SHORTEST DISTANCE BETWEEN MULTIPLE ORBITS AND GENERALIZED FRACTAL DIMENSIONS

VANESSA BARROS AND JÉRÔME ROUSSEAU*

Abstract. We consider rapidly mixing dynamical systems and link the decay of the shortest distance between multiple orbits with the generalized fractal dimension. We apply this result to multidimensional expanding maps and extend it to the realm of random dynamical systems. For random sequences, we obtain a relation between the longest common substring between multiple sequences and the generalized Rényi entropy. Applications to Markov chains and Gibbs states are given.

1. Introduction

Generalized fractal dimensions were originally introduced to characterize and measure the strangeness of chaotic attractors and, more generally, to describe the fractal structure of invariant sets in dynamical systems [25, 26, 27].

Given $k > 1$, the generalized fractal dimension (also known as L^q or HP dimensions) of a measure μ is defined (provided the limit exists) by:

$$D_k(\mu) = \lim_{r\to 0} \frac{\log \int_X \mu\left(B\left(x,r\right)\right)^{k-1} d\mu(x)}{(k-1)\log r}.$$

For the existence of these dimensions, their properties and relations with other dimensions, one can see e.g. [10, 19, 40, 41].

Since estimation of the generalized dimensions plays an important role in the description of dynamical systems, different numerical approaches and procedures have been developed to compute them (see e.g. [3, 7, 8, 12, 15, 39] and references within). In particular, we highlight [22] where Extreme Value Theory (EVT) was used as a tool to estimate the correlation dimension $D_2(\mu)$, and [15] for generalized dimensions. For a deeper discussion of EVT for dynamical systems we refer the reader to [21].

It is also worth mentioning the connection between generalized dimensions and the recurrence properties of the dynamics. Return time dimensions and generalized fractal dimensions were thoroughly compared in [29, 37]. Moreover, they appear in the rate function for the large deviations of the return time [15, 18].

In this communication we study, for a dynamical system (X, T, μ), the behaviour of the shortest distance between k orbits, i.e. for $(x_1, \ldots, x_k) \in X^k$:

$$m_n(x_1, \ldots, x_k) = \min_{i_1,\ldots,i_k=0,\ldots,n-1} \left(d(T^{i_1}x_1, \ldots, T^{i_k}x_k)\right), \tag{1}$$

Key words and phrases. Shortest distance, generalized fractal dimensions, decay of correlations, random dynamical systems, Rényi entropy, longest common substring, string matching.

* Corresponding author.

where $d(x_1,\dots,x_k)=\max_{i\neq j} d(x_i,x_j)$, and show a relation between this shortest distance and the generalized fractal dimensions.

Indeed, if the generalized dimension exists, then under some rapid mixing conditions on the system (X,T,μ), for $\mu\otimes\cdots\otimes\mu$-almost every $(x_1,\dots,x_k)\in X^k$, we have

$$\lim_{n\to+\infty}\frac{\log m_n(x_1,\dots,x_k)}{-\log n}=\frac{k}{(k-1)D_k(\mu)}. \tag{2}$$

In particular, we apply these results to the multidimensional expanding maps defined by Saussol [46]. Moreover, we also prove an annealed version of (2) for the shortest distance between k orbits of a random dynamical system.

These results extend and complement those in [11] and [17] where identity (2) (and its equivalent for random dynamical systems) was proved for two orbits ($k=2$).

Furthermore, it was shown in [11] that the problem of the shortest distance between orbits is a generalization of the longest common substring problem for random sequences, a problem thoroughly investigated in genetics, probability and computer science (see e.g [50]). More precisely, for α-mixing systems, they study the behaviour of the length of the longest common substring between two sequences x and y:

$$M_n(x,y)=\max\{m : x_{i+k}=y_{j+k} \text{ for } k=1,\dots,m \text{ and for some } 0\leq i,j\leq n-m\},$$

and generalized the work of Arratia and Waterman [6] where only independent irreducible and aperiodic Markov chains on a finite alphabet were considered. More recently, similar results for encoded sequences [17], random sequences in random environment [44], stationary determinantal process on the integer lattice [20] and also for the longest matching consecutive subsequence between two N-ary expansions [35] were obtained.

Following the ideas in [11], we extend here our study to the longest common substring between multiple sequences (previous results in this direction were obtained in [33, 34]). More precisely, for k sequences $x^1,\dots,x^k$, we define the length of the longest common substring by

$$M_n(x^1,\dots,x^k)$$
$$=\max\{m : x^1_{i_1+j}=...=x^k_{i_k+j} \text{ for } j=0,...,m-1 \text{ and for some } 0\leq i_1,...,i_k\leq n-m\}.$$

and link it to the generalized Rényi entropy (provided that it exists, see e.g. [1, 2, 30, 36]):

$$H_k=\lim_{n\to+\infty}\frac{\log\sum\mathbb{P}(C_n)^k}{-(k-1)n},$$

where the sum is taken over all n-cylinders C_n (see Section 4 for a precise definition).

Thus, we prove that for α-mixing systems with exponential decay (or ψ-mixing with polynomial decay), if the generalized Rényi entropy exists, then for $\mathbb{P}^k$-almost every $(x^1,\dots,x^k)$,

$$\lim_{n\to+\infty}\frac{M_n(x^1,\dots,x^k)}{\log n}=\frac{k}{(k-1)H_k}.$$

The paper is organized as follows. Our main results linking the shortest distance between multiple orbits and the generalized fractal dimensions are stated in Section 2 and proved in Section 7. An application of these results for multidimensional expanding maps is given in Section 5. Shortest distance between multiple observed orbits and random orbits are studied in Section 3. In Section 4, we study the longest common substring problem for multiple random sequences and its relation with the generalized Rényi entropy. These results are proved in Section 6.

2. Shortest distance between k orbits

Let (X,d) be a finite dimensional metric space and $\mathcal{A}$ its Borel σ-algebra. Let $(X,\mathcal{A},\mu,T)$ be a measure preserving system which means that $T: X\to X$ is a transformation on X and μ is a probability measure on $(X,\mathcal{A})$ such that μ is invariant by T, i.e., $\mu(T^{-1}A)=\mu(A)$ for all $A\in\mathcal{A}$. We will denote by μ^k the product measure $\mu\otimes\cdots\otimes\mu$.

We would like to study the behaviour of the shortest distance between k orbits:

$$m_n(x_1,\dots,x_k)=\min_{i_1,\dots,i_k=0,\dots,n-1}\left(d(T^{i_1}x_1,\dots,T^{i_k}x_k)\right)$$

where $d(x_1,\dots,x_k)=\max_{i\neq j} d(x_i,x_j)$.

Remark 2.1. *Other definitions could have been chosen for $d(x_1,\dots,x_k)$ without altering our results (see e.g.* [32, 49] *and references therein for examples of generalizations of the usual two-way distance). For example, we could have used $d_1(x_1,\dots,x_k)=\min_{z\in X}\max_i d(x_i,z)$, or $d_2(x_1,\dots,x_k)=\sqrt{\sum_{i\neq j} d(x_i,x_j)^2}$ but our results would have been the same since d, d_1, and d_2 are equivalent.*

We will show that the behaviour of m_n as $n\to\infty$ is linked with the generalized fractal dimension. Before stating the first theorem, we recall, for $k>1$, the definition of the lower and upper generalized fractal dimensions of μ:

$$\underline{D}_k(\mu)=\varliminf_{r\to0}\frac{\log\int_X\mu\left(B\left(x,r\right)\right)^{k-1}d\mu(x)}{(k-1)\log r}\qquad\text{and}\qquad\overline{D}_k(\mu)=\varlimsup_{r\to0}\frac{\log\int_X\mu\left(B\left(x,r\right)\right)^{k-1}d\mu(x)}{(k-1)\log r}.$$

When the limit exists we will denote the common value of $\underline{D}_k(\mu)$ and $\overline{D}_k(\mu)$ by $D_k(\mu)$.

