HAL
open science

# Almost sure invariance principle for sequential and non-stationary dynamical systems

Nicolai Haydn, Matthew Nicol, Andrew Török, Sandro Vaienti

## To cite this version:

Nicolai Haydn, Matthew Nicol, Andrew Török, Sandro Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. Transactions of the American Mathematical Society, 2017, 369 (8), pp. 5293-5316 10.1090/tran/6812 . hal-01127754

**HAL Id: hal-01127754**

**https://hal.science/hal-01127754v1**

Submitted on 8 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Almost sure invariance principle for sequential and non-stationary dynamical systems

Nicolai Haydn* Matthew Nicol † Andrew Török ‡
Sandro Vaienti §

August 17, 2014

## Contents

**1 Introduction** **2**

**2 Background and assumptions.** **7**

**3 ASIP for sequential expanding maps of the interval.** **8**

**4 Improvements of earlier work.** **10**

**5 ASIP for the shrinking target problem: expanding maps.** **11**

**6 ASIP for non-stationary observations on invertible hyperbolic systems.** **14**

*University of Southern California, Los Angeles. e-mail:<nhaysdn@math.usc.edu>.

†Department of Mathematics, University of Houston, Houston Texas, USA. e-mail: <nicol@math.uh.edu>.

‡Department of Mathematics, University of Houston, Houston Texas, USA. 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>.

**7 Further applications. 18**
7.1 $\beta$ transformations . . . . . 21
7.2 Perturbed expanding maps of the circle. . . . . 21
7.3 Covering maps: special cases . . . . . 24
7.3.1 One dimensional maps . . . . . 24
7.3.2 Multidimensional maps . . . . . 27
7.4 Covering maps: a general class . . . . . 29

**Abstract**

We establish almost sure invariance principles, a strong form of approximation by Brownian motion, for non-stationary time-series arising as observations on dynamical systems. Our examples include observations on sequential expanding maps, perturbed dynamical systems, non-stationary sequences of functions on hyperbolic systems as well as applications to the shrinking target problem in expanding systems.

# 1 Introduction

A recent breakthrough by Cuny and Merlevède [12] establishes conditions under which the almost sure invariance principle (ASIP) holds for reverse martingales. The ASIP is a matching of the trajectories of the dynamical system with a Brownian motion in such a way that the error is negligible in comparison with the Birkhoff sum. Limit theorems such as the central limit theorem, the functional central limit theorem and the law of the iterated logarithm transfer from the Brownian motion to time-series generated by observations on the dynamical system.

Suppose $\{U_j\}$ is a sequence of random variables on a probability space $(X, \mu)$ with $\mu(U_j) = 0$ for all $j$. We will say $(U_j)$ satisfies the ASIP if there is a sequence of independent centered Gaussian random variables $(Z_j)$ such that, enlarging our probability space if necessary,

$$\sum_{j=1}^{n} U_j = \sum_{j=1}^{n} Z_j + O(\sigma_n^{1-\gamma})$$

almost surely for some $\gamma > 0$ where

$$\sum_{j=1}^{n} E[Z_j^2] = \sigma_n^2$$

for some $0 < \eta < 1$.

If $(U_j)$ satisfies the ASIP then $(U_j)$ satisfies the (self-norming) CLT and

$$\frac{1}{\sigma_n} \sum_{j=1}^{n} U_j \to \mathrm{N}(0,1)$$

where the convergence is in distribution.

Furthermore if $(U_j)$ satisfies the ASIP then $(U_j)$ satisfies the law of the iterated logarithm and

$$\limsup_n [\sum_{j=1}^{n} U_j]/\sqrt{\sigma_n \log\log(\sigma_n)} = 1$$

while

$$\liminf_n [\sum_{j=1}^{n} U_j]/\sqrt{\sigma_n \log\log(\sigma_n)} = -1$$

In fact there is a matching of the Birkhoff sum $\sum_{j=1}^{n} U_j$ with a standard Brownian motion $B(t)$ observed at times $t_n = \sigma_n^2$ so that $\sum_{j=1}^{n} U_j = B(t_n)$ (plus negligible error) almost surely.

In the Gordin [14] approach to establishing the central limit theorem (CLT), reverse martingale difference schemes arise naturally.

We recall the definition of a reverse martingale difference scheme. Let $\mathcal{B}_i$, $i \geq 1$ be a decreasing sequence of $\sigma$-algebras, $\mathcal{B}_{i+1} \subset \mathcal{B}_i$. A sequence of square-integrable random variables $(X_i)$ is a (with respect to $(\mathcal{B}_i)$) if:

(1) $X_i$ is $\mathcal{B}_i$ measurable

(2) $E[X_i|\mathcal{B}_{i+1}] = 0$

If $X_i$ is a stationary reverse martingale difference scheme, under mild conditions, $\frac{1}{\sqrt{n}} \sum_{j=1}^{n} X_j$ satisfies the CLT.

To establish distributional limit theorems for stationary dynamical systems, such as the central limit theorem, it is possible to reverse time via the natural extension and use the martingale central limit theorem in backwards time to establish the CLT for the original system. This approach does not a priori work for the almost sure invariance principle, nor for other almost sure limit theorems. To circumvent this problem Melbourne and Nicol [24, 25] used results of Philipp and Stout [30] based upon the Skorokhod embedding theorem to establish the ASIP for Hölder functions on a class of non-uniformly hyperbolic systems, for example those modeled by Young Towers. Gouëzel [16] used spectral methods to give error rates in the ASIP for a wide class of dynamical systems, and his formulation does not require the assumption of a Young Tower. Rio and Merlevède [26] established the ASIP for a broader class of observations, satisfying only mild integrability conditions, on piecewise expanding maps of $[0, 1]$.

We will need the following theorem of Cuny and Merlevède:

**Theorem 1.1** *[12, Theorem 2.3] Let $(X_n)$ be a sequence of square integrable random variables adapted to a non-increasing filtration $(\mathcal{G}_n)_{n\in N}$. Assume that $E(X_n|\mathcal{G}_{n+1}) = 0$ a.s., that $\sigma_n^2 := \sum_{k=1}^n E(X_k^2) \to \infty$ and that $\sup_n E(X_n^2) < \infty$. Let $(a_n)_{n\in N}$ be a non-decreasing sequence of positive numbers such that $(a_n/\sigma_n^2)_{n\in N}$ is non-increasing and $(a_n/\sigma_n)_{n\in N}$ is non-decreasing. Assume that*

$$\textbf{(A)} \qquad \sum_{k=1}^{n} (E(X_k^2|\mathcal{G}_{k+1}) - E(X_k^2)) = o(a_n) \qquad P-a.s.$$

$$\textbf{(B)} \qquad \sum_{n\geq 1} a_n^{-v} E(|X_n|^{2v}) < \infty \quad \textit{for some } 1 \leq v \leq 2$$

*Then enlarging our probability space if necessary it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables with $E(Z_k^2) = E(X_k^2)$ such that*

$$\sup_{1\leq k\leq n} \left| \sum_{i=1}^{k} X_i - \sum_{i=1}^{k} Z_i \right| = o\big((a_n(|\log(\sigma_n^2/a_n)| + \log\log a_n))^{1/2}\big) \qquad P-a.s.$$

We use this result to provide sufficient conditions to obtain the ASIP for Hölder or BV observations on a large class of expanding sequential dynamical systems. We

also obtain the ASIP for some other classes of non-stationary dynamical systems, including ASIP limit laws for the shrinking target problem on a class of expanding maps and non-stationary observations on Axiom A dynamical systems.

In some of our examples the variance $\sigma_n^2$ grows linearly $\sigma_n^2 \sim n\sigma^2$ so that $S_n = \sum_{j=1}^n \phi_j \circ T^j$ is approximated by $\sum_{j=1}^n Z_j (= B(\sigma^2 n))$ where $Z_j$ are iid Gaussian all with variance $\sigma^2$ and $B(t)$ is standard Brownian motion. We will call this case a standard ASIP with variance $\sigma^2$.

In other settings, like the shrinking target problem, $\sigma_n^2$ does not grow linearly. In fact we don't know precisely its rate of increase, just that it goes to infinity. In these cases $S_n = \sum_{j=1}^n U_j$ is approximated by $\sum_{j=1}^n Z_j = B(\sigma_n^2)$ where the $Z_j$ are independent Gaussian but not with same variance, in fact $Z_j = B(\sigma_{j+1}^2) - B(\sigma_j^2)$ is a Brownian motion increment, the time difference (equivalently variance) of which varies with $j$.

Part of the motivation for this work is to extend our statistical understanding of physical processes from the stationary to the non-stationary setting, in order to better model non-equilibrium or time-varying systems. Non-equilibrium statistical physics is a very active field of research but ergodic theorists have until recently focused on the stationary setting. The notion of loss of memory for non-equilibrium dynamical systems was introduced and studied in the work of Ott, Stenlund and Young [28], but this notion only concerns the rate of convergence of initial distributions (in a metric on the space of measures) under the time-evolution afforded by the dynamics. In this paper we consider more refined statistics on a variety of non-stationary dynamical systems.

The term *sequential dynamical systems*, introduced by Berend and Bergelson [7], refers to a (non-stationary) system in which a sequence of concatenation of maps $T_k \circ T_{k-1} \circ \ldots \circ T_1$ acts on a space, where the maps $T_i$ are allowed to vary with $i$. The seminal paper by Conze and Raugi [11] considers the CLT and dynamical Borel-Cantelli lemmas for such systems. Our work is based to a large extent upon their work. In fact we show that the (non-stationary) ASIP holds under the same conditions as stated in [11, Theorem 5.1] (which implies the non-stationary CLT), provided a mild condition on the growth of the variance is satisfied.

We consider families $\mathcal{F}$ of non-invertible maps $T_\alpha$ defined on compact subsets $X$ of $\mathbb{R}^d$ or on the torus $\mathbb{T}^d$ (still denoted with $X$ in the following), and non-singular with respect to the Lebesgue or the Haar measure i.e. $m(A) \neq 0 \implies m(T(A)) \neq 0$. Such measures will be defined on the Borel sigma algebra $\mathcal{B}$. We will be mostly concerned with the case $d = 1$. We fix a family $\mathcal{F}$ and take a countable sequence of maps $\{T_k\}_{k\geq 1}$ from it: this sequence defines a *sequential dynamical system*. A *sequential orbit* of $x \in X$ will be defined by the concatenation

$$\mathcal{T}_n(x) := T_n \circ \cdots \circ T_1(x), \ n \geq 1 \tag{1.1}$$

We denote with $P_\alpha$ the Perron-Frobenius (transfer) operator associated to $T_\alpha$ defined by the duality relation

$$\int_M P_\alpha f \ g \ dm \ = \ \int_M f \ g \circ T_\alpha \ dm, \quad \text{for all } f \in \mathscr{L}^1_m, \ g \in \mathscr{L}^\infty_m$$

Note that here the transfer operator $P_\alpha$ is defined with respect to the reference measure $m$, in later sections we will consider the transfer operator defined by duality with respect to a natural invariant measure.

Similarly to (1.1), we define the composition of operators as

$$\mathcal{P}_n := P_n \circ \cdots \circ P_1, \ n \geq 1 \tag{1.2}$$

It is easy to check that duality persists under concatenation, namely

$$\int_M g(\mathcal{T}_n) \ f \ dm = \int_M g(T_n \circ \cdots \circ T_1) \ f \ dm \ = \ \int_M g( \ P_n \circ \cdots \circ P_1 f) \ dm \ = \int_M g \ (\mathcal{P}_n f) \ dm \tag{1.3}$$

To deal with probabilistic features of these systems, the martingale approach is fruitful. We now introduce the basic concepts and notations.

