| \leq Cx^{-\alpha-1} \quad \text{for all } x \in (0,1], \, n \geq 1. \quad (2.3)
\]

**Proof** By using again formula (2.1) for \( H_n \) (where \( \varphi_0 = 0 \)) and the bound \( P^n(1) \geq D_\alpha > 0 \) we are left with the pointwise estimate of

\[
P_n(\varphi_{n-1}P^{n-1}) + P_nP_{n-1}(\varphi_{n-2}P^{n-2}) + \ldots + P_nP_{n-1} \ldots P_1(\varphi_0P^0 1).
\]

By Corollary 2.6, for each \( k \geq 1 \) one can write \( \varphi_k P^k 1 = (\varphi - m(\varphi \circ T^k))P^k 1 = A_k - B_k \) where \( A_k, B_k \in C_2 \) with \( m(A_k), m(B_k) \) uniformly bounded by some constant \( C_{\alpha,K} < \infty \). Therefore, by the decay Theorem 1.2 (and ignoring the log-correction), there is a new constant \( C' \) depending only on \( \alpha \) and \( K \) such that

\[
||P_{k+1}^n(A_k - B_k)||_1 \leq C'(n-k)^{-\frac{1}{\alpha}+1}. \quad (2.4)
\]

We now recall the footnote to the proof of [18, Lemma 2.3]: if \( f \in C_2 \) with \( m(f) \leq M \) then

\[
|x^{\alpha+1} f(x) - y^{\alpha+1} f(y)| \leq a(1 + \alpha)M|x - y| \quad \text{for } 0 < x, y \leq 1. \quad (2.5)
\]

But a bound \( |g(x) - g(y)| \leq L|x - y| \) for the Lipschitz-seminorm \( |g|_{\text{Lip}} \) implies

\[
\|g\|_1 \geq C_L \|g\|_\infty. \quad (2.6)
\]
Combining the above observations and since \( m(\mathcal{P}^{n-k}_{k+1}(f)) = m(f) \), we obtain that 
\[
|X^{\alpha+1}p^{n-k}_{k+1}(A_k - B_k)_{Lip} - X^{\alpha+1}p^{n-k}_{k+1}(A_k)_{Lip} + X^{\alpha+1}p^{n-k}_{k+1}(B_k)_{Lip} | \leq L \text{ uniformly for } \ n \geq 1, \ 1 \leq k < n, \text{ and then }
\]

\[
\|X^{\alpha+1}p^{n-k}_{k+1}(A_k - B_k)\|_\infty \leq \frac{1}{C} \|X^{\alpha+1}p^{n-k}_{k+1}(A_k - B_k)\|_1 \leq C''(n-k)^{-\frac{1}{\alpha} + 1}
\]

for a new constant \( C'' \) depending only on \( \alpha, K, L, \) which implies that

\[
|\mathcal{P}^{n-k}_{k+1}(A_k - B_k)(x)| \leq x^{-\alpha} C''(n-k)^{-\frac{1}{\alpha} + 1}
\]

and therefore, for \( 0 < \alpha < 1/2 \),

\[
\left| \sum_{k=1}^{n-1} \mathcal{P}^{n-k}_{k+1}(A_k - B_k)(x) \right| \leq x^{-\alpha} C'' \sum_{k=1}^{n-1} (n-k)^{-\frac{1}{\alpha} + 1} \leq C x^{-\alpha}
\]

as desired.

We finish this Section by proving a type of Borel-Cantelli Lemma which is an unavoidable tool in proving non-stationary limit theorems.

**Theorem 2.10 (Strong Borel-Cantelli)** Suppose that for \( j \geq 1, \psi_j \in C^1([0, 1]) \) with uniformly bounded \( C^1 \)-norms.

(a) If \( 0 < \alpha < 1/2 \) then

\[
\sum_{j=1}^{n} \psi_j(T^j) - \sum_{j=1}^{n} m(\psi_j(T^j)) = O(n^{1/2}(\log \log n)^{3/2}) \text{ m-a.e.}
\]

and therefore, if \( \lim \inf_j m(\psi_j \circ T^j) > 0 \) then

\[
\frac{\sum_{j=1}^{n} \psi_j(T^j x)}{\sum_{j=1}^{n} m(\psi_j \circ T^j)} \to 1 \text{ m-a.e. x.}
\]

(b) If \( 0 < \alpha < 1 \) then

\[
\frac{1}{n} \left[ \sum_{j=1}^{n} \psi_j(T^j x) - \sum_{j=1}^{n} m(\psi_j \circ T^j) \right] \to 1 \text{ m-a.e. x.}
\]

**Proof** To prove the first statement in part (a) we will use the Gál-Koksma Theorem 6.1 in the Appendix. By adding the same constant to all the \( \psi_j \)'s, we can assume without loss of generally that \( \inf_j m(\psi_j \circ T^j) > 0 \). Thus, it suffices to give a linear upper bound for
\[ \mathbb{E}[(\sum_{j=1}^n \psi_j(T^j) - b_n)^2] \], where \( b_n := \sum_{j=1}^n m(\psi_j(T^j)) \); note that the same estimate can be derived for sums over \( m \leq j \leq n \). Expand

\[
\mathbb{E}[(\sum_{j=1}^n \psi_j \circ T^j - b_n)^2] = \sum_{j=1}^n \mathbb{E}[\psi_j \circ T^j - m(\psi_j \circ T^j)]^2 \\
+ 2 \sum_{i=1}^n \sum_{j>i}^n \mathbb{E}[(\psi_j \circ T^j - m(\psi_j \circ T^j))(\psi_i \circ T^i - m(\psi_i \circ T^i))]
\]

and use the decay to estimate the mixed terms. Denote \( \overline{\psi}_j = \psi_j - m(\psi_j \circ T^j) \). Then, for \( j > i \),

\[
|\mathbb{E}[(\psi_j(T^j) - m(\psi_j(T^j))) \cdot (\psi_i(T^i) - m(\psi_i(T^i))]| = |\mathbb{E}[\overline{\psi}_j \cdot T^j \cdot \overline{\psi}_i \circ T^i]| \\
= |\mathbb{E}[(\overline{\psi}_j \cdot T_{i+1}^{j-i} \cdot \overline{\psi}_i \cdot P^i(1))]| = |\mathbb{E}[(\overline{\psi}_j \cdot P_{i+1}^{j-i} \cdot \overline{\psi}_i \cdot P^i(1))]| \\
\leq \|\overline{\psi}_j\|_\infty \|P_{i+1}^{j-i} \cdot \overline{\psi}_i \cdot P^i(1)\|_1 \leq C(j-i)^{1-\frac{1}{\alpha}}
\]

where in the last inequality we used Corollary 2.6. Therefore

\[
\mathbb{E}[(\sum_{j=1}^n \psi_j(T^j) - b_n)^2] \\
\leq 2 \sum_{i=1}^n |(\psi_j(T^j) - m(\psi_j(T^j)))|_\infty m(\psi_i(T^i)) + 2C \sum_{i=1}^n \sum_{j>i}^n (j-i)^{1-\frac{1}{\alpha}} \leq nC',
\]

where the constants \( C, C' \) are independent of \( j \) and \( n \). The conclusion now follows from the Gál-Koksma Theorem 6.1.

