Devroye inequality for a class of non-uniformly hyperbolic dynamical systems

2005 Nonlinearity 18 2323

(http://iopscience.iop.org/0951-7715/18/5/023)

INSTITUTE OF PHYSICS PUBLISHING NONLINEARITY

Nonlinearity **18** (2005) 2323–2340 doi:10.1088/0951-7715/18/5/023

Devroye inequality for a class of non-uniformly hyperbolic dynamical systems

J-R Chazottes[1], **P Collet**[1] **and B Schmitt**[2]

[1] Centre de Physique Théorique, Ecole Polytechnique, CNRS UMR 7644, F-91128 Palaiseau Cedex, France
[2] Département de Mathématiques, Université de Bourgogne, Faculté des Sciences Mirande, BP 138, 21004 Dijon Cedex, France

E-mail: jeanrene@cpht.polytechnique.fr, collet@cpht.polytechnique.fr and schmittb@u-bourgogne.fr

Received 8 December 2004, in final form 6 June 2005
Published 15 July 2005
Online at stacks.iop.org/Non/18/2323

Recommended by L Bunimovich

Abstract
In this paper we prove an inequality which we call the 'Devroye inequality' for a large class of non-uniformly hyperbolic dynamical systems (M, f). This class, introduced by Young, includes families of piecewise hyperbolic maps (Lozi-like maps), scattering billiards (e.g. planar Lorentz gas), unimodal and Hénon-like maps. The Devroye inequality provides an upper bound for the variance of observables of the form $K(x, f(x), \ldots, f^{n-1}(x))$, where K is *any* separately Hölder continuous function of n variables. In particular, we can deal with observables which are not Birkhoff averages. We will show in Chazottes *et al* (2005 *Nonlinearity* **18** 2341–64) some applications of Devroye inequality to statistical properties of this class of dynamical systems.

Mathematics Subject Classification: 37D25, 37A50, 60E15

1. Introduction

This paper deals with variance estimates for a class of non-uniformly hyperbolic dynamical systems. This class was introduced by Young in an abstract way. It is strictly larger than axiom A since it encompasses families of piecewise hyperbolic maps like the Lozi maps; scattering billiards, like the planar periodic Lorentz gas; certain quadratic and Hénon maps. In this setting, she was able to prove the existence of Sinai–Ruelle–Bowen measures, exponential decay of correlations and the central limit theorem for Hölder continuous observables.

Very informally speaking, the strategy successfully carried out by Young for the above systems is to construct a new dynamical system over a horseshoe-like subset of the original system by using 'Markovian' return-times so as to obtain a 'tower Markov map'. Then one reduces this Markov extension to an 'expanding map' by quotienting out stable manifolds. On this reduced system, it is possible to define the transfer operator acting on a suitable function space giving back Hölder continuous observables in the original dynamical system. The crucial 'parameter' of this construction is the tail of Markovian return-times with respect to Lebesgue measure, see [7] for an informal description of this construction. For the above examples, this tail is exponentially small. In [5] the existence of a spectral gap is proved for the transfer operator for the quotiented tower map. From this follows an exponential decay of correlations for Hölder continuous observables in the original system.

In this paper we prove an inequality which we call the 'Devroye inequality'. In the context of iid random variables assuming values in a finite set, this inequality was first obtained by Devroye in [4]. This inequality provides an upper estimate for the variance of any Hölder continuous observable computed along orbit segments of length n in terms of the sum of the square of its Hölder constants. The two crucial features of this inequality are that it is valid for *any* n and for any separately Hölder continuous observable. In particular it applies to observables which are not necessarily time-averages of observables. We will show in [2] how to apply the Devroye inequality to obtain statistical properties for this class of dynamical systems.

In the setting of piecewise expanding maps of the interval a much stronger inequality holds, namely, an exponential inequality [3]. It immediately implies the Devroye inequality for Lipschitz observables. Our strategy to prove this inequality in the present setting shares the same global strategy as in [3], that is to exploit the spectral properties of the transfer operator, in particular its spectral gap. However, many crucial points have to be handled differently. In particular, some complications obviously arise due to the fact that we have to succeed in transferring information from the quotiented tower map back to the original system. In particular, we have to control the approximations we make to transform original observables into observables in the quotiented tower map.

Two open issues naturally appear after this work. The first one concerns the validity of the exponential inequality, proved in [3] for expanding maps of the interval, in the present setting. The second one is about dynamical systems with tails of Markovian return-times which are sub-exponential, in particular polynomial, as in [6]. Basic examples of such systems are maps of the interval with indifferent fixed points. We are not able, at present, to prove the Devroye inequality in the setting of [6]. For such systems, there is no spectral gap for the transfer operator and completely different techniques seem to be needed.

Outline of the paper. In section 2 we present, in a short self-contained way, the class of dynamical systems introduced by Young. In section 3 we state our main result, i.e. the Devroye inequality for the variance of separately Hölder continuous observables. Section 4 is devoted to a brief description of the tower Markov map and its quotiented version. In particular, we recall the spectral theory of the transfer operator. In section 5 we prove our main result.

2. A class of non-uniformly hyperbolic systems

In this section, we recall the essential features of the abstract class of dynamical systems in [5] that is to have a reasonably self-contained presentation and to fix the notation. For the complete set of assumptions and further details, we refer to [5].

Let M be a finite-dimensional, regular and compact Riemann manifold (endowed with a distance $d(\cdot,\cdot)$) and let f be a $C^{1+\epsilon}$ diffeomorphism ($\epsilon > 0$). We denote by m the Lebesgue measure on M.

Hyperbolic product structure. We assume that there is a set $\Lambda \subset M$ with a hyperbolic product structure in the following sense. For some $n \geqslant 1$, there exists a continuous family of d-dimensional unstable discs $\Gamma^u = \{\gamma^u\}$ and a continuous family of $(\dim M - d)$-dimensional stable discs $\Gamma^s = \{\gamma^s\}$ with

$$\Lambda = (\cup \gamma^u) \cap (\cup \gamma^s).$$

Recall that an unstable disc γ^u is defined by the property that for each $x, x' \in \gamma^u$,

$$\limsup_{n\to\infty} \frac{1}{n} \log d(f^{-n}(x), f^{-n}(x')) < 0$$

while a stable disc γ^s is defined via the same condition with forward iterations of f instead of backward ones.

For $x \in \Lambda$, writing $\gamma^u(x)$ for the element of Γ^u containing x, we assume that *each γ^u-disc meets each γ^s-disc at exactly one point*, and that the intersection is transversal with the angles bounded away from zero.

We assume that the Lebesgue measure m is compatible with the hyperbolic structure in the sense that for every $\gamma \in \Gamma^u$ we have $m_\gamma(\{\gamma \cap \Lambda\}) > 0$, where m_γ is the measure induced by m on γ.

Markovian return-times. We assume there are finitely many or countably many pairwise disjoint subsets $\Lambda_1, \Lambda_2, \ldots \subset \Lambda$, with a hyperbolic product structure and integers $R_i \geqslant R_0 > 1$ with the properties that:

1. $\cup_i \Lambda_i = \Lambda$, modulo zero Lebesgue sets in the unstable direction. The 'return-time map' $R : \cup_i \Lambda_i \to \mathbb{Z}^+$ is defined by $R|_{\Lambda_i} = R_i$ (with a slight abuse, R can be viewed as a Lebesgue almost everywhere defined function on Λ);
2. for each $x \in \Lambda_i$, we have $f^{R_i}(\gamma^s(x)) \subset \gamma^s(f^{R_i}(x))$ and $f^{R_i}(\gamma^u(x)) \supset \gamma^u(f^{R_i}(x))$;
3. for each n there are at most finitely many i with $R_i = n$.

These return times are used to construct the 'tower map' which is the Markov extension of $(\bigcup_{j=0}^{\infty} f^j(\Lambda), f)$ (see below).

From now on we assume that Markovian return-times have an exponential tail. This means that we assume there are $C > 0$ and $\theta < 1$ so that for some $\gamma \in \Gamma^u$

$$m_\gamma(\{x \in \Lambda \mid R \geqslant n\}) \leqslant C\theta^n. \tag{1}$$

Next, we recall two assumptions that we shall explicitly use in the following.

