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

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

June 6, 2014

Contents

[*]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>.

Abstract

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. In the stationary case 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 the dynamical system. For non-stationary systems self-norming limit theorems similarly transfer.

In the Gordin [14] approach to establishing the central limit theorem (CLT), reverse martingale difference schemes arise naturally. To establish distributional limit theorems for stationary dynamical systems, such as the central limit theorem, it is possible to reverse time 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, and 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 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*

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

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

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

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

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 for a class of expanding maps and Axiom A[1] 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

[1] **???????? have to change this**

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_α 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 $m(A) \neq 0 \implies m(T(A)) \neq 0$. 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* will be defined by the concatenation

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

We denote with P_α the Perron-Frobenius (transfer) operator associated to T_α defined by the duality relation

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

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 σ-algebra associated to the n-fold pull back of the Borel σ-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 σ-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:

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

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

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

In [11] the authors studied extensively a class of β 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 (DFLY) and (LB) conditions.

Property (DFLY) is a uniform Doeblin-Fortet-Lasota-Yorke inequality for concatenations of transfer operators; to introduce it 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 $\|\phi\|_\infty \le C\|\phi\|_\alpha$.

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.

Property (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

$$\|P_n \circ \cdots \circ P_1\ f\|_\alpha \le A\rho^n \|f\|_\alpha + B\|f\|_1 \tag{2.1}$$

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

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

2.1 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.

Let $\mathcal{V}$ be a Banach space with norm $\|.\|_\alpha$ such that $\|\phi\|_\infty \le C\|\phi\|_\alpha$. If (ϕ_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 $\tilde{\phi} = \phi - m(\phi)$ (we write $E[\phi]$ for the expectation of ϕ with respect to Lebesgue measure).

Theorem 2.1 *Let (ϕ_n) be a sequence in $\mathcal{V}$ such that $\sup_n \|\phi_n\|_\alpha < \infty$ and hence $\sup_n E|\phi_n|^4 < \infty$. Assume (DFLY) and (LB)) and $\sigma_n \ge n^{1/4+\delta}$ for some $0 < \delta < \frac{1}{4}$. Then (ϕ_n) satisfies the ASIP i.e. enlarging our probability space if necessary it is possible to find a sequence $(Z_k)_{k\ge 1}$ of independent centered Gaussian variables Z_k such that for any $\beta < \delta$*

$$\sup_{1\le k\le n} |\sum_{i=1}^k \phi_i - \sum_{i=1}^k Z_i| = o(\sigma_n^{1-\beta}) \qquad m - a.s.$$

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 $Q_n \phi = \frac{P_n(\phi \mathcal{P}_{n-1} 1)}{\mathcal{P}_n 1}$. In particular $Q_n T_n \phi = \phi$. Define h_n by $h_n = Q_n \tilde{\phi}_{n-1} + Q_n Q_{n-1} \tilde{\phi}_{n-2} + \cdots + Q_n Q_{n-1} \cdots Q_1 \tilde{\phi}_0$. Then

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

satisfies $Q_{n+1}\psi_n = 0$. For convenience let us put

$$U_n = \mathcal{T}_n \psi_n,$$

where, as before, $\mathcal{T}_n = T_n \circ \cdots \circ T_1$. As proven by Conze and Raugh [11], (U_n) is a sequence of reversed martingale differences for the filtration $(\mathcal{B}_n)$. Note that

$$\sum_{j=1}^{n} U_j = \sum_{j=1}^{n} \tilde{\phi}_j(\mathcal{T}_j) + h_1(\mathcal{T}_1) - h_n(\mathcal{T}_{n+1}) \tag{2.3}$$

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 L^∞ (and $\sigma_n \to \infty$). Thus we may use $\sigma_n = \sum_{j=1}^{n} E[U_i^2]$ as our variance.

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/σ_n^2 is non-increasing and a_n/σ_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

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

for some constant $c_1 > 0$. 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 $\eta > 0$. 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 $\sum_{n\ge 1} a_n^{-v} E(|U_n|^{2v}) < \infty$.