We define $\mathcal{B}_n := \mathcal{T}_n^{-1}\mathcal{B}$, the $\sigma$-algebra associated to the $n$-fold pull back of the Borel $\sigma$-algebra $\mathcal{B}$ whenever $\{T_k\}$ is a given sequence in the family $\mathcal{F}$. We set $\mathcal{B}_\infty = \bigcap_{n\geq 1} \mathcal{T}_n^{-1}\mathcal{B}$ the *asymptotic $\sigma$-algebra*; we say that the sequence $\{T_k\}$ is *exact* if $\mathcal{B}_\infty$ is trivial. We take $f$ either in $\mathscr{L}^1_m$ or in $\mathscr{L}^\infty_m$ whichever makes sense in the following expressions. It was proven in [11] that for $f \in \mathscr{L}^\infty_m$ the quotients $|\mathcal{P}_n f / \mathcal{P}_n 1|$ are bounded by $\|f\|_\infty$ on $\{\mathcal{P}_n 1 > 0\}$ and $\mathcal{P}_n f(x) = 0$ on the set $\{\mathcal{P}_n 1 = 0\}$, which allows

us to define $|\mathcal{P}_n f/\mathcal{P}_n 1| = 0$ on $\{\mathcal{P}_n 1 = 0\}$. We therefore have, the expectation being taken w.r.t. the Lebesgue measure:

$$\mathbb{E}(f|\mathcal{B}_k) \ = (\frac{\mathcal{P}_k f}{\mathcal{P}_k 1}) \circ \mathcal{T}_k \tag{1.4}$$

$$\mathbb{E}(\mathcal{T}_l f|\mathcal{B}_k) \ = (\frac{P_k \cdots P_{l+1}(f\mathcal{P}_l 1)}{\mathcal{P}_k 1}) \circ \mathcal{T}_k, \ 0 \leq l \leq k \leq n \tag{1.5}$$

Finally the martingale convergence theorem ensures that for $f \in \mathscr{L}^1_m$ there is convergence of the conditional expectations $(\mathbb{E}(f|\mathcal{B}_n))_{n\geq 1}$ to $\mathbb{E}(f|\mathcal{B}_\infty)$ and therefore

$$\lim_{n\to\infty} ||(\frac{\mathcal{P}_n f}{\mathcal{P}_n 1}) \circ \mathcal{T}_n - \mathbb{E}(f|\mathcal{B}_\infty)||_1 = 0,$$

the convergence being $m$-a.e.

# 2 Background and assumptions.

In [11] the authors studied extensively a class of $\beta$ transformations. We consider a similar class of examples and we will also provide some new examples for the theory developed in the next section. For each map we will give as well the properties needed the prove the ASIP; in particular we require two assumptions which we call, following [11], the (DEC) and (MIN) conditions.

To introduce them we first need to choose a suitable couple of adapted spaces. Due to the class of maps considered here, we will consider a Banach space $\mathcal{V} \subset \mathscr{L}^1_m$ $(1 \in \mathcal{V})$ of functions over $X$ with norm $||\cdot||_\alpha$, such that $\|\phi\|_\infty \leq C\|\phi\|_\alpha$.

For example we could let $\mathcal{V}$ be the Banach space of bounded variation functions over $X$ with norm $||\cdot||_{BV}$ given by the sum of the $\mathscr{L}^1_m$ norm and the total variation $|\cdot|_{bv}$. or we could take $\mathcal{V}$ to be the space of Lipschitz or Hölder functions with the usual norm.

**Property (DEC):** Given the family $\mathcal{F}$ there exist constants $\hat{C} > 0, \hat{\gamma} \in (0,1)$, such that for any $n$ and any sequence of operators $P_n, \cdots, P_1$ in $\mathcal{F}$ and any $f \in \mathcal{V}$ of zero

(Lebesgue) mean[1], we have

$$\|P_n \circ \cdots \circ P_1 f\|_\alpha \leq \hat{C}\hat{\gamma}^n \|f\|_\alpha \tag{2.1}$$

**Property (MIN)**: There exists $\delta > 0$ such that for any sequence $P_n, \cdots, P_1$ in $\mathcal{F}$ we have the uniform lower bound

$$\inf_{x \in M} P_n \circ \cdots \circ P_1 1(x) \geq \delta, \quad \forall n \geq 1. \tag{2.2}$$

# 3 ASIP for sequential expanding maps of the interval.

In this section we show that with an additional growth rate condition on the variance the assumptions of [11, Theorem 5.1] imply not just the CLT but the ASIP as well. We write $E[\phi]$ for the expectation of $\phi$ with respect to Lebesgue measure.

Let $\mathcal{V}$ be a Banach space with norm $\|.\|_\alpha$ such that $\|\phi\|_\infty \leq C\|\phi\|_\alpha$. If $(\phi_n)$ is a sequence in $\mathcal{V}$ define $\sigma_n^2 = E(\sum_{i=1}^n \tilde{\phi}_i(T_i \cdots T_1))^2$ where $\tilde{\phi}_n = \phi_n - m(\phi(T_n \cdots T_1))$.

**Theorem 3.1** *Let $(\phi_n)$ be a sequence in $\mathcal{V}$ such that $\sup_n \|\phi_n\|_\alpha < \infty$. Assume (DEC) and (MIN) and $\sigma_n \geq n^{1/4+\delta}$ for some $0 < \delta < \frac{1}{4}$. Then $(\phi_n \circ \mathcal{T}_n)$ satisfies the ASIP i.e. enlarging our probability space if necessary it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables $Z_k$ such that for any $\beta < \delta$*

$$\sup_{1 \leq k \leq n} |\sum_{i=1}^k \tilde{\phi}_i(T_i \cdots T_1) - \sum_{i=1}^k Z_i| = o(\sigma_n^{1-\beta}) \qquad m - a.s.$$

*Furthermore $\sum_{j=1}^n E[Z_i^2] = \sigma_n^2 + O(\sigma_n)$.*

**Proof** As above let $\mathcal{P}_n = P_n P_{n-1} \cdots P_1$ and define as in [11] the operators $Q_n \phi = \frac{P_n(\phi \mathcal{P}_{n-1} 1)}{\mathcal{P}_n 1}$.

[1]Actually, the definition of the (DEC) property in [11] is slightly more general since it requires the above property for functions in a suitable subspace, not necessarily that of functions with zero expectation.

It follows by a direct calculation that particular $Q_n T_n \phi = \phi$ and $E[\phi \circ \mathcal{T}_k | \mathcal{B}_k] = (Q_k \phi)(\mathcal{T}_k)$, so that $Q_k \phi = 0$ implies $E[\phi \circ \mathcal{T}_k | \mathcal{B}_k] = 0$.

With $h_n$ defined by

$$h_n = Q_n \tilde{\phi}_{n-1} + Q_n Q_{n-1} \tilde{\phi}_{n-2} + \cdots + Q_n Q_{n-1} \cdots Q_1 \tilde{\phi}_0$$

we then obtain that

$$\psi_n = \tilde{\phi}_n + h_n - T_{n+1} h_{n+1}$$

satisfies $Q_{n+1}\psi_n = 0$. For convenience let us put $U_n = \mathcal{T}_n \psi_n$, where, as before, $\mathcal{T}_n = T_n \circ \cdots \circ T_1$. Thus, as established by Conze and Raugi [11], $(U_n)$ is a sequence of reversed martingale differences for the filtration $(\mathcal{B}_n)$. Note that

$$\sum_{j=1}^{n} U_j = \sum_{j=1}^{n} \tilde{\phi}_j(\mathcal{T}_j) + h_1(\mathcal{T}_1) - h_n(\mathcal{T}_{n+1}) \tag{3.1}$$

and $\|h_n\|_\alpha$ is uniformly bounded. Hence

$$\begin{aligned} \left(\sum_{j=1}^{n} U_j\right)^2 &= \left(\sum_{j=1}^{n} \tilde{\phi}_j(\mathcal{T}_j)\right)^2 + (h_1(\mathcal{T}_1) - h_{n+1}(\mathcal{T}_{n+1}))^2 \\ &\quad + 2\left(\sum_{j=1}^{n} \tilde{\phi}_j(\mathcal{T}_j)\right)(h_1(\mathcal{T}_1) - h_{n+1}(\mathcal{T}_{n+1})) \end{aligned}$$

and integration yields

$$E\left(\sum_{j=1}^{n} U_j\right)^2 = \sigma_n^2 + \mathcal{O}(\sigma_n),$$

where we used that $h_n$ is uniformly bounded in $\mathscr{L}^\infty$ (and $\sigma_n \to \infty$). Since $\int U_j U_i = 0$ if $i \neq j$ one has $\sum_{j=1}^{n} E(U_j^2) = E\left(\sum_{j=1}^{n} U_j\right)^2 = \sigma_n^2 + \mathcal{O}(\sigma_n)$.

In Theorem 1.1, we will take $a_n$ to be $\sigma_n^{2-\epsilon}$, for some $\epsilon > 0$ sufficiently small ($\epsilon < 2\delta$ will do) so that $a_n^2 > n^{1/2+\delta'}$ for all large enough $n$, where $\delta' > 0$. Then $a_n/\sigma_n^2$ is non-increasing and $a_n/\sigma_n$ is non-decreasing. Furthermore Conze and Raugi show that $E[U_k^2 | \mathcal{B}_{k+1}] = \mathcal{T}_{k+1}(\frac{P_{k+1}(\psi_k^2 \mathcal{P}_k 1)}{\mathcal{P}_{k+1} 1})$ and in [11, Theorem 4.1] establish that

$$\int [\sum_{k=1}^{n} E(U_k^2 | \mathcal{B}_{k+1}) - E(U_k^2)]^2 \, dm \leq c_1 \sum_{k=1}^{n} E(U_k^2) \leq c_2 \sigma_n^2$$

for some constants $c_1, c_2 > 0$. This implies by the Gal-Koksma theorem (see e.g. [33]) that

$$\sum_{k=1}^{n} E(U_k^2|\mathcal{B}_{k+1}) - E(U_k^2) = o(\sigma_n^{1+\eta}) = o(a_n)$$

$m$ a.s. for any $\eta \in (0, 2-\varepsilon)$. Thus with our choice of $a_n$ we have verified Condition (A) of Theorem 1.1. Taking $v = 2$ in Condition (B) of Theorem 1.1 one then verifies that $\sum_{n\geq 1} a_n^{-v} E(|U_n|^{2v}) < \infty$.

Thus $U_n$ satisfies the ASIP with error term $o(\sigma_n^{1-\beta})$ for any $\beta < \delta$. This concludes the proof since in view of (3.1) and the fact that the terms involving $h_n$ telescope and $\|h_n\|_\alpha$ is uniformly bounded, the ASIP for $U_n$ implies the ASIP for $(\tilde{\phi}_j \circ \mathcal{T}_j)$. ∎

# 4 Improvements of earlier work.

We collect here examples for which a self-norming CLT was already proven, but actually a (self-norming) ASIP holds if the variance grows at the rate required by Theorem 3.1.

Conze and Raugi [11, Remark 5.2] show that for sequential systems formed by taking maps near a given $\beta$-transformation with $\beta > 1$, by which we mean maps $T_{\beta'}$ with $\beta' \in (\beta - \delta, \beta + \delta)$ for sufficiently small $\delta > 0$, the conditions (DEC) and (MIN) are satisfied and if $\phi$ is not a coboundary for $T_\beta$ then the variance for $\phi \in BV$ grows as $\sqrt{n}$.

Nándori, Szász and Varjú [27, Theorem 1] give conditions under which sequential systems satisfy a self-norming CLT. These conditions include (DEC) and (MIN) (the maps all preserve a fixed measure $\mu$, so one can use the transfer operator with respect to $\mu$), and their main condition gives the rate of growth for the variance (see [27, page 1220]). If this rate satisfies the requirement of Theorem 3.1, then for such systems the ASIP holds as well. Such cases follow from their Examples 1 and 2, where the maps are selected from the family $T_a(x) = ax (\mathrm{mod}\, 1)$, $a \geq 2$ integer, and Lebesgue as the invariant measure. Note however that their Example 2 includes sequential systems whose variance growth slower than any power of $n$, but still satisfy the self-norming CLT.