Theorem 2.2. *Let $(X,\mathcal{A},\mu,T)$ be a measure preserving system such that $\underline{D}_k(\mu)>0$. Then for μ^k-almost every $(x_1,\dots,x_k)\in X^k$,*

$$\varlimsup_{n\to+\infty}\frac{\log m_n(x_1,\dots,x_k)}{-\log n}\leq\frac{k}{(k-1)\underline{D}_k(\mu)}.$$

This general result can be applied to any dynamical system such that $\underline{D}_k(\mu)>0$. Even if the inequality in Theorem 2.2 can be strict (noting for example the trivial case when T is the identity), we will prove that an equality holds under some rapidly mixing conditions:

(H1) There exists a Banach space $\mathcal{C}$, such that for all ψ, $\phi\in\mathcal{C}$ and for all $n\in\mathbb{N}^*$, we have

$$\left|\int_X\psi.\phi\circ T^n\,d\mu-\int_X\psi d\mu\int_X\phi d\mu\right|\leq\|\psi\|_{\mathcal{C}}\|\phi\|_{\mathcal{C}}\theta_n,$$

with $\theta_n=a^n$ $(0\leq a<1)$ and where $\|\cdot\|_{\mathcal{C}}$ is the norm in the Banach space $\mathcal{C}$.

(H2) There exist $0<r_0<1$, $c\geq0$ and $\xi\geq0$ such that for every $p\in\{1,\dots,k\}$, for μ^{k-p}-almost every $x_{p+1},\dots,x_k\in X$ and any $0<r<r_0$, the function $\psi_p: X\to\mathbb{R}$, defined below, belongs to the Banach space $\mathcal{C}$ and verifies

$$\|\psi_p\|_{\mathcal{C}}\leq cr^{-\xi}.$$

Fixed $x_2,\dots,x_k\in X$, we define

$$\psi_1(x)=\prod_{j=2}^{k}\mathbb{1}_{B(x_j,r)}(x).\tag{3}$$

For $p > 1$, we fix $x_{p+1}, \dots, x_k \in X$, and set

$$\psi_p(x) = \bar{\psi}(x, x_{p+1}, \dots, x_k), \text{ where} \tag{4}$$

$$\bar{\psi}(x_p, x_{p+1}, \dots, x_k) = \prod_{l=p+1}^{k} \mathbb{1}_{B(x_l,r)}(x_p) \int_{X^{p-1}} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) \right] d\mu^{p-1}(x_1, \dots, x_{p-1}).$$

This condition imposes some regularity (with respect to the Banach space $\mathcal{C}$) on the indicator functions of intersections of balls and on the measure. More precisely, to satisfy (H2) the measure cannot change drastically around these intersections. When the Banach space $\mathcal{C}$ is the space of Hölder functions $\mathcal{H}^\alpha(X, \mathbb{R})$, we will replace our assumption (H2) by an assumption easier to interpret in Theorem 2.7.

We will also need some topological information on the space X.

Definition 2.3. *A separable metric space (X, d) is called tight if there exist $r_0 > 0$ and $N_0 \in \mathbb{N}$, such that for any $0 < r < r_0$ and any $x \in X$ one can cover $B(x, 2r)$ by at most N_0 balls of radius r.*

We emphasize that any subset of $\mathbb{R}^n$ with the Euclidian metric is tight, any subset of a Riemannian manifold of bounded curvature is tight and that if (X, d) admits a doubling measure then it is tight [28].

Now we can state our main result.

Theorem 2.4. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system, such that (X, d) is tight, satisfying (H1) and (H2) and such that $D_k(\mu)$ exists and is strictly positive. Then for μ^k-almost every $(x_1, \dots, x_k) \in X^k$,*

$$\lim_{n \to +\infty} \frac{\log m_n(x_1, \dots, x_k)}{-\log n} = \frac{k}{(k-1) D_k(\mu)}.$$

Now, we will apply this result to a short list of simple examples. Later, in Section 5, we use this theorem for a more complex family of examples (multidimensional piecewise expanding maps).

Denote by *Leb* the Lebesgue measure.

Example 2.5. *Theorem 2.4 can be applied to the following systems:*

1. *For $m \in \{2, 3, \dots\}$, let $T : [0,1] \to [0,1]$ be such that $x \mapsto mx \mod 1$ and $\mu = Leb$.*
2. *Let $T : (0,1] \to (0,1]$ be such that $T(x) = 2^n(x - 2^{-n})$ for $x \in (2^{-n}, 2^{-n+1}]$ and $\mu = Leb$.*
3. *(β-transformations) For $\beta > 1$, let $T : [0,1] \to [0,1]$ be such that $x \mapsto \beta x \mod 1$ and μ be the Parry measure (see [38]), which is an absolutely continuous probability measure with density ρ satisfying $1 - \frac{1}{\beta} \le \rho(x) \le (1 - \frac{1}{\beta})^{-1}$ for all $x \in [0,1]$.*
4. *(Gauss map) Let $T : (0,1] \to (0,1]$ be such that $T(x) = \left\{ \frac{1}{x} \right\}$ and $d\mu = \frac{1}{\log 2} \frac{dx}{1+x}$.*

In these examples it is easy to see that $D_k(\mu) = 1$. Moreover, (H1) and (H2) are satisfied with the Banach space $\mathcal{C} = BV$, the space of functions of bounded variation (see e.g. [24] Section 4.1 and [31, 42, 43]).

One can observe that Theorem 2.4 is an immediate consequence of Theorem 2.2 and the next theorem.

Theorem 2.6. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system, such that $\underline{D}_k(\mu) > 0$ and such that (X, d) is tight, satisfying (H1) and (H2). Then for μ^k-almost every $(x_1, \ldots, x_k) \in X^k$,*

$$\underline{\lim}_{n \to +\infty} \frac{\log m_n(x_1, \ldots, x_k)}{-\log n} \geq \frac{k}{(k-1)\overline{D}_k(\mu)}.$$

When the Banach space $\mathcal{C}$ is the space of Hölder functions $\mathcal{H}^\alpha(X, \mathbb{R})$ we can adapt our proof and (H2) can be replaced by the following condition:

(HA) There exist $r_0 > 0$, $\xi \geq 0$ and $\beta > 0$ such that for μ-almost every $x \in X$ and any $r_0 > r > \rho > 0$,

$$\mu(B(x, r + \rho) \backslash B(x, r - \rho)) \leq r^{-\xi} \rho^\beta.$$

This assumption is satisfied, for example, if the measure is Lebesgue or absolutely continuous with respect to Lebesgue with a bounded density.

Theorem 2.7. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system, such that $\underline{D}_k(\mu) > 0$ and such that (X, d) is tight, satisfying (H1) with $\mathcal{C} = \mathcal{H}^\alpha(X, \mathbb{R})$ and (HA). Then for μ^k-almost every $(x_1, \ldots, x_k) \in X^k$,*

$$\underline{\lim}_{n \to +\infty} \frac{\log m_n(x_1, \ldots, x_k)}{-\log n} \geq \frac{k}{(k-1)\overline{D}_k(\mu)}.$$

For example, one can apply this theorem to expanding maps of the interval with a Gibbs measure associated to a Hölder potential (see e.g. [47]) and C^2 endomorphism (of a d-dimensional compact Riemannian manifold) admitting a Young tower with exponential tail (see [23, Section 6] and [16]).

3. Observed orbits and random dynamical systems

In this section, we extend our analysis to the study of observation of orbits. Indeed, considering observations of systems (for example, temperature or pressure while studying climate) could be more significant than considering the whole system. From a more theoretical point of view, we will explain in Section 3.1 how the study of observed orbits allows us to study random dynamical systems.

Let (Y, d) be a metric space and for any $y_1, \ldots, y_k \in Y$, we define $d(y_1, \ldots, y_k) = \max_{i \neq j} d(y_i, y_j)$. Let $f : X \to Y$ be a measurable function (called the observation). We denote by $f_*\mu$ the pushforward measure, defined by $f_*\mu(A) = \mu(f^{-1}(A))$ for measurable subsets $A \subset Y$.