Thus U_n satisfies the ASIP with error term $o(\sigma_n^{1-\beta})$ for any $\beta < \delta$. This concludes the proof, in view of (2.3) and the fact that $\|h_n\|_\alpha$ is uniformly bounded. ∎

2.2 ASIP for the Shrinking Target Problem: expanding maps.

We now consider a fixed expanding map acting on the unit interval (for example a β-transformation, smooth expanding map, the Gauss map, Rychlik maps...) whose transfer operator is quasi compact in the bounded variation norm so that we have exponential decay of correlations in the bounded variation norm. Suppose $\phi_j = 1_{A_j}$ are indicator functions of a sequence of nested intervals A_j, where μ is the unique invariant measure for the map T.

We say that (T, X, μ) 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 $\|\psi\|_1 = \int |\psi| \, d\mu$.

Define $\sigma_n^2 = \mu(\sum_{i=1}^n \tilde{\phi}_i \circ T^i)^2$ where $\tilde{\phi} = \phi - \mu(\phi)$ and $E_n = \sum_{j=1}^n \mu(\phi_j)$.

Theorem 2.2 *Suppose (T, X, μ) is a dynamical system with exponential decay in the BV norm versus $\mathscr{L}^1(\mu)$. Suppose $\phi_j = 1_{A_j}$ are indicator functions of a sequence of nested sets A_j such that $\sup_n \|\phi_n\|_{BV} < \infty$ and $\frac{C_1}{n^\gamma} \leq \mu(A_n)$ ($C_1 > 0$) where $0 < \gamma < 1$. 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 $\beta < \frac{1-\gamma}{2}$*

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

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, $\sigma_n^2 \geq E_n \geq Cn^{1-\gamma}$ for some constant $C > 0$ (note that there is a typo in the statement of [21, Lemma 2.4] and $\limsup$ should be replaced with $\liminf$). We follow the proof of Theorem 2.1 based on [11, Theorem 5.1] taking $T_k = T$ for all k, m as the invariant measure μ and

$f_n = 1_{A_n}$. Note that conditions (DFLY) and (LB) 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 μ 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 $|P\phi|_\infty \le |\phi|_\infty$. 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 ($p > 1$)

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

where

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

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 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

$$\sum_n \frac{\mu(U_n^4)}{a_n^2} \leq c_5 \sum_n \frac{\mu(A_n) + q_{n+1}^4 \mu(A_{n-q_n})}{E_n^{2\alpha}} \leq c_6 \sum_n \frac{q_{n+1}^4 \mu(A_{n-q_n})}{E_{n-q_n}^{2\alpha}} \leq c_7 \sum_n \frac{q_{n+q_n+1}^4 \mu(A_n)}{E_n^{2\alpha}}.$$

Since

$$\frac{E_n^{2\alpha}}{\mu(A_n)} \geq \left(\sum_{j=1}^{n} (\mu(A_j)^{\frac{1}{2\alpha}} \right)^{2\alpha} \geq \left(\sum_{j=1}^{n} j^{-\frac{\gamma}{2\alpha}} \right)^{2\alpha} \geq c_8 n^{2\alpha-\gamma}$$

we obtain the majorisations

$$\sum_n \frac{\mu(U_n^4)}{a_n^2} \leq \sum_n q_{n+q_n+1}^4 n^{\gamma-2\alpha} \leq c_9 \sum_n n^{\gamma-2\alpha} \log^4 n$$

which converge if $\alpha > \frac{1+\gamma}{2}$. 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 $\beta < \frac{1-\gamma}{2}$

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 $\beta < \frac{1-\gamma}{2}$. ∎

Remark 2.3 We are unable with the present proof to obtain an ASIP in the case $\mu(A_n) = \frac{1}{n}$ ($\gamma = 1$) though a CLT has been proven [21, 11].

2.3 ASIP for the Shrinking Target Problem: Axiom A Systems

Suppose (T, X, μ) is an Axiom-A dynamical system. Let B_i be a sequence of nested balls (or rectangles) based at a point $p \in X$. Define $\sigma_n^2 = \mu[\sum_{i=1}^{n} [1_{B_i} - \mu(B_i)]^2$ and $E_n = \sum_{j=1}^{n} \mu(B_j)$. Chernov and Kleinbock [9] established the Strong Borel Cantelli property in the sense that for μ a.e. $x \in x$

$$\lim_{n\to\infty} \frac{1}{E_n} \sum_{j=1}^{n} (1_{B_n} \circ T^n)(x) \to 1$$

In this section we show the ASIP for the sequence of random variables $\{\sum_{j=1}^{n}(1_{B_n} \circ T^n)\}$.