Uniform contraction along γ^s-discs. There exist $C > 0$ and $0 < \alpha < 1$, such that for all $x \in \Lambda$, for each $x' \in \gamma^s(x)$, and all $n \in \mathbb{Z}^+$ we have

$$d(f^n(x), f^n(x')) \leqslant C\alpha^n. \tag{2}$$

The notion of *separation time* plays a central role. Let $s_0 : \Lambda \times \Lambda \to \mathbb{Z}^+ \cup \{\infty\}$ be such that

1. $s_0(x, x') = s_0(\tilde{x}, \tilde{x}')$ whenever $\tilde{x} \in \gamma^s(x)$ and $\tilde{x}' \in \gamma^s(x')$;
2. for each $n \in \mathbb{Z}^+$, the maximum number of orbits starting from Λ that are pairwise separated before time n is finite (where we say that x and x' are separated before time k if $s_0(x, x') < k$). This is related to condition 3 above;
3. for $x, x' \in \Lambda_i$ we have $s_0(x, x') \geqslant R_i + s_0(f^{R_i}(x), f^{R_i}(x'))$;
4. for $x \in \Lambda_i$, $x' \in \Lambda_j$ with $i \neq j$ but $R_i = R_j$ we have

$$s_0(x, x') < R_i - 1.$$

Backward contraction and distortion along γ^u-discs. The separation time s_0 on $\Lambda \times \Lambda$ is such that for all $x \in \Lambda$, each $x' \in \gamma^u(x)$ and all $0 \leqslant k \leqslant n < s_0(x, x')$:

1. $d(f^n(x), f^n(x')) \leqslant C\alpha^{s_0(x,x')-n}$;
2. $\log \prod_{i=k}^{n} \dfrac{\det Df^u(f^i(x))}{\det Df^u(f^i(x'))} \leqslant C\alpha^{s_0(x,x')-n}$.

We denote by f^u the restriction of f to the γ^u-discs.

Sinai–Ruelle–Bowen measure. It is proved in [5] that f admits a Sinai–Ruelle–Bowen measure supported on $\bigcup_{j=0}^{\infty} f^j(\Lambda)$, which we will denote by μ_M in the sequel.

3. Devroye inequality

A real-valued function of n variables $K : M^n \to \mathbb{R}$ is called *separately η-Hölder continuous* if the following Hölder constants $L_j = L_j(K)$, $1 \leqslant j \leqslant n$, are all finite

$$L_j := \sup_{x_1,x_2,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n} \sup_{\tilde{x}_j \neq x_j} \frac{|K(x_1,\ldots,x_{j-1},x_j,x_{j+1},\ldots,x_n) - K(x_1,\ldots,x_{j-1},\tilde{x}_j,x_{j+1},\ldots,x_n)|}{d(x_j,\tilde{x}_j)^\eta}. \tag{3}$$

It is convenient to define $L_j = 0$ for $j > n$ and $L_0 = 0$.

We can now formulate the main theorem of this paper. It provides, for any $n \geqslant 1$, an estimate on the variance of observables of the form $K(x, f(x), \ldots, f^{n-1}(x))$ where K is any separately Hölder continuous function.

Theorem 3.1 (Devroye inequality for the variance). *Let (M, f, μ_M) be the dynamical system defined above. Then, for any $0 < \eta \leqslant 1$, there exists a constant $D = D(\eta) > 0$ such that for any $n \geqslant 1$, for any separately η-Hölder continuous function K of n variables, we have*

$$\int \left(K(x, f(x), \ldots, f^{n-1}(x)) - \int K(y, f(y), \ldots, f^{n-1}(y))\, \mathrm{d}\mu_M(y) \right)^2 \mathrm{d}\mu_M(x) \leqslant D \sum_{j=1}^{n} L_j^2. \tag{4}$$

Examples of dynamical systems that fit the class of dynamical systems defined above include axiom A attractors; piecewise hyperbolic maps (Lozi-like mappings); billiards with convex scatterers (including planar periodic Lorentz gases); quadratic maps and Hénon-type attractors (for parameter sets with positive Lebesgue measure). We refer the reader to [1, 5, 7] for details.

4. Preparatory notions and results

To prove the Devroye inequality, we need to use the spectral gap for the transfer operator proved by Young. We use almost the same notations as in [5].

4.1. The tower map (F, Δ)

Let $F : \Delta \circlearrowleft$ be the 'tower map' as in [5]. More precisely, we have

$$\Delta := \{z := (x, q) : x \in \Lambda;\ q = 0, 1, \ldots, R(x) - 1\}$$

and

$$F(z) = F(x, q) := \begin{cases} (x, q+1) & \text{if } q+1 < R(x), \\ (f^R x, 0) & \text{if } q+1 = R(x). \end{cases}$$

There is a projection map $\pi : \Delta \to \bigcup_{j=0}^{\infty} f^j(\Lambda)$ such that $f \circ \pi = \pi \circ F$. There is an F-invariant measure μ_Δ related to μ_M via the equation $\mu_M := \mu_\Delta \circ \pi^{-1}$.

Markov partition for F. We denote by $\mathcal{M} = \{\Delta_{q,j}\}$ the *Markov partition for F* built explicitly in [5]. It is worth thinking of Δ as a disjoint union of sets Δ_q consisting of those pairs $(x, q) \in \Delta$ the second coordinate of which is q. We can picture Δ as a tower and refer to Δ_q as the qth level of the tower. In particular, Δ_q is a copy of $\{x \in \Lambda : R(x) > q\}$.

One needs to slightly modify the definition of the separation time $s_0(\cdot, \cdot)$ defined above to make it compatible with the Markov partition. Define, as in [5], for all pairs z, z' belonging to the same $\Delta_{q,j}$, the number

$$\begin{aligned} &s(z, z') := \text{the largest } n \geqslant 0 \text{ such that for all } i \leqslant n, \\ &F^i(z) \text{ lies in the same element of } \mathcal{M} \text{ as } F^i(z'). \end{aligned} \tag{5}$$

Note that restricted to $\Delta_0 \times \Delta_0$

$$s(\cdot, \cdot) \leqslant s_0(\cdot, \cdot). \tag{6}$$

The following consequence of the above definitions will be used repeatedly in the sequel.

Lemma 4.1. *There exists a constant $C > 0$ such that for any $y, y' \in \Delta$ such that there exists an integer q and two points $\tilde{y}, \tilde{y}' \in \Delta$ satisfying $s(\tilde{y}, \tilde{y}') \geqslant q$, $F^q(\tilde{y}) = y$, and $F^q(\tilde{y}') = y'$, then*

$$d(\pi(y), \pi(y')) \leqslant C\alpha^{\min(q, s(y,y'))}. \tag{7}$$

Proof. Without loss of generality, we can assume that $\tilde{y}, \tilde{y}' \in \Delta_0$ and $s(y, y') > 0$. Therefore, there exists an integer m such that $\tilde{y}, \tilde{y}' \in \Delta_{0,m}$. Let $Z := \gamma^s(\pi(\tilde{y})) \cap \gamma^u(\pi(\tilde{y}'))$. Notice that by assumption this intersection is not empty, and consists of exactly one point belonging to some Λ_p since this set has a hyperbolic product structure. Let z be the unique point in $\Delta_{0,m}$ such that $\pi(z) = Z$. Since Z is on the local stable manifold of $\pi(\tilde{y})$, it follows from the Markov property that for all $j \geqslant 0$, $F^j(z)$ and $F^j(\tilde{y})$ belong to the same atom of $\mathcal{M}$. This immediately implies that

$$s_0(Z, \pi(\tilde{y}')) \geqslant s(z, \tilde{y}') \geqslant q + s(y, y').$$

We now apply the 'backward contraction along γ^u-discs' for $Z \in \gamma^u(\pi(\tilde{y}'))$ and $n = q$. Using also the previous inequality we obtain

$$d(f^q(Z), \pi(y')) \leqslant C\alpha^{s(y,y')}.$$

On the other hand, from the 'uniform contraction along γ^s-discs', we have

$$d(f^q(Z), \pi(y)) \leqslant C\alpha^q.$$

The result follows from the triangle inequality. ■

4.2. *The quotiented tower map* $(\overline{F}, \overline{\Delta})$ *and the transfer operator*

Let $\overline{F} : \overline{\Delta} \circlearrowleft$ be the (non-invertible) expanding map obtained by quotienting out the γ^s-leaves from Δ. The projection will be denoted by $\overline{\pi} : \Delta \to \overline{\Delta}$, and we shall use the notations $\{\overline{\Delta}_q\}$, $\{\overline{\Delta}_{q,j}\}$, etc with the obvious meanings. Notice that $\overline{\mathcal{M}} = \{\overline{\Delta}_{q,j}\}$ is a Markov partition for $\overline{F}$.