We would like to study the behaviour of the shortest distance between k observed orbits:

$$m_n^f(x_1, \ldots, x_k) = \min_{i_1, \ldots, i_k = 0, \ldots, n-1} \left(d(f(T^{i_1} x_1), \ldots, f(T^{i_k} x_k)) \right)$$

Theorem 3.1. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system such that $\underline{D}_k(f_*\mu) > 0$. Then for μ^k-almost every $(x_1, \ldots, x_k) \in X^k$,*

$$\overline{\lim}_{n \to +\infty} \frac{\log m_n^f(x_1, \ldots, x_k)}{-\log n} \leq \frac{k}{(k-1)\underline{D}_k(f_*\mu)}.$$

We will assume that f is Lipschitz and as in Section 2, we prove that the equality holds under some rapidly mixing conditions:

(H1') For all $\psi,\ \phi \in \mathcal{H}^\alpha(Y, \mathbb{R})$ and for all $n \in \mathbb{N}^*$, we have

$$\left| \int_X \psi(f(x)).\phi(f(T^n x))\, d\mu(x) - \int_X \psi(f(x)) d\mu(x) \int_X \phi(f(x)) d\mu(x) \right| \leq \|\psi \circ f\|_{\mathcal{H}^\alpha} \|\phi \circ f\|_{\mathcal{H}^\alpha} \theta_n,$$

with $\theta_n = a^n$ $(0 \leq a < 1)$.

For simplicity, we only treat the case when the mixing property is satisfied for Hölder observables. However, we observe that one can adapt (H1) and (H2) to this setting to work with other Banach spaces.

Now we can state our version of Theorem 2.7 for observed orbits.

Theorem 3.2. *Let $(X, \mathcal{A}, \mu, T)$ be a measure preserving system and f a Lipschitz observation, such that $\underline{D}_k(f_*\mu) > 0$ and such that (Y, d) is tight, satisfying (H1') and such that $f_*\mu$ satisfies (HA). Then for μ^k-almost every $(x_1, \dots, x_k) \in X^k$,*

$$\underline{\lim}_{n\to+\infty} \frac{\log m_n^f(x_1, \dots, x_k)}{-\log n} \geq \frac{k}{(k-1)\overline{D}_k(f_*\mu)}.$$

Moreover, if $D_k(f_*\mu)$ exists, then for μ^k-almost every $(x_1, \dots, x_k) \in X^k$,

$$\lim_{n\to+\infty} \frac{\log m_n^f(x_1, \dots, x_k)}{-\log n} = \frac{k}{(k-1)D_k(f_*\mu)}.$$

3.1. **Shortest distance between multiple random orbits.** In this subsection, we will use the previous results to study the shortest distance between multiple orbits of a random dynamical system.

Let (X, d) be a tight metric space and let $(\Omega, \theta, \mathbb{P})$ be a probability measure preserving system, where Ω is a metric space and $B(\Omega)$ its Borelian σ-algebra.

Definition 3.3. *A random dynamical system $\mathcal{T} = (T_\omega)_{\omega\in\Omega}$ on X over $(\Omega, B(\Omega), \mathbb{P}, \theta)$ is generated by maps T_ω such that $(\omega, x) \mapsto T_\omega(x)$ is measurable and satisfies:*

$$T_\omega^0 = Id \text{ for all } \omega \in \Omega,$$

$$T_\omega^n = T_{\theta^{n-1}(\omega)} \circ \cdots \circ T_{\theta(\omega)} \circ T_\omega \text{ for all } n \geq 1.$$

The map $S : \Omega \times X \to \Omega \times X$ defined by $S(\omega, x) = (\theta(\omega), T_\omega(x))$ is the dynamics of the random dynamical systems generated by $\mathcal{T}$ and is called skew-product.

A probability measure μ is said to be an invariant measure for the random dynamical system $\mathcal{T}$ if it satisfies

1. *μ is S-invariant*
2. *$\pi_*\mu = \mathbb{P}$*

where $\pi : \Omega \times X \to \Omega$ is the canonical projection.

Let $(\mu_\omega)_\omega$ denote the decomposition of μ on X, that is, $d\mu(\omega, x) = d\mu_\omega(x)d\mathbb{P}(\omega)$. We denote by $\nu = \int \mu_\omega d\mathbb{P}$ the marginal of ν on X.

For $(\omega_1, x_1), \dots, (\omega_k, x_k)$, we define the shortest distance between k random orbits by

$$m_n^{\omega_1,\dots,\omega_k}(x_1, \dots, x_k) = \min_{i_1,\dots,i_k=0,\dots,n-1} \left(d\left(T_{\omega_1}^{i_1}(x_1), \dots, T_{\omega_k}^{i_k}(x_k)\right)\right).$$

Remark 3.4. *We observe that the technic developed here only allows us to obtain annealed results. Another object worth studying would be the quenched shortest distance*

$$m_n^\omega(x_1, \dots, x_k) = \min_{i_1,\dots,i_k=0,\dots,n-1} \left(d\left(T_\omega^{i_1}(x_1), \dots, T_\omega^{i_k}(x_k)\right)\right).$$

In this direction, the only known results are for 2 orbits and when the system is a random subshift of finite type [44].

As in the deterministic case, we will assume an exponential decay of correlations for the random dynamical system:

(H1R) (Annealed decay of correlations) For every $n \in \mathbb{N}^*$, and every ψ, $\phi \in \mathcal{H}^\alpha(X, \mathbb{R})$,

$$\left| \int_{\Omega\times X} \psi(T_\omega^n(x))\phi(x)\, d\mu(\omega, x) - \int_{\Omega\times X} \psi\, d\mu \int_{\Omega\times X} \phi\, d\mu \right| \leq \|\psi\|_{\mathcal{H}^\alpha}\|\phi\|_{\mathcal{H}^\alpha}\theta_n,$$

with $\theta_n = a^n$ $(0 \leq a < 1)$.

Theorem 3.5. *Let $\mathcal{T}$ be a random dynamical system on X over $(\Omega, B(\Omega), \mathbb{P}, \theta)$ with an invariant measure μ such that $\underline{D}_k(\nu) > 0$. Then for μ^k-almost every $(\omega_1, x_1, \ldots, \omega_k, x_k) \in (\Omega \times X)^k$,*

$$\overline{\lim_{n\to\infty}} \frac{\log m_n^{\omega_1,\ldots,\omega_k}(x_1, \ldots, x_k)}{-\log n} \leq \frac{k}{(k-1)\underline{D}_k(\nu)}.$$

Moreover, if the random dynamical system satisfies assumptions (H1R) and ν satisfies (HA), then

$$\underline{\lim_{n\to\infty}} \frac{\log m_n^{\omega_1,\ldots,\omega_k}(x_1, \ldots, x_k)}{-\log n} \geq \frac{k}{(k-1)\overline{D}_k(\nu)},$$

and if $D_k(\nu)$ exists, then

$$\lim_{n\to\infty} \frac{\log m_n^{\omega_1,\ldots,\omega_k}(x_1, \ldots, x_k)}{-\log n} = \frac{k}{(k-1)D_k(\nu)}\ .$$

Proof. Following the ideas in [45], it is enough to apply Theorem 3.1 and Theorem 3.2 for the dynamical system $(\Omega \times X, B(\Omega \times X), \mu, S)$ with the observation f defined by

$$\begin{aligned} f : \Omega \times X &\to X \\ (\omega, x) &\mapsto x. \end{aligned}$$

□

We now apply the above result to some simple non-i.i.d. random dynamical system and we observe that, as in [17], Theorem 3.5 could also be applied to randomly perturbed dynamical systems and random hyperbolic toral automorphisms.

Example 3.6 (Non-i.i.d. random expanding maps)**.** *Consider the two following linear maps*

$$\begin{aligned} T_1 : X &\to X \quad \textit{and} \quad & T_2 : X &\to X \\ x &\mapsto 2x & x &\mapsto 3x, \end{aligned}$$

where X is the one-dimensional torus $\mathbb{T}^1$. It is easy to see that T_1 and T_2 preserve the Lebesgue measure(Leb).

