Decay of correlations on towers with non-Hölder Jacobian and non-exponential return time

Jérôme Buzzi[1]
Véronique Maume-Deschamps[2]

[1] Centre de Mathématiques de l'Ecole Polytechnique
[2] Université de Bourgogne / Laboratoire de Topologie

April 2003

Abstract

We establish upper bounds on the rate of decay of correlations of tower systems with summable variation of the Jacobian and integrable return time. That is, we consider situations in which the Jacobian is not Hölder and the return time is only subexponentially decaying. We obtain a subexponential bound on the correlations, which is essentially the slowest of the decays of the variation of the Jacobian and of the return time.

Introduction

In this paper we study the speed of mixing, more precisely the rate of decay of correlations, of tower systems, a special class of countable Markov systems which naturally arise in the study of many dynamical systems by the procedure of **induction** – see [Y1]. Our goal is to provide a comprehensive statement in the following sense. There are two sources of loss of exponential speed: large return times and bad smoothness. By extending cone techniques, we deal simultaneously with both difficulties whereas previous works on decay of correlations [KMS, BFG, Po, Y1] considered only one of these two obstructions. We prove that, although the analysis becomes more difficult when both obstructions are present, they nevertheless operate independently: the speed is just the minimum of the speeds allowed 1) by the defect in smoothness if the statistics of return

AMS classifiction 1991: Primary 58F11
key words: absolutely continuous invariant measures; equilibrium states; decay of correlations; transfer operator; tower extension.

times were exponential; 2) by the statistics of return times if we had Hölder smoothness.

Let us state informally a corollary of our result:

THEOREM 0.1 *Consider a tower system F with a mixing invariant probability measure $\hat\mu$. Assume that the oscillation of the Jacobian on n-cylinders is bounded by $n^{-\alpha}$ and the probability of return time n decays like $n^{-\beta}$. Then, for sufficiently smooth observables, the rate of decay of correlations is:*

$$\left|\int \phi\cdot\psi\circ F^n\,d\hat\mu - \int\phi\,d\hat\mu\int\psi\,d\hat\mu\right| \le C\cdot K(\phi)\|\psi\|_{L^1}\cdot\frac{1}{n^{\min(\alpha,\beta-\varepsilon)-1}},$$

for any $\varepsilon>0$. $K(\phi)$ is some finite number depending only on ϕ ; $\|\psi\|_{L^1}$ is the L^1 norm w.r.t. the reference measure.

Remarks.
1. Our result allows returns which are not onto, which is quite convenient for applications.
2. The fact that the above bound depends on ψ only through its L^1-norm is important for the study of asymptotic laws of return times [C, CGS, Pa].
3. The loss of ε in the exponent is probably due to our method (in the Hölder-continuous case ($\alpha=\infty$ so to speak), L.-S. Young [Y1] obtained $\mathcal{O}(n^{-\beta-1})$).

An application to non-Hölder maps with an indifferent fixed point.
Fix $0<\gamma<1/2$ and $\alpha>1$ and consider the interval map $f:[0,1]\to[0,1]$ defined by $f(x)=2^{1+\gamma}(x+x^{1+\gamma})/(2^\gamma+1)$ for $x<1/2$ and $f(x)=\frac{3}{2}(x-\frac{1}{2})+\frac{(x-\frac{1}{2})(\log 2)^\alpha}{2|\log(x-\frac{1}{2})|^\alpha}$ for $x>1/2$. Our result implies a rate of correlation in $\mathcal{O}(n^{-\min(\alpha,1/\gamma-\varepsilon)+1})$ for arbitrarily small $\varepsilon>0$. Our approach is the first to our knowledge to be able to treat such maps.

Section 1 contains the precise statement of our results. We briefly recall definitions and properties of Birkhoff's cones and projective metrics (section 2) and the construction of the a.c.i.m., establishing regularity of the invariant density (section 3). We define a sequence of cones C_j of "Lipschitz" functions (w.r.t. to an ad-hoc metric) in section 4 and then establish that the transfer operator iterated some k_j times sends one cone into the next by a γ_j-contraction in section 5 for some semi-explicit $\gamma_j<1$. Finally in section 6, we deduce from this a convergence in the uniform norm at speed $\prod_{p=1}^{j}\gamma_p$, with j largest such that $k_1+\cdots+k_j\le n$, and make this estimate explicit in the exponential, stretched exponential and polynomial cases.

Indeed, using the notations of Section 1, one considers the tower defined by $\Delta_0=[\frac{1}{2},1]$, the map $f_0(x)=f^{R(x)}(x)$ with the return function $R(x)=\min\{n\ge1 : f^n(x)\ge\frac{1}{2}$ and $\#\{k<n : f^k(x)\ge\frac{1}{2}\}\ge\varepsilon_0 n\}$ for some small $\varepsilon_0>0$. Then one can prove that $\omega_n=\mathcal{O}(n^{-\alpha})$ and $\nu(\Delta_n)=\mathcal{O}(n^{-1/\gamma})$ and apply our main theorem.

Acknowledgments: The authors are grateful to *program ESF/PRODYN* which has partially supported the *International Conference on Dynamical Systems, Abbey of "La Bussière"* where part of this work was carried through.

1 Setting, statement of the results

Let us describe our tower model which follows Young's [Y1]. A tower is defined by:

- a basis, which is a probability space (Δ_0, m_0) together with a non-singular self-map f_0;
- a partition $\Delta_{0,j}$, $j \in \mathbb{N}$ such that $f_0 : \Delta_{0,j} \to f_0(\Delta_{0,j})$ is one-to-one and satisfies $f_0(\Delta_{0,j})$ is a union of some $\Delta_{0,k}$, for some k's;
- a return time, i.e., a function $R : \Delta_0 \to \mathbb{N}$, constant on each $\Delta_{0,j}$, $j \in \mathbb{N}$.

The tower Δ is then the disjoint union of the floors Δ_ℓ, $\ell \in \mathbb{N}$:

$$\Delta_\ell = \{(x, \ell) \mid x \in \Delta_0,\ R(x) > \ell\}.$$

It is endowed with the measure $\hat{\nu}$ which is just the restriction of the copy of m_0 on each floor. We will denote by $\Delta_{\ell,j}$, $\ell < R_{|\Delta_0^j}$ the copy of $\Delta_{0,j}$ inside Δ_ℓ:

$$\Delta_{\ell,j} = \{(x, \ell) \mid x \in \Delta_{0,j},\ R(x) > \ell\}.$$

The dynamic on the tower, $F\ : \Delta \longrightarrow \Delta$, is defined by

$$\begin{cases} F(x,\ell) &= (x, \ell+1) \text{ if } R(x) > \ell + 1 \\ &= (f_0(x), 0) \text{ otherwise.} \end{cases}$$

One can think of F as the unfolding of the underlying induction: in applications, F will be often conjugate to the original map, which f_0 is some (variable) power. We assume that the partition $\mathcal{R} = \{\Delta_{\ell,j}\}$ generates in the sense that $\bigvee_{i=0}^{\infty} F^{-i}\mathcal{R}$ is the partition into points mod $\hat{\nu}$. For $k \in \mathbb{N}$, the elements of the partition $\mathcal{R}^{(k)} = \bigvee_{i=0}^{k-1} F^{-i}\mathcal{R}$ are called **cylinders or k-cylinders**. We denote by $C_k(x)$ the element of $\mathcal{R}^{(k)}$ which contains x. Let JF be the Jacobian of F with respect to $\hat{\nu}$ (this Jacobian is well defined because of the non singularity of f_0). The modulus of continuity of JF will be controlled by the following, dynamics-dependent sequence:

$$\omega_n = \sup_{C \in \mathcal{R}^n} \sup_{x,y \in C} \log \frac{JF(x)}{JF(y)}.$$

For $x, y \in \Delta$ the **separation time** $s(x, y)$ is the largest integer $n \geq 0$ such that for all $0 \leq j \leq n$, $F^j(x)$ and $F^j(y)$ belong to the same atom of the partition $\mathcal{R}$. Set

$$d_0(x, y) = \sum_{j \geq s(x,y)+1} \omega_j.$$

Note that the metric d_0 is designed so that the family of functions:

$$\log JF^n = \log \prod_{i=0}^{n-1} JF \circ F^i.$$

are uniformly Lipschitz w.r.t. it.
Let us summarize our assumptions on the tower.

(A.I) **Summability of upper floors.**

$$\sum_{\ell \in \mathbb{N}} \hat{\nu}\left(\{x \in \Delta_0 \mid R(x) > \ell\}\right) = 1.$$

(A.II) **Generating Partition.** The partition $\mathcal{R}$ generates under F i.e.: the partition $\bigvee_{n=0}^{\infty} F^{-n}\mathcal{R}$ is the partition into points. In particular, d_0 defines a metric on Δ.