Let $\overline{m}$ be the reference measure on $\overline{\Delta}$ constructed in [5]. On each $\gamma \in \Gamma^u$, $\overline{m}_\gamma$ is absolutely continuous with respect to m_γ.

Before introducing the suitable Banach space on which will act the transfer operator, we recall the following facts established in [5].

Invariant measure for $\overline{F}$. The map $\overline{F} : \overline{\Delta} \circlearrowleft$ has an invariant probability measure $\mu_{\overline{\Delta}}$ of the form $\mathrm{d}\mu_{\overline{\Delta}} = \varphi \, \mathrm{d}\overline{m}$, where φ satisfies

$$c_0^{-1} \leqslant \varphi \leqslant c_0 \qquad \text{for some } c_0 > 0 \tag{8}$$

and

$$|\varphi(z) - \varphi(z')| \leqslant C\alpha^{s(z,z')/2} \qquad \forall z, z' \in \overline{\Delta}_{q,j}, \tag{9}$$

where α is defined in (2). This result of course motivates the choice of the function space.

Regularity of the Jacobian. In [5], it is explained how to give a 'differentiable structure' so that one can define the Jacobian $J\overline{F} = |\det D\overline{F}|$. We have the properties

$$J\overline{F} \equiv 1 \qquad \text{on } \overline{\Delta} \backslash \overline{F}^{-1}(\overline{\Delta}_0) \tag{10}$$

and

$$\left| \frac{J\overline{F}(y)}{J\overline{F}(y')} - 1 \right| \leqslant C\alpha^{s(\overline{F}(y), \overline{F}(y'))/2} \qquad \forall y, y' \in \Delta_{q,j} \cap \overline{F}^{-1}\overline{\Delta}_0. \tag{11}$$

Function space. For any σ such that $\sqrt{\alpha} < \sigma < 1$, let $X_\sigma = \{g : \overline{\Delta} \to \mathbb{R}, \|g\| < \infty\}$ where the norm $\|\cdot\|$ is defined as follows. Writing $g_{q,j} = g|\{\overline{\Delta}_{q,j}\}$ and letting $|\cdot|_p$ denote the L^p-norm with respect to the reference measure $\overline{m}$ we set

$$\|g\| := \|g\|_\infty + \|g\|_h,$$

where

$$\|g\|_\infty := \sup_{q,j} \|g_{q,j}\|_\infty, \qquad \|g\|_h := \sup_{q,j} \|g_{q,j}\|_h$$

and $\|g_{q,j}\|_\infty$ and $\|g_{q,j}\|_h$ are defined by

$$\|g_{q,j}\|_\infty := |g_{q,j}|_\infty \mathrm{e}^{-q\varepsilon},$$

where $\varepsilon > 0$ will be chosen adequately small later on and

$$\|g_{q,j}\|_h := \left(\operatorname{ess} \sup_{y,y' \in \overline{\Delta}_{q,j}} \frac{|g(y) - g(y')|}{\sigma^{s(y,y')}} \right) \mathrm{e}^{-q\varepsilon}.$$

It is easy to verify that $(X_\sigma, \|\cdot\|)$ is a Banach space (parametrized by ε).

The transfer operator associated with the dynamical system $\overline{F} : \overline{\Delta} \circlearrowleft$ and reference measure $\overline{m}$ is then defined by

$$\mathcal{P}g(y) = \sum_{y' : \overline{F}(y')=y} \frac{g(y')}{J\overline{F}(y')}.$$

The normalized transfer operator associated with $\mathcal{P}$ is defined as

$$\mathcal{L}g(y) = \frac{1}{\varphi(y)} \sum_{y' : \overline{F}(y')=y} \frac{\varphi(y')}{J\overline{F}(y')} g(y') \tag{12}$$

which satisfies $\mathcal{L}1 = 1$. The following spectral property of $\mathcal{L}$ is easily derived from [5].

Lemma 4.2. *For any $\sigma \in (\sqrt{\alpha}, 1)$, there exist constants $C > 0$ and $\rho \in (0, 1)$ such that, for all $g \in X_\sigma$ and for any integer n, we have*

$$\left\| \mathcal{L}^n g - \int g \, \mathrm{d}\mu_{\overline{\Delta}} \right\| \leqslant C\rho^n \|g\|. \tag{13}$$

5. Proof of Devroye inequality

Preparatory approximations for observables. Let $K : M^n \longrightarrow \mathbb{R}$ be a separately η Hölder function of n variables.

Let us use the short-hand notation

$$\mathbb{E}(K) = \int K(x, f(x), \ldots, f^{n-1}(x)) \, \mathrm{d}\mu_M(x)$$

and

$$\operatorname{var}(K) = \int (K(x, f(x), \ldots, f^{n-1}(x)) - \mathbb{E}(K))^2 \, \mathrm{d}\mu_M(x).$$

A standard computation gives:

$$\operatorname{var}(K) = \frac{1}{2} \int [K(x, \ldots, f^{n-1}(x)) - K(x', \ldots, f^{n-1}(x'))]^2 \, \mathrm{d}\mu_M(x) \, \mathrm{d}\mu_M(x').$$

Since by construction $\mu_M = \mu_\Delta \circ \pi^{-1}$ we also have

$$\operatorname{var}(K) = \frac{1}{2} \int [\tilde{K}(y, \ldots, F^{n-1}(y)) - \tilde{K}(y', \ldots, F^{n-1}(y'))]^2 \, \mathrm{d}\mu_\Delta(y) \, \mathrm{d}\mu_\Delta(y'), \tag{14}$$

where

$$\tilde{K}(y_1, \ldots, y_n) = K(\pi(y_1), \ldots, \pi(y_n)).$$

We now introduce a new piecewise constant observable V on Δ^n. We will write $\mathcal{M}(x)$ for the atom of the partition $\mathcal{M}$ containing x.

For a fixed integer p_0 large enough, we define the integer-valued function $\ell : \mathbb{N} \to \mathbb{N} \cup \{0\}$ by

$$\ell(k) := \begin{cases} k - 1 & \text{if } k \leqslant p_0 \log(1 + k), \\ p_0[\log(1 + k)] & \text{otherwise.} \end{cases} \tag{15}$$

We now define the function $V : \Delta^n \longrightarrow \mathbb{R}$ as follows.

$$V(y_1, \ldots, y_n) := \inf_{x_1,\ldots,x_n} \{\tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n)) : F^k(x_j) \in \mathcal{M}(y_{j-\ell(j)+k}),$$
$$k = 0, 1, \ldots, 2\ell(j),\ j = 1, \ldots, n\}. \tag{16}$$

If the above set is empty, then we set $V(y_1, \ldots, y_n) = 0$.

One can remark immediately that V factorizes through $\overline{\pi}$ in the sense that

$$V(y_1, \ldots, y_n) = U(\overline{\pi}(y_1), \ldots, \overline{\pi}(y_n)),$$

where $U : \overline{\Delta}^n \longrightarrow \mathbb{R}$ is defined as

$$U(z_1, z_2, \ldots, z_n) := \inf_{x_1,\ldots,x_n} \{\tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n)) :$$
$$F^k(x_j) \in \mathcal{M}(\overline{\pi}^{-1}(z_{j-\ell(j)+k})), k = 0, 1, \ldots, 2\ell(j),\ j = 1, \ldots, n\}. \tag{17}$$

If the above set is empty, then we set $U(z_1, \ldots, z_n) = 0$.

We have the following lemma which allows us to replace the observable $\tilde{K}$ by the piecewise constant observable V.