Theorem 2.4 *Suppose (T, X, μ) is an Axiom A dynamical system. Suppose 1_{B_j} are indicator functions of a sequence of nested balls or rectangles B_j in X such that $\frac{C_2}{n^{\gamma_2}} \le \mu(B_n) \le \frac{C_1}{n^{\gamma_1}}$ where $0 < \gamma_1 \le \gamma_2 < 1$. If $2\gamma_2 - \gamma_1 < 1$ then $(1_{B_n} \circ T^n)$ satisfies the ASIP i.e. enlarging our probability space if necessary it is possible to find a sequence $(Z_k)_{k\ge1}$ of independent centered Gaussian variables Z_k such that for small $\gamma > 0$*

$$\sup_{1\le k\le n} |\sum_{i=1}^{k} \phi_i \circ T^i - \sum_{i=1}^{k} Z_i| = o(\sigma_n^{1-\gamma}) \ \mu - a.s.$$

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

Remark 2.5 We are not able to establish an ASIP in the critical case $\mu(B_n) = \frac{1}{n}$.

Proof:

Let ϕ_j be Lipschitz approximations to $1_{B_j} - \mu(B_j)$ such that $\|\phi_j - [1_{B_j} - \mu(B_j)]\|_1 \le j^{-3}$. We arrange that $\mu(\phi_j) = 0$.

It is clear that if we define $\tilde{\sigma}_n^2 = \mu[(\sum_{i=1}^{n} \phi_i)^2]$ then $\sigma_n^2 = \tilde{\sigma}_n^2 + O(\sigma_n)$ so to simplify notation we will use σ_n in our notation instead of introducing $\tilde{\sigma}_n$.

From [21, Lemma 2.4] we see that for sufficiently large n, $\sigma_n^2 \ge E_n \ge Cn^{1-\gamma_2}$ for some constant $C > 0$.

The basic strategy is now the same as that of Field, Melbourne and Török [13]. We use a Markov partition to code (T, X, μ) by a 2-sided shift (σ, Ω, ν) in a standard way [8, 29]. We lift ϕ_j to the system (σ, Ω, ν) keeping the same notation for ϕ_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 ψ_j depends only on future coordinates and is Hölder of exponent $\sqrt{\beta}$ is ϕ_j is Hölder of exponent β. In fact $\|\psi_j\|_{\sqrt{\beta}} \le K\|\phi_j\|_\beta$ and similarly $\|v_j\|_{\sqrt{\beta}} \le K\|\phi_j\|_\beta$ for a uniform constant K.

There is a slight difference in this setting to the usual construction. Choose a Hölder map $G : X \to X$ that depends only on future coordinates we 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 this converges since $|\phi_k(\sigma^{k-n}x) - \phi_k(\sigma^{k-n}Gx)| \leq C\lambda^k\|\phi_k\|_{Lip}$ where $0 < \lambda < 1$. Furthermore as $\|\phi_k\|_{Lip} \leq Ck^4$ there exists a constant K and N such that for all

$$\sum_{j>K\log n} |\phi_k(\sigma^{k-n}x) - \phi_k(\sigma^{k-n}Gx)| \leq \sum_{j>K\log n} C\lambda^k k^4 < 1$$

Hence $\|v_n\|_\infty \leq C\log n$ and $\|v_n\|_\alpha \leq Cn^{\delta'}$ where C, δ' are uniform in n.

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 ψ_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 define $h_n = P\psi_{n-1} + P^2\psi_{n-2} + \cdots + P^n\psi_0$ and

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

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 $\mathcal{F}_n = \sigma^{-n}\mathcal{F}_\prime$. In fact $(UP)f = E[f|\sigma^{-1}\mathcal{F}_\prime]\circ\sigma$ while $(PU)f = f$ (this is easily checked, see [13, Remark 3.1.2] or [29]). It is important to note that as ϕ_j is Lipschitz V_n is Hölder of exponent 1/2 and $\|V_n\|_{\frac{1}{2}} \leq Cn^\kappa$ where C and κ are uniform constants.