(A.III) **Summable variation.** Let JF be the Jacobian of F with respect to $\hat{\nu}$. We assume that JF satisfies:

$$\sum_{n \in \mathbb{N}} \omega_n < \infty.$$

(A.IV) **Large image and Markov properties.** Each $F^R\Delta_{0,j}$ is a union of some $\Delta_{0,p}$, $p \in \mathbb{N}$, (**Markov property**) and (**Large image**):

$$\eta := \inf_{j \in \mathbb{N}} \hat{\nu}(F^R(\Delta_{0,j})) > 0.$$

Contrarily to [Y1] we do not assume the Bernoulli property: $f_0(\Delta_{0,j}) = \Delta_0$, but only the weaker Markov property above. The collection of sets $f_0(\Delta_{0,j})$ defines a partition $\mathcal{B}$ which is less refined than $\{\Delta_{0,j}\}_{j \in \mathbb{N}}$, so that it is in particular countable $\mathcal{B} = \{B_1, B_2, \dots\}$. Remark that, by an easy induction, if x, y are contained in the same element of $\mathcal{B}$, then the pre-images of all orders of x and y are paired in the following sense.

Given $x, y \in \Delta_0$, say that $x', y' \in \Delta$ are **paired pre-images** if $F^n x' = x$, $F^n y' = y$ and $F^k(x')$ nd $F^k(y')$ belong to the same element of $\mathcal{R}$ for all $0 \leq k < n$. Observe that (A.III) implies that in this situation we have:

$$\left| \frac{JF^n(x')}{JF^n(y')} - 1 \right| \leq C \cdot d_0(x, y), \qquad \text{with } C = \exp \sum_{j \geq 1} \omega_j. \tag{1.1}$$

This is "bounded distortion".

Remark *In [BM], we proved that multi-dimensional piecewise expanding maps in higher dimension are (under quite general hypothesis) conjugate to such a tower map.*

Let $\mathbf{L}(d_0)$ be the space of bounded functions on Δ that are **locally Lipschitz** with respect to the metric d_0, i.e., for some $K < \infty$, for all x, y in the same $B_{j,\ell}$,

$$|\varphi(x) - \varphi(y)| \leq K d_0(x, y).$$

$K(\varphi)$ is the smallest number K such that the above inequality is satisfied. Let $\|\varphi\|_{\mathbf{L}(d_0)} = K(\varphi) + \|\varphi\|_\infty$ be the norm on $\mathbf{L}(d_0)$.

To study the ergodic properties of F, we have to decompose it into topologically mixing components. Observe that $\mathcal{R}$ has a natural graph structure: $P \to Q$ iff $F(P) \supset Q$. Its (restricted) **spectral decomposition** is $\mathcal{P} = \mathcal{P}_t \cup \bigcup_i \bigcup_{j=0}^{p_i - 1} \mathcal{R}_j^{(i)}$, where:

- $\mathcal{P}_t$ is the set of **transient** elements of $\mathcal{P}$, i.e., elements P such that there exists a path from P going to some $Q \in \mathcal{P}$ and there is no path from Q to P (observe that we don't decompose this part into irreducible subsets). The elements that are not transient are called **recurrent**.
- for each i, $\bigcup_{j=0}^{p_i-1} \mathcal{R}_j^{(i)}$ is the set of $P \in \mathcal{R}$ such that there exist paths from P to Q and Q to P, for some fixed $Q = Q(i)$ (i.e., these unions are the irreducible components of $\mathcal{R}$ from which no arrows leave).
- if there is an arrow from $\mathcal{R}_j^{(i)}$ to $\mathcal{R}_l^{(k)}$ then $k = i$ and $l = j + 1 \bmod p_i$.

Finally, $\Delta_j^{(i)}$ is the union of the elements of $\mathcal{R}_j^{(i)}$. Observe that, up to trivialities, it is enough to study the dynamics of $F^{p_i} : \Delta_0^{(i)} \to \Delta_0^{(i)}$ for each i. We call this the **spectral reduction**.

Our main result is the following theorem.

Theorem 1.1 *Let $(\Delta, F, \hat\nu)$ be a tower system satisfying (A.I - IV). First, there exists an invariant probability measure absolutely continuous with respect to $\hat\nu$ (a $\hat\nu$-a.c.i.m. for short).*
Second, any $\hat\nu$-a.c.i.m. $\hat\mu$, up to the spectral reduction, is mixing, with the following speed estimate: for all $\varphi \in \mathbf{L}(d_0)$ and $\psi \in L^\infty(\Delta)$,

$$\left| \int_\Delta \varphi \circ F^n \cdot \psi \, d\hat\mu - \int_\Delta \varphi \, d\hat\mu \int_\Delta \psi \, d\hat\mu \right| \leq C \cdot \|\psi\|_{\mathbf{L}(d_0)} \|\varphi\|_{L^1(\hat\mu)} \cdot u_n \qquad \text{for all } n \geq 0$$

for some $C < \infty$ and a sequence $u = (u_n)_{n=0}^\infty$ converging to zero which can be made explicit:

- *if* $\omega_n = O(\rho^n)$ *for some* $0 < \rho < 1$ *and* $\hat{\nu}(\Delta_n) = O(\alpha^n)$ *for some* $0 < \alpha < 1$ *then* $u_n = \kappa^n$ *for some* $0 < \kappa < 1$*,*
- *if* $\omega_n = O(n^{-\alpha})$ *for some* $\alpha > 1$ *and* $\hat{\nu}(\Delta_n) = O(n^{-\beta})$ *for some* $\beta > 1$ *then* $u_n = n^{-\min(\alpha,\beta-\varepsilon)-1}$ *for all* $\varepsilon > 0$*.*
- *if* $\omega_n = O(e^{-n^\alpha})$ *and* $\hat{\nu}(\Delta_n) = O(e^{-n^\beta})$ *for some* $0 < \alpha, \beta < 1$*, then* $u_n = e^{-n^{\min(\alpha,\beta)-\varepsilon}}$ *for all* $\varepsilon > 0$*.*

2 Birkhoff's cones and projective metrics

The main tool for the proof of Theorem 1.1 will be the theory of cones and projective metrics of Garrett Birkhoff [Bi]. P. Ferrero and B. Schmitt [FS] applied it to estimate the correlation decay for random products of matrices. Recently this strategy has been used by many authors to obtain exponential decay of correlations (see for example [Li1]). We are closer to [KMS] and [M] which have used these techniques in a different way to obtain sub-exponential decay of correlations. Let us recall definitions and properties of cones and projective metrics (see [Li1] for a more complete presentation). Let B be a vector space and let $C \subset B$ be a **Birkhoff cone**, i.e., a cone with the following properties.

- C is convex,
- $C \cap -C = \{0\}$,
- if α_n is a sequence of real numbers such that $\alpha_n \to \alpha$ and $x - \alpha_n y \in C$ for all n, then $x - \alpha y \in C$. This property is called "integral closure".

Such a cone is endowed with the pseudo-metric δ_C on C defined in the following way (it is pseudo because it is not necessarily finite and it does not separate points). For $x, y \in C$,

$$\mu(x,y) = \inf\{\beta > 0 \text{ such that } \beta x - y \in C\}.$$

with the convention: $\mu(x,y) = \infty$ if the corresponding set is empty. Let $\delta_C(x,y) = \log \mu(x,y)\mu(y,x)$. We remark that δ_C satisfies the triangle inequality: if $\beta x - y \in C$ and $\tilde{\beta} y - z \in C$ then $\beta\tilde{\beta} x - z \in C$ since C is a convex cone, so $\mu(x,z) \leq \mu(x,y) \cdot \mu(y,z)$ and the triangle inequality follows. Finally, observe that δ_C is projective: $\delta(x,y) = 0 \iff x$ and y are colinear.

The usefulness of this projective metric is that it allows a 'geometric' proof of the contraction through the following result.

THEOREM 2.1 [Bi] *Let* C *and* C' *be two cones,* P *a linear operator* $P : C \to C'$*. Let* Γ *denote the diameter of* PC *in* C'*:*

$$\Gamma = \sup_{f,g \in C} \delta_{C'}(Pf, Pg) \leq \infty.$$

For any f, g in C, we have:

$$\delta_{C'}(Pf, Pg) \leq \tanh\left(\frac{\Gamma}{4}\right)\delta_C(f, g).$$

This theorem implies that a linear map between cones never increases distances and is in fact a contraction as soon as $\Gamma < \infty$.
The following result allows the translation of contraction w.r.t. cone metric to a contraction w.r.t. norm. A norm $\| \ \|$ on B is **adapted to** C, if for f and g in B such that both $f+g$ and $f-g$ belong to C, then $\|g\| \leq \|f\|$. $\rho : C \to \mathcal{R}_+$ is a **homogeneous form adapted to** C if, i) for any $\lambda > 0$; ii) $f \in C$, $\rho(\lambda f) = \lambda \rho(f)$ and if $f - g \in C$ implies $\rho(g) \leq \rho(f)$.