Lemma 5.1. *There is constant $C > 0$ such that for p_0 large enough (see (15)) we have*

$$\sup_{y\in\Delta} |\tilde{K}(y, \ldots, F^{n-1}(y)) - V(y, \ldots, F^{n-1}(y))| \leqslant C \sum_{k=1}^{n} \frac{L_k}{k}.$$

Proof. Given $y \in \Delta$, let $x_1, \ldots, x_j, \ldots, x_n$ be a sequence such that for any $1 \leqslant j \leqslant n$

$$F^k(x_j) \in \mathcal{M}(F^{j-\ell(j)+k-1}(y)) \qquad \text{for } k = 0, 1, \ldots, 2\ell(j). \tag{18}$$

We have the identity

$$\tilde{K}(y, \ldots, F^{n-1}(y)) - \tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n))$$
$$= \sum_{p=-1}^{n-2} (\tilde{K}(y, \ldots, F^{p+1}(y), F^{\ell(p+3)}(x_{p+3}), \ldots, F^{\ell(n)}(x_n))$$
$$- \tilde{K}(y, \ldots, F^{p}(y), F^{\ell(p+2)}(x_{p+2}), \ldots, F^{\ell(n)}(x_n))),$$

where terms with indices out of range are absent. Therefore using (3) yields

$$|\tilde{K}(y, \ldots, F^{n-1}(y)) - \tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n))|$$
$$\leqslant \sum_{q=1}^{n} L_q (d(\pi(F^{\ell(q)}(x_q)), f^{q-1}(\pi(y))))^{\eta}.$$

Using (18) and lemma 4.1 we obtain

$$d(\pi(F^{\ell(q)}(x_q)), f^{q-1}(\pi(y))) \leqslant C\alpha^{\ell(q)}.$$

The lemma follows by choosing p_0 large enough in the definition of $\ell(q)$ in (15). ■

We can now give an approximation of the variance of K in terms of the piecewise constant observable U defined on the quotiented tower $\overline{\Delta}$.

Lemma 5.2. *We have the following approximation*

$$\operatorname{var}(K) \leqslant \int [U(z, \ldots, \overline{F}^{n-1}(z)) - U(z', \ldots, \overline{F}^{n-1}(z'))]^2 \, \mathrm{d}\mu_{\overline{\Delta}}(z) \, \mathrm{d}\mu_{\overline{\Delta}}(z') + \mathcal{O}(1) \sum_{j=1}^{n} L_j^2,$$

where U is defined in (17).

Proof. Using (17) and the fact that $\mu_{\overline{\Delta}} = \mu_\Delta \circ \overline{\pi}^{-1}$ we have

$$\int (V(y, \ldots, F^{n-1}(y)) - V(y', \ldots, F^{n-1}(y')))^2 \, \mathrm{d}\mu_\Delta(y) \, \mathrm{d}\mu_\Delta(y')$$
$$= \int (U(z, \ldots, \overline{F}^{n-1}(z)) - U(z', \ldots, \overline{F}^{n-1}(z')))^2 \, \mathrm{d}\mu_{\overline{\Delta}}(z) \, \mathrm{d}\mu_{\overline{\Delta}}(z').$$

To alleviate notation let us set

$$U_n(z) = U(z, \ldots, \overline{F}^{n-1}(z)),$$
$$V_n(y) = V(y, \ldots, F^{n-1}(y)),$$
$$\tilde{K}_n(y) = \tilde{K}(y, \ldots, F^{n-1}(y)).$$

By (14) we have

$$\begin{aligned}
\operatorname{var}(K) &= \frac{1}{2} \int (\tilde{K}_n(y) - V_n(y) + V_n(y) + \tilde{K}_n(y') - V_n(y') + V_n(y'))^2 \, \mathrm{d}\mu_\Delta(y) \, \mathrm{d}\mu_\Delta(y') \\
&\leqslant \int (V_n(y) - V_n(y'))^2 \, \mathrm{d}\mu_\Delta(y) \, \mathrm{d}\mu_\Delta(y') \\
&\quad + \int ((\tilde{K}_n(y) - V_n(y)) - (\tilde{K}_n(y') - V_n(y')))^2 \, \mathrm{d}\mu_\Delta(y) \, \mathrm{d}\mu_\Delta(y').
\end{aligned}$$

Therefore

$$\operatorname{var}(K) \leqslant \int (U_n(z) - U_n(z'))^2 \, \mathrm{d}\mu_{\overline{\Delta}}(z) \, \mathrm{d}\mu_{\overline{\Delta}}(z') + 4 \int (\tilde{K}_n(y) - V_n(y))^2 \, \mathrm{d}\mu_\Delta(y).$$

We now use lemma 5.1 to estimate $|\tilde{K}_n(y) - V_n(y)|$ and the Cauchy–Schwarz inequality, i.e.

$$\sum_{k=1}^{n} \frac{L_k}{k} \leqslant \mathcal{O}(1) \left(\sum_{k=1}^{n} L_k^2 \right)^{1/2}.$$

This ends the proof of lemma 5.2. ∎

Martingale procedure. As suggested by the previous lemma we will give an upper bound to the integral

$$\int (U_n(z) - U_n(z'))^2 \, \mathrm{d}\mu_{\overline{\Delta}}(z) \, \mathrm{d}\mu_{\overline{\Delta}}(z').$$

To do that we will use the spectral properties of the normalized transfer operator $\mathcal{L}$ associated to $\overline{F}$, which is defined in (12).

We now define an extension of $\mathcal{L}$, also denoted by $\mathcal{L}$. It maps a function $\kappa(x_1, \ldots, x_n)$ of n variables on $\overline{\Delta}$ to a function of $(n-1)$ variables, and is given by

$$\mathcal{L}\kappa(x_1, \ldots, x_{n-1}) = \frac{1}{\varphi(x_1)} \sum_{y : \overline{F}(y) = x_1} \frac{\varphi(y)}{J\overline{F}(y)} \kappa(y, x_1, \ldots, x_{n-1}).$$

It can be immediately verified that if the function of one variable v is given by

$$v(x) = \kappa(x, \overline{F}(x), \ldots, \overline{F}^{n-1}(x)),$$

then $\mathcal{L}v(x) = \mathcal{L}\kappa(x, \overline{F}(x), \ldots, \overline{F}^{n-2}(x))$. Moreover, if κ is a function of n variables and $k < n$ we have

$$\mathcal{L}^k\kappa(x_1, \ldots, x_{n-k}) = \frac{1}{\varphi(x_1)} \sum_{y\,:\,\overline{F}^k(y)=x_1} \frac{\varphi(y)}{J\overline{F}^k(y)} \kappa(y, \overline{F}(y), \ldots, \overline{F}^{k-1}(y), x_1, \ldots, x_{n-k}).$$

For $k \geqslant n$, we can use the same definition noting that a function of n variables is also a function of k variables not depending on the last $(n - k)$ variables.

The extended transfer operator inherits the main properties of the basic one. In particular the probability measure $\mu_{\overline{\Delta}}$ is $\overline{F}$-invariant, i.e.

$$\int \mathcal{L}\kappa(x, \ldots, \overline{F}^{n-2}(x))\, \mathrm{d}\mu_{\overline{\Delta}}(x) = \int \kappa(x, \ldots, \overline{F}^{n-1}(x))\, \mathrm{d}\mu_{\overline{\Delta}}(x). \tag{19}$$

The following lemma (reminiscent of a Martingale-difference argument) will allow us to use lemma 4.2 later on.

Lemma 5.3. *The following identity holds for any* $p \geqslant 0$

$$\begin{aligned}\int (U_n(y) - U_n(y'))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y') = 2\sum_{k=0}^{n-2} \int (\mathcal{L}^k U(y, \ldots, \overline{F}^{n-k-1}(y)) \\ - \mathcal{L}^{k+1} U(\overline{F}(y), \ldots, \overline{F}^{n-k-1}(y)))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y) \\ + 2\sum_{k=0}^{p} \int (\mathcal{L}^k S_n(y) - \mathcal{L}^{k+1} S_n(y))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y) \\ + \int (\mathcal{L}^{p+1} S_n(y) - \mathcal{L}^{p+1} S_n(y'))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y'),\end{aligned} \tag{20}$$

where $S_n(y) = \mathcal{L}^{n-1} U_n(y)$ *is a function which depends only on one variable.*