For (b), note that for \( 1/2 \leq \alpha < 1 \) the above computation still gives

\[ \mathbb{E}[(\sum_{j=1}^n \psi_j(T^j) - b_n)^2] \leq Cn^{3-\frac{1}{\alpha}} \]

which implies that

\[ \sum_{j=1}^n \psi_j(T^j) - b_n = O(n^{1-\eta}) \text{ a.s.} \]

for some \( \eta > 0 \), see the standard Lemma 2.11.

\[ \textbf{Lemma 2.11} \] Assume the random variables \( X_n \) have mean zero, and there are \( M < \infty, \gamma < 2 \) such that

\[ \|X_n\|_\infty \leq M, \quad \text{Var} \left( \sum_{k=1}^n X_k \right) \leq Cn^\gamma \quad \text{for all } n. \]
Then
\[ \sum_{k=1}^{n} X_k = O(n^\eta) \text{ a.s. for } \eta > \frac{\gamma + 1}{3}. \]

Proof Denote \( S_n := \sum_{k=1}^{n} X_k \). From Tchebycheff’s inequality,
\[ P(|S_n| > n^{1-\delta}) \leq \frac{\text{Var}(S_n)}{(n^{1-\delta})^2} \leq C n^{\gamma - 2\delta - 2}. \]

Pick \( \delta > 0 \) so that \( \gamma - 2\delta - 2 < 0 \) and \( \omega > 0 \) such that \( \omega(2 - \gamma + 2\delta) > 1 \). Then, for the subsequence \( n_k := k^\omega \),
\[ \sum_{k} P(|S_{n_k}| > n_k^{1-\delta}) < \infty \]
so, by Borel-Cantelli,
\[ |S_{n_k}| = O(n_k^{1-\delta}) \text{ a.s.} \quad (2.7) \]
Using (2.7), one has a.s.: if \( n_k \leq n < n_{k+1} \) for some \( k \), then
\[ |S_n| \leq |S_{n_k}| + [n_{k+1} - n_k] \sup \|X_\ell\|_\infty \leq O(n_k^{1-\delta}) + Ck^{\omega-1}M \leq O(n^{1-\delta}) + C(n^{1/\omega})^{\omega-1}M \]
therefore \( |S_n| = O(n^\eta) \text{ a.s. with} \)
\[ \eta = \max \left\{ 1 - \delta, \frac{\omega - 1}{\omega} \right\}. \]
Optimize over \( \delta \) and \( \omega \) to get the claimed lower bound on \( \eta \).

3 Central Limit Theorem

We assume in this section that \( 0 < \alpha < 1/2 \) (note that in the stationary case the CLT holds only in this range). With our approach we can only prove the non-stationary CLT for a lower upper bound on \( \alpha \), which will be stated later.

We define scaling constants \( \sigma_n^2 = \mathbb{E}[(\sum_{j=1}^{n} \varphi_j \circ T_j)^2] \). This sequence of constants play the role of non-stationary variance. As we pointed out in the Introduction, giving estimates on the growth and non-degeneracy of \( \sigma_n \) in this non-stationary setting is more difficult than in the usual stationary case.
Theorem 3.1 (CLT for \(C^1\) functions) \(\) Let \(\varphi\) be a \(C^1([0,1])\) function, and define \(S_n\) as in (1.4),
\[
S_n := \sum_{k=1}^{n} \varphi_k \circ T_{\beta_k} \circ \cdots \circ T_{\beta_1}. 
\]
Assume that
\[
\sigma_n^2 := \text{Var}(S_n) = E[(\sum_{i=1}^{n} \varphi_i \circ T_i)^2] \approx n^\beta. 
\]
Then, provided \(\alpha < 1/8\) and \(\beta > 2/3\) (see (3.6) for other cases),
\[
\frac{S_n}{\sigma_n} \rightarrow^d \mathcal{N}(0,1). 
\]

Following the approach of Gordin we will express \(S_n = \sum_{j=1}^{n} \varphi_j \circ T_j\) as the sum of a (non-stationary) martingale difference array and a controllable error term and then use the following Theorem from Conze and Raugi [6, Theorem 5.8], which is a modification of a result of B. M. Brown [5] from martingale differences to reverse martingale differences.

**Theorem 3.2 ([6, Theorem 5.8])** Let \((X_i, \mathcal{F}_i)\) be a sequence of differences of square integrable reversed martingales, defined on a probability space \((\Omega, \mathcal{F}, \mathcal{P})\). For \(n \geq 0\) let
\[
S_n = X_0 + \cdots + X_{n-1}, \quad \sigma_n^2 = \sum_{k=0}^{n-1} E[X_k^2], \quad V_n = \sum_{k=0}^{n-1} E[X_k^2 | \mathcal{F}_{k+1}]. 
\]
Assume the following two conditions hold:

(i) the sequence of random variables \((\sigma_n^{-2} V_n)_{n \geq 1}\) converges in probability to 1.

(ii) For each \(\varepsilon > 0\), \(\lim_{n \to \infty} \sigma_n^{-2} \sum_{k=0}^{n-1} E[X_k^2 \mathbf{1}_{\{|X_k| > \varepsilon \sigma_n\}}] = 0.\)

Then
\[
\lim \sup_{n \to \infty} \alpha \in \mathbb{R} \left| P \left[ \frac{S_n}{\sigma_n} < a \right] - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{x^2}{2}} \, dx \right| = 0. 
\]

**Proof of Theorem 3.1**

Let us take the quantity \(H_n\) defined in (2.1) and then the function \(\psi_n\) given in (2.2)
\[
\psi_n := \varphi_n + H_n - H_{n+1} \circ T_{n+1}. 
\]
We note that \(\psi_n \circ T^n\) is a reverse martingale difference scheme, uniformly bounded in \(L^{r_1}(m)\), for some \(r_1\) verifying the second item in Corollary 2.8; in particular we will take \(r_1\) as the exponent for which \(H_{n+1} \circ T^{n+1}\) is bounded in \(L^{r_1}(m)\). That is, \(1 \leq r_1 < \frac{1}{2\alpha} - \frac{1}{2} \).
We will verify conditions (i) and (ii) of Theorem 3.2. For condition (ii) we begin by noticing that the functions \( \psi_n \circ T^n \) have a uniformly bounded \( L^2 \)-norm if the same is true for \( H_{n+1} \circ T_{n+1} \); this holds provided \( 2 < \frac{1}{\alpha} - \frac{1}{2} \iff 0 < \alpha < \frac{1}{3} \). By Minkowski’s inequality, \( \| \psi_n \circ T^n \|_{L^2(m)} \) will be bounded uniformly in \( n \) by some constant \( \hat{C} \). Then we have by Hölder’s and Tchebycheff’s inequality
\[
\sigma_n^{-2} \sum_{k=0}^{n-1} \mathbb{E}[\psi_k^2 \mathbf{1}_{\{ |\psi_k| > \varepsilon \sigma_n \}}] \leq \sigma_n^{-2} \hat{C} \sum_{k=0}^{n-1} m( |\psi_k| > \varepsilon \sigma_n ) \leq \sigma_n^{-2} \hat{C}^2 \frac{n}{\varepsilon \sigma_n}.
\]
We note at this point that by prescribing a growth of the variance as \( \sigma_n^2 \approx n^\beta \) we need \( \beta > 2/3 \).

The hard part lies in establishing (i). This is in contrast with the stationary setting where condition (i) is usually a straightforward consequence of the ergodic theorem.