The following skew product gives the dynamics of the random dynamical system:

$$\begin{aligned} S : \Omega \times X &\to \Omega \times X \\ (\omega, x) &\mapsto (\theta(\omega), T_\omega(x)), \end{aligned}$$

with $\Omega = [0, 1]$, $T_\omega = T_1$ if $\omega \in [0, 2/5)$ and $T_\omega = T_2$ if $\omega \in [2/5, 1]$ where ω is the following piecewise linear map:

$$\theta(\omega) = \begin{cases} 2\omega & \text{if } \omega \in [0, 1/5) \\ 3\omega - 1/5 & \text{if } \omega \in [1/5, 2/5) \\ 2\omega - 4/5 & \text{if } \omega \in [2/5, 3/5) \\ 3\omega/2 - 1/2 & \text{if } \omega \in [3/5, 1]. \end{cases}$$

The associated skew-product S is $Leb \otimes Leb$-invariant. It is easy to check that Lebesgue measure satisfies (HA). Moreover, by [9] *the skew product S has an exponential decay of correlations. Since in this example $\nu = Leb$, we have $D_k(\nu) = 1$ and Theorem 3.5 implies that for Leb^{2k}-almost every $(\omega_1, x_1, \dots, \omega_k, x_k) \in ([0,1] \times \mathbb{T}^1)^k$,*

$$\lim_{n\to\infty} \frac{\log m_n^{\omega_1,\dots,\omega_k}(x_1,\dots,x_k)}{-\log n} = \frac{k}{k-1}.$$

4. Longest common substring between k random sequences

It was shown in [11] that studying the shortest distance between orbits for a symbolic dynamical system coincides with studying the length of the longest common substring between sequences.

Thus we will consider the symbolic dynamical systems $(\Omega, \mathbb{P}, \sigma)$, where $\Omega = \mathcal{A}^{\mathbb{N}}$ for some alphabet $\mathcal{A}$, σ is the (left) shift on Ω and $\mathbb{P}$ is a σ-invariant probability measure. For k sequences $x^1, \dots, x^k \in \Omega$, we are interested in the behaviour of

$$M_n(x^1,...,x^k) = \max\{m : x^1_{i_1+j} = ... = x^k_{i_k+j} \text{ for } j = 0,...,m-1 \text{ and for some } 0 \le i_1,...,i_k \le n-m\}.$$

We will show that the behaviour of M_n is linked with the generalized Rényi entropy of the system.

For $y \in \Omega$ we denote by $C_n(y) = \{z \in \Omega : z_i = y_i \text{ for all } 0 \le i \le n-1\}$ the *n-cylinder* containing y. Set $\mathcal{F}_0^n$ as the sigma-algebra over Ω generated by all n-cylinders.

For $k > 1$, we recall the definition of the lower and upper generalized Rényi entropy:

$$\underline{H}_k(\mathbb{P}) = \varliminf_{n\to+\infty} \frac{\log \sum \mathbb{P}(C_n)^k}{-(k-1)n} \qquad \text{and} \qquad \overline{H}_k(\mathbb{P}) = \varlimsup_{n\to+\infty} \frac{\log \sum \mathbb{P}(C_n)^k}{-(k-1)n},$$

where the notation $\sum \mathbb{P}(C_n)^k$ means $\sum_{y\in\mathcal{A}^n} \mathbb{P}(C_n(y))^k$. When the limit exists, we will denote it by $H_k(\mathbb{P})$.

We say that a system $(\Omega, \mathbb{P}, \sigma)$ is *α-mixing* if there exists a function $\alpha : \mathbb{N} \to \mathbb{R}$ satisfying $\alpha(g) \to 0$ when $g \to +\infty$ and such that for all $m, n \in \mathbb{N}$, $A \in \mathcal{F}_0^n$ and $B \in \mathcal{F}_0^m$:

$$\left|\mathbb{P}(A \cap \sigma^{-g-n}B) - \mathbb{P}(A)\mathbb{P}(B)\right| \le \alpha(g).$$

It is said to be *α-mixing with an exponential decay* if the function $\alpha(g)$ decreases exponentially fast to 0.

We say that our system is *ψ-mixing* if there exists a function $\psi : \mathbb{N} \to \mathbb{R}$ satisfying $\psi(g) \to 0$ when $g \to +\infty$ and such that for all $m, n \in \mathbb{N}$, $A \in \mathcal{F}_0^n$ and $B \in \mathcal{F}_0^m$:

$$\left|\mathbb{P}(A \cap \sigma^{-g-n}B) - \mathbb{P}(A)\mathbb{P}(B)\right| \le \psi(g)\mathbb{P}(A)\mathbb{P}(B).$$

Now we are ready to state our next result.

Theorem 4.1. *If $\underline{H}_k(\mathbb{P}) > 0$, then for $\mathbb{P}^k$-almost every $(x^1, \dots, x^k) \in \Omega^k$,*

$$\varlimsup_{n\to+\infty} \frac{M_n(x^1,\dots,x^k)}{\log n} \le \frac{k}{(k-1)\underline{H}_k(\mathbb{P})}. \tag{5}$$

Moreover, if the system is α-mixing with an exponential decay or if it is ψ-mixing with $\psi(g) = g^{-a}$ for some $a > 0$ then, for $\mathbb{P}^k$-almost every $(x^1, \dots, x^k) \in \Omega^k$,

$$\varliminf_{n\to+\infty} \frac{M_n(x^1, \dots, x^k)}{\log n} \geq \frac{k}{(k-1)\overline{H}_k(\mathbb{P})}. \tag{6}$$

Therefore, if the generalized Rényi entropy exists, then for $\mathbb{P}^k$-almost every $(x^1, \dots, x^k) \in \Omega^k$,

$$\lim_{n\to+\infty} \frac{M_n(x^1, \dots, x^k)}{\log n} = \frac{k}{(k-1)H_k(\mathbb{P})}.$$

This theorem can be applied, for example, to Markov chains and Gibbs states:

Example 4.2 (Markov chains)**.** *If $(\Omega, \mathbb{P}, \sigma)$ is an irreducible and aperiodic Markov chain on a finite alphabet $\mathcal{A}$, then it is ψ-mixing with an exponential decay (see e.g.* [14]*). If we denote by P the associated stochastic matrix (with entries P_{ij}), then the matrix $P(k)$ whose entries are $P_{ij}(k) = P_{ij}^k$ has, by the Perron-Frobenius theorem, a single largest eigenvalue λ_k. Moreover, the generalized Rényi entropy exists and $H_k(\mathbb{P}) = -\log \lambda_k/(k-1)$* [30]*. Thus, for $\mathbb{P}^k$-almost every $(x^1, \dots, x^k) \in \Omega^k$,*

$$\lim_{n\to+\infty} \frac{M_n(x^1, \dots, x^k)}{\log n} = \frac{k}{-\log \lambda_k}.$$

Example 4.3 (Gibbs states)**.** *Let $\mathbb{P}$ be a Gibbs state of a Hölder-continuous potential ϕ. Then, the system is ψ-mixing with an exponential decay* [13, 48]*. Moreover, the generalized Rényi entropy exists and $H_k(\mathbb{P}) = (1/(k-1))\,(kP(\phi) - P(k\phi))$ where $P(\phi)$ is the pressure of the potential ϕ* [30]*. Thus, for $\mathbb{P}^k$-almost every $(x^1, \dots, x^k) \in \Omega^k$,*

$$\lim_{n\to+\infty} \frac{M_n(x^1, \dots, x^k)}{\log n} = \frac{k}{kP(\phi) - P(k\phi)}.$$

5. Multidimensional piecewise expanding maps

In this section, we apply Theorem 2.4 to a family of maps defined by Saussol [46]: multidimensional piecewise uniformly expanding maps. It was observed in [4] that these maps generalize Markov maps which also contain one-dimensional piecewise uniformly expanding maps.