Proof. We can write

$$\begin{aligned}&\int (U_n(y) - U_n(y'))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y') \\ &\quad = \int \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y')(U_n(y) - \mathcal{L}U_n(\overline{F}(y)) \\ &\qquad + \mathcal{L}U_n(\overline{F}(y)) - U_n(y') + \mathcal{L}U_n(\overline{F}(y')) - \mathcal{L}U_n(\overline{F}(y')))^2 \\ &\quad = 2\int (U_n(y) - \mathcal{L}U_n(\overline{F}(y)))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y) + \int (\mathcal{L}U_n(\overline{F}(y)) \\ &\qquad - \mathcal{L}U_n(\overline{F}(y')))^2\, \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y') - 2\left(\int (U_n(y) - \mathcal{L}U_n(\overline{F}(y)))\, \mathrm{d}\mu_{\overline{\Delta}}(y)\right)^2 \\ &\qquad + 2\int \mathrm{d}\mu_{\overline{\Delta}}(y)\, \mathrm{d}\mu_{\overline{\Delta}}(y')(U_n(y) - \mathcal{L}U_n(\overline{F}(y)) + \mathcal{L}U_n(\overline{F}(y')) - U_n(y')) \\ &\qquad \times (\mathcal{L}U_n(\overline{F}(y)) - \mathcal{L}U_n(\overline{F}(y'))).\end{aligned}$$

The term before last is equal to zero using the $\overline{F}$-invariance of $\mu_{\overline{\Delta}}$ and (19). Similarly the last term vanishes using the $\overline{F}$-invariance of $\mu_{\overline{\Delta}}$ and the identity

$$\int U_n(y)\mathcal{L}U_n(\overline{F}(y))\,\mathrm{d}\mu_{\overline{\Delta}}(y) = \int (\mathcal{L}U_n(y))^2\,\mathrm{d}\mu_{\overline{\Delta}}(y)$$

which follows at once from (19). Lemma 5.3 follows by iterating this inequality. ■

We now need to estimate

$$\mathcal{L}^{k+1}U(\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y)) - \mathcal{L}^k U(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y)).$$

We will use a decomposition of U into a sum of terms.

For $0 \leqslant k \leqslant n-1$ and $0 \leqslant l \leqslant k$, we define the function U_l^k on $\overline{\Delta}^3$ by

$$U_l^k(u,s,y) = \inf_{E_1(u,l)\cap E_2(s,l,k)\cap E_3(y,k,n)} \{\tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2),\ldots,F^{\ell(n)}(x_n))\}, \tag{21}$$

where with the notation $x_1^n := (x_1,\ldots,x_n)$

$$E_1(u,l) := \{x_1^n \mid F^q(x_j) \in \mathcal{M}(\overline{\pi}^{-1}(\overline{F}^{q+j-\ell(j)-1}(u))) \text{ for } 0 \leqslant q \leqslant 2\ell(j) \text{ and } 1 \leqslant j \leqslant l\},$$

$$E_2(s,l,k) := \{x_1^n \mid F^q(x_j) \in \mathcal{M}(\overline{\pi}^{-1}(\overline{F}^{j-l-\ell(j-l)+q-1}(s))) \text{ for } 0 \leqslant q \leqslant 2\ell(j-l) \text{ and } l+1 \leqslant j \leqslant k+1\},$$

$$E_3(y,k,n) := \{x_1^n \mid F^q(x_j) \in \mathcal{M}(\overline{\pi}^{-1}(\overline{F}^{j-k-\ell(j-k-1)+q-2}(y))) \text{ for } 0 \leqslant q \leqslant 2\ell(j-k-1) \text{ and } k+2 \leqslant j \leqslant n\}.$$

It is convenient to set

$$E_1(u,0) = \overline{\Delta}^n, \qquad E_2(s,k+1,k) = \overline{\Delta}^n, \qquad E_3(y,n-1,n) = \overline{\Delta}^n.$$

We define for $0 \leqslant l \leqslant k$

$$v_l^k\big(\xi,y\big) = \int U_l^k(\xi,s,y)\,\mathrm{d}\mu_{\overline{\Delta}}(s). \tag{22}$$

Note that $v_0^k\big(\xi,y\big)$ does not depend on ξ. We have obviously for $k \geqslant 1$

$$\begin{aligned} U(\xi,\overline{F}(\xi),\ldots,\overline{F}^{n-1}(\xi)) &= U(\xi,\overline{F}(\xi),\ldots,\overline{F}^{n-1}(\xi)) - v_k^k(\xi,\overline{F}^{k+1}(\xi)) + v_0^k(\overline{F}^{k+1}(\xi)) \\ &\quad + \sum_{l=0}^{k-1}(v_{l+1}^k(\xi,\overline{F}^{k+1}(\xi)) - v_l^k(\xi,\overline{F}^{k+1}(\xi))). \end{aligned}$$

For $k=0$, the same formula holds without the sum.

By an easy computation one gets

$$\begin{aligned} \mathcal{L}^k U(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y)) &= \mathcal{L}^k(U - v_{k-1}^k)(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y)) \\ &\quad + v_0^k(\overline{F}(y)) + \sum_{l=0}^{k-1}\mathcal{L}_1^{k-l}w_l^k(y,\overline{F}(y)), \end{aligned} \tag{23}$$

where $\mathcal{L}_1$ acts only on the first variable, i.e.

$$\mathcal{L}_1^{k-l}w_l^k(y,y') = \frac{1}{\varphi(y)}\sum_{\overline{F}^{k-l}(z)=y}\frac{\varphi(z)}{J\overline{F}^{k-l}(z)}w_l^k(z,y')$$

and w_l^k is defined by

$$w_l^k(u, y) = \frac{1}{\varphi(u)} \sum_{\overline{F}^l(z)=u} \frac{\varphi(z)}{J\overline{F}^l(z)} (v_{l+1}^k(z, y) - v_l^k(z, y)).$$

Regularity estimates. We now estimate the various terms. We will use several times, the following elementary lemma whose proof is left to the reader.

Lemma 5.4. *Let* Ω_1, Ω_2 *be two sets and* Ψ *a real-valued function on* $\Omega_1 \times \Omega_2$. *Let* Υ_1, Υ_1' *be two subsets of* Ω_1. *Then*

$$\left| \inf_{\omega_1 \in \Upsilon_1, \omega_2 \in \Omega_2} \Psi(\omega_1, \omega_2) - \inf_{\omega_1 \in \Upsilon_1', \omega_2 \in \Omega_2} \Psi(\omega_1, \omega_2) \right| \leqslant \sup_{\omega_1 \in \Upsilon_1, \omega_1' \in \Upsilon_1', \omega_2 \in \Omega_2} |\Psi(\omega_1, \omega_2) - \Psi(\omega_1', \omega_2)|.$$

To apply this lemma we will use the following sequence of sets

$$\mathcal{E}(u, m) := \{x \in \Delta \mid F^q(x) \in \mathcal{M}(\overline{\pi}^{-1}(\overline{F}^{q+m-\ell(m)-1}(u))) \text{ for } 0 \leqslant q \leqslant 2\ell(m)\},$$

where $u \in \overline{\Delta}$, m is an integer. It is useful to observe that

$$E_1(u, l) = \bigtimes_{j=1}^{l} \mathcal{E}(u, j), \qquad E_2(s, l, k) = \bigtimes_{j=l+1}^{k+1} \mathcal{E}(s, j - l)$$

$$E_3(y, k, n) = \bigtimes_{j=k+2}^{n} \mathcal{E}(y, j - k - 1).$$

We denote by $\mathrm{diam}(M)$ the diameter of M.