Theorem 2.2 [Bi], [Li1]. *Let C be a Birkhoff cone, let $\| \ \|$ and ρ be adapted to C. For any f and g in C such that $\rho(f) = \rho(g) \neq 0$ we have:*

$$\|f - g\| \leq (e^{\delta(f,g)} - 1)\min(\|f\|, \|g\|).$$

3 Construction of a $\hat{\nu}$-a.c.i.m.

As usual, the **transfer operator** acting on bounded functions is defined by:

$$\mathcal{L}_0 f(x) = \sum_{Fy=x} \frac{1}{JF(y)} f(y).$$

The measure $\hat{\nu}$ is conformal for $\mathcal{L}_0$ in the following sense: for any bounded function f,

$$\int \mathcal{L}_0 f d\hat{\nu} = \int f d\hat{\nu}.$$

For $s \in \mathbb{N}$, the **s-cylinders** are the non empty sets of the form: $\bigcap_{i=0}^{s-1} F^{-i} A_i$ with $A_i \in \mathcal{R}$. For $k \in \mathbb{N}$ and $x \in \Delta$, $C_k(x)$ denotes the k-cylinder which contains x.
The following lemmas are technical tools to study $\mathcal{L}_0^n$.

Lemma 3.1 *There exists $C < \infty$ such that for any $\ell \in \mathbb{N}$ and any $x \in \Delta_\ell$ and $k \in \mathbb{N}$ with $F^k x \in \Delta_0$,*

$$C^{-1}\hat{\nu}(C_k(x)) \leq \frac{1}{JF^k(x)} \leq C\hat{\nu}(C_k(x)).$$

Proof:Let $x \in \Delta_\ell$ such that $F^k(x) \in \Delta_0$. The Markov property and the large image property (A.IV) imply that $\hat{\nu}(F^k C_k(x)) \geq \eta > 0$. The bounded distortion property (1.1) gives:

$$\begin{array}{rcl} C^{-1}\dfrac{\hat\nu(C_k(x))}{\hat\nu(F^kC_k(x))} & \le \dfrac{1}{JF^k(x)} \le & C\dfrac{\hat\nu(C_k(x))}{\hat\nu(F^kC_k(x))} \\ C^{-1}\dfrac{\hat\nu(C_k(x))}{1} & \le \dfrac{1}{JF^k(x)} \le & C\dfrac{\hat\nu(C_k(x))}{\eta} \end{array}$$

The Lemma is proved. ∎

LEMMA 3.2 *There exists $K<\infty$ such that:*

- *for all $x\in\Delta$, all $n\in\mathbb{N}$, $\mathcal{L}_0^n\mathbf{1}(x)\le K$.*
- *for all x,y in a given $B_{j,\ell}$ and all $n\in\mathbb{N}$,:*

$$|\mathcal{L}_0^n\mathbf{1}(x)-\mathcal{L}_0^n\mathbf{1}(y)|\le Kd_0(x,y). \tag{3.1}$$

Proof: The upper bound $\mathcal{L}_0^n\mathbf{1}\le K$ follows from Lemma 3.1, by writing:

$$\mathcal{L}_0^n\mathbf{1}(x)=\sum_{x'\in F^{-n}x}\frac{1}{JF^n(x')}\le C\sum_{x'\in F^{-n}x}\hat\nu(C_n(x'))\le C \tag{3.2}$$

Let x and y belong to one $B_{j,\ell}$. Their preimages by F^n are paired, i.e., if $F^nx'=x$, there is exactly one $y'\in C_n(x')$ such that $F^ny'=y$. So, using (A.IV), we get:

$$\begin{aligned} |\mathcal{L}_0^n\mathbf{1}(x)-\mathcal{L}_0^n\mathbf{1}(y)| &= \left|\sum_{F^nx'=x}JF^n(x')^{-1}-\sum_{F^ny'=y}JF^n(y')^{-1}\right| \\ &= \left|\sum_{F^nx'=x}JF^n(x')^{-1}\left(\frac{JF^n(x')}{JF^n(y')}-1\right)\right| \\ &\le C\mathcal{L}_0^n\mathbf{1}(x)d_0(x,y) \\ &\le KCd_0(x,y) \end{aligned}$$

(3.1) is proved. ∎

COROLLARY 3.3 *F admits a $\hat\nu$ a.c.i.m.*

Proof: By Lemma 3.2, the sequence $\frac{1}{n}\sum_{i=0}^{n-1}\mathcal{L}_0^i\mathbf{1}$ is relatively compact for the topology of uniform convergence on compact subsets (this is Arzela-Ascoli theorem on separable spaces). Each limit point h of this sequence is a non zero fixed point for $\mathcal{L}_0$ (by Lebesgue's dominated theorem, $\hat\nu(h)=1$), so that $\hat\mu=h\hat\nu$ is a $\hat\nu$-a.c.i.m. ∎

The system $(\Delta, F, \hat{\nu}, \mathcal{R})$ has a Markov structure in the sense that for each $P \in \mathcal{R}$, $F(P)$ is a union of atom of $\mathcal{R}$. According to [ADU], we will say that F is **aperiodic** if:

$$\forall\ P,\ P' \in \mathcal{R}\ \exists N \in \mathbb{N} \text{ such that } \hat{\nu}(F^{-n}P \cap P') > 0\ \forall\ n \geq N. \tag{3.3}$$

The existence of a $\hat{\nu}$-a.c.i.m. implies that the recurrent part is non empty (it contains the support of $\hat{\mu}$). Up to the spectral reduction, **we may and shall assume that F is aperiodic**. We remark that aperiodicity implies that any s-cylinder has positive $\hat{\nu}$-measure. The following lemma implies that any s-cylinder also has $\hat{\mu}$ positive measure.

LEMMA 3.4 *If F is aperiodic then $h(x) > 0$ for all $x \in \Delta$. Moreover,* $\inf \hat{\mu}[F^k(C_k(x))] > 0$ *where the inf is taken on all $k \in \mathbb{N}$ and x such that $F^k x \in \Delta_0$.*

Proof: Theorems 2.5 and 3.2 in [ADU] imply that if F is aperiodic then $\mathcal{L}_0^n \mathbf{1} \to h$ uniformly on compact sets. Let K be given by Lemma 3.2. We have for $j = 1, \ldots,$ any $\ell, n \in \mathbb{N}$, $x, y \in B_{\ell,j}$, their paired preimages will be denoted by x' and y',

$$\begin{aligned} \mathcal{L}_0^n \mathbf{1}(x) &= \sum_{F^n x' = x} JF^n(x')^{-1} = \sum_{F^n y' = y} JF^n(y')^{-1} \frac{JF^n(y')}{JF^n(x')} \\ &\leq (Cd_0(x,y) + 1) \sum_{F^n y' = y} JF^n(y')^{-1} \leq K \mathcal{L}_0^n \mathbf{1}(y) \text{ using (1.1).} \end{aligned} \tag{3.4}$$

Taking the limit when n goes to infinity implies: for $x, y \in B_{\ell,j}$,

$$h(x) \leq K h(y). \tag{3.5}$$

So, for all (j, ℓ), either $h \equiv 0$ on $B_{\ell,j}$ or $h > 0$ on $B_{\ell,j}$. But $h|_{B_{j,\ell}} \equiv 0$ implies that $\hat{\nu}(B_{j,\ell}) = 0$, a contradiction to the aperiodicity. This concludes the proof of the first part of the lemma.

To prove the second part, let us remark first that $\hat{\nu}[F^k C_k(x)] \geq \eta > 0$ for all k and x such that $F^k x \in \Delta_0$. Also, the Markov property implies that there exists finitely many integers $i_1, \ldots, i_p$ such that each $F^k C_k(x)$ contains at least one Δ_{0,i_j}, $j = 1, \ldots, p$. This implies the announced result using that $h > 0$. ■

Let us note that Lemma 3.4 implies that $\hat{\mu}(P) > 0$ for any cylinder P. The following lemma is a direct consequence of mixing.