Let $N \geq 1$ be an integer. We will work in the Euclidean space $\mathbb{R}^N$. We denote by $B_\epsilon(x)$ the ball with center x and radius ϵ. For a set $E \subset \mathbb{R}^N$, we write

$$B_\epsilon(E) := \{y \in \mathbb{R}^N : d(y, E) \leq \epsilon\}.$$

Definition 5.1 (Multidimensional piecewise expanding systems)**.** *Let X be a compact subset of $\mathbb{R}^N$ with $\overline{X^\circ} = X$ and $T : X \to X$. The system (X, T) is a multidimensional piecewise expanding system if there exists a family of at most countably many disjoint open sets $U_i \subset X$ and V_i such that $\overline{U_i} \subset V_i$ and maps $T_i : V_i \to \mathbb{R}^N$ satisfying for some $0 < \alpha \leq 1$, for some small enough $\epsilon_0 > 0$, and for all i:*

(1) $T|_{U_i} = T_i|_{U_i}$ and $B_{\epsilon_0}(TU_i) \subset T_i(V_i)$;

(2) $T_i \in C^1(V_i), T_i$ is injective and $T_i^{-1} \in C^1(T_iV_i)$. Moreover, there exists a constant c, such that for all $\epsilon \leq \epsilon_0, z \in T_iV_i$ and $x, y \in B_\epsilon(z) \cap T_iV_i$ we have

$$|\det D_xT_i^{-1} - \det D_yT_i^{-1}| \leq c\epsilon^\alpha |\det D_zT_i^{-1}|;$$

(3) $Leb(X \setminus \bigcup_i U_i) = 0$;

(4) *there exists* $s = s(T) < 1$ *such that for all* $u, v \in TV_i$ *with* $d(u,v) \le \epsilon_0$ *we have* $d(T_i^{-1}u, T_i^{-1}v) \le sd(u,v)$;
(5) *let* $G(\epsilon, \epsilon_0) := \sup_x G(x, \epsilon, \epsilon_0)$ *where*

$$G(x,\epsilon,\epsilon_0) = \sum_i \frac{Leb(T_i^{-1}B_\epsilon(\partial TU_i) \cap B_{(1-s)\epsilon_0}(x))}{m(B_{(1-s)\epsilon_0}(x))},$$

then the number $\eta = \eta(\delta) := s^\alpha + 2\sup_{\epsilon\le\delta} \frac{G(\epsilon)}{\epsilon^\alpha}\delta^\alpha$ *satisfies* $\sup_{\delta\le\epsilon_0}\eta(\delta) < 1$.

We will prove that the multidimensional piecewise expanding systems satisfy the conditions of Theorem 2.4.

Proposition 5.2. *Let* (X,T) *be a topologically mixing multidimensional piecewise expanding map and* μ *be its absolutely continuous invariant probability measure. Then for* μ^k*-almost every* $(x_1, \ldots, x_k) \in X^k$,

$$\lim_{n\to+\infty} \frac{\log m_n(x_1,\ldots,x_k)}{-\log n} = \frac{k}{(k-1)N}.$$

Proof. First of all, we define the Banach space involved in the mixing conditions. Let $\Gamma \subset X$ be a Borel set. We define the oscillation of $\varphi \in L^1(Leb)$ over Γ as

$$osc(\varphi,\Gamma) = \operatorname*{ess-sup}_\Gamma(\varphi) - \operatorname*{ess-inf}_\Gamma(\varphi).$$

Now, given real numbers $0 < \alpha \le 1$ and $0 < \epsilon_0 < 1$ consider the following α-seminorm

$$|\varphi|_\alpha = \sup_{0<\epsilon\le\epsilon_0} \epsilon^{-\alpha} \int_X osc(\varphi, B_\epsilon(x))dx.$$

We observe that $X \ni x \mapsto osc(\varphi, B_\epsilon(x))$ is a measurable function (see [46]) and

$$\text{supp}(\text{osc}(\varphi, B_\epsilon(x))) \subset B_\epsilon(\text{supp } \varphi).$$

Let V_α be the space of $L^1(Leb)-$functions such that $|\varphi|_\alpha < \infty$ endowed with the norm

$$\|\varphi\|_\alpha = \|\varphi\|_{L^1(Leb)} + |\varphi|_\alpha.$$

Then $(V_\alpha, \|\cdot\|_\alpha)$ is a Banach space which does not depend on the choice of ϵ_0 and $V_\alpha \subset L^\infty$ (see [46]).

Saussol [46] proved that for a piecewise expanding map $T : X \longrightarrow X$, where $X \subset \mathbb{R}^N$ is a compact set, there exists an absolutely continuous invariant probability measure μ with density $h \in V_\alpha$ which enjoys exponential decay of correlations against L^1 observables on V_α. More precisely, for all $\psi \in V_\alpha$, $\phi \in L^1(\mu)$ and $n \in \mathbb{N}^*$, we have

$$\left|\int_X \psi.\phi\circ T^n\, d\mu - \int_X \psi d\mu \int_X \phi d\mu\right| \le \|\psi\|_\alpha \|\phi\|_1 \theta_n,$$

with $\theta_n = a^n$ $(0 \le a < 1)$. This means that the system (X,T,μ) satisfies the condition (H1) with $\mathcal{C} = V_\alpha$.

It remains to show that the system also satisfies the conditions (H2) (with $r_0 = \epsilon_0$). To this end we need to estimate for each $p \in \{1, \cdots, k\}$, the norm $\|\psi_p\|_\alpha$, where the functions ψ_p were defined in (3) and (4). Since $\psi_p \in L^1(Leb)$ we just need to estimate its α-seminorm.

Since

$$\text{supp } osc(\psi_p, B_\epsilon(\cdot)) \subset B_\epsilon(X),$$

we infer that

$$|\psi_p|_\alpha = \sup_{0<\epsilon\le\epsilon_0} \epsilon^{-\alpha} \int_{B_\epsilon(X)} osc(\psi_p, B_\epsilon(x))dx.$$

For $p = 1$ the computation is similar to the one leading to (20) in [11] so we will only treat the case $p \geq 2$.

Let $0 < \epsilon \leq \epsilon_0$. First of all, suppose that $r \leq \epsilon$. Since the density h belongs to $V_\alpha \subset L^\infty$, we have $h \leq c$ for some constant $c > 0$. Thus, we observe that

$$\begin{aligned} osc(\psi_p, B_\epsilon(x)) &\leq \operatorname*{ess-sup}_{y\in B(x,\epsilon)\cap X} \psi_p(y) \leq \operatorname*{ess-sup}_{y\in B(x,\epsilon)\cap X} \int_{X^{p-1}} \left[\prod_{j=1}^{p-1} \mathbb{1}_{B(x_j,r)}(y)\right] d\mu^{p-1}(x_1,\dots,x_{p-1}) \\ &= \operatorname*{ess-sup}_{y\in B(x,\epsilon)\cap X} \mu(B(y,r))^{p-1} \leq C_0^{p-1} c^{p-1} \epsilon^{N(p-1)}, \end{aligned}$$

where C_0 denotes the Lebesgue measure of the unit ball in $\mathbb{R}^N$. Now, using the fact that $B_\epsilon(X) \subset B_{\epsilon_0}(X)$ which is a compact set, we conclude that

$$|\psi_p|_\alpha \leq \sup_{0<\epsilon\le\epsilon_0} \epsilon^{-\alpha} C_0^{p-1} c^{p-1} \epsilon^{N(p-1)} Leb(B_\epsilon(X)) \leq C\epsilon_0^{N(p-1)-\alpha},$$

where $C = C_0^{p-1} c^{p-1} Leb(X_{\epsilon_0})$.

For simplicity of notation, from now on we write $d\mu^{j-i+1}(i,j)$ instead of $d\mu^{j-i+1}(x_i,\dots,x_j)$.