The first term we have to estimate is

$$\sup_{\xi \in \overline{\Delta}} |U(\xi, \overline{F}(\xi), \ldots, \overline{F}^{n-1}(\xi)) - v_k^k(\xi, \overline{F}^{k+1}(\xi))|$$

$$= \sup_{\xi \in \overline{\Delta}} \left| U(\xi, \overline{F}(\xi), \ldots, \overline{F}^{n-1}(\xi)) - \int U_k^k(\xi, \overline{F}^{\ell(k+1)}(s), \overline{F}^{k+1}(\xi))\, \mathrm{d}\mu_{\overline{\Delta}}(s) \right|,$$

where we have used the invariance of the measure. We know apply lemma 5.4 by taking

$$\Omega_1 = \Delta^{n-k}, \qquad \Omega_2 = \bigtimes_{j=1}^{k} \mathcal{E}(\xi, j)$$

$$\Upsilon_1 = \bigtimes_{p=k+1}^{n} \mathcal{E}(\xi, p), \qquad \Upsilon_1' = \mathcal{E}(s, 1) \times \bigtimes_{j=k+2}^{n} \mathcal{E}(\overline{F}^{k+1}(\xi), j - k - 1)$$

and

$$\Psi(\omega_1, \omega_2) = \tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n)),$$

where

$$\omega_1 = (x_{k+1}, \ldots, x_n), \qquad \omega_2 = (x_1, \ldots, x_k).$$

We have

$$|\Psi(\omega_1, \omega_2) - \Psi(\omega_1', \omega_2)|$$
$$\leqslant \sum_{p=k}^{n-1} |\tilde{K}(F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x_p), F^{\ell(p+1)}(x_{p+1}'), F^{\ell(p+2)}(x_{p+2}'), \ldots, F^{\ell(n)}(x_n'))$$
$$-\tilde{K}(F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x_p), F^{\ell(p+1)}(x_{p+1}), F^{\ell(p+2)}(x_{p+2}'), \ldots, F^{\ell(n)}(x_n'))|.$$

For $p = k$, we get the upper bound $L_{k+1}(\mathrm{diam}\,(M))^\eta$ by using (3). For $p \geqslant k + 1$, we apply lemma 4.1 with $y = F^{\ell(p+1)}(x_{p+1})$, $y' = F^{\ell(p+1)}(x_{p+1}')$, $q = \ell(p - k)$,

$\tilde{y} = F^{\ell(p+1)-\ell(p-k)}(x_{p+1})$, $\tilde{y}' = F^{\ell(p+1)-\ell(p-k)}(x'_{p+1})$, observing that $s(y, y') \geqslant \ell(p-k)$ (this follows from the definition of the sets $\mathcal{E}$). We finally obtain

$$|\Psi(\omega_1, \omega_2) - \Psi(\omega'_1, \omega_2)| \leqslant L_{k+1}(\mathrm{diam}(M))^\eta + C^\eta \sum_{p=k+1}^{\infty} L_{p+1}\alpha^{\eta\ell(p-k)}.$$

Hence

$$\sup_{\xi\in\overline{\Delta}} |U(\xi, \overline{F}(\xi), \ldots, \overline{F}^{n-1}(\xi)) - v_k^k(\xi, \overline{F}^{k+1}(\xi))| \leqslant B_{k+1}, \tag{24}$$

where for any $q \in \mathbb{N}$

$$B_q := L_q(\mathrm{diam}(M))^\eta + C^\eta \sum_{p=q}^{\infty} L_{p+1}\alpha^{\eta\ell(p-q+1)}. \tag{25}$$

To estimate $v_l^k - v_{l-1}^k$, we use, as in [3] the invariance of the measure $\mu_{\overline{\Delta}}$ to write

$$v_l^k(\xi, y) - v_{l-1}^k(\xi, y) = \int (U_l^k(\xi, \overline{F}^{\ell(l+1)}(\overline{F}(s)), y) - U_{l-1}^k(\xi, \overline{F}^{\ell(l)}(s), y))\, \mathrm{d}\mu_{\overline{\Delta}}(s).$$

To estimate the integrand, we apply lemma 5.4 by taking

$$\Omega_1 = \Delta^{n-k+l-2}, \qquad \Omega_2 = \bigtimes_{j=1}^{l-1}\mathcal{E}(\xi, j) \times \bigtimes_{j=k+2}^{n}\mathcal{E}(y, j-k-1)$$
$$\Upsilon_1 = \mathcal{E}(\xi, l) \times \bigtimes_{j=l+1}^{k+1}\mathcal{E}(\overline{F}(s), j-l), \qquad \Upsilon'_1 = \bigtimes_{j=l}^{k+1}\mathcal{E}(s, j-l+1)$$

and

$$\Psi(\omega_1, \omega_2) = \tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n)),$$

where

$$\omega_1 = (x_l, \ldots, x_{k+1}), \qquad \omega_2 = (x_1, \ldots, x_{l-1}, x_{k+2}, \ldots, x_n).$$

We have

$$\begin{aligned}
|\Psi(\omega_1, \omega_2) - \Psi(\omega'_1, \omega_2)| \leqslant \sum_{p=l}^{k+1} |&\tilde{K}(F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x'_p), F^{\ell(p+1)}(x'_{p+1}), \ldots,\\
&F^{\ell(k+1)}(x'_{k+1}), F^{\ell(k+2)}(x_{k+2}), \ldots, F^{\ell(n)}(x_n))\\
&-\tilde{K}(F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x_p), F^{\ell(p+1)}(x'_{p+1}), \ldots,\\
&F^{\ell(k+1)}(x'_{k+1}), F^{\ell(k+2)}(x_{k+2}), \ldots, F^{\ell(n)}(x_n))|.
\end{aligned}$$

For $p = l$, we get the upper bound $L_l(\mathrm{diam}\,(M))^\eta$ by using (3). For $p \geqslant l+1$, we apply lemma 4.1 with the two points $F^{\ell(p)}(x_p)$, $F^{\ell(p)}(x'_p)$, $q = \ell(p-l)$, $\tilde{y} = F^{\ell(p)-\ell(p-l)}(x_p)$, $\tilde{y}' = F^{\ell(p)-\ell(p-l)}(x'_p)$, observing that $s(F^{\ell(p)}(x_p), F^{\ell(p)}(x'_p)) \geqslant \ell(p-l)$. We finally obtain

$$|\Psi(\omega_1, \omega_2) - \Psi(\omega'_1, \omega_2)| \leqslant B_l,$$

where B_l is defined in (25).

It follows that for any $1 \leqslant l \leqslant k$

$$\sup_{\xi, y} |v_l^k(\xi, y) - v_{l-1}^k(\xi, y)| \leqslant B_l. \tag{26}$$

This immediately implies for any $0 \leqslant l \leqslant k-1$

$$\left|w_l^k\right|_\infty \leqslant B_{l+1}. \tag{27}$$

We now have to estimate the regularity of w_l^k with respect to its first variable. This is the content of the following lemma.

Lemma 5.5. *There is a constant $C > 0$ such that for any z, z' and y in $\overline{\Delta}$, for any integers k, l with $0 \leqslant l \leqslant k-1$ and for any separately η-Hölder continuous observable K,*

$$|w_{l+1}^k(z, \overline{F}(y)) - w_l^k(z', \overline{F}(y))| \leqslant C\alpha^{\eta s(z,z')/2}\left(B_l + (\mathrm{diam}(M))^\eta \sum_{j=0}^{l-1} \alpha^{\eta(l-j)/2} L_j\right).$$

Proof. It is convenient to distinguish two cases. The first case corresponds to $s(z, z') = 0$. We then use the estimate (27) and the result follows. We now consider the case $s(z, z') > 0$. Using the Markov property of the map $\overline{F}$ on $\overline{\Delta}$, we can write in this case

$$w_l^k(z, y) - w_l^k(z', y) = \sum_{\zeta \in \bigvee_{j=1}^{l} \overline{F}^{-j}\overline{\mathcal{M}}} 1_{\zeta \cap \overline{F}^{-l}(z)}(\xi) 1_{\zeta \cap \overline{F}^{-l}(z')}(\xi')$$
$$\times \left(\frac{\varphi(\xi)}{\varphi(z) J\overline{F}^l(\xi)} (v_l^k(\xi, y) - v_{l-1}^k(\xi, y)) - \frac{\varphi(\xi')}{\varphi(z') J\overline{F}^l(\xi')} (v_l^k(\xi', y) - v_{l-1}^k(\xi', y)) \right). \tag{28}$$

We first observe that using properties (9) and (11), and the fact that $s(\xi, \xi') \geqslant s(z, z')$, we get for some uniform constant $C_1 > 0$

$$\left| \frac{\varphi(\xi)}{\varphi(z) J\overline{F}^l(\xi)} - \frac{\varphi(\xi')}{\varphi(z') J\overline{F}^l(\xi')} \right| \leqslant C_1 \alpha^{s(\xi,\xi')/2} \frac{\varphi(\xi)}{\varphi(z) J\overline{F}^l(\xi)}.$$

Therefore, using the estimate (26) we obtain

$$\left| \left(\frac{\varphi(\xi)}{\varphi(z) J\overline{F}^l(\xi)} - \frac{\varphi(\xi')}{\varphi(z') J\overline{F}^l(\xi')} \right) (v_l^k(\xi, y) - v_{l-1}^k(\xi, y)) \right|$$
$$\leqslant (1 + C_1)\alpha^{s(z,z')/2} B_{l+1} \frac{\varphi(\xi)}{\varphi(z) J\overline{F}^l(\xi)}. \tag{29}$$