LEMMA 3.5 *There exists positive numbers A and B such that for any $f,\ g \in L^2(\hat{\nu})$, with $\hat{\mu}(f) > 0$, $\hat{\mu}(g) > 0$, there exists n_0 such that for $n \geq n_0$,*

$$A \leq \frac{\hat{\mu}(f \circ F^n \cdot g)}{\hat{\mu}(f)\hat{\mu}(g)} \leq B.$$

We shall now construct a sequence of cones C_j and a sequence of integers k_j such that $\mathcal{L}^{k_j}$ maps C_{j-1} into C_j with uniformly bounded diameter, where $\mathcal{L}$ is the normalized transfer operator defined as follows: $\mathcal{L}f = \frac{1}{h}\mathcal{L}_0(fh)$. Because of Lemma 3.4, $\mathcal{L}$ is well defined. Moreover it satisfies: $\mathcal{L}\mathbf{1} = \mathbf{1}$. The Jacobian of F^n with respect to $\hat{\mu}$ is:

$$\frac{JF^n \cdot h \circ F^n}{h}$$

Let x and y belong to the same $B_{\ell,j}$, x' and y' be their paired preimages by F^n. Following eqs. (3.4 - 3.5), we get, for some $C' > C$:

$$\begin{aligned}(1 - C'd_0(x', y')) &\leq \tfrac{h(x')}{h(y')} \leq (1 + C'd_0(x', y')),\\ (1 - C'd_0(x, y)) &\leq \tfrac{h\circ F^n(x')}{h\circ F^n(y')} \leq (1 + C'd_0(x, y)).\end{aligned}$$

We deduce that the Jacobian of F^n with respect to $\hat{\mu}$ satisfies a bounded distortion inequality like (1.1) with an appropriate constant that we will continue to denote by C. From now on, we abuse notations and JF will be the Jacobian of F with respect to the invariant measure $\hat{\mu}$. We remark that the proof of Lemma 3.1 and Lemma 3.4 give for some $C > 0$:

$$C^{-1}\hat{\mu}(C_k(x)) \leq \frac{1}{JF^k(x)} \leq C\hat{\mu}(C_k(x)). \tag{3.6}$$

4 The cones

4.1 Auxiliary definitions

In what follows $\hat{\mu}$ is a mixing a.c.i.m. on Δ.

We need first some auxiliary definitions. We set for convenience $D = 5$.
Let $(v_n)_{n\in\mathbb{N}}$ be such that:

$$\sum_{n\geq 1} v_n \cdot \hat{\nu}(\Delta_n) < \infty \text{ and } v_n \to \infty.$$

We also assume that v_n, and for each $k \in \mathbb{N}$, v_n/v_{n+k} are non-decreasing functions of n. We define $\hat{\mu}_v := v \cdot \hat{\mu}$ where we have introduced the function $v = \sum_{\ell\geq 0} v_\ell \cdot 1_{\Delta_\ell}$. Let $R_0(p) = \sum_{k>p} \omega_k(g)$. We pick an integer s so large that

$$R_0(s) \leq 10^{-5}.$$

Let $P_\infty = \bigcup_{\substack{i\geq t\\ \ell\geq 0}} \Delta_{\ell,i} \cup \bigcup_{\substack{\ell\geq t\\ i\geq 0}} \Delta_{\ell,i}$ with the parameter t chosen so large that:

$$\frac{\hat{\mu}_v(P_\infty)}{\eta} \leq 10^{-5},$$

where $\eta = \inf_j \hat{\nu}(F^R(\Delta_{0,j})) > 0$.
Let $\mathcal{Q}_1$ be the finite collection of s-cylinders covering $\Delta \setminus P_\infty$. Let $\mathcal{Q}$ be the finite partition of Δ defined as $\mathcal{Q}_1 \cup \{P_\infty\}$. Let k_0 be such that for all $k \geq k_0$, for all $P, Q \in \mathcal{Q}$:

$$\frac{7}{8} \leq \frac{\hat{\mu}(F^{-k}P \cap Q)}{\hat{\mu}(P)\hat{\mu}(Q)} \leq \frac{9}{8}$$

$$\frac{7}{8} \leq \frac{\hat{\mu}_v(F^{-k}P \cap Q)}{\hat{\mu}(P)\hat{\mu}_v(Q)} \leq \frac{9}{8}$$

Such a k_0 exists as $(F, \hat{\mu})$ is mixing, $\mathcal{Q}$ is finite and the function v is in $L^1(\hat{\mu})$.

4.2 Definition of the distances d_j

We set $d_j(x,y) = R_j(s(x,y))$ where the functions $R_j(\cdot)$ are defined inductively in the following way. Recall that $R_0(\cdot)$ and k_0 have been defined above. Assuming that $R_{j-1}(\cdot)$ is defined we set:

$$k_j = \min\{k \geq k_0 : R_{j-1}(s+k) \leq D^{-1}R_0(s)\}$$

and

$$R_j(p) = D[R_0(p) + R_{j-1}(p + k_j)].$$

We observe that, $\mathcal{Q}_1$ being a collection of s-cylinders, its d_j-diameter is bounded by $R_j(s) = D[R_0(p) + R_{j-1}(p + k_\ell)] \leq (D+1)R_0(s)$, a number independent of j. We introduce the auxiliary values $q(j) = k_1 + \cdots + k_j$.

4.3 Definition of the cones

As stated in the introduction, we are going to prove Theorem 1.1 by cone techniques. Let us explain a bit how to construct the cones and how sub exponential decay of correlations may be obtained.

We start by recalling the classical way of using cones (see [FS] and [Li1] for details). To get exponential decay of correlations, it is sufficient to find a cone C and an integer k such that $\mathcal{L}^k$ maps C into itself and the diameter Γ of $\mathcal{L}^k C$ into C is finite. If the fixed point h of $\mathcal{L}$ belongs to C then Theorem 2.1 gives, for any integer j:

$$\delta_C(\mathcal{L}^{kj} f,\ h) \leq \gamma^{j-1}\Gamma \text{ where } \gamma = \tanh\frac{\Gamma}{4} < 1.$$

Hence Theorem 2.2 gives that for $f \in C$, $\|\mathcal{L}^p f\ -\ hm(f)\|$ goes to zero exponentially fast for $\|\ \|$ an adapted norm, provided $f \mapsto m(f)$ is adapted. Then one has to extend this result from the cone to a Banach space.

The starting point of the construction of cones is usually a Lasota-Yorke inequality (it will be done in section 5.2). If the metric d_0 is not of exponential type (i.e., $d_0(x,y) \le \beta^{s(x,y)}$ with $0 < \beta < 1$), then we cannot obtain a Lasota-Yorke inequality. This is why we have to introduce the sequence of metric d_j and a sequence of cones. Roughly speaking, for any integer j, we will consider a cone C_j of functions f that are locally Lipschitz for the metric d_j and the Lipschitz constant of which is controlled (**see condition 2. below**). Thanks to the definition of the metric d_j and to the Lasota-Yorke inequality, if f belongs to C_j then $\mathcal{L}^{k_j} f$ will be locally Lipschitz with respect to the metric d_{j+1} and we will control its Lipschitz constant. This will imply that $\mathcal{L}^{k_j}$ maps C_j into C_{j+1} with finite diameter Γ. Then, using Theorems 2.1 and 2.2, we get that for f in C_0, $\mathcal{L}^{k_1+\cdots+k_j} f$ goes to $m(f)h$ at rate γ^j (with $\gamma = \frac{\tanh \Gamma}{4}$) in any adapted norm, provided that h belongs to all the cones C_j and that $f \mapsto m(f)$ is adapted. This is the philosophy of the construction.
A source of difficulty is the following. To ensure that a cone C satisfies properties of section 2 and more specifically the condition $C \cap -C = \{0\}$, some positivity for the functions in the cone is needed. On the other hand, if $C \subset \{f \ge 0\} =: C_+$ then for any f, g in C, $\theta_C(f,g) \ge \theta_{C_+}(f,g)$ (use Theorem 2.1 with $P = Id$) and

$$\theta_{C_+}(f,g) = \frac{\sup f}{\inf f} \cdot \frac{\sup g}{\inf g}.$$

Since the functions of the cones are only *locally* Lipschitz, we will have a good control on $\frac{\sup f}{\inf f}$ on each floor Δ_ℓ but not on the whole space Δ. Observe that because of the definition of $\mathcal{L}$, we cannot hope to control globally Lipschitz constant (just try to compute $|\mathcal{L}f(x) - \mathcal{L}f(y)|$ for $x \in \Delta_0$ and $y \in \Delta_\ell$, $\ell > 0$) and we have to restrict ourselves to locally Lipschitz functions. This problem is solved by considering the finite partition $\mathcal{Q}$ of Δ which is decomposed into finitely many s-cylinders (the "compact" part) and the complementary of the union of these s-cylinders (the "non compact" part). Then we require the positivity of some kind of conditional expectation of f with respect to this finite partition (**see condition 1. below**). This together with the control of the local Lipschitz constant leads to a good control of f on the atoms of the compact part. Then, we requir! e ! ! another kind of control on the non compact part (**see conditions 3 and 4 below**).