Now suppose $r > \epsilon$. Observe that $y \in B(x,\epsilon)$ implies $B(x, r-\epsilon) \subset B(y,r) \subset B(x, r+\epsilon)$. Thus, if $x_p \in B(x,\epsilon)$, we infer that

$$\begin{aligned} &\prod_{l=p+1}^{k} \mathbb{1}_{B(x_p,r)}(x_l) \int_{X^{p-1}} \left[\prod_{j=1}^{p-1}\prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l)\right] d\mu^{p-1}(1,p-1) \\ &= \prod_{l=p+1}^{k} \mathbb{1}_{B(x_p,r)}(x_l) \int_{X^{p-1}} \left[\prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k} \mathbb{1}_{B(x_j,r)}(x_l)\right] \left[\prod_{j=1}^{p-1} \mathbb{1}_{B(x_p,r)}(x_j)\right] d\mu^{p-1}(1,p-1) \\ &\leq \prod_{l=p+1}^{k} \mathbb{1}_{B(x,r+\epsilon)}(x_l) \int_{X^{p-1}} \left[\prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k} \mathbb{1}_{B(x_j,r)}(x_l)\right] \left[\prod_{j=1}^{p-1} \mathbb{1}_{B(x,r+\epsilon)}(x_j)\right] d\mu^{p-1}(1,p-1). \end{aligned}$$

Then, we deduce that

$$\begin{aligned} &\operatorname*{ess-sup}_{y\in B(x,\epsilon)\cap X} \psi_p(y) \\ &\leq \prod_{l=p+1}^{k} \mathbb{1}_{B(x,r+\epsilon)}(x_l) \int_{X^{p-1}} \prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k} \mathbb{1}_{B(x_j,r)}(x_l) \prod_{j=1}^{p-1} \mathbb{1}_{B(x,r+\epsilon)}(x_j) d\mu^{p-1}(1,p-1). \end{aligned} \tag{7}$$

Using similar ideas, one can prove that

$$\begin{aligned}&\underset{y\in B(x,\epsilon)\cap X}{\text{ess-inf}}\,\psi_p(y) \qquad (8)\\ \geq\;& \prod_{l=p+1}^{k}\mathbb{1}_{B(x,r-\epsilon)}(x_l)\int_{X^{p-1}}\prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k}\mathbb{1}_{B(x_j,r)}(x_l)\prod_{j=1}^{p-1}\mathbb{1}_{B(x,r-\epsilon)}(x_j)d\mu^{p-1}(1,p-1).\end{aligned}$$

From (7) and (8) we find that

$$\begin{aligned}&\underset{y\in B(x,\epsilon)\cap X}{\text{ess-sup}}\,\psi_p(y)-\underset{\tilde{y}\in B(x,\epsilon)\cap X}{\text{ess-inf}}\,\psi_p(\tilde{y})\\ \leq\;& \prod_{l=p+1}^{k}\mathbb{1}_{B(x,r+\epsilon)}(x_l)\int_{X^{p-1}}\prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k}\mathbb{1}_{B(x_j,r)}(x_l)A_1^{p-1}(x)d\mu^{p-1}(1,p-1)\\ &+A_{p+1}^{k}(x)\int_{X^{p-1}}\prod_{j=1}^{p-1}\prod_{\substack{l=j+1\\ l\neq p}}^{k}\mathbb{1}_{B(x_j,r)}(x_l)\prod_{j=1}^{p-1}\mathbb{1}_{B(x,r-\epsilon)}(x_j)d\mu^{p-1}(1,p-1),\end{aligned}$$

where for $i<j$ we define $A_i^j(x)=\prod_{\theta=i}^{j}\mathbb{1}_{B(x,r+\epsilon)}(x_\theta)-\prod_{\theta=i}^{j}\mathbb{1}_{B(x,r-\epsilon)}(x_\theta)$.

Therefore

$$\int_{B_\epsilon(X)}osc(\psi_p,B_\epsilon(x))dx\leq\int_{B_\epsilon(X)}\Big(\int_{X^{p-1}}A_1^{p-1}(x)d\mu^{p-1}(1,p-1)+A_{p+1}^{k}(x)\Big)dx. \qquad (9)$$

Since μ is absolutely continuous we observe that

$$\begin{aligned}\int_{X^{p-1}}A_1^{p-1}(x)d\mu^{p-1}(1,p-1)&=\mu(B(x,r+\epsilon))^{p-1}-\mu(B(x,r-\epsilon))^{p-1}\\ &\leq(p-1)\left(\mu(B(x,r+\epsilon))-\mu(B(x,r-\epsilon)\right)\leq c(p-1)Leb(D(x)), \qquad (10)\end{aligned}$$

where $D(x)=B(x,r+\epsilon)\setminus B(x,r-\epsilon)$. For the second term in (9), we have

$$\begin{aligned}\int_{B_\epsilon(X)}A_{p+1}^{k}(x)dx&\leq Leb\left(\bigcap_{l=p+1}^{k}B(x_l,r+\epsilon)\setminus\bigcap_{l=p+1}^{k}B(x_l,r-\epsilon)\right)\\ &\leq\sum_{l=p+1}^{k}Leb\left(B(x_l,r+\epsilon)\setminus B(x_l,r-\epsilon)\right)=\sum_{l=p+1}^{k}Leb(D(x_l)). \qquad (11)\end{aligned}$$

One can see that for any $y\in\mathbb{R}^N$

$$Leb(D(y))\leq 2C_0\epsilon\sum_{k=0}^{N-1}\binom{N}{k}\leq 2^{N+1}C_0\epsilon. \qquad (12)$$

Finally, from (9) - (12) we deduce that

$$|\psi|_\alpha\;\leq\;\sup_{0<\epsilon\leq\epsilon_0}\epsilon^{-\alpha}\left(c(p-1)2^{N+1}C_0\epsilon Leb(B_\epsilon(X))+(k-p)2^{N+1}C_0\epsilon\right)\leq C_1\epsilon_0^{1-\alpha}, \qquad (13)$$

where $C_1=c(p-1)2^{N+1}C_0Leb(X_{\epsilon_0})+(k-p)2^{N+1}C_0$.

Thus, from (5) and (13), we obtain that (H2) is satisfied. Moreover, one can show easily that $D_k(\mu)=N$ and the proposition is proved. □

6. Proof of the symbolic case

In this section we prove the symbolic case (Theorem 4.1). We emphasize that even if the proof is based on the ideas of Theorem 7 in [11], the generalisation is not immediate and some extra care is needed. In particular, one needs to choose carefully between several different (and equivalent) definitions for S_n (see (14) and (25)) so the proof goes smoothly when using the mixing assumptions. We will focus on these extensions rather than the technical details that are similar to the ideas in [11].

We also observe that next section is dedicated to Theorems 2.2 and 2.6 whose proofs follow the lines of the proof of Theorem 4.1 but are more complex and technical. Thus, this section can be seen as a warm-up to Section 7.

We will assume that the system is α-mixing with an exponential decay, the ψ-mixing case can be easily deduced using the same ideas.

Proof of Theorem 4.1-(5). First, for $\varepsilon > 0$ and $k_n > 0$ let us define

$$S_n(x^1, \dots, x^k) = \sum_{i_1,\dots,i_k=0}^{n-1} \prod_{l=2}^{k} \mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l), \tag{14}$$

and observe that

$$S_n(x^1, \dots, x^k) \geq 1 \Longleftrightarrow M_n(x^1, \dots, x^k) \geq k_n. \tag{15}$$

Next we compute the expectation of S_n. Since $\mathbb{P}$ is a σ-invariant probability measure we infer that

$$\begin{aligned}\mathbb{E}(S_n) &= \sum_{i_1,\dots,i_k=0}^{n-1} \int_\Omega \left[\prod_{l=2}^{k} \int_\Omega \mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l) d\mathbb{P}(x^l)\right] d\mathbb{P}(x_1) \\ &= \sum_{i_1,\dots,i_k=0}^{n-1} \int_\Omega \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{k-1} d\mathbb{P}(x_1) = n^k \int_\Omega \mathbb{P}\left(C_{k_n}(x^1)\right)^{k-1} d\mathbb{P}(x_1).\end{aligned}$$

Using the partition of k_n*-cylinders*, we infer that

$$\mathbb{E}(S_n) = n^k \sum_{C_{k_n}} \int_{C_{k_n}} \mathbb{P}\left(C_{k_n} \cap C_{k_n}(x^1)\right)^{k-1} d\mathbb{P}(x^1) = n^k \sum_{C_{k_n}} \mathbb{P}\left(C_{k_n}\right)^k. \tag{16}$$

Now we are ready to prove (5).