It remains to estimate

$$\frac{\varphi(\xi')}{\varphi(z') J\overline{F}^l(\xi')} [(v_l^k(\xi, y) - v_{l-1}^k(\xi, y)) - (v_l^k(\xi', y) - v_{l-1}^k(\xi', y))]$$

for ξ and ξ' in the same atom of $\bigvee_{j=0}^{l+s(z,z')-1} \overline{F}^{-j}\overline{\mathcal{M}}$. Coming back to the definition (22) of v_l^k, we get

$$v_l^k(\xi, y) - v_l^k(\xi', y) = \int (U_l^k(\xi, s, y) - U_l^k(\xi', s, y))\, \mathrm{d}\mu_{\overline{\Delta}}(s).$$

We are going to prove that if ξ and ξ' belong to the same atom of $\bigvee_{j=0}^{l-1} \overline{F}^{-j}\overline{\mathcal{M}}$, with $\overline{F}^l(\xi) = z$ and $\overline{F}^l(\xi') = z'$, we have

$$\sup_{s,y} |U_l^k(\xi, s, y) - U_l^k(\xi', s, y)| \leqslant C(\mathrm{diam}(M))^\eta \alpha^{\eta s(z,z')/2} \sum_{j=1}^{l} \alpha^{\eta(l-j-1)/2} L_j, \tag{30}$$

where $C > 0$ is a uniform constant.

First observe that if $s(z, z') \geqslant \ell(l)$ then it follows immediately from definition (21) that $U_l^k(\xi, s, y) = U_l^k(\xi', s, y)$. Hence the estimate is true in this case.

Now assume that $s(z, z') < \ell(l)$. Let $p_* = p_*(l)$ be the largest integer such that $p_* + \ell(p_*) < l$.

Observe that for any $1 \leqslant j \leqslant l - \ell(l)$ we have $\mathcal{E}(\xi, j) = \mathcal{E}(\xi', j)$, since by assumption, ξ and ξ' belong to the same atom of $\bigvee_{j=0}^{l-1} \overline{F}^{-j}\overline{\mathcal{M}}$.

We now apply lemma 5.4 by taking

$$\Omega_1 = \Delta^{l-p_*}, \qquad \Omega_2 = \bigtimes_{j=1}^{p_*}\mathcal{E}(\xi, j) \times \bigtimes_{j=l+1}^{k+1}\mathcal{E}(s, j) \times \bigtimes_{j=k+2}^{n}\mathcal{E}(y, j-k-1)$$

$$\Upsilon_1 = \bigtimes_{j=p_*+1}^{l+1}\mathcal{E}(\xi, j), \qquad \Upsilon_1' = \bigtimes_{j=p_*+1}^{l+1}\mathcal{E}(\xi', j)$$

and

$$\Psi(\omega_1, \omega_2) = \tilde{K}(F^{\ell(1)}(x_1), F^{\ell(2)}(x_2), \ldots, F^{\ell(n)}(x_n)),$$

where

$$\omega_1 = (x_{p_*+1}, \ldots, x_l), \qquad \omega_2 = (x_1, \ldots, x_{p_*}, x_{l+1}, \ldots, x_n).$$

We have

$$\begin{aligned}|\Psi(\omega_1, \omega_2) - \Psi(\omega_1', \omega_2)| \leqslant \sum_{p=p_*+1}^{l} |\tilde{K}(&F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x_p'),\\ &F^{\ell(p+1)}(x_{p+1}'), \ldots, F^{\ell(k+1)}(x_{k+1}'), F^{\ell(k+2)}(x_{k+2}), \ldots, F^{\ell(n)}(x_n))\\ &- \tilde{K}(F^{\ell(1)}(x_1), \ldots, F^{\ell(p)}(x_p), F^{\ell(p+1)}(x_{p+1}'), \ldots, F^{\ell(k+1)}(x_{k+1}'),\\ &F^{\ell(k+2)}(x_{k+2}), \ldots, F^{\ell(n)}(x_n))|.\end{aligned}$$

We apply lemma 4.1 with the two points $F^{\ell(p)}(x_p)$, $F^{\ell(p)}(x_p')$, $q = \ell(p)$, $\tilde{y} = x_p$, $\tilde{y}' = x_p'$, observing that $s(x_p, x_p') \geqslant \min(l + s(z, z'), p + \ell(p)) - p$. We finally obtain

$$|\Psi(\omega_1, \omega_2) - \Psi(\omega_1', \omega_2)| \leqslant \sum_{p=p_*+1}^{l} L_p \alpha^{\eta(\min(\ell(p), l-p+s(z,z')))}.$$

We claim that there exists a number c_0 such that for any l, z, z' and p such that $l \geqslant p \geqslant p_* + 1$, one has $\min(\ell(p), l - p + s(z, z')) \geqslant s(z, z')/2 + (l - p)/2 - c_0$. This is obvious if $l - p + s(z, z') \leqslant \ell(p)$. From the definition of p_* it follows that there exists a constant $c_1 > 0$ such that $\ell(p_*) \geqslant \ell(l) - c_1$. This implies (since $l \geqslant p \geqslant p_*$) $\ell(p) \geqslant \ell(p_*) \geqslant \ell(l) - c_1 \geqslant s(z, z')/2 + \ell(l)/2 - c_1$. This follows from the monotonicity of ℓ and the assumption $s(z, z') < \ell(l)$. On the other hand, from the definition of p_*, we have $\ell(l) \geqslant \ell(p_* + 1) \geqslant l - p_* - 1 \geqslant l - p$. Therefore we get the estimate (30).

It immediately follows from the definition that

$$|v_l^k(\xi, y) - v_l^k(\xi', y)| \leqslant C^\eta (\mathrm{diam}(M))^\eta \alpha^{\eta s(z,z')/2} \sum_{j=1}^{l} \alpha^{\eta(l-j-1)/2} L_j. \tag{31}$$

Using the estimate (29) we get

$$\left|\frac{\varphi(\xi)}{\varphi(z)J\overline{F}^l(\xi)}(v_l^k(\xi,y)-v_{l-1}^k(\xi,y))-\frac{\varphi(\xi')}{\varphi(z')J\overline{F}^l(\xi')}(v_l^k(\xi',y)-v_{l-1}^k(\xi',y))\right|$$
$$\leqslant (1+C_1)\alpha^{s(z,z')/2}B_{l+1}\frac{\varphi(\xi)}{\varphi(z)J\overline{F}^l(\xi)}+C^\eta(\mathrm{diam}(M))^\eta\alpha^{\eta s(z,z')/2}$$
$$\times\frac{\varphi(\xi')}{\varphi(z')J\overline{F}^l(\xi')}\sum_{j=1}^{l}\alpha^{\eta(l-j-1)/2}L_j.$$

The lemma follows from relation (28) by summing over ζ and using the identity $\mathcal{L}1=1$. ■

It follows immediately from the estimate (27) and lemma 5.5 that for fixed y, as a function of u, $w_l^k(u,\overline{F}(y))$ belongs to the space X_σ, where $\sigma=\alpha^{\eta/2}$, with an X_σ-norm satisfying uniformly in y and k

$$\|w_l^k(\cdot,\overline{F}(y))\|\leqslant\mathcal{O}(1)\left(B_{l+1}+(\mathrm{diam}(M))^\eta\sum_{j=1}^{l}\alpha^{\eta(l-j)/2}L_j\right).$$

Using lemma 4.2, we get for some constants $C>0$ and $0<\rho<1$ independent of K, l and k,

$$\|\mathcal{L}_1^{k-l}w_l^k(\cdot,\overline{F}(y))-a_{k,l}(\overline{F}(y))\|\leqslant C\Gamma_l^k, \tag{32}$$

where

$$\Gamma_l^k=\rho^{k-l}\left(B_{l+1}+(\mathrm{diam}(M))^\eta\sum_{j=1}^{l}\alpha^{\eta(l-j)/2}L_j\right) \tag{33}$$

and

$$a_{k,l}(y')=\int w_l^k(u,y')\,\mathrm{d}\mu_{\overline{\Delta}}(u).$$

Final estimates. We start by estimating the first term in (20). We observe that

$$\mathcal{L}^{k+1}U(\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y))-\mathcal{L}^kU(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y))$$
$$=\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}[\mathcal{L}^kU(u,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y))$$
$$-\mathcal{L}^kU(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y))],$$

where we have used the fact that $\mathcal{L}1=1$.