# 5 ASIP for the shrinking target problem: expanding maps.

We now consider a fixed expanding map $(T, X, \mu)$ acting on the unit interval equipped with a unique ergodic absolutely continuous invariant probability measure $\mu$. Examples to which our results apply include $\beta$-transformations, smooth expanding maps, the Gauss map, and mixing Rychlik-type maps. We will define the transfer operator with respect to the natural invariant measure $\mu$, so that $\int (Pf)g\, d\mu = \int fg(T)\, d\mu$ for all $f \in \mathscr{L}^1(\mu)$, $g \in \mathscr{L}^\infty(\mu)$.

We assume that the transfer operator $P$ is quasi compact in the bounded variation norm so that we have exponential decay of correlations in the bounded variation norm and $\|P^n\phi\|_{BV} \leq C\theta^n\|\phi\|_{BV}$ for all $\phi \in BV(X)$ such that $\int \phi d\mu = 0$ (here $C > 0$ and $0 < \theta < 1$ are constants independent of $\phi$).

We say that $(T, X, \mu)$ has exponential decay in the BV norm versus $\mathscr{L}^1(\mu)$ if there exist constants $C > 0$, $0 < \theta < 1$ so that for all $\phi \in BV$, $\psi \in \mathscr{L}^1(\mu)$ such that $\int \phi\, d\mu = \int \psi\, d\mu = 0$:

$$\left| \int \phi\psi \circ T^n\, d\mu \right| \leq C\theta^n \|\phi\|_{BV} \|\psi\|_1$$

where $\|\psi\|_1 = \int |\psi|\, d\mu$. If the density $\frac{d\mu}{dm}$ of $\mu$ with respect to Lebesgue measure $m$ is strictly bounded below, $\frac{d\mu}{dm} \geq c > 0$ for some constant $c$ then $\|P^n\phi\|_{BV} \leq C\theta^n\|\phi\|_{BV}$ implies exponential decay in the BV norm versus $\mathscr{L}^1(\mu)$ [**?**, Proposition 13]. Suppose $\phi_j = 1_{A_j}$ are indicator functions of a sequence of nested intervals $A_j$, where $\mu$ is the unique invariant measure for the map $T$.

The variance is given by $\sigma_n^2 = \mu(\sum_{i=1}^n \tilde{\phi}_i \circ T^i)^2$, where $\tilde{\phi} = \phi - \mu(\phi)$ and $E_n = \sum_{j=1}^n \mu(\phi_j)$.

**Theorem 5.1** *Suppose $(T, X, \mu)$ is a dynamical system with exponential decay in the BV norm versus $\mathscr{L}^1(\mu)$ and whose transfer operator $P$ satisfies $\|P^n\phi\|_{BV} \leq C\theta^n\|\phi\|_{BV}$ for all $\phi \in BV(X)$ such that $\int \phi d\mu = 0$. Suppose $\phi_j = 1_{A_j}$ are indicator functions of a sequence of nested sets $A_j$ such that $\sup_n \|\phi_n\|_{BV} < \infty$, $\mu(A_n) \geq n^{-\gamma}$ for some $\gamma > 0$ and $E_n := \sum_{j=1}^n = \mu(A_n)$ diverges. Then $(\phi_n \circ T^n)_{n\geq 1}$ satisfies the*

*ASIP i.e. enlarging our probability space if necessary it is possible to find a sequence $(Z_k)_{k\geq 1}$ of independent centered Gaussian variables $Z_k$ such that for all $\beta < \frac{1}{2}$*

$$\sup_{1\leq k\leq n} |\sum_{i=1}^{k} \tilde{\phi}_i \circ T^i - \sum_{i=1}^{k} Z_i| = o(\sigma_n^{1-\beta}) \qquad \mu - a.s.$$

*Furthermore* $\sum_{i=1}^{n} E[Z_i^2] = \sigma_n^2 + O(\sigma_n)$.

**Proof** From [21, Lemma 2.4] we see that for sufficiently large $n$, $\sigma_n^2 \geq E_n$ (note that there is a typo in the statement of [21, Lemma 2.4] and lim sup should be replaced with lim inf). We follow the proof of Theorem 3.1 based on [11, Theorem 5.1] taking $T_k = T$ for all $k$, $m$ as the invariant measure $\mu$ and $f_n = 1_{A_n}$. Note that conditions (DEC) and (MIN) are satisfied automatically under the assumption that we have exponential decay of correlations in BV norm versus $\mathscr{L}^1$ and the transfer operator $P$ is defined with respect to the invariant measure $\mu$ in the usual way by $\int (Pf)g\, d\mu = \int fg(T)\, d\mu$ for all $f \in \mathscr{L}^1(\mu)$, $g \in \mathscr{L}^\infty(\mu)$. Hence $P1 = 1$ and in particular $|P\phi|_\infty \leq |\phi|_\infty$. We write $P^n$ for the $n$-fold composition of the linear operator $P$. Let $\tilde{\phi}_i = \phi_i - \mu(\phi_i)$. As before define $h_n = \sum_{j=1}^{n} P^j \tilde{\phi}_{n-j}$ and write

$$\psi_n = \tilde{\phi}_n + h_n - h_{n+1} \circ T.$$

Again, for convenience we put

$$U_n = \psi_n \circ T^n$$

so that $(U_n)$ is a sequence of reversed martingale differences for the filtration $(\mathcal{B}_n)$. As in the case of sequential expanding maps one shows that $\sum_{i=1}^{n} E[U_i^2] = \sigma_n^2 + O(\sigma_n)$. Condition (A) of Theorem 1.1 holds exactly as before.

In order to estimate $\mu(|U_n|^4)$ observe that by Minkovski's inequality ($p > 1$)

$$\|h_n\|_p \leq \sum_{j=1}^{n-1} \|P^j \tilde{\phi}_{n-j}\|_p,$$

where

$$\|P^j \tilde{\phi}_{n-j}\|_p \leq \|P^j \tilde{\phi}_{n-j}\|_{BV} \leq c_1 \vartheta^j \|\tilde{\phi}_{n-1}\|_{BV} \leq c_2 \vartheta^j$$

for all $n$ and $j < n$. For small values of $j$ we use the estimate (as $|\tilde{\phi}_{n-j}|_\infty \le 1$)

$$\int \left|P^j \tilde{\phi}_{n-j}\right|^p \le \int \left|P^j \tilde{\phi}_{n-j}\right| \le \int P^j(\phi_{n-j}+\mu(A_{n-j})) = \int \phi_{n-j}\circ T^j + \mu(A_{n-j}) = 2\mu(A_{n-j}).$$

If we let $q_n$ be the smallest integer so that $\vartheta^{q_n} \le (\mu(A_{n-q_n}))^{\frac{1}{p}}$, then

$$\|h_n\|_p \le \sum_{j=1}^{q_n} (2\mu(A_{n-j}))^{\frac{1}{p}} + \sum_{j=q_n}^{n} c_2 \vartheta^j \le c_3 q_n \left(\mu(A_{n-q_n})\right)^{\frac{1}{p}}.$$

A similar estimate applies to $h_{n+1}$. Note that $q_n \le c_4 \log n$ for some constant $c_4$. Let us put $p = 4$; then factoring out yields

$$\int \psi_n^4 = \mathcal{O}(\mu(A_n)) + \|h_n - h_{n+1}T\|_4^4 = \mathcal{O}(\mu(A_n)) + \mathcal{O}(q_{n+1}^4 \mu(A_{n-q_n})).$$

Let $\alpha < 1$ (to be determined below) and put $a_n = E_n^\alpha$, where $E_n = \sum_{j=1}^n \mu(A_j)$. Then

$$\begin{aligned}
\sum_n \frac{\mu(U_n^4)}{a_n^2} &\le c_5 \sum_n \frac{\mu(A_n) + q_{n+1}^4 \mu(A_{n-q_n})}{E_n^{2\alpha}} \\
&\le c_6 \sum_n \frac{q_{n+1}^4 \mu(A_{n-q_n})}{E_{n-q_n}^{2\alpha}} \\
&\le c_7 \sum_n \int_{E_{n-q_n-1}}^{E_{n-q_n}} \frac{\log^4 x}{x^{2\alpha}}\, dx \\
&\le 2c_7 \int_{E_1}^{\infty} \frac{\log^4 x}{x^{2\alpha}}\, dx
\end{aligned}$$

converges if $\alpha > \frac{1}{2}$. We have thus verified Condition (B) of Theorem 1.1 with the value $v = 2$.

Thus $U_n$ satisfies the ASIP with error term $o(E_n^{\frac{1-\beta}{2}}) = o(\sigma_n^{1-\beta})$ for any $\beta < \frac{1}{2}$

Finally

$$\sum_{j=1}^n U_j = \sum_{j=1}^n \tilde{\phi}_j(T^j) + h_1(T_1) - h_n(T^n)$$

and as $|h_n|$ is uniformly bounded we conclude that $(\phi_j(T^j))$ satisfies the ASIP with error term $o(\sigma_n^{1-\beta})$ for all $\beta < \frac{1}{2}$. ∎

**Remark 5.2** The bound $\mu(A_n) \ge n^{-\gamma}$, $\gamma \in \mathbb{R}$, weakens the conditions under which the CLT has been proven in [21, 11], namely $\mu(A_n) \ge \frac{C}{n}$ for some $n$.

# 6 ASIP for non-stationary observations on invertible hyperbolic systems.

In this section we will suppose that $B_\alpha$ is the Banach space of $\alpha$-Hölder functions on a compact metric space $X$ and that $(T, X, \mu)$ is an ergodic measure preserving transformation. Suppose that $P$ is the $\mathscr{L}^2$ adjoint of the Koopman operator $U$, $U\phi = \phi \circ T$, with respect to $\mu$. First we consider the non-invertible case and suppose that $\|P^n\phi\|_\alpha \leq C\vartheta^n\|\phi\|_\alpha$ for all $\alpha$-Hölder $\phi$ such that $\int \phi\, d\mu = 0$ where $C > 0$ and $0 < \vartheta < 1$ are uniform constants. Under this assumption we will establish the ASIP for sequences of uniformly Hölder functions satisfying a certain variance growth condition. Then we will give a corollary which establishes the ASIP for sequences of uniformly Hölder functions on an Axiom A system satisfying the same variance growth condition.

The main difficulty in this setting is establishing a strong law of large numbers with error (Condition (A)) for the squares $(U_j^2)$ of the martingale difference scheme. We are not able to use the Gal-Koksma lemma in the same way as we did in the setting of decay in bounded variation norm. Nevertheless our results, while clearly not optimal, point the way to establishing strong statistical properties for non-stationary time series of observations on hyperbolic systems.