Define $k_n = \frac{1}{(k-1)\underline{H}_k - \varepsilon}(k \log n + \log\log n)$. From (15), (16) and Markov's inequality, we find that

$$\mathbb{P}^k\left(M_n(x^1, \dots, x^k) \geq k_n\right) \leq \mathbb{E}(S_n) = n^k \sum_{C_{k_n}} \mathbb{P}\left(C_{k_n}\right)^k.$$

By the definition of the lower entropy and the definition of k_n, for n large enough, we have

$$\mathbb{P}^k\left(M_n(x^1, \dots, x^k) \geq k_n\right) \leq \frac{1}{\log n}.$$

Finally, choosing a subsequence $n_\ell = \lceil e^{\ell^2} \rceil$, we know that

$$\mathbb{P}^k\left(M_n(x^1, \dots, x^k) \geq k_{n_\ell}\right) \leq \frac{1}{\log n_\ell} \leq \frac{1}{\ell^2}.$$

Thus $\sum_\ell \mathbb{P}^k\left(M_n(x^1,\dots,x^k)\geq k_{n_\ell}\right)<+\infty$. By Borel-Cantelli Lemma, we know that

$$\frac{M_{n_\ell}(x^1,\dots,x^k)}{\log n_\ell}\leq\frac{1}{(k-1)\underline{H}_k-\varepsilon}\left(k+\frac{\log\log n_\ell}{\log n_\ell}\right),$$

for $\mathbb{P}^k$-almost every $(x^1,\dots,x^k)\in\Omega^k$ and ℓ large enough. Since $\varepsilon>0$ can be chosen arbitrarily small we obtain that

$$\overline{\lim_{\ell\to+\infty}}\frac{M_{n_\ell}(x^1,\dots,x^k)}{\log n_\ell}\leq\frac{k}{(k-1)\underline{H}_k}.$$

Observing that $(n_\ell)_\ell$ is increasing, $(M_n)_n$ is increasing and $\lim\limits_{\ell\to+\infty}\frac{\log n_\ell}{\log n_{\ell+1}}=1$, we infer that the last inequality holds if we replace n_l by n, and (5) is proved. □

Proof of Theorem 4.1-(6). Let $b<0$ to be choosen later. To prove (6), we set

$$k_n=\frac{1}{(k-1)\overline{H}_k+\varepsilon}(k\log n+b\log\log n).$$

From (15) and Chebychev's inequality we infer that

$$P^k\Big(M_n(x^1,\dots,x^k)<k_n\Big)=\mathbb{P}^k\Big(S_n(x^1,\dots,x^k)=0\Big)\leq\frac{\operatorname{var}(S_n)}{\mathbb{E}(S_n)^2}. \tag{17}$$

In order to bound $\frac{\operatorname{var}(S_n)}{\mathbb{E}(S_n)^2}=\frac{\mathbb{E}(S_n^2)-\mathbb{E}(S_n)^2}{\mathbb{E}(S_n)^2}$ we need to analyse the term

$$\mathbb{E}(S_n^2)=\sum_{\substack{i_1,\dots,i_k=0,\dots,n-1\\ i'_1,\dots,i'_k=0,\dots,n-1}}\int_{\Omega^k}\prod_{l=2}^k\mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}^k(x^1,\dots,x^k). \tag{18}$$

We will split this sum into two cases depending on the relative position of i_l and i'_l.

Let $g=g(n)=\log(n^{2k+1})$. First of all, we observe that if $|i_l-i'_l|>g+k_n$ then the $\alpha-$mixing condition gives that

$$\int_\Omega\mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}(x^l)\leq\alpha(g)+\mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right). \tag{19}$$

If otherwise $|i_l-i'_l|\leq g+k_n$ then Hölder's inequality infers that

$$\int_\Omega\mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}(x^l)\leq\mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{1/2}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{1/2}. \tag{20}$$

Now suppose that $i_l - i'_l > g + k_n$ for every $l \in \{1, \ldots, k\}$ (the case $i'_l - i_l > g + k_n$ can be treated identically). From (19) we find that

$$\begin{aligned}
&\int_{\Omega^k} \prod_{l=2}^{k} \mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}^k(x^1,\ldots,x^k) \qquad (21)\\
&= \int_\Omega \left[\prod_{l=2}^{k}\int_\Omega \mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}(x^l)\right]d\mathbb{P}(x_1)\\
&\leq \int_\Omega \left(\alpha(g) + \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)\right)^{k-1} d\mathbb{P}(x^1)\\
&\leq (2^{k-1}-1)\alpha(g) + \int_\Omega \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{k-1}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{k-1} d\mathbb{P}(x^1).
\end{aligned}$$

To conclude the first case we use the partition $\{C_{k_n} \cap \sigma^{-(i_1-i'_1)}C'_{k_n}\}_{C_{k_n},C'_{k_n}}$ of Ω to infer that

$$\begin{aligned}
&\int_\Omega \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{k-1}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{k-1} d\mathbb{P}(x^1) \qquad (22)\\
&= \sum_{C_{k_n},C'_{k_n}} \int_{C_{k_n}\cap\sigma^{-(i_1-i'_1)}C'_{k_n}} \mathbb{P}\left(C_{k_n}(\sigma^{i_1-i'_1}x^1)\right)^{k-1}\mathbb{P}\left(C_{k_n}(x^1)\right)^{k-1} d\mathbb{P}(x^1)\\
&= \sum_{C_{k_n},C'_{k_n}} \mathbb{P}(C_{k_n}\cap\sigma^{-(i_1-i'_1)}C'_{k_n})\mathbb{P}\left(C_{k_n}\right)^{k-1}\mathbb{P}\left(C'_{k_n}\right)^{k-1} \leq \alpha(g) + \left(\sum_{C_{k_n}}\mathbb{P}\left(C_{k_n}\right)^k\right)^2.
\end{aligned}$$

Next, for $p \in \{1, \ldots, k\}$, assume that we have p pairs of close indices and $k-p$ pairs of distant indices. We will firstly treat the case where $|i_1 - i'_1| \leq g + k_n$. Without loss of generality we can assume $|i_2 - i'_2| \leq g + k_n, \ldots, |i_p - i'_p| \leq g + k_n$, $|i_{p+1} - i'_{p+1}| > g + k_n, \ldots, |i_k - i'_k| > g + k_n$. From (19) and (20) we deduce that

$$\begin{aligned}
&\int_{\Omega^k} \prod_{l=2}^{k} \mathbb{1}_{C_{k_n}(\sigma^{i_1}x^1)}(\sigma^{i_l}x^l)\mathbb{1}_{C_{k_n}(\sigma^{i'_1}x^1)}(\sigma^{i'_l}x^l)d\mathbb{P}^k(x^1,\ldots,x^k) \qquad (23)\\
&\leq \int_\Omega \left(\alpha(g) + \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)\right)^{k-p} \times\\
&\qquad\qquad \times \left(\mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{1/2}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{1/2}\right)^{p-1} d\mathbb{P}(x^1)\\
&\leq (2^{k-p}-1)\alpha(g) + \int_\Omega \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{k-(p+1)/2}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{k-(p+1)/2} d\mathbb{P}(x^1).
\end{aligned}$$

Using Hölder's inequality and the invariance of $\mathbb{P}$, we obtain

$$\begin{aligned}
&\int_\Omega \mathbb{P}\left(C_{k_n}(\sigma^{i_1}x^1)\right)^{k-(p+1)/2}\mathbb{P}\left(C_{k_n}(\sigma^{i'_1}x^1)\right)^{k-(p+1)/2} d\mathbb{P}(x^1)\\
&\leq \int_\Omega \mathbb{P}\left(C_{k_n}(x^1)\right)^{2k-(p+1)} d\mathbb{P}(x^1) = \sum_{C_{k_n}}\mathbb{P}\left(C_{k_n}\right)^{2k-p} \leq \left(\sum_{C_{k_n}}\mathbb{P}\left(C_{k_n}\right)^k\right)^{(2k-p)/k}, \qquad (24)
\end{aligned}$$

where the last inequality came from the fact that $x \mapsto x^{k/(p+k)}$ is a countably subadditive function.