Once we have established (i) and (ii) it follows that \( \lim_{n \to \infty} \frac{1}{\sigma_n} \sum_{j=1}^{n} \psi_j \circ T^j \to N(0,1) \) in distribution. Finally, since \( \sum_{j=1}^{n} \varphi_j \circ T^j = H_{n+1} \circ T^{n+1} \) is bounded in \( L^r, r \geq 2 \) (Corollary 2.8), \( \lim_{n \to \infty} \frac{1}{\sigma_n} \sum_{j=1}^{n} \varphi_j \circ T^j \to N(0,1) \) in distribution as well.

For (i), we first prove that
\[
\frac{1}{\sigma_n^2} \sum_{j=1}^{n} \psi_j^2 \circ T^j \to 1 \quad \text{in probability as } n \to \infty.
\]
and then show that in our setting this implies (i) (see Theorem 3.5).

We follow [14, Lemma 3.3 and proof of Theorem 3.1 (II)], which uses an argument of Peligrad [19]. Since \( \psi_j = \varphi_j + H_j - H_{j+1} \circ T_{n+1} \),
\[
\psi_j^2 = \varphi_j^2 + 2 \varphi_j H_j + H_j^2 + H_{j+1}^2 \circ T_{n+1} - 2 H_{j+1} \circ T_{n+1} (\varphi_j + H_j) = \varphi_j^2 + 2 \varphi_j H_j + H_j^2 + H_{j+1}^2 \circ T_{n+1} - 2 H_{j+1} \circ T_{n+1} (\varphi_j + H_{j+1} \circ T_{n+1}) = \varphi_j^2 + (H_j^2 - H_{j+1}^2 \circ T_{n+1}) - 2 \varphi_j \cdot H_{j+1} \circ T_{n+1} + 2 \varphi_j H_j.
\]

Therefore
\[
\sum_{j=1}^{n} \psi_j \circ T^j = (H_1^2 \circ T_1 - H_{n+1}^2 \circ T_{n+1}) - \left[ \sum_{j=1}^{n} \psi_j \circ T^j \cdot H_{j+1} \circ T^{j+1} \right] + \left[ \sum_{j=1}^{n} \varphi_j^2 \circ T^j \right] + 2 \left[ \sum_{j=1}^{n} (\varphi_j \cdot H_j) \circ T^j \right].
\]
By the $L^r$ uniform boundedness of $H_n \circ T^n$ (Corollary 2.8), $\frac{1}{\sigma_n^2} H_{n+1}^2 \circ T^{n+1} \to 0$ in probability.

Next we show that

$$\frac{1}{\sigma_n^2} \left[ \sum_{j=1}^{n} \psi_j \circ T^j \cdot H_{j+1} \circ T^{j+1} \right] \to 0 \text{ in probability.} \quad (3.1)$$

Define

$$H_j^\varepsilon := H_j 1_{\{|H_j| \leq \varepsilon \sigma_n\}}.$$ 

By Lemma 2.2,

$$U_n^2 := \int \left( \sum_{j=1}^{n} [\psi_j \circ T^j \cdot H_j^\varepsilon \circ T^{j+1}] \right)^2 = \int \sum_{j=1}^{n} [\psi_j \circ T^j \cdot H_j^\varepsilon \circ T^{j+1}]^2.$$

Hence, using Lemma 2.1 for the equal below,

$$U_n^2 \leq \varepsilon^2 \sigma_n^2 \sum_{j=1}^{n} \int \psi_j^2 \circ T^j$$

$$= \varepsilon^2 \sigma_n^2 \left[ \int (\sum_{j=1}^{n} \varphi_j \circ T^j)^2 + \int H_1^2 \circ T^1 - \int H_{n+1}^2 \circ T^{n+1} \right] \leq \varepsilon^2 \sigma_n^4. \quad (3.2)$$

For any $a > \varepsilon$ we obtain, using Tchebycheff’s inequality in the third and fourth lines below, the inequality (3.2), and our uniform $L^r$ bound on $H_j \circ T^j$ (Corollary 2.8), given by the constant $\hat{D}$

$$m \left( \left| \frac{1}{\sigma_n^2} \sum_{j=1}^{n} \psi_j \circ T^j \cdot H_{j+1} \circ T^{j+1} \right| > a \right)$$

$$\leq m \left( \max_{1 \leq j \leq n} |H_{j+1} \circ T^{j+1}| > \varepsilon \sigma_n \right) + m \left( \left| \frac{1}{\sigma_n^2} \sum_{j=1}^{n} \psi_j \circ T^j \cdot H_j^\varepsilon \circ T^{j+1} \right| > a \right)$$

$$\leq \sum_{j=1}^{n} m(|H_{j+1} \circ T^{j+1}| > \varepsilon \sigma_n) + \frac{1}{a^2 \sigma_n^4} U_n^2$$

$$\leq \frac{n}{(\varepsilon \sigma_n)^r} \left( \max_{1 \leq j \leq n} \int |H_{j+1} \circ T^{j+1}|^r \right) + \frac{\varepsilon^2}{a^2} \leq \frac{n \hat{D}}{(\varepsilon \sigma_n)^r} + \frac{\varepsilon^2}{a^2}.$$

Take $a = \sqrt{\varepsilon}$; if we use that $\sigma_n^2 \approx n^\beta$, then $\beta > \frac{2}{r}$ with $1 \leq r < \frac{1}{2\alpha} - \frac{1}{2}$, that is $\beta > \frac{4\alpha}{1-\alpha}$, allows us to obtain (3.1). We defer to the end of this proof the discussion about the possible choices for $\alpha, \beta$. 

18
Finally, we show that
\[
\frac{1}{\sigma_n^2} \sum_{j=1}^{n} (\varphi_j^2 + 2\varphi_j H_j) \circ \mathcal{T}^j \to 1 \quad \text{in probability.} \tag{3.3}
\]

We know from our Strong Borel-Cantelli Theorem \ref{thm:strong-borel-cantelli} that
\[
\sum_{j=1}^{n} \varphi_j^2 \circ \mathcal{T}^j = \sum_{j=1}^{n} \mathbb{E}[\varphi_j^2 \circ \mathcal{T}^j] + o(n^{1+\varepsilon}) \quad \text{m-a.e.} \tag{3.4}
\]

We will show in Lemma \ref{lem:sum-bound} that
\[
\frac{1}{\sigma_n^2} \left( \sum_{j=1}^{n} (\varphi_j H_j) \circ \mathcal{T}^j - \sum_{j=1}^{n} \mathbb{E}[(\varphi_j H_j) \circ \mathcal{T}^j] \right) \to 0 \quad \text{in probability.} \tag{3.5}
\]

In view of Lemma \ref{lem:bound}, equations (3.3) and (3.5) imply
\[
\frac{1}{\sigma_n^2} \left[ \sum_{j=1}^{n} \varphi_j^2 \circ \mathcal{T}^j + 2 \sum_{j=1}^{n} (\varphi_j H_j) \circ \mathcal{T}^j \right] \to 1 \quad \text{in probability.} \tag{3.6}
\]

\textbf{Lemma 3.3} For \(\alpha < 1/8\) and the variance growing as \(\sigma_n^2 \approx n^\beta, \beta > 2/3\), we have:
\[
\frac{1}{\sigma_n^2} \left( \sum_{j=1}^{n} (\varphi_j H_j) \circ \mathcal{T}^j - \sum_{j=1}^{n} \mathbb{E}[(\varphi_j H_j) \circ \mathcal{T}^j] \right) \to 0 \quad \text{in probability.} \tag{3.7}
\]