**Theorem 6.1** *Suppose $\{\phi_j\}$ is a sequence of $\alpha$-Hölder functions such that $\int \phi_j\, d\mu = 0$ and $\sup_j \|\phi_j\|_\alpha \leq C_1$ for some constant $C_1 < \infty$.*

*Let $\sigma_n^2 = \int (\sum_{j=1}^n \phi_j \circ T^j)^2 d\mu$ and suppose that $\sigma_n^2 \geq C_2 n^\delta$ for some $\delta > \frac{\sqrt{17}-1}{4}$ and a constant $C_2 < \infty$. Then there is a sequence of centered independent Gaussian random variables $(Z_j)$ such that, enlarging our probability space if necessary,*

$$\sum_{j=1}^n \phi_j \circ T^j = \sum_{j=1}^n Z_j + \mathcal{O}(\sigma_n^{1-\beta})$$

*$\mu$ almost surely for any $\beta < \frac{\sqrt{17}-1}{4\delta}$.*

*Furthermore $\sum_{i=1}^n E[Z_i^2] = \sigma_n^2 + \mathcal{O}(\sigma_n)$.*

**Proof** Define $h_n = P\phi_{n-1} + P^2\phi_{n-2} + \cdots + P^n\phi_0$ and put

$$\psi_n = \phi_n + h_n - h_{n+1} \circ T.$$

Note $P\psi_n = 0$ and that $\|h_n\| = \mathcal{O}(1)$ for $n > 1$ by the same argument as in the proof of Theorem 5.1. The sequence $U_n = \psi_n \circ T^n$ is a sequence of reversed martingale differences with respect to the filtration $\mathcal{F}_n$, where $\mathcal{F}_n = T^{-n}\mathcal{F}_0$. We will take $a_n = \sigma_n^{2\eta}$ where $\eta > 0$ will be determined below. Since $\|\psi_j\|_\alpha = \mathcal{O}(1)$ and consequently $\|U_j\|_\alpha = \mathcal{O}(1)$ we conclude that

$$\sum_n \frac{\mu(U_n^4)}{a_n^2} \le c_1 \sum_n \frac{1}{\sigma_n^{4\eta}} \le c_2 \sum_n \frac{1}{n^{2\eta\delta}} < \infty$$

provided $\eta > \frac{1}{2\delta}$. In this case Condition (B) of Theorem 1.1 is satisfied for $v = 2$.

In order to verify Condition (A) of Theorem 1.1 let us observe that $E[U_j^2|\mathcal{F}_{j+1}] = E[\psi_j^2 \circ T^j|\mathcal{F}_{j+1}] = P^{j+1}(\psi_j \circ T^j) \circ T^{j+1} = (P^{j+1}U^j\psi_j^2) \circ T^{j+1} = (P\psi_j^2) \circ T^{+1}$. We now shall prove a strong law of large numbers with rate for the sequence $E[U_j^2|\mathcal{F}_{j+1}]$. For simplicity of notation we denote $E[U_j^2|\mathcal{F}_{j+1}]$ by $\hat{U_j}^2$.

Let us write $S_n = \sum_{j=1}^n [\hat{U_j}^2 - \mu(U_j^2)]$ for the LHS of condition (A) in Theorem 1.1. Then $\rho_n^2 = \int S_n^2\, d\mu = \int (\sum_{j=1}^n \hat{U_j}^2 - E[U_j^2])^2\, d\mu$ satisfies by decay of correlations the estimate $\rho_n^2 = \mathcal{O}(n)$, where we used that $\|\hat{U_j}^2\|_\alpha = \mathcal{O}(1)$. Hence by Chebyshev's inequality

$$P\left(|S_n| > \frac{\sigma_n^{2\eta}}{\log n}\right) \le \frac{\rho_n^2}{\sigma_n^{4\eta}} \log^2 n \le c_3 n^{-(2\eta\delta-1)} \log^2 n$$

as $\sigma_n^2 = \mathcal{O}(n^\delta)$. Since $\delta$ is never larger than 2, we have $2\eta\delta - 1 \le 1$. Then along a subsequence $f(n) = [n^\omega]$ for $\omega > \omega_0 = \frac{1}{2\eta\delta-1} \ge 1$ we can apply the Borel-Cantelli lemma since $P\Big(|S_{f(n)}| > \sigma_{f(n)}^{2\eta}/\log f(n)\Big)$ is summable as $\sum_n n^{-\omega(2\eta\delta-1)} \log^2 n < \infty$. Hence by Borel-Cantelli for $\mu$ a.e. $x \in X$, $|S_{f(n)}(x)| > \frac{\sigma_{f(n)}^{2\eta}}{\log f(n)}$ only finitely often.

In order to control the gaps note that $[(n+1)^\omega] - [n^\omega] = \mathcal{O}(n^{\omega-1})$ and let $k \in (f(n), f(n+1))$. Since along the subsequence $S_{f(n)} = o(\sigma_{f(n)}^{2\eta})$ we conclude that $S_k = o(\sigma_{f(n)}^{2\eta}) + \mathcal{O}(n^{\omega-1})$ as there are at most $n^{\omega-1}$ terms $\hat{U_j}^2 - E[U_j^2] = \mathcal{O}(1)$ in the range $j \in (f(n), k]$.

Choosing $\omega > \omega_0$ close enough to $\omega_0$ we conclude that

$$S_k = o\Big(\sigma_{f(n)}^{2\eta} + n^{\omega-1}\Big) = o\Big(\sigma_n^{2\eta} + \sigma_n^{(\omega-1)\frac{2}{\delta}}\Big) = o\big(\sigma_k^{2\eta}\big)\,,$$

for $\eta > \eta_0$ where $\eta_0$ satisfies $2\eta_0 = (\omega_0 - 1)\frac{2}{\delta} = \frac{2-2\eta\delta}{2\eta\delta-1}\frac{2}{\delta}$ which implies $\eta_0 = \frac{\gamma_0}{\delta}$, with $\gamma_0 = \frac{\sqrt{17}-1}{4}$.

This concludes the proof of Condition (A) with $a_n = \sigma_n^{2\eta}$. Also note that $\eta_0$ is larger than $\frac{1}{2\delta}$ which ensures Condition (B). Thus $\{U_j\}$ satisfies the ASIP with error $\mathcal{O}(\sigma_n^{1-\beta})$ for $0 < \beta < \beta_0 = 1 - \eta_0 = 1 - \frac{\gamma_0}{\delta}$ and hence so does $\{\phi_j \circ T^j\}$. In particular we must require $\delta$ to be bigger than $\gamma_0$ (which is slightly larger than $\frac{3}{4}$). ∎

We now state a corollary of this theorem for a sequence of non-stationary observations on Axiom A dynamical systems.

**Corollary 6.2** *Suppose $(T, X, \mu)$ is an Axiom-A dynamical system, where $\mu$ is a Gibbs measure. Suppose $\{\phi_j\}$ is a sequence of $\alpha$-Hölder functions such that $\int \phi_j \, d\mu = 0$ and $\sup_j \|\phi_j\|_\alpha < \infty$ for some constant $C$. Let $\sigma_n^2 = \int (\sum_{j=1}^n \phi_j \circ T^n)^2 d\mu$ and suppose that $\sigma_n^2 \geq Cn^\delta$ for some $\delta > \frac{\sqrt{17}-1}{4}$ and a constant $C < \infty$. Then there is a sequence of centered independent Gaussian random variables $(Z_j)$ and a $\gamma > 0$ such that, enlarging our probability space if necessary,*

$$\sum_{j=1}^n \phi_j \circ T^j = \sum_{j=1}^n Z_j + O(\sigma_n^{1-\beta})$$

*$\mu$ almost surely for any $\beta < \frac{\sqrt{17}-1}{4\delta}$.*

*Furthermore $\sum_{i=1}^n E[Z_i^2] = \sigma_n^2 + O(\sigma_n)$.*

**Proof** The assumption $\sigma_n^2 \geq Cn^\delta$ for some $\delta > \frac{\sqrt{17}-1}{4}$ agrees with Theorem 6.1. The basic strategy is now the standard technique of coding first by a two sided shift and then reducing to a non-invertible one-sided shift. There is a good description in Field, Melbourne and Török [13]. We use a Markov partition to code $(T, X, \mu)$ by a 2-sided shift $(\sigma, \Omega, \nu)$ in a standard way [8, 29]. We lift $\phi_j$ to the system $(\sigma, \Omega, \nu)$ keeping the same notation for $\phi_j$ for simplicity. Using the Sinai trick [13, Appendix A] we may write

$$\phi_j = \psi_j + v_j - v_{j+1} \circ \sigma$$

where $\psi_j$ depends only on future coordinates and is Hölder of exponent $\sqrt{\alpha}$ if $\phi_j$ is of exponent $\alpha$. In fact $\|\psi_j\|_{\sqrt{\alpha}} \leq K$ and similarly $\|v_j\|_{\sqrt{\alpha}} \leq K$ for a uniform constant $K$.

There is a slight difference in this setting to the usual construction. Pick a Hölder map $G : X \to X$ that depends only on future coordinates (e.g. a map which locally substitutes all negative coordinates by a fixed string) and define

$$v_n(x) = \sum_{k \geq n} \phi_k(\sigma^{k-n}x) - \phi_k(\sigma^{k-n}Gx).$$

It is easy to see that the sum converges since $|\phi_k(\sigma^{k-n}x) - \phi_k(\sigma^{k-n}Gx)| \leq C\lambda^k \|\phi_k\|_\alpha$ (where $0 < \lambda < 1$) and that $\|v_n\|_\alpha \leq C_2$ for some uniform $C_2$.

Since

$$\phi_n - v_n + v_{n+1} \circ \sigma = \phi_n(Gx) + \sum_{k>n} [\phi_k(\sigma^{k-n}Gx) - \phi_k(\sigma^{k-n}G\sigma x)]$$

defining $\psi_n = \phi_n - v_n + v_{n+1} \circ \sigma$ we see $\psi_n$ depends only on future coordinates.

We let $\mathcal{F}_0$ denote the $\sigma$-algebra consisting of events which depend on past coordinates. This is equivalent to conditioning on local stable manifolds defined by the Markov partition. Symbolically $\mathcal{F}_0$ sets are of the form $(* * * * .\omega_0\omega_1 \ldots)$ where $*$ is allowed to be any symbol.

Finally using the transfer operator $P$ associated to the one-sided shift $\sigma(x_0x_1 \ldots x_n \ldots) = (x_1x_2 \ldots x_n \ldots)$ we are in the set-up of Theorem 6.1. As before we define $h_n = P\psi_{n-1} + P^2\psi_{n-2} + \cdots + P^n\psi_0$ and put

$$V_n = \psi_n + h_n - h_{n+1} \circ T$$

The sequence $U_n = V_n \circ T^n$ is a sequence of reversed martingale differences with respect to the filtration $\mathcal{F}_n$, where $\mathcal{F}_n = \sigma^{-n}\mathcal{F}_0$. In fact $(UP)f = E[f|\sigma^{-1}\mathcal{F}_0] \circ \sigma$ while $(PU)f = f$ (this is easily checked, see [13, Remark 3.1.2] or [29]).

Thus $U_n$ satisfies the ASIP with error term $o(\sigma_n^{1-\beta})$ for $\beta \in (0, 1 - \frac{\gamma_0}{\delta})$. Hence $\psi_n \circ T^n$ satisfies the ASIP with error term $o(\sigma_n^{1-\beta})$.

Finally

$$\sum_{j=0}^{n} \phi_j = \sum_{j=0}^{n} \psi_j(T^j) + [v_0 - v_n \circ \sigma^{n+1}]$$

as the sum telescopes. As $|v_n| \leq C$ we have the ASIP with error term $o(\sigma_n^{1-\beta})$ for the sequence $\{\phi_n \circ T^n\}$. This concludes the proof. ∎

# 7 Further applications.

We consider here maps for which conditions (DEC) and (MIN) are satisfied, but in order to guarantee the unboundedness of the variance when $\phi$ is not a coboundary, we will see that further assumptions are needed. We follow here again [11], especially Sect. 5. We begin with looking for a useful sufficient condition to get the (DEC) condition; we adapt it to the class of maps we are going to introduce. These maps will be defined on the unit interval or on compact subspaces of $\mathbb{R}^n$. The adapted spaces will be denoted with $\mathcal{V}$ and $\mathcal{L}^1_m$ being $\mathcal{V}$ the space of bounded variation functions in the case of one-dimensional maps, and the space of quasi-Hölder functions for maps defined on compact subsets of $\mathbb{R}^n$, $n > 1$. In both cases we will denoted with $||\cdot||_\alpha$ the norm on $\mathcal{V}$: this norm will be again the sum of a suitable seminorm and of the $\mathcal{L}^1_m$ norm.
We say that a transfer operator $P$ acting on $\mathcal{V}$ is *exact* if, acting on functions $f \in \mathcal{V}$ of zero Lebesgue mean, it verifies: $\lim_{n\to\infty} ||P^n f||_1 = 0$.
We now begin to list the assumptions we need: we will see that the maps in $\mathcal{F}$ will be close, in a sense we will describe below, to a given map $T_0$. Call $P_0$ the transfer operator associated to $T_0$.