The cone $C_j(a,b,c)$ is the set of all real functions f on Δ satisfying the following conditions:

1. $a \cdot E_{\hat\mu}(f) \le E_{\hat\mu}(f|\mathcal{Q}) \le 6b \cdot E_{\hat\mu}(f)$.
2. for all $x, y \in \Delta$ with $\mathcal{B}(x) = \mathcal{B}(y)$,

$$|f(x) - f(y)| \le 12b \cdot E_{\hat\mu}(f) \cdot d_j(x,y).$$

3. for all $\ell \le q(j)$,
$$\sup_{\mathcal{P}_\infty \cap \Delta_\ell} |f| \le 90c \cdot v_\ell \cdot E_{\hat\mu}(f).$$

4. for all $\ell > q(j)$,
$$\sup_{\mathcal{P}_\infty \cap \Delta_\ell} |f| \le 90c \cdot v_{q(j)} \cdot E_{\hat\mu}(f).$$

5 Contraction of the cones

The purpose of this section is to prove the following proposition.

PROPOSITION 5.1 *We have*

$$\mathcal{L}^{k_j} C_j(0,1,1) \subset C_{j+1}\left(\frac{4}{5}, \frac{1}{5}, \max\left(\frac{1}{5}, \frac{v_{q(j)}}{v_{q(j+1)}}\right)\right)$$

and $\mathcal{L}^{k_j} : C_j(0,1,1) \to C_{j+1}(0,1,1)$ *admits, w.r.t. cone metrics, a contraction coefficient less than:*

$$\max\left(\frac{1}{5}, \frac{v_{q(j)}}{v_{q(j+1)}}\right) =: \gamma_j.$$

Let us prove that if $\mathcal{L}^{k_j} C_j(0,1,1) \subset C_{j+1}\left(\frac{4}{5}, \frac{1}{5}, \max\left(\frac{1}{5}, \frac{v_{q(j)}}{v_{q(j+1)}}\right)\right)$ then we have the announced estimation on the contraction rate.
This will follow if we prove that for all $f, g \in C_{j+1}(4/5, 1/5, \max(1/5, v_{q(j)}/v_{q(j+1)}))$ with the normalization $E_{\hat\mu}(f) = E_{\hat\mu}(g) = 1$ we have:

$$\alpha f - g \in C_{j+1}(0,1,1)$$

for

$$\alpha = \max\left(\frac{1+D^{-1}}{1-D^{-1}}, \frac{1+v_{q(j)}/v_{q(j+1)}}{1-v_{q(j)}/v_{q(j+1)}}\right). \tag{5.1}$$

Indeed, in that case, we have that the diameter Γ_j of $\mathcal{L}^{k_j} C_j(0,1,1)$ into $C_{j+1}(0,1,1)$ is less than

$$2 \log \max\left(\frac{1+D^{-1}}{1-D^{-1}}, \frac{1+v_{q(j)}/v_{q(j+1)}}{1-v_{q(j)}/v_{q(j+1)}}\right).$$

and then

$$\tanh \frac{\Gamma_j}{4} \le \max\left(\frac{1}{D}, \frac{v_{q(j)}}{v_{q(j+1)}}\right) = \max\left(\frac{1}{5}, \frac{v_{q(j)}}{v_{q(j+1)}}\right).$$

The upper bound in the cone condition (1) for $\alpha f - g$ is, for all $P \in \mathcal{Q}$,

$$\alpha \ge \frac{6E_{\hat\mu}(g) - E_{\hat\mu}(g|P)}{6E_{\hat\mu}(f) - E_{\hat\mu}(f|P)}.$$

The right hand side is bounded by:

$$\frac{6-0}{6-\frac{6}{5}} = \frac{5}{4} \leq 6/4.$$

The lower bound in this condition is, for all $P \in \mathcal{Q}$,

$$\alpha \geq \frac{E_{\hat{\mu}}(g|P)}{E_{\hat{\mu}}(f|P)}.$$

The right hand side is upper bounded by:

$$\frac{6/5 \cdot E_{\hat{\mu}}(g)}{4/5 \cdot E_{\hat{\mu}}(f)} = \frac{6}{4}.$$

Thus, both bounds in condition (1) are implied by eq. (5.1).

The cone condition (2) is, for all $x, y \in \Delta$ with $\mathcal{B}(x) = \mathcal{B}(y)$

$$\alpha \geq \frac{12E_{\hat{\mu}}(g) + |g(x) - g(y)|}{12E_{\hat{\mu}}(f) - |f(x) - f(y)|}$$

The right hand side is bounded by:

$$\frac{1 + D^{-1}}{1 - D^{-1}} = 6/4.$$

Thus, condition (2) is implied by eq. (5.1).

The cone condition (3) is implied by eq. (5.1) as can be seen by practically identical computations.

The cone condition (4) is satisfied iff, for all $x \in P_\infty \cap \Delta_\ell$, $\ell > q(j+1)$,

$$\alpha \geq \frac{90v_{q(j+1)}E_{\hat{\mu}}(g) + |g(x)|}{90v_{q(j+1)}E_{\hat{\mu}}(f) - |f(x)|}.$$

But the right hand side is bounded by:

$$\frac{1 + (v_{q(j)}/v_{q(j+1)})}{1 - (v_{q(j)}/v_{q(j+1)})}.$$

Thus, condition (4) is implied by eq. (5.1) and this concludes the proof that the claim implies the stated contraction coefficient.

5.1 Contraction of the first condition

f is an arbitrary function in $C_j(0, 1, 1)$ for the remainder of section 5.

Let $P \in \mathcal{Q}$. We first prove the lower bound:

$$\begin{aligned}
E_{\hat\mu}(\mathcal{L}^{k_j} f|P) &= \frac{1}{\hat\mu(P)} \int_P \mathcal{L}^{k_j} f \, d\hat\mu = \frac{1}{\hat\mu(P)} \int_\Delta 1_P \cdot \mathcal{L}^{k_j} f \, d\hat\mu \\
&= \frac{1}{\hat\mu(P)} \int_\Delta 1_P \circ F^{k_j} \cdot f \, d\hat\mu = \frac{1}{\hat\mu(P)} \int_{F^{-k_j}P} f \, d\hat\mu \\
&\geq \sum_{P' \in \mathcal{Q}_1} \frac{1}{\hat\mu(P)} \int_{F^{-k_j}P \cap P'} f \, d\hat\mu + \frac{1}{\hat\mu(P)} \int_{F^{-k_j}P \cap P_\infty} f \, d\hat\mu \\
&\geq \sum_{P' \in \mathcal{Q}_1} \frac{\hat\mu(F^{-k_j}P \cap P')\hat\mu(P')}{\hat\mu(P)\hat\mu(P')} \Big\{ E_{\hat\mu}(f|P') - 12(D+1)R_0(s)E_{\hat\mu}(f) \Big\} \\
&\qquad - \sum_{\ell \geq 0} \frac{\hat\mu(F^{-k_j}P \cap P_\infty \cap \Delta_\ell)}{\hat\mu(P)} \cdot 90 v_{\min(\ell, q(j))} E_{\hat\mu}(f),
\end{aligned}$$

using $\operatorname{diam}_{d_j}(\mathcal{Q}_1) \leq (D+1)R_0(s)$, conditions (2)-(4). We continue (obviously: $v_\ell \geq v_{\min(\ell,q(j))}$):

$$\begin{aligned}
E_{\hat\mu}(\mathcal{L}^{k_j} f|P) &\geq \sum_{P' \in \mathcal{Q}_1} \frac{\hat\mu(F^{-k_j}P \cap P')}{\hat\mu(P)\hat\mu(P')} \hat\mu(P') \Big\{ E_{\hat\mu}(f|P') - 12(D+1)R_0(s)E_{\hat\mu}(f) \Big\} \\
&\qquad - \frac{\hat\mu_v(F^{-k_j}P \cap P_\infty)}{\hat\mu(P)\hat\mu_v(P_\infty)} \hat\mu_v(P_\infty) \cdot 90 E_{\hat\mu}(f) \\
&\geq \sum_{P' \in \mathcal{Q}_1} \frac{7}{8} \hat\mu(P') E_{\hat\mu}(f|P') - \sum_{P' \in \mathcal{Q}_1} \frac{7}{8} \hat\mu(P') \cdot 12(D+1)R_0(s)E_{\hat\mu}(f) \\
&\qquad - \frac{9}{8} \hat\mu_v(P_\infty) \cdot 90 E_{\hat\mu}(f) \\
&\geq \frac{7}{8} \Big\{ E_{\hat\mu}(f) - \int_{P_\infty} f \, d\hat\mu \Big\} - 12 \frac{9}{8}(D+1)R_0(s)E_{\hat\mu}(f) \\
&\qquad - 90 \frac{9}{8} \hat\mu_v(P_\infty) E_{\hat\mu}(f).
\end{aligned}$$