We obtain, using equations (23) and (24) and observing that $v_0^k(\overline{F}(u))=v_0^k(\overline{F}(y))$, the following estimate

$$\int(\mathcal{L}^{k+1}U(\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y))-\mathcal{L}^kU(y,\overline{F}(y),\ldots,\overline{F}^{n-k-1}(y)))^2\,\mathrm{d}\mu_{\overline{\Delta}}(y)$$
$$\leqslant\mathcal{O}(1)B_{k+1}^2+\mathcal{O}(1)\int\left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\right.$$
$$\left.\times\left(\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y))-\mathcal{L}_1^{k-l}w_l^k(y,\overline{F}(y)))\right)\right)^2\mathrm{d}\mu_{\overline{\Delta}}(y). \tag{34}$$

Since $\overline{F}(u) = \overline{F}(y)$, we have

$$\left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y)) - \mathcal{L}_1^{k-l}w_l^k(y,\overline{F}(y)))\right)^2$$
$$= \left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y)) - a_{k,l}(\overline{F}(u)))\right.$$
$$\left. - \frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(y,\overline{F}(y)) - a_{k,l}(\overline{F}(y)))\right)^2$$
$$\leqslant 2\left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y)) - a_{k,l}(\overline{F}(u)))\right)^2$$
$$+ 2\left(\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(y,\overline{F}(y)) - a_{k,l}(\overline{F}(y)))\right)^2. \tag{35}$$

We now estimate separately the integral of each term.

We define the integer-valued function $q(y)$ by

$$q(y) = q \qquad \text{if } y \in \overline{\Delta}_{q,j}.$$

We have from (32) and the definition of the norm in X_σ the following estimate uniform in $\overline{F}(y)$

$$|\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y)) - a_{k,l}(\overline{F}(y))| \leqslant \mathcal{O}(1)\mathrm{e}^{\varepsilon q(u)}\Gamma_l^k. \tag{36}$$

We have, since $\overline{F}(u) = \overline{F}(y)$,

$$\int\left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\sum_{l=0}^{k-1}(\mathcal{L}_1^{k-l}w_l^k(u,\overline{F}(y)) - a_{k,l}(\overline{F}(u)))\right)^2 \mathrm{d}\mu_{\overline{\Delta}}(y)$$
$$\leqslant \mathcal{O}(1)\left(\sum_{\ell=1}^{k-1}\Gamma_l^k\right)^2\int\left(\frac{1}{\varphi(\overline{F}(y))}\sum_{\overline{F}(u)=\overline{F}(y)}\frac{\varphi(u)}{J\overline{F}(u)}\mathrm{e}^{\varepsilon q(u)}\right)^2 \mathrm{d}\mu_{\overline{\Delta}}(y)$$
$$= \mathcal{O}(1)\left(\sum_{\ell=1}^{k-1}\Gamma_l^k\right)^2\int\left(\frac{1}{\varphi(y)}\sum_{\overline{F}(u)=y}\frac{\varphi(u)}{J\overline{F}(u)}\mathrm{e}^{\varepsilon q(u)}\right)^2 \mathrm{d}\mu_{\overline{\Delta}}(y)$$

by the invariance of the measure $\mu_{\overline{\Delta}}$. Using the Cauchy–Schwarz inequality and the property $\mathcal{L}1 = 1$, the last integral is bounded above by

$$\int\frac{1}{\varphi(y)}\sum_{\overline{F}(u)=y}\frac{\varphi(u)}{J\overline{F}(u)}\mathrm{e}^{2\varepsilon q(u)}\,\mathrm{d}\mu_{\overline{\Delta}}(y) = \int \mathrm{e}^{2\varepsilon q(y)}\,\mathrm{d}\mu_{\overline{\Delta}}(y).$$

The integral of the last term in the estimate (35) is bounded by the same quantity. By the invariance of the measure $\mu_{\overline{\Delta}}$, we have

$$\int \mathrm{e}^{2\varepsilon q(y)}\,\mathrm{d}\mu_{\overline{\Delta}}(y) = \sum_j\sum_{l=0}^{R_j-1}\mathrm{e}^{2\varepsilon l}\mu_{\overline{\Delta}}(\overline{\Delta}_{l,j}) = \sum_j \mu_{\overline{\Delta}}(\overline{\Delta}_{0,j})\sum_{l=0}^{R_j-1}\mathrm{e}^{2\varepsilon l}$$
$$\leqslant \frac{1}{\mathrm{e}^{2\varepsilon}-1}\sum_j \mu_{\overline{\Delta}}(\overline{\Delta}_{0,j})\mathrm{e}^{2\varepsilon R_j}.$$

Since φ is bounded on $\overline{\Delta}_0$ (see (8)) we get

$$\sum_j \mu_{\overline{\Delta}}(\overline{\Delta}_{0,j}) \mathrm{e}^{2\varepsilon R_j} \leqslant \mathcal{O}(1) \sum_n \mathrm{e}^{2\varepsilon n} m(R \geqslant n)$$

and this quantity is finite if ε is small enough by using (1) and property (I)-(ii) in [5, section 3.2]. Collecting all the bounds we get the following upper bound for the first term in (20):

$$2 \sum_{k=1}^{n-1} \int (\mathcal{L}^k U(y, \ldots, \overline{F}^{n-k-1}(y)) - \mathcal{L}^{k+1} U(\overline{F}(y), \ldots, \overline{F}^{n-k-1}(y)))^2 \, \mathrm{d}\mu_{\overline{\Delta}}(y)$$
$$\leqslant \mathcal{O}(1) \sum_{k=1}^{\infty} \left(\sum_{l=1}^{k-1} \Gamma_l^k \right)^2 .$$

Choosing p_0 large enough in the definition (15) of ℓ, we have

$$B_l \leqslant L_l (\mathrm{diam}(M))^\eta + C^\eta \sum_{j=l}^{\infty} \frac{L_{j+1}}{(j-l+1)^2}.$$

Using several times Young's inequality, one easily gets

$$\sum_{k=1}^{\infty} \left(\sum_{l=1}^{k-1} \Gamma_l^k \right)^2 \leqslant \mathcal{O}(1) \sum_{j=1}^{n} L_j^2.$$

Since a separately Hölder continuous function of n variables can also be considered as a separately Hölder continuous function of $n + k$ ($k > 0$) with $L_j = 0$ for $j > n$, the same estimate holds for the second term in (20).

We now prove that the third term in (20) tends to zero when $p \to \infty$. Using lemma 5.5 and estimate (31) with $k = n$ and $l = n - 1$ we observe that $S_n = \mathcal{L}^n U_n$ belongs to the Banach space X_σ. The result follows at once using lemma 4.2.

This ends the proof of theorem 3.1.

Acknowledgments

BS acknowledges the kind hospitality of the CPhT at Ecole Polytechnique. The authors acknowledge the anonymous referee for a very careful reading of the paper.

References

[1] Benedicks M and Young L-S 2000 Markov extensions and decay of correlations for certain Hénon maps *Géométrie Complexe et Systèmes Dynamiques* (*Orsay, 1995*) *Astérisque* **261** 13–56

[2] Chazottes J-R, Collet P and Schmitt B 2005 Statistical consequences of Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems *Nonlinearity* **18** 2341–64

[3] Collet P, Martinez S and Schmitt B 2002 Exponential inequalities for dynamical measures of expanding maps of the interval *Probab. Theory Relat. Fields* **123** 301–22

[4] Devroye L 1991 Exponential inequalities in nonparametric estimation *Nonparametric Functional Estimation and Related Topics (Spetses, 1990)* (*NATO Advanced Science Institute Series C, Mathematical and Physical Sciences* vol 335) (Dordrecht: Kluwer Academic) pp 31–44

[5] Young L-S 1998 Statistical properties of dynamical systems with some hyperbolicity *Ann. Math.* **147** 585–650

[6] Young L-S 1999 Recurrence times and rates of mixing *Israel J. Math.* **110** 153–88

[7] Chernov N and Young L-S 2000 Decay of correlations for Lorentz gases and hard balls *Hard Ball Systems and the Lorentz Gas* (*Encyclopedia Mathematical Sciences* vol 101) (Berlin: Springer) pp 89–120