**Uniform Doeblin-Fortet-Lasota-Yorke inequality (DFLY)**: Given the family $\mathcal{F}$ there exist constants $A, B < \infty, \rho \in (0,1)$, such that for any $n$ and any sequence of operators $P_n, \cdots, P_1$ in $\mathcal{F}$ and any $f \in \mathcal{V}$ we have

$$\|P_n \circ \cdots \circ P_1 f\|_\alpha \leq A\rho^n \|f\|_\alpha + B\|f\|_1 \tag{7.1}$$

The bound (7.1) will be true in particular when applied to $P_0^n$, we moreover suppose that:

**Exactness property**: The operator $P_0$ has a spectral gap, which implies that there are two constants $C_1 < \infty$ and $\gamma_0 \in (0,1)$ so that

$$\textbf{(Exa)} \qquad ||P_0^n f||_\alpha \leq C_1 \gamma_0^n ||f||_\alpha$$

for all $f \in BV$ of zero (Lebesgue) mean and $n \geq 1$. By the very definition of the $\alpha$ norm, we immediately have that the operator $P_0$ is exact.

Then one considers the following distance between two operators $P$ and $Q$ acting on $BV$:

$$d(P,Q) = \sup_{f \in BV,\, \|f\|_{BV} \leq 1} ||Pf - Qf||_1.$$

The useful criterion to verify the (DEC) condition is given in Proposition 2.10 in [11], and in our setting it reads: *if $P_0$ is exact, then there exists $\delta_0 > 0$, such that the set $\{P \in \mathcal{F}; d(P, P_0) < \delta_0\}$ satisfies the (DEC) condition.*

By induction on the Doeblin-Fortet-Lasota-Yorke inequality for compositions we immediately have

$$\textbf{(DS)} \qquad d(P_r \circ \cdots \circ P_1, P_0^r) \leq M \sum_{j=1}^{r} d(P_j, P_0), \tag{7.2}$$

with $M = 1 + A\rho^{-1} + B$.

According to [11, Lemma 2.13], (DS) and (Exa) imply that there exists a constant $C_2$ such that

$$\|P_n \circ \cdots \circ P_1 \phi - P_0^n \phi\|_1 \leq C_2 \|\phi\|_{BV} (\sum_{k=1}^{p} d(P_{n-k+1}, P_0) + (1-\gamma_0)^{-1} \gamma_0^p)$$

for all integers $p \leq n$ and all functions $\phi$ of bounded variation.

**Lipschitz continuity property**: Assume that the maps (and their transfer operators) are parametrized by a sequence of numbers $\varepsilon_k$, $k \in \mathbb{N}$, such that $\lim_{k\to\infty} \varepsilon_k = \varepsilon_0$, $(P_{\varepsilon_0} = P_0)$. We assume that there exists a constant $C_3$ so that

$$\textbf{(Lip)} \qquad d(P_{\varepsilon_k}, P_{\varepsilon_j}) \leq C_3 |\varepsilon_k - \varepsilon_j|, \qquad \text{for all } k, j \geq 0.$$

We will restrict in the following to the subclass $\mathcal{F}_{exa}$ of maps, and therefore of operators, for which

$$\mathcal{F}_{exa} := \{P_{\varepsilon_k} \in \mathcal{F};\ |\varepsilon_k - \varepsilon_0| < C_3^{-1}\delta_0\}.$$

The maps in $\mathcal{F}_{exa}$ will therefore verify the (DEC) condition; in order to guarantee the unboundeness of the variance we need something stronger, namely:

**Convergence property**: We require algebraic convergence of the parameters, that is, there exist a constant $C_4$ and $\kappa > 0$ so that

$$\textbf{(Conv)} \qquad |\varepsilon_n - \varepsilon_0| \leq \frac{C_4}{n^\kappa} \qquad \forall n \geq 1.$$

With these last assumptions, we get a polynomial decay for (7.2) of the type $O(n^{-\kappa})$ and in particular we obtain the same algebraic convergence in $\mathscr{L}^1$ of $P_n \circ \cdots \circ P_1\phi$ to $h \int \phi\, dm$, where $h$ is the density of the absolutely continuous mixing measure of the map $T_0$. This convergence is necessary to establish the growth of the variance $\sigma_n^2$.

Finally, we also require

**Positivity property**: The density $h$ for the limiting map $T_0$ is strictly positive, namely

$$\textbf{(Pos)} \qquad \inf_x h(x) > 0.$$

The relevance of these four properties is summarised by the following result:

**Lemma 7.1** *[11, Lemma 5.7] Assume the assumptions (Exa), (Lip), (Conv) and (Pos) are satisfied. If $\phi$ is not a coboundary for $T_0$ then $\sigma_n^2/n$ converges as $n \to \infty$ to $\sigma^2$ which moreover is given by*

$$\sigma^2 = \int \hat{P}[G\phi - \hat{P}G\phi]^2(x)h(x)\ dx,$$

*where $\hat{P}\phi = \frac{P_0(h\phi)}{h}$ is the normalized transfer operator of $T_0$ and $G\phi = \sum_{k\geq 0} \frac{P_0^k(h\phi)}{h}$.*

**Warning** In the next sections we will give several examples of maps satisfying [Lemma 5.7]. The family of maps will be parametrized by a small positive number

$\varepsilon$ (or a vector with small positive components). When we will write sentences like *The maps in $\mathcal{F}$ verify condition (Lip)*, we will tacitly suppose that we restrict to $\mathcal{F}_{exa}$ having previously proved that the transfer operator $P_0$ is exact. This will impose restriction on the choice of $\varepsilon$ (less than a constant times $\delta_0$, see above), and in this case we will use the terminology *for $\varepsilon$ small enough*. In particular we could eventually take the sequences $\varepsilon_k$ with $k$ larger than 1 in the Theorems stated below to get the variance.

## 7.1 $\beta$ transformations

Let $\beta > 1$ and denote by $T_\beta(x) = \beta x \bmod 1$ the $\beta$-transformation on the unit circle. Similarly for $\beta_k \geq 1 + c > 1$, $k = 1, 2, \ldots$, we have the transformations $T_{\beta_k}$ of the same kind, $x \mapsto \beta_k x \bmod 1$. Then $\mathcal{F} = \{T_{\beta_k} : k\}$ is the family of functions we want to consider here. The property (DEC) was proved in [11, Theorem 3.4 (c)] and condition (MIN) in [11, Proposition 4.3]. Namely, for any $\beta > 1$ there exist $a > 0, \delta > 0$ such that whenever $\beta_k \in [\beta - a, \beta + a]$, then $P_k \circ \cdots \circ P_1\, 1(x) \geq \delta$, where $P_\ell$ is the transfer operator of $T_{\beta_\ell}$. The invariant density of $T_\beta$ is bounded below, and continuity (Lip) is precisely the content of Sect. 5 in [11]. We therefore obtain (see [11, Corollary 5.4]):

**Theorem 7.2** *Assume that $|\beta_n - \beta| \leq n^{-\theta}$, $\theta > 1/2$. Let $\phi \in BV$ be such that $m(hf) = 0$, where $m$ is the Lebesgue measure and $\phi$ is not a coboundary for $T_\beta$, so $\sigma^2 \neq 0$. Then the random variables*

$$W_n = \phi + T_{\beta_1}\phi + \cdots + T_{\beta_1} T_{\beta_2} \ldots T_{\beta_{n-1}}\phi$$

*satisfy a standard ASIP with variance $\sigma^2$.*

## 7.2 Perturbed expanding maps of the circle.

We consider a $C^2$ expanding map $T$ of the circle $\mathbb{T}$; let us put $A_k = [v_k, v_{k+1}]; k = 1, \cdots, m, v_{m+1} = v_1$ the closed intervals such that $TA_k = \mathbb{T}$ and $T$ is injective over $[v_k, v_{k+1})$. The family $\mathcal{F}$ then consists of the perturbed maps $T_\varepsilon$ which are given by the

translations (*additive noise*): $T_\varepsilon(x) = T(x)+\varepsilon, \mod 1$, where $\varepsilon \in (-1,1)$. We observe that the intervals of local injectivity $[v_k, v_{k+1}),\ k = 1, \cdots, m$, of $T_\varepsilon$ are independent of $\varepsilon$. We call $\mathcal{A}$ the partition $\{A_k : k\}$ into intervals of monotonicity. We assume there exist constants $\Lambda > 1$ and $C_1 < \infty$ so that

$$\inf_{x\in\mathbb{T}} |DT(x)| \geq \Lambda; \qquad \sup_{\varepsilon\in(-1,1)} \sup_{x\in\mathbb{T}} \left| \frac{D^2 T_\varepsilon(x)}{DT_\varepsilon(x)} \right| \leq C_1. \tag{7.3}$$

**Lemma 7.3** *The maps $\mathcal{F} = \{T_\varepsilon\}$ for $\varepsilon$ small enough satisfy the conditions of Lemma 7.1.*

**Proof** (I) (DFLY) It is well known that any such map $T_\varepsilon$ satisfying (7.3) verifies a Doeblin-Fortet-Lasota-Yorke inequality $||P_\varepsilon f||_{BV} \leq \rho||f||_{BV} + B||f||_1$ where $\rho \in (0,1)$ and $B < \infty$ are independent of $\varepsilon$ ($P_\varepsilon$ is the associated transfer operator of $T_\varepsilon$). For any concatenation of maps one consequently has

$$\|\mathcal{P}_n f\|_{BV} \leq \rho^k \|f\|_{BV} + \frac{B}{1-\rho}\|f\|_1,$$

where $\mathcal{P}_n = P_{\varepsilon_k} \circ \cdots \circ P_{\varepsilon_1}$.