\textbf{Proof} Write \(S_n = \sum_{j=1}^{n} (\varphi_j H_j) \circ \mathcal{T}^j\) and \(E_n = \sum_{j=1}^{n} \mathbb{E}[(\varphi_j H_j) \circ \mathcal{T}^j]\) and estimate
\[
\mathbb{E}(|S_n - E_n| > \sigma_n^2 \varepsilon) = \mathbb{E}(|S_n - E_n|^2 > \sigma_n^4 \varepsilon^2) \leq \frac{1}{\sigma_n^4 \varepsilon^2} \mathbb{E}(|S_n - E_n|^2).
\]

When we estimate \(\mathbb{E}(|S_n - E_n|^2)\) we have, as usual, the diagonal terms and a double summation of off-diagonal terms:
\[
\mathbb{E}(|S_n - E_n|^2) = \sum_{j=1}^{n} \mathbb{E}([(\varphi_j H_j) \circ \mathcal{T}^j - m((\varphi_j H_j) \circ \mathcal{T}^j)^2])
+ 2 \sum_{j=1}^{n} \sum_{i=1}^{j-1} \int [(\varphi_j H_j) \circ \mathcal{T}^j - m((\varphi_j H_j) \circ \mathcal{T}^j)][(\varphi_i H_i) \circ \mathcal{T}^i - m((\varphi_i H_i) \circ \mathcal{T}^i)]dx.
\]

The sum of diagonal terms is \(O(n)\) as \((\varphi_j H_j) \circ \mathcal{T}^j \in L^2(m)\) with uniformly bounded norm if \(\alpha < 1/5\).
We note that by prescribing a growth of the variance as $\sigma_n^2 \approx n^\beta$, the exponent $\beta$ must verify $\beta > 1/2$.

We now consider
\[
\sum_{j=1}^{n} \sum_{i=1}^{j-1} \int \left[ (\varphi_j H_j) \odot T^j - m((\varphi_j H_j) \odot T^j) \right] \left[ (\varphi_i H_i) \odot T^i - m((\varphi_i H_i) \odot T^i) \right] dx
\]
\[
= \sum_{j=1}^{n} \sum_{i=1}^{j-1} \int \left[ \varphi_j H_j - m((\varphi_j H_j) \odot T^j) \right] T^j \cdot \left[ \varphi_i H_i - m((\varphi_i H_i) \odot T^i) \right] T^i dx
\]
\[
= \sum_{j=1}^{n} \sum_{i=1}^{j-1} \int \left[ \varphi_j H_j - m((\varphi_j H_j) \odot T^j) \right] \cdot P_i^{j-i} \left[ H_i \varphi_i P^i 1 - m((\varphi_i H_i) \odot T^i) P^i 1 \right] dx.
\]

We will prove in Lemma 3.4 below that $||P_i^{j-i} [P^i 1 \varphi_i - P^i (\varphi_i H_i) T^i]||_2 \leq \frac{C^*}{(j-i)^{\alpha^*}}$, where $C^*$ is a constant depending only on $\alpha$ and the $C^1$ norm of $\varphi$ (and uniform in $i$ and $j$).

Here the numerator $i$ comes about as $1 \leq i \leq j - 1$, and $\alpha^* = \frac{1-p}{2\alpha}$ follows from the decay Theorem 1.2 and Lemma 2.7, provided $\alpha < 1/2$. Note also that $||(\varphi_j H_j) - m((\varphi_j H_j) \odot T^j)||_2$ is uniformly bounded in $j$ provided $\alpha < 1/4$, see Corollary 2.8.

We have to show that each row summation satisfies
\[
\left| \sum_{i=1}^{j-1} \int \left[ (\varphi_j H_j) - m((\varphi_j H_j) \odot T^j) \right] P_i^{j-i} \left[ P^i 1 \varphi_i - P^i (\varphi_i H_i) T^i \right] dx \right| \leq j^{1+\chi}
\]
where $n^{1+\chi} = o(\sigma_n^4)$ otherwise the double summation contributes a term which is too large.

So we divide the sum into two parts, with $0 < \delta < 1$
\[
\sum_{j=1}^{j-\delta} \int \left[ (\varphi_j H_j) - m((\varphi_j H_j) \odot T^j) \right] P_i^{j-i} \left[ P^i 1 \varphi_i - P^i (\varphi_i H_i) T^i \right] dx
\]
\[
+ \sum_{i=1}^{j-\delta} \int \left[ (\varphi_j H_j) - m((\varphi_j H_j) \odot T^j) \right] P_i^{j-i} \left[ P^i 1 \varphi_i - P^i (\varphi_i H_i) T^i \right] dx.
\]

The first sum we bound by $C^* j^\delta$ using $L^2$ bounds without decay. The second uses our decay estimate (see Lemma 3.4) and we get $\sum_{i=1}^{j-\delta} \frac{C^*}{(j-i)^{\alpha^*}} \leq C^* j^{1-(\alpha^* - 1)\delta} = C^* j^{1+\delta-\alpha^* \delta}$ provided
\(\alpha^* > 1 \iff 0 < \alpha < 1/2\). Then \(|\sum_{i=1}^{j-1} \int [((\varphi_j H_j) - m((\varphi_j H_j) \circ T_j)) \mathcal{P}_{i+1}^j i \varphi_i - \mathcal{P}_{i+1}^j i m((\varphi_i H_i) \circ T_i)] dx| \leq C(j^\delta + j^{1+\delta - \alpha \delta})\) which is lowest for \(\delta = 1/\alpha^*\). We obtain

\[
|\sum_{j=1}^{n} \sum_{j=1}^{n-1} \int [((\varphi_j H_j) \circ T_j - m((\varphi_j H_j) \circ T_j))((\varphi_i H_i) \circ T_i - m((\varphi_i H_i) \circ T_i))] dx| \leq C^* n^{1+1/\alpha^*} = C^* n^{1/(1-2\alpha)}
\]

so

\[
E(|S_n - E_n|^2) \leq C n^{1/(1-2\alpha)}.
\]

By dividing for \(\sigma_n^4\) and asking again for a growth like \(\sigma_n^2 \approx n^\beta\) we have now that \(\beta > \frac{1}{2(1-2\alpha)}\). This estimate allows us to show that \(\frac{1}{\sigma_n^2} \left( \sum_{j=1}^{n} ((\varphi_j H_j) \circ T_j - \sum_{j=1}^{n} E[((\varphi_j H_j) \circ T_j)] \right) \to 0\) in probability.

We now collect the various inequalities involving \(\beta\), which is the scaling of \(\sigma_n^2 \approx n^\beta\), and \(\alpha\):

- for our proof of condition (ii) in Brown’s Theorem 3.2 we need \(\beta > \frac{2}{3}\) and \(\alpha < \frac{1}{5}\);
- in Peligrad’s argument we have \(\beta > \frac{4\alpha}{1-\alpha}\);
- in Lemma 3.3, using that \(\alpha < \frac{1}{5}\), we have \(\beta > \frac{1}{2}\) and \(\beta > \frac{1}{2(1-2\alpha)}\).