Observe that:

$$\begin{aligned}
\int_{P_\infty} f \, d\hat\mu &= \int_{P_\infty} \frac{f}{v} \, d\hat\mu_v \leq \sum_{\ell \geq 0} \hat\mu_v(P_\infty \cap \Delta_\ell) \cdot 90 \frac{v_{\min(\ell,q(j))}}{v_\ell} E_{\hat\mu}(f) \\
&\leq 90 \hat\mu_v(P_\infty) E_{\hat\mu}(f).
\end{aligned}$$

Hence,

$$\begin{aligned}
E_{\hat\mu}(\mathcal{L}^{k_j} f|P) &\geq \left\{ \frac{7}{8} - 90 \left(\frac{9}{8} + \frac{7}{8} \right) \hat\mu_v(P_\infty) - 12 \frac{9}{8}(D+1)R_0(s) \right\} E_{\hat\mu}(f) \\
&\geq \frac{4}{5} E_{\hat\mu}(f).
\end{aligned}$$

Similarly, we get the upper bound,

$$\begin{aligned}E_{\hat\mu}(\mathcal{L}^{k_j}f|P) &\le \left\{\frac{9}{8}+90\left(\frac{9}{8}+\frac{7}{8}\right)\hat\mu_v(P_\infty)+12\frac{9}{8}(D+1)R_0(s)\right\}E_{\hat\mu}(f)\\ &\le 6D^{-1}E_{\hat\mu}(f).\end{aligned}$$

5.2 Contraction of the second condition

Let $x, y \in \Delta_\ell$ with $\mathcal{B}(x)=\mathcal{B}(y)$. First assume that $\ell \ge k_j$. Setting $x^- = (x, \ell-k_j)$, $y^- = (y, \ell-k_j) \in \Delta$ (with a slight abuse of notation), we have

$$\begin{aligned}|\mathcal{L}^{k_j}f(x)-\mathcal{L}^{k_j}f(y)| &= |f(x^-)-f(y^-)| \le 12 d_j(x^-,y^-)E_{\hat\mu}(f)\\ &= 12R_j(s(x,y)+k_j) \le 12D^{-1}d_{j+1}(x,y).\end{aligned}$$

Now assume that $\ell < k_j$. We have $|\mathcal{L}^{k_j}f(x)-\mathcal{L}^{k_j}f(y)| = |\mathcal{L}^r f(x^0)-\mathcal{L}^r f(y^0)|$ with $r = k_j-\ell$ and $x^0=(x,0)$, $y^0=(y,0)$ (with the same abuse). Hence it is enough to bound $|\mathcal{L}^r f(x)-\mathcal{L}^r f(y)|$ for $r \le k_j$ and $x, y \in \Delta_0$ with $\mathcal{B}(x)=\mathcal{B}(y)$. As $\mathcal{B}(x)=\mathcal{B}(y)$, the pre-images by F^r of x and y can be paired (i.e., to each pre-image x' of x corresponds a pre-image y' of y defined by the same inverse branch). Thus,

$$\begin{aligned}|\mathcal{L}^r f(x)-\mathcal{L}^r f(y)| &\le \textstyle\sum_{x'\in F^{-r}x}\left|\frac{f(x')}{JF^r(x')}-\frac{f(y')}{JF^r(y')}\right|\\ &\le \textstyle\sum_{x'\in F^{-r}x}\frac{1}{JF^r(x')}|f(x')-f(y')|+\\ &\quad+\textstyle\sum_{x'\in F^{-r}x}|f(y')|\frac{1}{JF^r(x')}\left|\frac{JF^r(x')}{JF^r(y')}-1\right|\\ &\le 12R_j(r+s(x,y))E_{\hat\mu}(f)\\ &\quad+Cd_0(x,y)\left(\sum_{\substack{x'\in P_\infty\\ x'\in F^{-r}x}}\frac{|f(y')|}{JF^r(x')}+\sum_{\substack{x'\notin P_\infty\\ x'\in F^{-r}x}}\frac{|f(y')|}{JF^r(x')}\right)\end{aligned}$$

recall $\mathcal{L}1=1$ and C is defined in (A.III). We have

$$\begin{aligned}\sum_{\substack{x'\in P_\infty\\ x'\in F^{-r}x}} &\le \sum_{\ell\ge0}\sum_{\substack{x'\in P_\infty\cap\Delta_\ell\\ x'\in F^{-r}x}} 90v_{\min(\ell,q(j))}\frac{1}{JF^r(x')}E_{\hat\mu}(f)\\ &\le 180\frac{K}{\eta}\hat\mu_v(P_\infty)E_{\hat\mu}(f)\end{aligned}$$

where K is given by the bounded distortion and η by the large image property (we have used that $\int_{F^rC_r(x')}(JF^r)^{-1}\,d\hat\mu = \hat\mu(C_r(x')) \ge \eta/K(JF^r(x'))^{-1}$).
We also have

$$\begin{aligned}\sum_{\substack{x'\notin P_\infty\\ x'\in F^{-r}x}} &\le \left\{E_{\hat\mu}(f|\mathcal{Q})(x')+12R_j(s)E_{\hat\mu}(f)\right\}\\ &\le (12R_j(s)+1)E_{\hat\mu}(f).\end{aligned}$$

Hence,

$$\begin{aligned}|\mathcal{L}^{k_j} f(x) - \mathcal{L}^{k_j} f(y)| &\leq \Big\{12R_j(k_j - \ell + s(x^0, y^0)) + d_0(x^0, y^0)MC(12R_j(s) \\ &\quad +1 + 180\frac{K}{\eta}\hat{\mu}_v(P_\infty))\Big\} E_{\hat{\mu}}(f) \\ &\leq 12\,\{R_j(k_j + s(x,y)) + R_0(s(x,y))(MCR_j(s) + 1/12 \\ &\quad +15\frac{K}{\eta}\hat{\mu}_v(P_\infty))\Big\} E_{\hat{\mu}}(f)\end{aligned}$$

Now, $CR_j(s) + 15(K/\eta)\hat{\mu}_v(P_\infty) < 1/2$ so that

$$|\mathcal{L}^{k_j} f(x) - \mathcal{L}^{k_j} f(y)| \leq 12D^{-1}R_{j+1}(s(x,y))E_{\hat{\mu}}(f) = 12D^{-1}d_{j+1}(x,y)E_{\hat{\mu}}(f).$$

5.3 Contraction of the third condition

Let $x \in \Delta_\ell \cap P_\infty$. First assume $0 \leq \ell \leq k_j$. As for the second condition it is enough to consider $\mathcal{L}^r f(x)$ with $0 \leq r \leq k_j$ and $x \in \Delta_0$. We have

$$\begin{aligned}|\mathcal{L}^r f(x)| &\leq \sum_{x' \in F^{-r}x} \frac{1}{JF^r(x')}|f(x')| \\ &\leq \sum_{\substack{x' \in F^{-r}x \\ x' \notin P_\infty}} \frac{1}{JF^r(x')}\,(12R_j(s)E_{\hat{\mu}}(f) + 6E_{\hat{\mu}}(f)) \\ &\quad + \sum_{\ell \geq 0} \sum_{\substack{x' \in F^{-r}x \\ x' \in P_\infty \cap \Delta_\ell}} \frac{1}{JF^r(x')} 90 v_{\min(\ell, q(j))} E_{\hat{\mu}}(f) \\ &\leq (2R_j(s) + 6)E_{\hat{\mu}}(f) + \sum_{\ell \geq 0} \sum_{\substack{x' \in F^{-r}x \\ x' \in P_\infty \cap \Delta_\ell}} \frac{K}{\eta}\hat{\mu}(C_r(x')) \cdot 90 v_{\min(\ell, q(j))} E_{\hat{\mu}}(f) \\ &\leq (12R_j(s) + 6)E_{\hat{\mu}}(f) + \sum_{\ell \geq 0} \sum_{\substack{x' \in F^{-r}x \\ x' \in P_\infty \cap \Delta_\ell}} \frac{K}{\eta}\hat{\mu}_v(C_r(x')) \cdot 90 v_{\min(\ell, q(j))} E_{\hat{\mu}}(f) \\ &\leq \left((12R_j(s) + 6) + 90\frac{K}{\eta}\hat{\mu}_v(P_\infty)\right) E_{\hat{\mu}}(f) \\ &\leq 90D^{-1}E_{\hat{\mu}}(f)\end{aligned}$$

Now assume $k_j \leq \ell \leq q(j+1) = q(j) + k_j$ and let $x^- = (x, \ell - k_j)$. We have:

$$\begin{aligned}|\mathcal{L}^{k_j} f(x)| = |f(x^-)| &\leq 90 v_{\ell - k_j} E_{\hat{\mu}}(f) \\ &\leq 90\frac{v_{\ell - k_j}}{v_\ell} v_\ell E_{\hat{\mu}}(f) \\ &\leq 90\frac{v_{q(j)}}{v_{q(j+1)}} v_\ell E_{\hat{\mu}}(f).\end{aligned}$$

using that $\ell \mapsto v_\ell / v_{\ell+k}$ is increasing for any k.
We get the claimed contraction by $\max(1/5, \frac{v_{q(j)}}{v_{q(j+1)}})$.