(II) (MIN) In order to obtain the lower bound property (MIN) we have to consider an upper bound for concatenations of operators. Since each $T_\varepsilon$ has $m$ intervals of monotonicity we have (where $\mathcal{T}_n = T_{\varepsilon_n} \circ \cdots \circ T_{\varepsilon_1}$ as before)

$$\mathcal{P}_n 1(x) = \sum_{k_n,\cdots,k_1=1}^{m} \frac{1}{|D\mathcal{T}_n(T^{-1}_{k_1,\varepsilon_1} \circ \cdots T^{-1}_{k_n,\varepsilon_n}(x))|} \times \mathbf{1}_{\mathcal{T}_n A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}}(x) \tag{7.4}$$

where $T^{-1}_{k_l,\varepsilon_l}, k_l \in [1,m]$, denotes the local inverse of $T_{\varepsilon_l}$ restricted to $A_{k_l}$ and

$$A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n} = T^{-1}_{k_1,\varepsilon_1} \circ \cdots \circ T^{-1}_{k_{n-1},\varepsilon_{n-1}} A_{k_n} \cap \cdots \cap T^{-1}_{k_1,\varepsilon_1} A_{k_2} \cap A_{k_1} \tag{7.5}$$

is one of the $m^n$ intervals of monotonicity of $\mathcal{T}_n$. Since those images satisfy[2]

$$\mathcal{T}_n A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n} = T_{\varepsilon_n}(A_{k_n} \cap T_{\varepsilon_{n-1}} A_{k_{n-1}} \cap \cdots \cap T_{\varepsilon_{n-1}} \circ \cdots \circ T_{\varepsilon_1} A_{k_1}) \tag{7.6}$$

[2]This can be proved by induction; for instance for $n = 3$ we have $T_{\varepsilon_3}T_{\varepsilon_2}T_{\varepsilon_1}(T^{-1}_{k_1,\varepsilon_1}T^{-1}_{k_2,\varepsilon_2}A_{k_3} \cap T^{-1}_{k_1,\varepsilon_1}A_{k_2} \cap A_{k_1}) = T_{\varepsilon_3}T_{\varepsilon_2}T_{\varepsilon_1}[T^{-1}_{k_1,\varepsilon_1}(T^{-1}_{k_2,\varepsilon_2}A_{k_3} \cap A_{k_2} \cap T_{\varepsilon_1}A_{k_1})] = T_{\varepsilon_3}T_{\varepsilon_2}(T^{-1}_{k_2,\varepsilon_2}A_{k_3} \cap A_{k_2} \cap T_{\varepsilon_1}A_{k_1}) = T_{\varepsilon_3}T_{\varepsilon_2}[T^{-1}_{k_2,\varepsilon_2}(A_{k_3} \cap T_{\varepsilon_2}A_{k_2} \cap T_{\varepsilon_2}T_{\varepsilon_1}A_{k_1})] = T_{\varepsilon_3}(A_{k_3} \cap T_{\varepsilon_2}A_{k_2} \cap T_{\varepsilon_2}T_{\varepsilon_1}A_{k_1})$.

and each branch is onto, we have that the inverse image is the full interval. By the Mean Value Theorem there exists a point $\xi_{k_1,\cdots,k_n}$ in the interior of the connected interval $A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$ such that $|D\mathcal{T}_n(\xi_{k_1,\cdots,k_n})|^{-1} = |A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}|$, where $|A|$ denotes the length of the connected interval $A$. In order to get distortion estimates, let us take two points $u, v$ in the closure of $A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$. Then ($\mathcal{T}_0$ is the identity map)

$$\begin{aligned}
\left|\frac{D\mathcal{T}_n(u)}{D\mathcal{T}_n(v)}\right| &= \exp\left(\log|D\mathcal{T}_n(u)| - \log|D\mathcal{T}_n(v)|\right) \\
&= \exp\sum_{j=1}^{n}\left(\log\left|DT_{\varepsilon_j}\circ\mathcal{T}_{j-1}(u)\right| - \log\left|DT_{\varepsilon_j}\circ\mathcal{T}_{j-1}(v)\right|\right) \\
&= \exp\sum_{j=1}^{n}\frac{|D^2T_{\varepsilon_j}(\iota_k)|}{|DT_{\varepsilon_j}(\iota_j)|}\,|\mathcal{T}_{j-1}(u) - \mathcal{T}_{j-1}(v)|
\end{aligned}$$

for some points $\iota_j$ in $\mathcal{T}_{j-1}A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$. Using the second bound in (7.3) and the fact that $|\mathcal{T}_{j-1}(u) - \mathcal{T}_{j-1}(v)| \leq \Lambda^{-(j-1)}$ we finally have

$$|D\mathcal{T}_n(u)/D\mathcal{T}_n(v)| \leq e^{\frac{C_1}{1-\Lambda}}$$

which in turn implies that

$$\mathcal{P}_n 1(x) \geq e^{-\frac{C_1}{1-\Lambda}}$$

and this independently of any choice of the $\varepsilon_k, k = 1, \cdots, n$ and of $n$.

(III) The strict positivity condition (Pos) holds since the map $T$ is Bernoulli and for such maps it is well known that its invariant densities are uniformly bounded from below away from zero [1].

(IV) The continuity condition (Lip) follows the same proof as in the next section and therefore we refer to that. ∎

We now conclude by Lemma 7.1 the following result:

**Theorem 7.4** *Let $\mathcal{F}$ be a family of functions as described in this section. Then for any function $\phi$ which is not a coboundary for $T_\beta$ we have that the random variables*

$$W_n = \sum_{j=0}^{n-1}\phi\circ\mathcal{T}_j$$

*satisfy a standard ASIP with variance $\sigma^2$.*

## 7.3 Covering maps: special cases

### 7.3.1 One dimensional maps

The next example concerns piecewise uniformly expanding maps $T$ on the unit interval. The family $\mathcal{F}$ will consist of maps $T_\varepsilon$, which are constructed with *local* additive noise starting from $T$, which in turn satisfies:

- (i) $T$ is locally injective on the open intervals $A_k, k = 1, \ldots, m$, that give a partition $\mathcal{A} = \{A_k : k\}$ of the unit interval $[0, 1] = M$ (up to zero measure sets).
- (ii) $T$ is $C^2$ on each $A_k$ and has a $C^2$ extension to the boundaries. Moreover there exist $\Lambda > 1, C_1 < \infty$, such that $\inf_{x\in M} |DT(x)| \geq \Lambda$ and $\sup_{x\in M} \left|\frac{D^2T(x)}{DT(x)}\right| \leq C_1$.

At this point we give the construction of the family $\mathcal{F}$ of maps $T_\varepsilon$ by defining them locally on each interval $A_k$. On each interval $A_k$ we put $T_\varepsilon(x) = T(x) + \varepsilon$ where $|\varepsilon| < 1$ and we extend by continuity to the boundaries. We restrict to values of $\varepsilon$ so that the image $T_\varepsilon(A_k)$ stays in the unit interval; this we achieve for a given $\varepsilon$ by choosing the sign of $\varepsilon$ so that the image of $A_k$ remains in the unit interval; if not we do *not* move the map. The sign will consequently vary with each interval.

We add now new the new assumption. Assume there exists a set $\mathcal{J}$ so that:

- (iii) $\mathcal{J} \subset T_\varepsilon A_k$ for all $T_\varepsilon \in \mathcal{F}$ and $k = 1, \ldots, m$.
- (iv) The map $T$ send $\mathcal{J}$ on $[0, 1]$ and therefore it will not be affected there by the addition of $\varepsilon$. In particular it will exist $1 \geq L' > 0$ such that $\forall k = 1, \ldots, q$ we have $|T(\mathcal{J}) \cap A_k| > L'$.

**Lemma 7.5** *The maps $T_\varepsilon$ satsify the conditions (DFLY), (MIN), (Pos) and (Lip).*

**Proof** (I) The condition (DFLY) follows from assumption (ii).

(II) In order to prove the lower bound condition (MIN) we begin by observing that, thanks to (iv), the union over the $m^n$ images of the intervals of monotonicity of *any* concatenation of $n$ maps, still covers $M$. Assumption (iii) above does not require that each branch of the maps in $\mathcal{F}$ be onto; instead, and thanks again to (7.6),

we see that each image $\mathcal{T}_n A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$ will have at least length $L = \Lambda L'$, so that the reciprocal of the derivative of $\mathcal{T}_n$ over $A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$ will be of order $L^{-1}|A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}|$ (as before $\mathcal{T}_n = T_{\varepsilon_n} \circ \cdots \circ T_{\varepsilon_1}$). By distortion we make it precise by multiplying by the same distortion constant $e^{\frac{C_1}{1-\Lambda}}$ as above. In conclusion we have

$$P_{\varepsilon_n} \circ \cdots \circ P_{\varepsilon_1} 1(x) \geq L^{-1} e^{-\frac{C_1}{1-\Lambda}}$$

(III) To show strict positivity of the invariant density $h$ for the map $T$ we use Assumption (iv). Since $h$ is of bounded variation, it will be strictly positive on an open interval $J$, where $\inf_{x\in J} h(x) \geq h_*$ where $h_* > 0$. We now choose a partition element $R_n$ of the join $\mathcal{A}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}$, such that $R_n \subset J$. This is possible by choosing $n$ large enough since the partition $\mathcal{A}$ is generating. By iterating $n$ times forward we achieve that $\mathcal{T}_n R_n$ covers $\mathcal{J}$ and therefore after $n+1$ iterations the image of $\mathcal{J}$ will cover the entire unit interval. Then for any $x$ in the unit interval:

$$h(x) = P^{n+1}h(x) \geq h(T_w^{-(n+1)}(x))\|DT^{n+1}\|_\infty^{-1} \geq h_*\|DT^{n+1}\|_\infty^{-1},$$

where $T_w^{-(n+1)}$ is one of the inverse branches of $T^{n+1}$ which sends $x$ into $R_n$.

(IV) To prove the continuity property (Lip) we must estimate the difference $||P_{\varepsilon_1} f - P_{\varepsilon_2} f||_1$ for all $f$ in BV. We will adapt for that to the one-dimensional case a similar property proved in the multidimensional setting in Proposition 4.3 in [3] We have

$$\begin{aligned} P_{\varepsilon_1} f(x) - P_{\varepsilon_2} f(x) &= E_1(x) + \sum_{l=1}^{m} (f \cdot \mathbf{1}_{U_n^c})(T^{-1}_{\varepsilon_1,l} x) \left[ \frac{1}{|DT_{\varepsilon_1}(T^{-1}_{\varepsilon_1,l} x)|} - \frac{1}{|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x)|} \right] + \\ &\quad + \sum_{l=1}^{m} \frac{1}{|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x)|} [(f \cdot \mathbf{1}_{U_n^c})(T^{-1}_{\varepsilon_1,l} x) - (f \cdot \mathbf{1}_{U_n^c})(T^{-1}_{\varepsilon_2,l} x)] \\ &= E_1(x) + E_2(x) + E_3(x) \end{aligned}$$

The term $E_1$ comes from those points $x$ which we omitted in the sum because they have only one pre-image in each interval of monotonicity. The total error $E_1 = \int E_1(x)\,dx$ is then estimated by $|E_1| \leq 4m|\varepsilon_1 - \varepsilon_2| \cdot \|\hat{P}_\varepsilon f\|_\infty$. But $\|\hat{P}_\varepsilon f\|_\infty \leq \|f\|_\infty \sum_{l=1}^{m} \frac{|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x')|}{|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x)|} \frac{1}{|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x')|}$, where $x'$ is the point so that $|DT_{\varepsilon_2}(T^{-1}_{\varepsilon_2,l} x')| \cdot |A_l| \geq$

$\eta$, and $\eta$ is the minimum of the length $T(A_k), k = 1, \ldots, m$. Due to the bounded distortion property, the first ratio inside the summation is bounded by some constant $D_c$; therefore

$$E_1 \leq 4m|\varepsilon_1 - \varepsilon_2| \cdot \|f\|_\infty \frac{D_c}{\eta} \sum_{l=1}^{m} |A_l| \leq 4m|\varepsilon_1 - \varepsilon_2| \cdot \|f\|_\infty \frac{D_c}{\eta}$$

We now bound $E_2$. For any $l$, the term in the square bracket (we drop this index in the derivatives in the next formulas), will be equal to $\frac{D^2T(\xi)}{[DT(\xi)]^2}|T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)|$, where $\xi$ is an interior point of $A_l$. The first factor is uniformly bounded by $C_1$. Since $x = T_{\varepsilon_1}(T_{\varepsilon_1}^{-1}(x)) = T((T_{\varepsilon_1}^{-1}(x)) + \varepsilon_1 = T((T_{\varepsilon_2}^{-1}(x)) + \varepsilon_2 = T_{\varepsilon_2}(T_{\varepsilon_2}^{-1}(x))$, we obtain $|T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| = |\varepsilon_1 - \varepsilon_2||DT(\xi')|^{-1}$, for some $\xi' \in A_l$. We now use distortion to replace $\xi'$ with $T_{\varepsilon_1,l}^{-1}x$ and get

$$\begin{aligned}
\int |E_2(x)|\, dx &\leq |\varepsilon_1 - \varepsilon_2| C_1 D_c \int \sum_{l=1}^{m} |f(T_{\varepsilon_1,l}^{-1})| \frac{1}{|DT_{\varepsilon_1}(T_{\varepsilon_1,l}^{-1}x)|}\, dx \\
&= |\varepsilon_1 - \varepsilon_2| C_1 D_c \int P_{\varepsilon_1}(|f|)(x) dx \\
&= |\varepsilon_1 - \varepsilon_2| C_1 D_c \|f\|_1.
\end{aligned}$$