These give

\[
\alpha < \frac{1}{5}, \quad \beta > \max \left\{ \frac{2}{3}, \frac{4\alpha}{1-\alpha}, \frac{1}{2(1-2\alpha)} \right\}
\]

which are all satisfied if \(\alpha < \frac{1}{5}, \beta > \frac{2}{3}, \) or \(\alpha < \frac{1}{5}, \beta \geq 1\).

To conclude the proof we need the statement of Lemma 3.4, whose proof is in the Appendix, and of Theorem 3.5, which allows us to get the convergence in probability of the conditional expectations from condition (i) in Brown’s Theorem.

**Lemma 3.4** For \(1 \leq p < 1/\alpha\)

\[
\|\mathcal{P}_k^n \left( [\mathcal{P}_{i}^1 H_i \varphi_i - \mathcal{P}_{i}^1 m((\varphi_i H_i) \circ T_i)] \right) \|_p \leq i C_{\alpha,p} C_{\varphi} n^{-\frac{1}{p\alpha} + 1} (\log n)^{\frac{1}{\alpha} - \frac{1}{p\alpha}}
\]

**Theorem 3.5** The following inference holds:

\[
\frac{1}{\sigma_n^2} \sum_{j=1}^{n} \psi_n^2 \circ T^n \to^p 1 \implies \frac{1}{\sigma_n^2} \sum_{j=1}^{n} E[\psi_n^2 \circ T^n | B_{n+1}] \to^p 1.
\]
**Proof** To do this we will use Burckholder’s inequality (Theorem 2.10 of [8]).

We will show that

\[ \frac{1}{\sigma^2_n} \sum_{j=1}^{n} (\psi_n^2 \circ T^n - \mathbb{E}[\psi_n^2 \circ T^n | B_{n+1}]) \to 0 \quad \text{in probability.} \]

First define \( V_n = \psi_n^2 \circ T^n - \mathbb{E}[\psi_n^2 \circ T^n | B_{n+1}] \) and note that \( \mathbb{E}[V_n | B_{n+1}] = 0 \).

We define a martingale, reading from left to right,

\[ S_1 = V_n, S_2 = V_n + V_{n-1}, V_n + V_{n-1} + V_{n-2} + ... + V_1 \]

with filtration

\[ F_0 = B_{n+1}, F_1 = B_n, F_2 = B_{n-1}, ..., F_n = B_0 = B. \]

Then \( V_n \) is \( F_1 \) measurable as \( \psi_n^2 \circ T^n \) is \( B_n \) measurable, since \( \mathbb{E}[\psi_n^2 \circ T^n | B_{n+1}] \) is \( B_{n+1} \) measurable and \( B_{n+1} \subset B_n \ E[\psi_n^2 \circ T^n | B_{n+1}] \) is \( F_1 \) measurable. Similarly \( V_i \) is \( F_{n-i+1} \) measurable. This implies \( S_i \) is \( F_i \) measurable.

Note that \( \mathbb{E}[V_{n-1} | F_1] = \mathbb{E}[V_{n-1} | B_n] = 0 \) so

\[ \mathbb{E}[S_{i+1} | F_i] = \mathbb{E}[V_{n-i} | F_i] + S_i = S_i. \]

Hence \((S_i, F_i)\) is a martingale.

By Burckholder’s inequality taking \( p = 2 \) we have

\[ \mathbb{E}|S_n|^2 \leq C_1 \mathbb{E}(\sum_{j=1}^{n} V_i^2) \leq C_2 \sigma_n^2 \]

where \( C_2 \) is a universal constant.

Hence \( P(|S_n| > \sigma_n \varepsilon) = P(|S_n|^2 > \sigma_n^4 \varepsilon^2) \leq \frac{C_2}{\varepsilon^2 \sigma_n^2} \) by Chebyshev.

\[ \Box \]

4 Central Limit Theorem for nearby maps

**Theorem 4.1** Given \( \beta \in (0, 1/5) \) and \( \phi \in C^1([0, 1]) \) if \( \phi \) is not a coboundary (up to a constant) for \( T_\beta \) there exists \( \varepsilon > 0 \) such that for all parameters \( \beta_k \in (\beta - \varepsilon, \beta + \varepsilon) \) the variance grows linearly for any sequential system formed from concatenation of the maps \( T_{\beta_k} \).

Therefore, by Theorem 3.1 and (3.6), the CLT holds.
Proof

Recall the quantities defined by a concatenation of different maps.

\[ H_n = \frac{1}{P_n} \left[ P_n(\varphi_{n-1}P^{n-1}1) + P_nP_{n-1}(\varphi_{n-2}P^{n-2}1) + \cdots + P_1P_{n-1}P_1(\varphi_0P^01) \right] \]

and

\[ \psi_n := \varphi_n + H_n - H_{n+1} \circ T_{n+1}. \]

First assume that the maps all coincide with \( T_\beta \) so that \( P_\beta^n1 \to h_\beta \) (at a polynomial rate in \( L^2 \)), \( P_nP_{n-1} \cdots P_{n-k} = P_\beta^k \), where \( h_\beta \) is the invariant density for \( T_\beta \) and \( P_\beta \) is the transfer operator for \( T_\beta \) with respect to Lebesgue measure. Furthermore \( \varphi_n = \varphi - m(\varphi(T_\beta^n)) \to \varphi - \int \varphi h_\beta dx \). Denote the \( H_n \) corresponding to this situation by \( H_{\beta,n} \).

Note the terms \( P_nP_{n-1} \cdots P_{n-j}(\varphi_{n-j-1}P^{n-j-1}1) \) decay at a polynomial rate in \( L^2 \), \( ||P_nP_{n-1} \cdots P_{n-j}(\varphi_{n-j-1}P^{n-j-1}1)||_2 \leq \frac{C}{\tau^j} \) for some \( \tau > 1 \) for \( \beta < 1/4 \), by Proposition 1.3 and Lemma 2.4. Note that \( C \) and \( \tau \) may be taken as uniform over all \( T_{\beta_k} \) if \( \beta_k \) is close to \( \beta \).

Combining this with the fact that \( P_\beta^n1 \to h_\beta \) in \( L^2 \) (and hence \( \frac{1}{P_\beta^n1} \to \frac{1}{h_\beta} \) in \( L^2 \) as both \( h_\beta \) and \( P_\beta^n1 \) are bounded below by \( \delta > 0 \), we see that given \( \varepsilon > 0 \) there exists an \( N \) such that for all \( n > N \), \( H_{\beta,n} = \frac{1}{h_\beta} \left[ P_\beta(h_\beta \varphi - \int \varphi h_\beta dx) + P_\beta^2(h_\beta \varphi - \int \varphi h_\beta dx) + \cdots + P_\beta^N(h_\beta \varphi - \int \varphi h_\beta dx) \right] + \gamma(\beta,n) \) where \( ||\gamma(\beta,n)||_2 < \varepsilon \). We define \( G_{\beta,N} = \frac{1}{h_\beta} \left[ P_\beta(h_\beta \varphi - \int \varphi h_\beta dx) + P_\beta^2(h_\beta \varphi - \int \varphi h_\beta dx) + \cdots + P_\beta^N(h_\beta \varphi - \int \varphi h_\beta dx) \right] \) so that \( H_{\beta,n} = G_{\beta,N} + \gamma(\beta,n) \).