We now use results of Viana [36, Proposition 4.4]. Let $\mathcal{G}_0 \subset \mathcal{B}$ be the σ-algebra consisting of those subsets which are unions of local stable leaves (recall that we have a Markov partition which canonically defines the local stable foliation). Define $\mathcal{G}_n = T^{-n}\mathcal{G}_0$. This is a non-increasing sequence of σ-algebras and equivalent via the symbolic representation to the σ-algebra $\mathcal{F}_n$.

Viana shows that for any Hölder function f of exponent β, $\|E(f|\mathcal{G}_n)\|_2 \le C\|f\|_\beta \theta^n$ for some $0 < \theta < 1$ where C and θ are independent of f [36, Remark 4.2].

Since T preserves μ, $\|E[U_n^2 - \mu(U_n^2)|\mathcal{G}_n]\|_2 = \|E[V_n^2 - \mu(V_n^2)|\mathcal{G}_n]\|_2 \le Cn^\delta\theta^n$ for some uniform constants C, δ.

This shows that condition A of Cuny and Merlevede holds for any sequence a_n which grows at most at a polynomial rate by Chebychev and Borel-Cantelli. More precisely writing $F_n = V_n^2 - \mu(V_n^2)$ we see $\mu(|F_n| > a_n) \le \frac{1}{a_n^2}E[F_n^2]$ which is summable for any non-decreasing a_n. Thus Condition (A) of Cuny and Merlevede [12] holds.

We now take $v = 2$ in Condition (B) and estimate

$$\sum_{n\ge 1} a_n^{-2}\mu(|U_n|^4) \le \sum_n Cn^{-2(1-\gamma_2-\epsilon')-\gamma_1}$$

which converges under our assumption $2\gamma_2 - \gamma_1 < 1$ by taking ϵ' sufficiently small.

Thus U_n satisfies the ASIP with error term $o(\sigma_n^{1-\gamma})$ for small $\gamma > 0$. Hence $\psi_n \circ T^n$ satisfies the ASIP with error term $o(\sigma_n^{1-\gamma})$ for small $\gamma > 0$.

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 $|v_n| \le C\log n$ by changing γ slightly we have the ASIP with error term $o(\sigma_n^{1-\gamma})$ for the sequence $\{\phi_n \circ T^n\}$ and hence for $1_{B_n} \circ T^n - \mu(B_n)\}$. This concludes the proof.

2.4 Improvements of published results

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 2.1.

Conze and Raugi [11, Remark 5.2] show that for sequential systems formed by taking maps near a given β-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 (DFLY) and (LB) are satisfied and if ϕ is not a coboundary for T_β 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 (DFLY) and (LB) (the maps all preserve a fixed measure μ, so one can use the transfer operator with respect to μ), and their main condition gives the rate of growth for the variance (see [27, page 1220]). If this rate satisfies the requirement of Theorem 2.1, 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.

2.5 Further assumptions for growth of the variance

We consider here maps for which conditions (DFLY) and (LB) are satisfied, but in order to guarantee the unboundedness of the variance when ϕ is not a coboundary, we need to introduce new assumptions; we follow here again [11], especially Sect. 5. First of all, all 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. Then one considers the following distance between two operators P and Q acting on BV:

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

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

$$\textbf{(DS)} \qquad d(P_r \circ \cdots \circ P_1, P_0^r) \leq M \sum_{j=1}^{r} d(P_j, P_0), \tag{2.4}$$

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

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

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

for all $f \in BV$ of zero (Lebesgue) mean and $n \ge 1$.

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

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

for all integers $p \le n$ and all functions ϕ of bounded variation.

Lipschitz continuity property: Assume that the maps (and their transfer operators) are parametrized by a sequence of numbers ε_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

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

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

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

With this last assumption and (Lip), we get a polynomial decay for (2.4) 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 σ_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 2.6 *[11, Lemma 5.7] Assume the assumptions (Exa), (Lip), (Conv) and (Pos) are satified. If ϕ is not a coboundary for T_0 then σ_n^2/n converges as $n \to \infty$ to σ^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 $G\phi = \sum_{k\geq 0} \frac{P_0^k(h\phi)}{h}$.