To bound the third error term we use formula (3.11) in [11]

$$\int \sup_{|y-x|\leq t} |f(y) - f(x)| dx \leq 2t\mathrm{Var}(f).$$

and again use the fact that $|T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| = |\varepsilon_1 - \varepsilon_2||DT(\xi')|^{-1}$, for some $\xi' \in A_l$. Integrating $E_3(x)$ yields

$$\int |E_3(x)| dx \leq \ 2m\Lambda^{-1}|\varepsilon_1 - \varepsilon_2|\mathrm{Var}(f\mathbf{1}_{U_n^c}) \leq 10m\Lambda^{-1}\ |\varepsilon_1 - \varepsilon_2|\mathrm{Var}(f)$$

Combining the three error estimates we conclude that there exists a constant $\tilde{C}$ such that

$$||P_{\varepsilon_1} f - P_{\varepsilon_2} f||_1 \leq \tilde{C}|\varepsilon_1 - \varepsilon_2|\|f\|_{BV}.$$

∎

**Theorem 7.6** *Let $\mathcal{F}$ be the family of maps defined above and consisting of the sequence $\{T_{\varepsilon_k}\}$, where the sequence $\{\varepsilon_k\}_{k\geq 1}$ satisfies $|\varepsilon_k| \leq k^{-\theta}$, $\theta > 1/2$. If $\phi$ is not a coboundary for $T$, then*

$$W_n = \sum_{j=0}^{n-1} \phi \circ \mathcal{T}_j$$

*satisfies a standard ASIP with variance $\sigma^2$.*

### 7.3.2 Multidimensional maps

We give here a multidimensional version of the maps considered in the preceding section; these maps were extensively investigated in [34, 20, 3, 2, 21] and we defer to those papers for more details. Let $M$ be a compact subset of $\mathbb{R}^N$ which is the closure of its non-empty interior. We take a map $T : M \to M$ and let $\mathcal{A} = \{A_i\}_{i=1}^m$ be a finite family of disjoint open sets such that the Lebesgue measure of $M \setminus \bigcup_i A_i$ is zero, and there exist open sets $\tilde{A}_i \supset \overline{A_i}$ and $C^{1+\alpha}$ maps $T_i : \tilde{A}_i \to \mathbb{R}^N$, for some real number $0 < \alpha \leq 1$ and some sufficiently small real number $\varepsilon_1 > 0$ such that

1. $T_i(\tilde{A}_i) \supset B_{\varepsilon_1}(T(A_i))$ for each $i$, where $B_\varepsilon(V)$ denotes a neighborhood of size $\varepsilon$ of the set $V$. The maps $T_i$ are the local extensions of $T$ to the $\tilde{A}_i$.

2. there exists a constant $C_1$ so that for each $i$ and $x, y \in T(A_i)$ with $\mathrm{dist}(x, y) \leq \varepsilon_1$,

$$|\det DT_i^{-1}(x) - \det DT_i^{-1}(y)| \leq C_1 |\det DT_i^{-1}(x)| \mathrm{dist}(x, y)^\alpha;$$

3. there exists $s = s(T) < 1$ such that $\forall x, y \in T(\tilde{A}_i)$ with $\mathrm{dist}(x, y) \leq \varepsilon_1$, we have

$$\mathrm{dist}(T_i^{-1}x, T_i^{-1}y) \leq s\, \mathrm{dist}(x, y);$$

4. each $\partial A_i$ is a codimension-one embedded compact piecewise $C^1$ submanifold and

$$s^\alpha + \frac{4s}{1-s} Z(T) \frac{\gamma_{N-1}}{\gamma_N} < 1, \tag{7.7}$$

where $Z(T) = \sup_x \sum_i \#\{\text{smooth pieces intersecting } \partial A_i \text{ containing } x\}$ and $\gamma_N$ is the volume of the unit ball in $\mathbb{R}^N$.

Given such a map $T$ we define locally on each $A_i$ the map $T_\varepsilon$ by $T_\varepsilon(x) := T(x) + \varepsilon$ where now $\varepsilon$ is an $n$-dimensional vector with all the components of absolute value less than one. As in the previous example the translation by $\varepsilon$ is allowed if the image $T_\varepsilon A_i$ remains in $M$: in this regard, we could play with the sign of the components of $\varepsilon$ or do not move the map at all. As in the one dimensional case, we shall also make the following assumption on $\mathcal{F}$. We assume that there exists a set $\mathcal{J}$ satisfying:

(i) $\mathcal{J} \subset T_\varepsilon A_k$ for all $\forall\, T_\varepsilon \in \mathcal{F}$ and for all $k = 1, \ldots, m$.

(ii) $T\mathcal{J}$ is the whole $M$, which in turn implies that there exists $1 \geq L' > 0$ such that $\forall k = 1, \ldots, q$ and $\forall T_\varepsilon \in \mathcal{F}$, diameter$(T_\varepsilon(\mathcal{J}) \cap A_k) > L'$.

As $\mathcal{V} \subset \mathscr{L}^1(m)$ we use the space of quasi-Hölder functions, for which we refer again to [34, 20].

**Theorem 7.7** *Assume $T : M \to M$ is a map as above such that it has only one absolutely continuous invariant measure, which is also mixing. If conditions (i) and (ii) hold, let $\mathcal{F}$ be the family of maps consisting of the sequence $\{T_{\varepsilon_k}\}$, where the sequence $\{\varepsilon_k\}_{k\geq 1}$ satisfies $||\varepsilon_k|| \leq k^{-\theta}$, $\theta > 1/2$. If $\phi$ is not a coboundary for $T$, then*

$$W_n = \sum_{j=0}^{n-1} \phi \circ \mathcal{T}_j$$

*satisfies a standard ASIP with variance $\sigma^2$.*

**Proof** The transfer operator is suitably defined on the space of quasi-Hölder functions, and on this functional space it satisfies a Doeblin-Fortet-Lasota-Yorke inequality. The proof of the lower bound condition (MIN) follows the same path taken in the one-dimensional case in Section 7.3.1 using the distortion bound on the determinants and Assumption (ii) which ensures that the images of the domains of local injectivity of any concatenation have diameter large enough. The positivity of the density follows by the same argument used for maps of the unit interval since the space of quasi-Hölder functions has the nice property that a non-identically zero function in such a space is strictly positive on some ball [34]. Finally, Lipschitz continuity has been proved for additive noise in Proposition 4.3 in [3]. ∎

## 7.4 Covering maps: a general class

We now present a more general class of examples which were introduced in [6] to study metastability for randomly perturbed maps. As before the family $\mathcal{F}$ will be constructed around a given map $T$ which is again defined on the unit interval $M$. We therefore begin to introduce such a map $T$.

**(A1)** There exists a partition $\mathcal{A} = \{A_i : i = 1, \ldots, m\}$ of $M$, which consists of pairwise disjoint intervals $A_i$. Let $\bar{A}_i := [c_{i,0}, c_{i+1,0}]$. We assume there exists $\delta > 0$ such that $T_{i,0} := T|_{(c_{i,0}, c_{i+1,0})}$ is $C^2$ and extends to a $C^2$ function $\bar{T}_{i,0}$ on a neighbourhood $[c_{i,0} - \delta, c_{i+1,0} + \delta]$ of $\bar{A}_i$ ;

**(A2)** There exists $\beta_0 < \frac{1}{2}$ so that $\inf_{x \in I \setminus \mathcal{C}_0} |T'(x)| \geq \beta_0^{-1}$, where $\mathcal{C}_0 = \{c_{i,0}\}_{i=1}^m$.

We note that Assumption **(A2)**, more precisely the fact that $\beta_0^{-1}$ is strictly bigger than 2 instead of 1, is sufficient to get the uniform Doeblin-Fortet-Lasota-Yorke inequality (7.10) below, as explained in Section 4.2 of [17]. We now construct the family $\mathcal{F}$ by choosing maps $T_\varepsilon \in \mathcal{F}$ close to $T_{\varepsilon=0} := T$ in the following way:
Each map $T_\varepsilon \in \mathcal{F}$ has $m$ branches and there exists a partition of $M$ into intervals $\{A_{i,\varepsilon}\}_{i=1}^m$, $A_{i,\varepsilon} \cap A_{j,\varepsilon} = \emptyset$ for $i \neq j$, $\bar{A}_{i,\varepsilon} := [c_{i,\varepsilon}, c_{i+1,\varepsilon}]$ such that

(i) for each $i$ one has that $[c_{i,0} + \delta, c_{i+1,0} - \delta] \subset [c_{i,\varepsilon}, c_{i+1,\varepsilon}] \subset [c_{i,0} - \delta, c_{i+1,0} + \delta]$; whenever $c_{1,0} = 0$ or $c_{q+1}, 0 = 1$, we *do not move* them with $\delta$. In this way we have established a one-to-one correspondence between the unperturbed and the perturbed extreme points of $A_i$ and $A_{i,\varepsilon}$. (The quantity $\delta$ is from Assumption (A1) above.)

(ii) The map $T_\varepsilon$ is locally injective over the closed intervals $\overline{A_{i,\varepsilon}}$, of class $C^2$ in their interiors, and expanding with $\inf_x |T'_\varepsilon x| > 2$. Moreover there exists $\sigma > 0$ such that $\forall T_\varepsilon \in \mathcal{F}, \forall i = 1, \cdots, m$ and $\forall x \in [c_{i,0} - \delta, c_{i+1,0} + \delta] \cap \overline{A_{i,\varepsilon}}$ where $c_{i,0}$ and $c_{i,\varepsilon}$ are two (left or right) *corresponding points* we have:

$$|c_{i,0} - c_{i,\varepsilon}| \leq \sigma \tag{7.8}$$

and

$$|\bar{T}_{i,0}(x) - T_{i,\varepsilon}(x)| \leq \sigma. \tag{7.9}$$

Under these assumptions and by taking, with obvious notations, a concatenation of $n$ transfer operators, we have the uniform Doeblin-Fortet-Lasota-Yorke inequality, namely there exist $\eta \in (0,1)$ and $B < \infty$ such that for all $f \in BV$, all $n$ and all concatenations of $n$ maps of $\mathcal{F}$ we have

$$||P_{\varepsilon_n} \circ \cdots \circ P_{\varepsilon_1} f||_{BV} \leq \eta^n ||f||_{BV} + B||f||_1. \tag{7.10}$$