Now suppose \( \varphi \) is not a coboundary for \( T_\beta \). Denote by \( \bar{P}_\beta \) the transfer operator for \( T_\beta \) with respect to the invariant measure \( d\mu_\beta = h_\beta dx \). Then \( \bar{P}_\beta^n(\varphi) = \frac{1}{h_\beta} P_\beta^n(h_\beta \varphi) \) where \( P_\beta \) is the transfer operator for \( T_\beta \) with respect to Lebesgue measure.

Hence \( \frac{1}{h_\beta} \left[ P_\beta(h_\beta \varphi - \int \varphi h_\beta dx) + P_\beta^2(h_\beta \varphi - \int \varphi h_\beta dx) + \cdots + P_\beta^N(h_\beta \varphi - \int \varphi h_\beta dx) \right] = \sum_{k=1}^N \bar{P}_\beta^k(\varphi - \int \varphi d\mu_\beta) \). If \( \varphi \) is not a coboundary then \( \sum_{k=1}^\infty \bar{P}_\beta^k(\varphi - \int \varphi d\mu_\beta) \) converges to a coboundary \( \bar{H}_\beta \) so that

\[ \varphi = \bar{\psi}_\beta + \bar{H}_\beta \circ T_\beta - \bar{H}_\beta \]
defines a martingale difference sequence \( \{\bar{\psi}_\beta \circ T_\beta^n\} \), where \( \bar{\psi}_\beta \neq 0 \) in \( L^2 \) (as \( \varphi \) is not a coboundary for \( T_\beta \)). Suppose \( ||\bar{\psi}_\beta||_2 > \eta \).

Choose \( N \) large enough that for all \( n > N \), \( ||H_{\beta,n} - H_{\beta,n+1} \circ T_\beta||_2 < \frac{\eta}{20} \) and \( ||\bar{H}_\beta - \sum_{k=1}^N \bar{P}_\beta^k(\varphi - \int \varphi d\mu_\beta)||_2 < \frac{\eta}{20} \). Then \( ||\psi(\beta,n)||_2 > \frac{\eta}{2} \) for all \( n > N \).

Now we consider a concatenation of maps \( T_{\beta_k} \) where \( \beta_k \) is close to \( \beta \). The idea is to break \( H_n \) into a sum of \( N \) terms uniformly close to \( G(\beta,N) \) (no matter what the sequence of maps) and a small error.
Choose all $\beta_k$’s sufficiently close to $\beta$ that when we form a concatenation of the maps $T_{\beta_k}$ we have

$$\|G_{\beta,N} = \frac{1}{p^n_1} \left[ P_n(\varphi_{n-1}p^{n-1}1) + P_nP_{n-1}(\varphi_{n-2}p^{n-2}1) + \ldots \right.$$ $$+ P_nP_{n-1} \cdots P_{n-N}(\varphi_{n-N-1}p^{n-N-1}1) \right] \|_2 < \frac{\eta}{20}.$$  

We can do this as we have fixed $N$ and the finite terms are continuous in $L_2$ as $\beta_k \to \beta$, see [16, Theorem 5.1] and Lemmas 2.4, 2.7.

Recall we also have $\|\gamma(\beta,n)\|_2 < \frac{\eta}{20}$ for all $n \geq N$.

Using the uniform contraction ($\tau$ and $C$ are uniform for $T\beta$ where $\beta$ is in a small neighborhood of $\beta$) we have

$$\|H_n = \frac{1}{p^n_1} \left[ P_n(\varphi_{n-1}p^{n-1}1) + P_nP_{n-1}(\varphi_{n-2}p^{n-2}1) + \ldots \right.$$ $$+ P_nP_{n-1} \cdots P_{n-N}(\varphi_{n-N-1}p^{n-N-1}1) \right] \|_2 < \frac{\eta}{20}$$

for all $n > N$. Then $\|\psi_n\|_2 > \frac{\eta}{10}$ for all $n > N$ and we have linear growth of variance for the concatenation of maps as $\sigma_n^2 = \sum_{k=1}^{n} E[\psi_n \circ T^k]^2$.

5 Random compositions of intermittent maps

Suppose $S = \{T_{\alpha_1}, \ldots, T_{\alpha_l}\}$ is a finite number of intermittent type maps as in Section 1, with $\alpha_i < \frac{1}{4}$. We will take an iid selection of maps from $S$ according to a probability vector $p = (p_1, \ldots, p_l)$ where the probability of choosing map $T_{\alpha_i}$ is $p_i$. This induces a Bernoulli measure $\nu$ on the shift space $\Omega := \{1, \ldots, l\}^N$, where $(i_1, i_2, \ldots, i_n, \ldots)$ corresponds to the sequence of maps: first apply $T_{\alpha_{i_1}}$, then $T_{\alpha_{i_2}}$ and so on. Writing elements of $\omega \in \Omega$ as sequences $\omega := (\omega_0, \omega_1, \ldots, \omega_n, \ldots)$ the shift operator $S : \Omega \to \Omega$, $(S\omega)_i = \omega_{i+1}$ preserves the measure $\nu$.

This random system also induces a Markov process on $[0,1]$ with the transition probability function $P(x, A) = \sum_{i=1}^{l} p_{\alpha_i} 1_A(T_{\alpha_i}(x))$. A measure $\mu$ is invariant for the Markov process if $P^\ast \mu = \mu$. In this setting Bahsoun and Bose [4] have shown (among other results) that there is a unique absolutely continuous invariant measure $\mu$ and that if $\varphi : [0,1] \to \mathbb{R}$ is a Hölder function then $\varphi$ satisfies an annealed CLT for this random dynamical system in
the sense that if \( \int \varphi d \mu = 0 \) then
\[
(\nu \times \mu) \{(\omega, x) : \frac{1}{n} \sum_{j=1}^{n} \varphi(T(S_j)\omega_0 \ldots T(\omega_{j-1})x) \in A\} \rightarrow \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{-\frac{x^2}{2\sigma^2}} dx
\]
for some \( \sigma^2 \geq 0 \). In fact the result of Bahsoun and Bose [4] also shows that this convergence is with respect to \((\nu \times m)\) where \( m \) is Lebesgue measure on \([0, 1]\).

This follows from a well known result by Eagleson [9] which states the equivalence of the convergence in distribution for measures which are absolutely continuous one with respect to the other.

We will show that almost every realization of choices of concatenations of maps, i.e. with respect to the product measure \( \nu \), satisfies a self-norming CLT if:

\((*)\) \( \varphi \) is not a coboundary for all maps i.e. there exists an \( i \) such that \( \varphi \neq \psi \circ T_{\alpha_i} \) for any measurable (hence Hölder by standard Livsic theory) function \( \psi \).

First we show that if we take a random composition of a finite number of intermittent type maps we obtain linear growth of the variance almost surely under assumption \((*)\).

**Lemma 5.1** If \( \varphi \) is not a coboundary for all maps, i.e. there exists an \( i \) such that \( \varphi \neq \psi \circ T_{\alpha_i} \) for any measurable \( \psi \), then for \( \nu \)-almost every sequence of maps \( T_j \)
\[
\sigma_n^2 := \int \left( \sum_{j=1}^{n} \varphi \circ T_i - m(\varphi \circ T_i) \right)^2 dx
\]
grows at a linear rate in that \( \sigma_n^2 \geq Cn \) for sufficiently large \( n \) for some \( C > 0 \).