2.6 β transformations

Let $\beta > 1$ and denote by $T_\beta(x) = \beta x \bmod 1$ the β-transformation on the unit circle. Similarly for $\beta_k \geq 1 + c > 1$, $k = 1, 2, \ldots$, we have the transformations T_{β_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 (DFLY) was proved in [11, Theorem 3.4 (c)] and condition (LB) 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_ℓ is the transfer operator of T_{β_ℓ}. The invariant density of T_β is bounded below, and continuity (Lip) is precisely the content of Sect. 5 in [11]. We therefore obtain (see [11, Corollary 5.4]):

Theorem 2.7 *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 ϕ is not a coboundary for T_β, 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 σ^2.

2.7 Expanding maps on 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 famile $\mathcal{F}$ then consists of the perturbed maps T_ε which are given by the translations (*additive noise*): $T_\varepsilon(x) = T(x)+\varepsilon,\ \bmod 1$, where $\varepsilon \in (-1,1)$. We observe that the intervals of local injectivity $[v_k, v_{k+1}),\ k = 1, \cdots, m$, of T_ε are independent of ε. 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

$$\inf_{x\in\mathbb{T}} |DT(x)| \geq \Lambda; \qquad \sup_{\varepsilon\in(-1,1)} \sup_{x\in\mathbb{T}} \left|\frac{D^2T_\varepsilon(x)}{DT_\varepsilon(x)}\right| \leq C_1. \tag{2.5}$$

Lemma 2.8 *The maps $\mathcal{F} = \{T_\varepsilon :\ \ |\varepsilon| < 1\}$ satisfy the conditions of Lemma 2.6.*

Proof (I) (DFLY) It is well known that any such map T_ε satisfying (2.5) 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 ε (P_ε is the associated transfer operator of T_ε). 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) (LB) In order to obtain the lower bound property (LB) we have to consider an upper bound for concatenations of operators. Since each T_ε has m intervals of monotonicity we have (where $\mathcal{T}_n = T_{\varepsilon_n} \circ \cdots \circ T_{\varepsilon_1}$ as before)

$$\mathcal{P}_n 1(x) = \sum_{k_n,\cdots,k_1=1}^{m} \frac{1}{|D\mathcal{T}_n(T^{-1}_{k_1,\varepsilon_1} \circ \cdots T^{-1}_{k_n,\varepsilon_n}(x))|} \times \mathbf{1}_{\mathcal{T}_n A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}}(x) \tag{2.6}$$

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

$$A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n} = T^{-1}_{k_1,\varepsilon_1} \circ \cdots \circ T^{-1}_{k_{n-1},\varepsilon_{n-1}} A_{k_n} \cap \cdots \cap T^{-1}_{k_1,\varepsilon_1} A_{k_2} \cap A_{k_1} \tag{2.7}$$

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

$$\mathcal{T}_n A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n} = T_{\varepsilon_n}(A_{k_n} \cap T_{\varepsilon_{n-1}} A_{k_{n-1}} \cap \cdots \cap T_{\varepsilon_{n-1}} \circ \cdots \circ T_{\varepsilon_1} A_{k_1}) \tag{2.8}$$

[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)

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

for some points ι_j in $\mathcal{T}_{j-1}A^{\varepsilon_1,\cdots,\varepsilon_n}_{k_1,\cdots,k_n}$. Using the second bound in (2.5) and the fact that $|\mathcal{T}_{j-1}(u) - \mathcal{T}_{j-1}(v)| \le \Lambda^{-(j-1)}$ we finally have

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

which in turn implies that

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

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 2.6 the following result:

Theorem 2.9 *Let $\mathcal{F}$ be a family of functions as described in this section. Then for any function ϕ which is not a coboundary for T_β we have that the random variables*

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

satisfies an ASIP.