5.4 Contraction of the fourth condition

Finally we take $x \in P_\infty \cap \Delta_\ell$ with $\ell > q(j+1)$. We have

$$\begin{aligned} |\mathcal{L}^{k_j} f(x)| = |f(x^-)| &\leq 90 v_{q(j)} E_{\hat{\mu}}(f) \\ &\leq 90 \frac{v_{q(j)}}{v_{q(j+1)}} v_{q(j+1)} E_{\hat{\mu}}(f). \end{aligned}$$

and this gives the contraction by $\frac{v_{q(j)}}{v_{q(j+1)}}$.

6 Conclusion

To conclude the proof of Theorem 1.1, we need to derive from the projective metric bound obtained above, bounds on the correlations. The following lemma is standard when using Birkhoff's cones (see [KMS] page 687, [M] Lemmas 3.9-3.10). Let $|||\ |||_j$ be the norm on bounded functions defined by:

$$\begin{aligned} |||f|||_j \quad = \quad \max \Bigg[& \max(90 v_{q(j)}, 12 D R_0(s) + 6) \left| \int_\Delta f d\hat{\mu} \right|, \\ & \sup_{P \in \mathcal{Q}} \hat{\mu}(P)^{-1} \left| \int_P f d\hat{\mu} \right|, \ \|f\|_\infty \Bigg]. \end{aligned}$$

Lemma 6.1 *The norms $|||\ |||_j$ and the homogeneous form $f \mapsto \hat{\mu}(f)$ are adapted to the cones $C_j(0,1,1)$*
For any $f \in \mathbf{L}(d_0)$, there exists $R(f) > 0$ such that $f + R(f)\mathbf{1} \in C_0(0,1,1)$ and $R(f) \leq C \|f\|_{\mathbf{L}(d_0)}$.

Sketch of proof : It is clear that the homogeneous form $f \mapsto \hat{\mu}(f)$ is adapted. To prove that $|||\ |||_j$ is also adapted, let us consider f and g such that $f+g$ and $f-g$ are in $C_j(0,1,1)$. The first condition in the definition of the cone gives:

$$\forall\ P \in \mathcal{Q}, \left| \frac{1}{\hat{\mu}(P)} \int_P g d\hat{\mu} \right| \leq \frac{1}{\hat{\mu}(P)} \int_P f d\hat{\mu},$$

and

$$\left| \int_\Delta g d\hat{\mu} \right| \leq \int_\Delta f d\hat{\mu}.$$

The last three conditions give

$$\|g\|_\infty \le \max[90 v_{q(j)}, 12DR_0(s)+6] \int_\Delta f d\hat{\mu}.$$

Hence, we have $|||g|||_j \le |||f|||_j$.
To prove the second point of the lemma, we may assume that $f \ge 0$. To have that $f + R(f)\mathbf{1} \in C_0(0,1,1)$, it suffices that:

- $\forall P \in \mathcal{Q}$, $R(f) \ge \frac{\frac{1}{\hat{\mu}(P)}\int_\Delta P f d\hat{\mu} - 6\int_\Delta f d\hat{\mu}}{5}$, so that condition 1. is satisfied,
- $R(f) \ge \frac{L(f)}{12}$, so that condition 2. is satisfied,
- $R(f) \ge \frac{\sup f}{90 v_{q(0)} - 1}$, so that conditions 3 and 4 are satisfied,

so we may choose $R(f) \le \text{const}||f||_{\mathbf{L}(d_0)}$. ■

Let us conclude the proof of Theorem 1.1.
Let $f \in C_0(0,1,1)$, By Proposition 5.1, for any ℓ, $\mathcal{L}^{k_1+\cdots+k_\ell} f$ and $\mathcal{L}^{k_1+\cdots+k_\ell}\mathbf{1} = \mathbf{1}$ belong to C_ℓ (we remark that $\mathbf{1} \in C_0(0,1,1)$). Applying $\ell-1$ times Theorem 2.1, we get:

$$\delta_{C_\ell}(\mathcal{L}^{k_1+\cdots+k_\ell} f, \mathbf{1}) \le \prod_{j=2}^{\ell} \gamma_j \cdot \delta_1(\mathcal{L}^{k_1} f, \mathbf{1}) \le \prod_{j=2}^{\ell} \gamma_j \cdot \Gamma_1,$$

where $\gamma_j =$ is given by Proposition 5.1. Since the norm $|||\ |||_j$ is adapted to the cones $C_j(0,1,1)$ and is greater than the uniform norm, Theorem 2.2 gives for f in $C_0(a,b,c)$ with $\hat{\mu}(f) = 1$,

$$\|\mathcal{L}^{k_1+\cdots k_\ell} f - \mathbf{1}\|_\infty \le \text{const} \prod_{j=2}^{\ell} \gamma_j.$$

For $n \in \mathbb{N}$, let $\ell(n)$ be defined by:

$$n = k_1 + \cdots k_{\ell(n)} + r \text{ with } r < k_{\ell(n)+1},$$

we have

$$\|\mathcal{L}^n f - \mathbf{1}\|_\infty \le \|\mathcal{L}^r \mathbf{1}\|_\infty \ \|\mathcal{L}^{k_1+\cdots+k_\ell} f - \mathbf{1}\|_\infty \le \text{const} \left(\prod_{j=2}^{\ell(n)} \gamma_j\right).$$

For any function $f \in \mathbf{L}(d_0)$, applying the above inequality to $\frac{f+R(f)\mathbf{1}}{\hat{\mu}(f)+R(f)}$ gives

$$\|\mathcal{L}^n f - \hat{\mu}(f)\|_\infty \le \text{const} \cdot \left(\prod_{j=2}^{\ell(n)} \gamma_j\right) \|f\|_{\mathbf{L}(d_0)}.$$

The decay of correlations follows: for $f \in L$ and $g \in L^1(\hat{\mu})$,

$$\begin{aligned} \left| \int_\Delta g \circ F^n \cdot f d\hat{\mu} - \int_\Delta f d\hat{\mu} \int_\Delta g d\hat{\mu} \right| &= \left| \int_\Delta g[\mathcal{L}^n(f) - \hat{\mu}(f)] d\hat{\mu} \right| \\ &\leq \text{const} \cdot \left(\prod_{j=2}^{\ell(n)} \gamma_j \right) \|f\|_{\mathbf{L}(d_0)} \|g\|_1. \end{aligned}$$

Set $\prod_{j=2}^{\ell(n)} \gamma_j = u_n$.

We now prove that u_n has the announced behavior for exponential, stretched exponential or polynomial sequences ω_n and $\hat{\nu}(\Delta_n)$.

Estimating u_n

To estimate the rate of mixing u_n, we have to analyze the asymptotic behavior of the sequences k_j and γ_j. Recall that:

$$k_{\ell+1} = \inf\{k \geq k_0 : R_\ell(s+k) \leq \frac{R_0(s)}{D}\}$$

and observe that an easy induction gives:

$$R_\ell(n) = D^\ell R_0(k_1 + \cdots + k_\ell + n) + \sum_{i=1}^{\ell} D^{\ell+1-i} R_0(k_{i+1} + \cdots + k_\ell + n). \quad (6.1)$$

Recall $R_0(m) = \sum_{k>m} \omega_k$.

From these remarks we get the following results.

- If there exists $0 < \rho < 1$ such that $\omega_n = O(\rho^n)$ then $R_0(n) = O(\rho^n)$ and one can take $k_j = p$ provided p is such that $\rho^p < \frac{1}{D(D+1)}$, i.e. $p > \dfrac{-\log[D(D+1)]}{\rho}$. Assume also that $\hat{\nu}(\Delta_n) = O(\alpha^n)$ for some $0 < \alpha < 1$. Then one may choose $v_n = {\alpha'}^{-n}$ provided $0 < \alpha < \alpha' < 1$. We have: $\gamma_j = \max\left(\frac{1}{D}, {\alpha'}^p\right) =: \kappa < 1$ and $u_n = \kappa^{\frac{n}{p}}$.
- If there exists $\alpha > 1$ such that $\omega_n = O(n^{-\alpha})$ then $R_0(n) = O(n^{-\alpha+1})$ and $R_\ell(s+k_{\ell+1}) = O\left(R_0(s+k_{\ell+1}) \cdot D^\ell\right) = O\left(\frac{D^\ell}{k_{\ell+1}^{\alpha-1}}\right)$ so, if $k_{\ell+1} \sim \text{const} D^{\frac{\ell+1}{\alpha-1}}$, it satisfies $R_\ell(s+k_{\ell+1}) \leq \frac{1}{DR_0(s)}$. So $\ell(n) \sim (\alpha-1)\frac{\log n}{\log D} + \text{const}$. Assume also that $\hat{\nu}(\Delta_n) = O(n^{-\beta})$ for some $\beta > 1$. Then one may choose $v_n = n^\gamma$ provided $0 < \gamma < \beta - 1$. We have: $\gamma_j = \max\left(\frac{1}{D}, D^{-\frac{\gamma}{\alpha-1}}\right)$.