In order to deal with lower bound condition (MIN), we have to restrict the class of maps just defined. This class was first introduced in an unpublished, but circulating, version of [6]. A similar class has also been used in the recent paper [4]: both are based on the adaptation to the sequential setting of the covering conditions introduced formerly by Collet [10] and then generalized by Liverani [22]. In the latter, the author studied the Perron-Frobenius operator for a large class of uniformly piecewise expanding maps of the unit interval $M$; two ingredients are needed in this setting. The first is that such an operator satifies the Doeblin-Fortet-Lasota-Yorke inequality on the pair of adapted spaces $BV \subset \mathscr{L}^1(m)$. The second is that the cone of functions

$$\mathcal{G}_a = \{g \in BV;\ g(x) \neq 0;\ g(x) \geq 0, \forall x \in M;\ \text{Var}\ g \leq a \int_M g\, dm\}$$

for $a > 0$ is invariant under the action of the operator. By using the inequality (7.10) with the norm $\|\cdot\|_{BV}$ replaced by the total variation Var and using the notation (1.2) for the arbitrary concatenation of $n$ operators associated to $n$ maps in $\mathcal{F}$ we see immediately that

$$\forall n,\ \overline{P}_n \mathcal{G}_a \subset \mathcal{G}_{ua}$$

with $0 < u < 1$, provided we choose $a > B(1-\eta)^{-1}$. The next result from [22] is Lemma 3.2 there, which asserts that given a partition, mod-0, $\mathcal{P}$ of $M$, if each element $p \in \mathcal{P}$ is a connected interval with Lebesgue measure less than $1/2a$, then for each $g \in \mathcal{G}_a$, there exists $p_0 \in \mathcal{P}$ such that $g(x) \geq \frac{1}{2} \int_M g\, dm$, $\forall x \in p_0$. Before continuing we should stress that contrarily to the interval maps investigated above, the domain

of injectivity are now (slightly) different from map to map, and in fact we used the notation $A_{i,\varepsilon_k}$ to denote the $i$ domain of injectivity of the map $T_{\varepsilon_k}$. Therefore the sets (7.5) will be now denoted as

$$A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n} = T^{-1}_{k_1,\varepsilon_1} \circ \cdots \circ T^{-1}_{k_{n-1},\varepsilon_{n-1}} A_{k_n,\varepsilon_n} \cap \cdots \cap T^{-1}_{k_1,\varepsilon_1} A_{k_2,\varepsilon_2} \cap A_{k_1\varepsilon_1}$$

Since we have supposed that $\inf_{T_\varepsilon \in \mathcal{F}, i=1,\ldots,m, x\in A_{i,\varepsilon}} |DT_\varepsilon(x)| \geq \beta_0^{-1} > 2$, it follows that the previous intervals have all lengths bounded by $\beta_0^n$ independently of the concatenation we have chosen. We are now ready to strengthen the assumptions on our maps by requiring the following condition:

**Covering Property:** There exist $n_0$ and $N(n_0)$ such that:
(i) The partition into sets $A^{\varepsilon_1,\cdots,\varepsilon_{n_0}}_{k_1,\cdots,k_{n_0}}$ has diameter less than $\frac{1}{2au}$.
(ii) For any sequence $\varepsilon_1,\ldots,\varepsilon_{N(n_0)}$ and $k_1,\ldots,k_{n_0}$ we have

$$T_{\varepsilon_{N(n_0)}} \circ \cdots \circ T_{\varepsilon_{n_0+1}} A^{\varepsilon_1,\cdots,\varepsilon_{n_0}}_{k_1,\cdots,k_{n_0}} = M$$

We now consider $g = 1$ and note that for any $l$, $\overline{P}_l 1 \in \mathcal{G}_{ua}$. Then for any $n \geq N(n_0)$, we have (from now on using the notation (1.2), we mean that the particular sequence of maps used in the concatenation is irrelevant), $\overline{P}^n 1 = \overline{P}^{N(n_0)} \overline{P}^{n-N(n_0)} 1 := \overline{P}^{N(n_0)} \hat{g}$, where $\hat{g} = \overline{P}^{n-N(n_0)}\, 1$. By looking at the structure of the sequential operators (7.4), we see that for any $x \in M$ (apart at most finitely many points for a given concatenation, which is irrelevant since what one really needs is the $\mathscr{L}^\infty_m$ norm in the condition (MIN)), there exists a point $y$ in a set of type $A^{\varepsilon_1,\cdots,\varepsilon_{n_0}}_{k_1,\cdots,k_{n_0}}$, where $\hat{g}(y) \geq \frac{1}{2}\int_m \hat{g}\, dm$, and such that $T_{\varepsilon_{N(n_0)}} \circ \cdots \circ T_{\varepsilon_1} y = x$. This immediately implies that

$$\overline{P}^n 1 \geq \frac{1}{2\beta_M^{N(n_0)}}, \quad \forall\, n \geq N(n_0),$$

which is the desired result together with the obvious bound $\overline{P}^l 1 \geq \frac{m^{N(n_0)}}{\beta_M}$, for $l < N(n_0)$, and where $\beta_M = \sup_{T_\varepsilon \in \mathcal{F}} \max |DT_\varepsilon|$. The positivity condition (Pos) for the density will follow again along the line used before, since the covering condition holds in particular for the map $T$ itself. About the continuity (Lip): looking carefully at the proof of the continuity for the expanding map of the intervals, one sees that it

extends to the actual case if one gets the following bounds:

$$\left.\begin{array}{r} |T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| \\ |DT_{\varepsilon_1}(x) - DT_{\varepsilon_2}(x)| \end{array}\right\} = O((|\varepsilon_1 - \varepsilon_2|) \tag{7.11}$$

where the point $x$ is in the same domain of injectivity of the maps $T_{\varepsilon_1}$ and $T_{\varepsilon_2}$, the comparison of the *same* functions and derivative in two *different* points being controlled controlled by the condition (7.8). The bounds (7.11) follow easily by adding to (7.8), (7.9) the further assumptions that $\sigma = O(\varepsilon)$ and requiring a continuity condition for derivatives like (7.9) and with $\sigma$ again being of order $\varepsilon$. With these requirement we can finally state the following theorem

**Theorem 7.8** *Let $\mathcal{F}$ be the family of maps constructed above and consisting of the sequence $\{T_{\varepsilon_k}\}$, where the sequence $\{\varepsilon_k\}_{k\geq 1}$ satisfies $|\varepsilon_k| \leq k^{-\theta}$, $\theta > 1/2$. If $\phi$ is not a coboundary for $T$, then*

$$W_n = \sum_{j=0}^{n-1} \phi \circ \mathcal{T}_j$$

*satisfies a standard ASIP with variance $\sigma^2$.*

**Acknowledgement**. AT was partially supported by the Simons Foundation grant 239583. MN was partially supported by NSF grant DMS 1101315. SV was supported by the ANR- Project *Perturbations* and by the PICS (Projet International de Coopération Scientifique), *Propriétés statistiques des systèmes dynamiques detérministes et aléatoires*, with the University of Houston, n. PICS05968. NH and SV thank the University of Houston for the support and kind hospitality during the preparation of this work. SV also acknowledges discussions with R. Aimino about the content of this work. MN thanks Christophe Cuny for very helpful comments which improved this work.

# References

[1] R. Adler, L. Flatto, Geodesic flows, interval maps and symbolic dynamics, Bull. Amer. Math. Soc., **25**, 1992.

[2] J. Alves, J. Freitas, S. Luzzatto, S. Vaienti, From rates of mixing to recurrence times via large deviations, *Advances in Mathematics*, **228** (2011), 1203–123.

[3] H. Aytach, J. Freitas, S. Vaienti, Laws of rare events for deterministic and random dynamical systems, to appear on *Trans. Amer Math. Soc.*

[4] R. Aimino, J. Rousseau, Concentration inequalities for sequential dynamical systems of the unit interval, in preparation.

[5] W. Bahsoun, Ch. Bose, Y. Duan, Decay of correlation for random intermittent maps, arXiv:1305.6588

[6] W. Bahsoun, S. Vaienti, Escape rates formulae and metastability for randomly perturbed maps, *Nonlinearity*, **26** (2013) 1415–1438.

[7] D. Berend, and V. Bergelson, Ergodic and Mixing Sequences of Transformations, *Ergod. Th. & Dynam. Sys.*, **4** (1984), 353–366.

[8] R. Bowen, Equilibrium states and the ergodic theory of Anosov diffeomorphisms, *Lecture Notes in Mathematics*, Vol. 470, 1975.

[9] N. Chernov and D. Kleinbock, Dynamical Borel-Cantelli lemmas for Gibbs measures, *Israel J. Math.*, **122** (2001), 1–27.

[10] P. Collet, An estimate of the decay of correlations for mixing non Markov expanding maps of the interval, preprint (1984).

[11] J.-P. Conze, A. Raugi, Limit theorems for sequential expanding dynamical systems on $[0, 1]$, *Ergodic theory and related fields*, **89121**, Contemp. Math., 430, Amer. Math. Soc., Providence, RI, 2007

[12] C. Cuny and F. Merlevède, Strong invariance principles with rate for "reverse" martingales and applications. *Preprint 2014*.

[13] M. Field, I. Melbourne, A. Török, Decay of Correlations, Central Limit Theorems and approximation by Brownian Motion for compact Lie group extensions. *Ergodic Theory and Dynamical Systems*, **23** (2003), 87–110

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

[15] S. Gouëzel, Central limit theorem and stable laws for intermittent maps *Probab. Theory Relat. Fields*, **128** (2004), 82–122,

[16] S. Gouëzel, Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.*, **38(4)** (2010), 1639–1671.

[17] Gonzalez Tokman, C., Hunt, B. R. and Wright, P.: Approximating invariant densities of metastable systems, *Ergodic Theory Dynam. Systems*, **31(5)** (2011), 1345–1361.

[18] C. Gupta, W. Ott, A. Török, Memory loss for time-dependent piecewise expanding systems in higher dimension, *Mathematical Research Letters*, **20** (2013), 155–175

[19] H Hu, Decay of correlations for piecewise smooth maps with indifferent fixed points, *Ergodic Theory and Dynamical Systems*, **24** (2004), 495–524

[20] H. Hu, S. Vaienti, Absolutely continuous invariant measures for non-uniformly expanding maps, *Ergodic Theory and Dynamical Systems*, **29** (2009), 1185–1215.

[21] N. Haydn, M. Nicol, S. Vaienti and L. Zhang. Central limit theorems for the shrinking target problem, to appear on *Journal of Statistical Physics.*

[22] C. Liverani, Decay of Correlations in Piecewise Expanding maps, *Journal of Statistical Physics*, **78** (1995), 1111–1129.

[23] C. Liverani, B. Saussol, S. Vaienti, A probabilistic approach to intermittency, *Ergodic theory and dynamical systems*, **19** (1999), 671–685.

[24] I. Melbourne and M. Nicol, Almost sure invariance principle for nonuniformly hyperbolic systems. *Commun. Math. Phys.*, **260** (2005), 131–146.

[25] I. Melbourne and M. Nicol, A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probab.*, (2009), 478–505.

[26] Merlevède, F. and Rio, E. Strong approximation of partial sums under dependence conditions with application to dynamical systems. *Stochastic Process. Appl.* **122** (2012), 386–417.

[27] P. Nándori, D. Szász and Y. Varjú. A central limit theorem for time-dependent dynamical systems. **Vol 22(1)** (2006).

[28] W. Ott, M. Stenlund, L.-S. Young, Memory loss for time-dependent dynamical systems, *Math. Res. Lett.*, **16** (2009), 463–475.

[29] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque* **187–188** Société Mathématique de France, Montrouge, 1990.

[30] W. Philipp and W. Stout "Almost sure invariance principles for partial sums of weakly dependent random vectors", Memoirs AMS 1975.

[31] O. Sarig, Subexponential decay of correlations, *Invent. Math.*, **150** (2002), 629–653.

[32] W. Shen, S. Van Strien, On stochastic stability of expanding circle maps with neutral fixed points, *Dynamical Systems, An International Journal*, **28**, (2013)

[33] Vladimir G. Sprindzuk, *Metric theory of Diophantine approximations*, V. H. Winston and Sons, Washington, D.C., 1979, Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics. MR MR548467 (80k:10048).

[34] M. Stenlund, Non-stationary compositions of Anosov diffeomorphisms, *Nonlinearity* **24** (2011) 2991–3018

[35] M. Stenlund, L-S. Young, H. Zhang, Dispersing billiards with moving scatterers, *Comm. Math. Phys.*, to appear (2013)

[36] M. Viana, Stochastic dynamics of deterministic systems, Brazillian Math. Colloquium 1997, IMPA.