**Proof**
Under our assumption \( \varphi \) is not a coboundary for one of the maps, say \( T_{\alpha_1} \).

We will construct a martingale decomposition using the transfer operator \( Q_{\alpha} \) corresponding to the invariant measure \( \mu_{\alpha} \) for \( T_{\alpha} \). The invariant measure \( \mu_{\alpha} \) has a density \( h_{\alpha} \).

The coboundary function is defined by \( H_{\alpha} = \sum_{j=1}^{\infty} Q_{\alpha}^j [\varphi - \int \varphi d\mu_{\alpha}] \) where \( Q_{\alpha} \) is the adjoint operator of the Koopman operator \( U_{\varphi} = \varphi \circ T_{\alpha} \) with respect to the invariant measure \( d\mu_{\alpha} = h_{\alpha} dx \) for \( T_{\alpha} \).

When we do the usual decomposition \( \varphi = \psi_{\alpha_1} + H_{\alpha_1} - H_{\alpha_1} \circ T_{\alpha_1} \) then the martingale difference function \( \psi_{\alpha_1} \) is bounded below from zero in \( L^2 \). Suppose \( \|\psi_{\alpha_1}\|_2 > \rho > 0 \).
It is known that $Q^n_{\alpha}(\varphi) = \frac{1}{h_{\alpha}}P^n_{\alpha}(h_{\alpha}\varphi)$ where $P_{\alpha}$ is the adjoint of the Koopman operator of $T_{\alpha}$ with respect to Lebesgue measure. Furthermore $P^n_{\alpha}1 \rightarrow h_{\alpha}$ (at a polynomial rate in $L^2$) and since $\Pi - j1$ lies in the cone $C_2$ and $\int \Pi_j1dx = 1$,

$$P^k_{\alpha}[h_{\alpha} - \Pi_j1] \rightarrow 0$$

in $L^2$ at a uniform polynomial rate, in fact $\|P^k_{\alpha}[h_{\alpha} - \Pi_j1]\|_2 \leq C\frac{1}{k^{1+\eta}}$ where $C$ and $\eta$ are uniform over $\Pi_j1$.

Now we consider the quantities defined by a concatenation of different maps. We will use the notation from previous sections.

$$H_n = \frac{1}{P^n_{\alpha}1}[P_n(\varphi_{n-1}P^{n-1}1) + P_nP_{n-1}(\varphi_{n-2}P^{n-2}1) + \cdots + P_nP_{n-1}\cdots P_1(\varphi_0P^01)]$$

and $\psi_n := \varphi_n + H_n - H_{n+1} \circ T_{n+1}$.

We will first consider what happens when we have a sequence of $k$ maps $T_{\alpha}$ applied one after the other. We will suppose we have concatenated $n$ maps and then apply $k$ $T_{\alpha}$ maps in turn.

Then $\varphi_{n+k} = \varphi - \int \varphi(T^k_{\alpha}T_{n}\cdots T_1)dx = \varphi - \int \varphi P^k_{\alpha} \Pi_n1dx = \varphi - \int \varphi h_{\alpha} dx + \int \varphi P^k_{\alpha}[h_{\alpha} - \Pi_n1]dx$ where $\|P^k_{\alpha}[h_{\alpha} - \Pi_n1]\|_2 \leq C\frac{1}{k^{1+\eta}}$.

We are considering here $n$ fixed and $k$ increasing.

Note the terms $P_nP_{n-1}\cdots P_{n-j}(\varphi_{n-j-1}P^{n-j-1}1)$ decay at a polynomial rate in $L^2$, in fact $\|P_nP_{n-1}\cdots P_{n-j}(\varphi_{n-j-1}P^{n-j-1}1)\|_2 \leq C\frac{1}{j^{1+\eta}}$. Note that $C$ and $\eta$ may be taken as uniform over all choices of $T_{\alpha}$ in the concatenation.

Combining this with the fact that $P^k_{\alpha}\Pi_n1 \rightarrow h_{\alpha}$ in $L^2$ (and hence $\frac{1}{P^n_{\alpha}1} \rightarrow \frac{1}{h_{\alpha}}$ in $L^2$ as both $h_{\alpha}$ and $P^k_{\alpha}\Pi_n1$ are bounded below by $\delta > 0$), we see that given $\rho > 0$ there exists an $r$ such that for all $m > n + rk$, $H_m = \frac{1}{h_{\alpha}}[P_{\alpha}(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx) + P^2_{\alpha}(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx) + \cdots + P^r_{\alpha}(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx)] + \gamma(m, \alpha)$ where $\|\gamma(m, \alpha)\|_2 < \frac{\rho}{20}$.

Now $\frac{1}{h_{\alpha}}[P(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx) + P^2(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx) + \cdots + P^k(h_{\alpha}\varphi - \int \varphi h_{\alpha} dx)] = \sum_{j=1}^k Q^j_{\alpha}[\varphi - \int \varphi d\mu_{\alpha}].$ The infinite sum $\sum_{j=1}^{\infty} Q^j_{\alpha}[\varphi - \int \varphi d\mu_{\alpha}]$ converges to $H_{\alpha}$ at a polynomial rate.

We choose $k$ large enough that $\|H_{\alpha} - \sum_{j=1}^{\infty} Q^j_{\alpha}[\varphi - \int \varphi d\mu_{\alpha}]\|_2 \leq \frac{\rho}{20}$.

Recall

$$\varphi = \psi_{\alpha} + H_{\alpha} \circ T_{\alpha} - H_{\alpha}$$

26
defines a martingale difference sequence \( \{ \psi_{\alpha_1} \circ T_{\alpha_1}^j \} \), where \( \psi_{\alpha_1} \) is bounded away from zero in \( L^2 \) (as \( \varphi \) is not a coboundary for \( T_{\alpha_1} \)). We assumed \( \| \psi_{\alpha_1} \|_2 > \rho \).

We have shown that if we choose \( k \) and \( r \) large enough then \( \| H_m - H_{\alpha_1} \|_2 < \frac{\rho}{10} \) for all \( m > n + rk \) and hence as

\[
\psi_m := \varphi_m + H_m - H_{m+1} \circ T_{m+1}
\]

we see that \( \| \psi_m - H_{\alpha_1} \|_2 \leq \frac{\rho}{5} \) and hence \( \| \psi_m \|_2 > \frac{\rho}{2} \).

This implies linear growth in the random composition setting as almost all choices of maps will have \( rk \) long sequences of the map \( T_{\alpha_1} \) at a fixed frequency. In fact the only way we won’t obtain linear growth almost surely is if the function \( \varphi \) is a coboundary for all the maps \( T_{\alpha_i} \).

The next theorem is an immediate consequence of the previous lemma and Theorem 3.1 (see (3.6) for the bound on \( \alpha \)).

**Theorem 5.2** If \( \alpha_i < 1/5 \) for all \( 1 \leq i \leq \ell \) and \( \varphi \) is not a coboundary for all maps then \( \sigma_n^2 \geq Cn \) for some \( C > 0 \) and hence \( \varphi \) satisfies a CLT for \( \nu \) almost every sequence of maps.