If $\beta \le \alpha$ then we can choose $\gamma < \beta - 1 \le \alpha - 1$ and then $\gamma_j = D^{-\frac{\gamma}{\alpha-1}}$ and $u_n = O(n^{-\gamma})$.
If $\alpha < \beta$ then we can choose $\beta - 1 > \gamma > \alpha - 1$ and then $\gamma_j = \frac{1}{D}$ and $u_n = O(n^{-\alpha+1})$.
Finally, we get that $u_n = O(n^{-\min[\alpha-1,\beta-1-\varepsilon]})$, for all $\varepsilon > 0$.

- If $\omega_n = O(e^{-n^\alpha})$ and $\hat{\nu}(\Delta_n) = O(e^{-n^\beta})$ for some $0 < \alpha, \beta < 1$, set $k_\ell := \ell^{\frac{1}{\alpha}-1}$. We obtain $q(\ell) = k_1 + \cdots + k_\ell \sim \ell^{\frac{1}{\alpha}}$ so $\ell(n) \sim n^\alpha$. An easy estimate gives that $R_0(m) \le e^{-m^{\alpha-\varepsilon}}$ for $\varepsilon > 0$ and all large m. Now, by eq. (6.1),

$$R_\ell(s + k_{\ell+1}) \le (\ell+1)D^\ell R_0(k_{\ell+1}) \le (\ell+1)D^\ell e^{-\ell^{(\alpha-\varepsilon)(\frac{1}{\alpha}-1)}} \to 0 \text{ as } \ell \to \infty.$$

Hence, $R_\ell(s + k_{\ell+1}) \le R_0(s)/D$ is satisfied and the choice of k_ℓ is correct for all large ℓ. Let us compute:

$$\begin{aligned} \gamma_\ell &= \max\left\{\frac{1}{D}, \frac{q(\ell+1)^2}{q(\ell)^2} \exp\left(\ell^{\frac{\beta}{\alpha}} - (\ell+1)^{\frac{\beta}{\alpha}}\right)\right\} \\ &= \max\left\{\frac{1}{D}, \left(1 + \frac{1}{\ell}\right)^{\frac{2}{\alpha}} \exp\left(-\ell^{\frac{\beta}{\alpha}} \cdot \frac{\beta}{\alpha}\frac{1}{\ell} + \dots\right)\right\}. \end{aligned}$$

If $\beta > \alpha$, then the second term of the above maximum goes to zero and therefore $\gamma_\ell = \frac{1}{D}$ for large ℓ. We compute the contraction coefficient at time n:

$$u_n = D^{-\ell(n)} = D^{-Cn^\alpha} = e^{-C'n^\alpha} \le e^{-n^{\alpha-\varepsilon}}.$$

If $\beta < \alpha$, then the second term of the above maximum goes to one and therefore sets the value of γ_ℓ for large ℓ. We compute:

$$u_n = \prod_{j=1}^{\ell(n)} e^{-j^{\frac{\beta}{\alpha}}} \le e^{-\ell(n)^{\frac{\beta}{\alpha}}} = e^{-Cn^\beta} \le e^{-n^{\beta-\varepsilon}}.$$

If $\alpha = \beta$, then we change, for instance, α to $\alpha' > \beta$, arbitrarily close, and apply the previous case.

References

[ADU] J. Aaronson, M. Denker & M. Urbanski *Ergodic theory for Markov fibered systems and parabolic rational maps.* Trans. Amer. Math. Soc. (1993), **337** (2), 495-548.

[Ba,Y] V.Baladi, L.-S. Young *On the spectra of randomly perturbed expanding maps.* Comm. Math. Phys. (1993) **156**, 2, 355-385.

[Bi] G. Birkhoff *Extensions of Jentzch's theorem.* Trans. Amer. Math. Soc. (1957), **85**, 219-227.

[BFG] X. Bressaud, R. Fernandez & A. Galves *Speed of decay of correlations for non Holderian dynamics. A coupling approach* Electronic Journal of Probability 4 (1999), **3**.

[Bu1] J. Buzzi *Markov extensions for multi-dimensional dynamical systems.* Israel J. Math. **112** (1999), 357-380.

[Bu2] J. Buzzi *Absolutely continuous invariant measures for generic piecewise affine and expanding maps* Int. J. Chaos & Bif. (1999), **9** (9), 1743-1750.

[Bu3] J. Buzzi *Thermodynamical formalism for piecewise invertible dynamical systems*, Proc. Sympos. Pure Math. **68**, 749–783, Amer. Math. Soc., 2001.

[BM] J. Buzzi, V. Maume-Deschamps *Decay of correlations for piecewise invertible maps in higher dimensions.* Israel J. Math. (2002) **131**, 203–220.

[BPS] J. Buzzi, F. Paccaut, B. Schmitt *Conformal measures for multidimensional expanding maps.* Ergod. th. dynam. syst. (2001) **21**, 1035–1049.

[Ch] N. Chernov *Statistical properties of piecewise smooth hyperbolic systems in high dimensions.* Discrete Contin. Dynam. Systems (1999), **5**, 2, 425-448.

[C] P. Collet *Statistics of closest return for some non uniformly hyperbolic systems.* Ergodic Theory Dynam. Systems, (2001), **21**, no. 2, 401-420.

[CGS] P. Collet, A. Galves, B. Schmitt *Unpredictability of the occurrence time of a long laminar period in a model of temporal intermitency.* Ann. Inst. H. Poincaré Phys. Théor. (1992), **57**, 3, 319-331.

[D] D. Dolgopyat *On dynamics of mostly contracting diffeomorphisms.* Comm. Math. Phys. (2000), **213**, 1, 181-201.

[FS] P. Ferrero, B. Schmitt *Ruelle Perron Frobenius theorems and projective metrics.* Colloque Math. Soc. J. Bolyai Random Fields. Estergom (Hungary) (1979).

[G,S] A. Galves & B. Schmitt *Inequalities for hitting time in mixing dynamical systems.* Random and Computational Dynamics (1997), **5**, 4, 337-347.

[Ho] F. Hofbauer *Piecewise invertible dynamical systems.* Probab. Theory Relat. Fields **72** (1986), no. 3, 359–386.

[KN] G. Keller, T. Nowicki *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps.* Comm. Math. Phys. (1992), **149**, 1, 31-69.

[KMS] A. Kondah, V. Maume & B. Schmitt *Vitesse de convergence vers l'état d'équilibre pour des dynamiques markoviennes non höldériennes.* Ann. Inst. Poincaré Sec. Prob. Stat. (1997) **33** (6) 675-695.

[Li1] C. Liverani *Decay of correlations.* Ann. of Math. (1995), **142** (2), 239-301

[Li2] C. Liverani *Central limit theorem for deterministic systems.* Proceedings of the International Congress on Dynamical Systems, Montevideo 95, Research Notes in Mathematics series, Pittman, (1997).

[LSV] C. Liverani, B. Saussol, S. Vaienti *Conformal measure and decay of correlations for covering weighted systems* Ergodic th. & dynam. syst. **18**, 1998, 1399–1420.

[M] V. Maume-Deschamps *Projective metrics and mixing properties on towers.* To appear in Trans. A.M.S.

[Pa] F. Paccaut *Statistics of return times for weighted maps of the interval* Ann. Inst. H. Poincaré Probab. Statist. **36**, 2000, no. 3, 339–366. .

[Po] M. Pollicott *Rates of mixing for potentials of summable variation* Trans. Amer. Math. Soc. **352**, 2000, no. 2, 843–853.

[Sz] L. Schwartz *Topologie générale et analyse fonctionnelle.* Hermann (1970).

[Y1] L.-S. Young *Recurrence times and rates of mixing* Israel J. Math **110**, (1999), 153–188.

e-mail: `buzzi@math.polytechnique.fr`
Centre de Mathématiques de l'Ecole Polytechnique
U.M.R. 7640 du C.N.R.S.
Ecole Polytechnique
91128 Palaiseau Cedex / FRANCE
e-mail: `vmaume@topolog.u-bourgogne.fr`
Laboratoire de Topologie
Université de Bourgogne
9,Avenue Alain Savary - B.P. 47870
21078 Dijon Cedex FRANCE