2.8 Covering maps: special cases

2.8.1 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_ε, 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_ε 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 ε so that the image $T_\varepsilon(A_k)$ stays in the unit interval; this we achieve for a given ε by choosing the sign of ε 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_w so that:

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

Lemma 2.10 *The maps T_ε satsify the conditions (DFLY), (LB), (Pos) and (Lip).*

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

(II) In order to prove the lower bound condition (LB) 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 (2.8), 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 A_w and therefore after $n+1$ iterations the image of A_w 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

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

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 $\|\hat{P}_\varepsilon f\|_\infty \leq \|f\|_\infty \sum_{l=1}^m \frac{|DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1}x')|}{|DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1}x)|} \frac{1}{|DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1}x')|}$, where x' is the point so that $|DT_{\varepsilon_2}(T_{\varepsilon_2,l}^{-1}x')| \cdot |A_l| \geq$

η, and η is the minimum of the length $T(A_k), k = 1, \ldots, m$. The first ratio inside the summation is bounded by the distortion constant $D_c = \Lambda\lambda^{-1}$; 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 ξ 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 ξ' 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

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

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 2.11 *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 ϕ is not a cobaundary for T, then*

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

satisfies an ASIP.

2.8.2 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 ε 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 ∂A_i is a codimension-one embedded compact piecewise C^1 submanifold and

   $$s^\alpha + \frac{4s}{1-s} Z(T) \frac{\gamma_{N-1}}{\gamma_N} < 1, \tag{2.9}$$

   where $Z(T) = \sup_x \sum_i \#\{\text{smooth pieces intersecting } \partial A_i \text{ containing } x\}$ and γ_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_ε by $T_\varepsilon(x) := T(x) + \varepsilon$ where now ε is an n-dimensional vector with all the components of absolute value less than one. As in the previous example the translation by ε 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 ε 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_w satisfying:

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

(ii) TA_ω 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$(T_\varepsilon(A_w) \cap A_k) > L'$.

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 2.12 *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 ϕ is not a cobaundary for T, then*

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

satisfies an ASIP.

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 (LB) follows the same path taken in the one-dimensional case in Section 2.8.1 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 inner 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, the closeness has been proved for additive noise in Proposition 4.3 in [3]. ∎

2.9 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 $\mathcal{A} = \{A_i : i = 1, \dots, m\}$ 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 β_0^{-1} is strictly bigger than 2 instead of 1, is sufficient to get the uniform Doeblin-Fortet-Lasota-Yorke inequality (2.12) 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 δ. 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 δ is from Assumption (A1) above.)

(ii) The map T_ε 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:

$$|c_{i,0} - c_{i,\varepsilon}| \leq \sigma \tag{2.10}$$

and

$$|\bar{T}_{i,0}(x) - T_{i,\varepsilon}(x)| \leq \sigma. \tag{2.11}$$

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

$$||P_{\varepsilon_n} \circ \cdots \circ P_{\varepsilon_1} f||_{BV} \leq \eta^n ||f||_{BV} + B||f||_1. \tag{2.12}$$

In order to deal with lower bound condition (LB), 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 (2.12) 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,ε_k} to denote the i domain of injectivity of the map T_{ε_k}. Therefore the sets (2.7) will be now denoted as

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

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 β_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 $A_{k_1,\cdots,k_{n_0}}^{\varepsilon_1,\cdots,\varepsilon_{n_0}}$ 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

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

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 (2.6), 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 $\mathscr{L}_m^\infty$ norm in the condition (LB)), there exists a point y in a set of type $A_{k_1,\cdots,k_{n_0}}^{\varepsilon_1,\cdots,\varepsilon_{n_0}}$, 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:

$$\left.\begin{array}{r} |T_{\varepsilon_1}^{-1}(x) - T_{\varepsilon_2}^{-1}(x)| \\ |DT_{\varepsilon_1}(x) - DT_{\varepsilon_2}(x)| \end{array}\right\} = O((|\varepsilon_1 - \varepsilon_2|) \tag{2.13}$$

where the point x is in the same domain of injectivity of the maps T_{ε_1} and T_{ε_2}, the comparison of the *same* functions and derivative in two *different* points being controlled controlled by the condition (2.10). The bounds (2.13) follow easily by adding to (2.10), (2.11) the further assumptions that $\sigma = O(\varepsilon)$ and requiring a continuity condition for derivatives like (2.11) and with σ again being of order ε. With these requirement we can finally state the following theorem

Theorem 2.13 *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 ϕ is not a cobaundary for T, then*

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

satisfies an ASIP.

Acknowledgement. AT was partially supported by the Simons Foundation grant 239583. 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.

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.