6 Appendices

6.1 Gál-Koksma Theorem.

We recall the following result of Gál and Koksma as formulated by W. Schmidt [20, 21] and stated by Sprindzuk [22]:

**Theorem 6.1** Let \((\Omega, \mathcal{B}, \mu)\) be a probability space and let \( f_k(\omega) \), \( (k = 1, 2, \ldots) \) be a sequence of non-negative \( \mu \) measurable functions and \( g_k, h_k \) be sequences of real numbers such that \( 0 \leq g_k \leq h_k \leq 1 \), \( (k = 1, 2, \ldots) \). Suppose there exists \( C > 0 \) such that

\[
\int \left( \sum_{m < k \leq n} (f_k(\omega) - g_k) \right)^2 d\mu \leq C \sum_{m < k \leq n} h_k
\]

for arbitrary integers \( m < n \). Then for any \( \varepsilon > 0 \)

\[
\sum_{1 \leq k \leq n} f_k(\omega) = \sum_{1 \leq k \leq n} g_k + O(\Theta^{1/2}(n) \log^{3/2+\varepsilon}(n))
\]

for \( \mu \text{-a.e. } \omega \in \Omega \), where \( \Theta(n) = \sum_{1 \leq k \leq n} h_k \).
6.2 Proof of Lemma 3.4

Proof For simplicity of notation we discuss only the case $k = 1$; the general case is the same, since we use the $n$ Perron-Frobenius maps in $\mathcal{P}_k^n$ only for the decay given by Theorem 1.2.

The idea is to write $[\mathcal{P}^i \mathbf{1}_k \varphi_i - \mathcal{P}^i \mathbf{1}_n ((\varphi_i \mathbf{H}_i) \circ \mathcal{T}^i)]$ as a difference of $2i$ functions in the cone of the same integral. By writing explicitly $\mathbf{H}_i$ we get

$$[\mathcal{P}^i \mathbf{1}_k \varphi_i - \mathcal{P}^i \mathbf{1}_n ((\varphi_i \mathbf{H}_i) \circ \mathcal{T}^i)] = \sum_{k=1}^{i} \prod_{j=0}^{k-1} P_{i-j}((\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1}) \varphi_i - \mathcal{P}^i \mathbf{1}_n ((\varphi_i \mathbf{H}_i) \circ \mathcal{T}^i)) =$$

$$\sum_{k=1}^{i} \prod_{j=0}^{k-1} P_{i-j}((\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1}) \varphi_i - \mathcal{P}^i \mathbf{1}_n ((\varphi_i \mathbf{H}_i) \circ \mathcal{T}^i))$$

$$\sum_{k=1}^{i} \left[ \varphi_i \mathcal{P}^k_{i-k+1}((\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1}) - \mathcal{P}^i \mathbf{1}_n ((\varphi_i \mathbf{H}_i) \circ \mathcal{T}^i)) \right].$$

Call $C_{k,i} := m((\varphi_i \mathcal{P}^k_{i-k+1}((\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1}) \circ \mathcal{T}^i))$; then consider the quantity

$$(*) := \varphi_i \mathcal{P}^k_{i-k+1}((\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1}) - \mathcal{P}^i \mathbf{1}_n (\mathcal{P}_k \circ \mathcal{T}^i)).$$

Since $\varphi_{i-k} \in C^1$ and $\mathcal{P}^{i-k} \mathbf{1} \in C_2$ we can write by Lemma 2.4

$$\varphi_{i-k} \mathcal{P}^{i-k} \mathbf{1} = F_{i-k} - G_{i-k}$$

with $F_{i-k}, G_{i-k} \in C_2$. By the invariance of the cone, the functions $h_{i-k}^{(1)} := \mathcal{P}^k_{i-k+1} F_{i-k}; h_{i-k}^{(2)} := \mathcal{P}^{k}_{i-k+1} G_{i-k}$ are still in the cone, and we rewrite $(*)$ as

$$(*) = \varphi_i h_{i-k}^{(1)} - \varphi_i h_{i-k}^{(2)} - C_{i,k} \mathcal{P}^i \mathbf{1}.$$

Although the functions (in the cone), $F_{i-k}, G_{i-k}$ are not of zero mean, we can still apply Lemma 2.4 and split the product of $\varphi_i$ with them into the differences of two new functions belonging to the cone, namely

$$\varphi_i h_{i-k}^{(1)} = M_{i-k}^{(1)} - M_{i-k}^{(2)}, \varphi_i h_{i-k}^{(2)} = N_{i-k}^{(1)} - N_{i-k}^{(2)}$$

with $M_{i-k}^{(1,2)}, N_{i-k}^{(1,2)} \in C_2$. We finally have

$$(*) = [M_{i-k}^{(1)} + N_{i-k}^{(2)}] - [M_{i-k}^{(2)} + N_{i-k}^{(1)} + C_{i,k} \mathcal{P}^i \mathbf{1}] := R_{i,k} - S_{i,k}$$
where the functions $R_{i,k}, S_{i,k}$ are in the cone and have the same expectation. Before continuing, let us summarize what we got

$$[\mathcal{P}^i 1 H_i \varphi_i - \mathcal{P}^i 1 m((\varphi_i H_i) \circ T^i)] = \sum_{k=1}^{i} (R_{i,k} - S_{i,k}).$$

By taking the power $\mathcal{P}^n$ on both sides we have by our Theorem 1.2 on the loss of memory and Proposition 1.3

$$\|\mathcal{P}^n ([\mathcal{P}^i 1 H_i \varphi_i - \mathcal{P}^i 1 m((\varphi_i H_i) \circ T^i)]) \|_p \leq \sum_{k=1}^{i} C_{\alpha,p} (\|R_{i,k}\|_1 + \|S_{i,k}\|_1) n^{-\frac{1}{p\alpha} + 1} (\log n)^{\frac{1}{\alpha p} - \frac{1}{\alpha p\alpha}}.$$

From Lemma 2.4, one observes that if we have $\varphi \in C^1([0,1])$ and $H \in C_2$ the splitting $\varphi H = A - B$, with $A, B \in C_2$ is such that the functions $A, B$ depend only on the $C^1$ norm of $\varphi$ and the integrals $m(H), m(\varphi H)$. In our case since $\varphi_i(x) = \varphi(x) - m(\varphi \circ T^i)$, we have that $\|\varphi_i\|_{C^1} \leq \|\varphi\|_{C^1}$; moreover, at each application of Lemma 2.4, the function $H$ is either $\mathcal{P}^i 1$ or obtained by applying $\mathcal{P}^\ell$ to a function obtained in the previous step and which only depends upon $\|\varphi\|_{C^1}$; in conclusion the norms $\|R_{i,k}\|_1, \|S_{i,k}\|_1$ are bounded by a function $C_\varphi$ which only depends on the choice of the observable $\varphi$. We finally get

$$\|\mathcal{P}^n (\mathcal{P}^i 1 H_i \varphi_i - m((\varphi_i H_i) \circ T^i)) \|_p \leq i C_{\alpha,p} C_\varphi n^{-\frac{1}{p\alpha} + 1} (\log n)^{\frac{1}{\alpha p} - \frac{1}{\alpha p\alpha}}.$$

**Acknowledgments**

SV was supported by the ANR-Project Perturbations and by the PICS (Projet International de Coopération Scientifique), Propriétés statistiques des systèmes dynamiques déterministes et aléatoires, with the University of Houston, n. PICS05968. SV thanks the University of Houston for supporting his visits during the preparation of this work. S.V thanks the Leverhulme Trust for support thorough the Network Grant IN-2014-021, and for useful conversations with S. Galatolo and W. Bahsoun. MN was supported by NSF grant DMS 1101315. AT was partially supported by the Simons Foundation grant 239583. The authors warmly thank R. Aimino for useful suggestions concerning Theorem 2.10.
References