If $|i_1 - i'_1| > g + k_n$ then since we have $p \geq 1$ pairs of close indices, there exists at least one $j \in \{2, \ldots, k\}$ such that $|i_j - i'_j| \leq g + k_n$. In this case, the estimations (23) and (24) could be done similarly using the following equivalent definition of S_n

$$S_n(x^1, \ldots, x^k) = \sum_{i_1, \ldots, i_k = 0}^{n-1} \prod_{\substack{l=1 \\ l \neq j}}^{k} \mathbb{1}_{C_{k_n}(\sigma^{i_j} x^j)}(\sigma^{i_l} x^l). \tag{25}$$

Now, gathering the estimates (18) and (21)- (24) we conclude that

$$\begin{aligned}
\operatorname{var}(S_n) &\leq n^{2k} 3^k \alpha(g) + \sum_{p=1}^{k} \left[\binom{k}{p} n^{2k-p} (g + k_n)^p \left(\sum_{C_{k_n}} \mathbb{P}\left(C_{k_n}\right)^k \right)^{(2k-p)/k} \right] \\
&= n^{2k} 3^k \alpha(g) + \sum_{p=1}^{k} \left[\binom{k}{p} (g + k_n)^p \left(\mathbb{E}(S_n)\right)^{(2k-p)/k} \right]. \qquad (26)
\end{aligned}$$

Thus, (26) together with (17) gives us

$$\mathbb{P}^k \left(M_n(x^1, \ldots, x^k) < k_n \right) \leq \frac{n^{2k} 3^k \alpha(g)}{\mathbb{E}(S_n)^2} + \sum_{p=1}^{k} \frac{\binom{k}{p} (g + k_n)^p}{(\mathbb{E}(S_n))^{p/k}}.$$

By the definitions of k_n and (16), we observe that for n large enough we have $\mathbb{E}(S_n) \geq (\log n)^{-b}$, and since $g = \log\left(n^{2k+1}\right)$, we infer that

$$\frac{n^{2k} 3^k \alpha(g)}{\mathbb{E}(S_n)^2} = \mathcal{O}\left(\frac{1}{\log n} \right).$$

We can choose $b \ll -1$ so that

$$\mathbb{P}^k \left(M_n(x^1, \ldots, x^k) < k_n \right) = \mathcal{O}\left(\frac{1}{\log n} \right).$$

To conclude the proof it suffices to take a subsequence n_ℓ and use Borel-Cantelli Lemma as in the proof of (5).

□

7. Proofs of the main results

In this section we adapt the proof of Theorem 4.1 for multiple orbits (Theorems 2.2 and 2.6). In order to do that, one must replace M_n by $-\log m_n$ and the cylinders $C_k(x)$ by balls $B(x, e^{-k})$. However, one major drawback is that for cylinders we have that $x \in C_n(y)$ implies that $C_n(y) = C_n(x)$ but, when working with balls, $x \in B(y, r)$ does not imply that $B(y, r) = B(x, r)$. This simple fact prohibits us to define S_n as in the previous section, in particular in view of (25). To overcome this problem we will need to define S_n as

$$S_n(x_1, \ldots, x_k) = \sum_{i_1, \ldots, i_k = 0}^{n-1} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j, r_n)}(T^{i_l} x_l) \tag{27}$$

which will complexify our proofs. In particular, we will need to use the following lemma in the proof of both theorems.

Lemma 7.1.
$$(k-1)\underline{D}_k(\mu) = \varliminf_{r\to 0} \frac{\log \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) d\mu^k(1,k)}{\log r}$$
and
$$(k-1)\overline{D}_k(\mu) = \varlimsup_{r\to 0} \frac{\log \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) d\mu^k(1,k)}{\log r}$$

Proof. First of all, one can observe that for every $(x_1, \dots, x_k) \in X^k$
$$\prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) \leq \prod_{l=2}^{k} \mathbb{1}_{B(x_1,r)}(x_l).$$
Thus,
$$\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) d\mu^k(1,k) \leq \int_X \mu\left(B\left(x,r\right)\right)^{k-1} d\mu(x). \tag{28}$$
Moreover, one can observe that if $\{x_i, x_j\} \subset B(x_1, r/2)$ then $x_i \in B(x_j, r)$. Therefore
$$\prod_{l=2}^{k} \mathbb{1}_{B(x_1,r/2)}(x_l) \leq \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l)$$
for every $(x_1, \dots, x_k) \in X^k$, which implies that
$$\int_X \mu\left(B\left(x, \frac{r}{2}\right)\right)^{k-1} d\mu(x) \leq \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r)}(x_l) d\mu^k(1,k). \tag{29}$$
Using (28) and (29) and the fact that $\lim_{r\to 0} \frac{\log(r/2)}{\log r} = 1$ we get the result. □

Proof of Theorem 2.2. As in the proof of Theorem 4.1-(6) it suffices to show that
$$\mu^k\left(m_n(x_1, \dots, x_k) < r_n\right) = \mathcal{O}\left(\frac{1}{\log n}\right).$$

For $\varepsilon > 0$, let us define
$$k_n = \frac{1}{(k-1)\underline{D}_k(\mu) - \varepsilon}(k \log n + \log\log n) \quad \text{and} \quad r_n = e^{-k_n}.$$
Defining $S_n(x_1, \dots, x_k)$ as in (27), it is easy to see that for every $(x_1, \dots, x_k) \in X^k$
$$m_n(x_1, \dots, x_k) < r_n \iff S_n(x_1, \dots, x_k) \geq 1, \tag{30}$$
where m_n was defined in (1). Then, from (30) and Markov's inequality, we deduce that
$$\begin{aligned} \mu^k\left(m_n(x_1, \dots, x_k) < r_n\right) &\leq \mathbb{E}(S_n) = \\ &= \sum_{i_1,\dots,i_k=0}^{n-1} \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^k(1,k) \\ &\leq n^k \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^k(1,k), \end{aligned}$$

since μ is invariant.

By Lemma 7.1 and the definition of k_n, for n large enough, we infer that

$$\mu^k\left(m_n(x_1,\dots,x_k)<r_n\right)\le n^k r_n^{(k-1)\underline{D}_k(\mu)-\varepsilon}=\frac{1}{\log n},$$

and this is the desired conclusion. □

Before proving Theorem 2.6 we state a few facts in order to simplify the calculations. At first let us recall the notion of (λ,r)-grid partition.

Definition 7.2. *Let $0<\lambda<1$ and $r>0$. A partition $\{Q_i\}_{i=1}^{\infty}$ of X is called a (λ,r)-grid partition if there exists a sequence $\{y_i\}_{i=1}^{\infty}$ such that for any $i\in\mathbb{N}$*

$$B(y_i,\lambda r)\subset Q_i\subset B(y_i,r).$$

The following technical lemma will be used during the proof. One can observe that in the symbolic case, this lemma corresponds to (24). Moreover, this lemma is a generalization of Lemma 14 in [11].

Lemma 7.3. *Let $p\in\{1,\dots,k-1\}$. Under the hypotheses of Theorem 2.6, there exists a constant $K>0$ such that for n large enough*

$$\int_{X^{k-p}}\prod_{j=p+1}^{k-1}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j,r_n)}(x_l)\left(\int_{X^p}\prod_{j=1}^{p}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j,r_n)}(x_l)d\mu^p(1,p)\right)^2 d\mu^{k-p}(p+1,k)$$
$$\le K\left(\int_{X^k}\prod_{j=1}^{k-1}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j,r_n)}(x_l)d\mu^k(1,k)\right)^{(p+k)/k}=K\left(\frac{\mathbb{E}(S_n)}{n^k}\right)^{\frac{p+k}{k}},$$

where $d\mu^{j-i+1}(i,j)$ denotes $d\mu^{j-i+1}(x_i,\dots,x_j)$, for $i<j$.

Proof. By Proposition 2.1 in [28], there exist $0<\lambda<\frac{1}{2}$ and $R>0$ such that for any $0<r<R$ there exists a (λ,r)-grid partition.

Given r_0 as in definition 2.3 let us choose n large enough so that $r_n<\min\{R,r_0/2\}$. Let $\{Q_i\}_{i=1}^{\infty}$ be a $(\lambda,\frac{r_n}{2})$-grid partition and $\{y_i\}_{i=1}^{\infty}$ be such that

$$B\left(y_i,\lambda\frac{r_n}{2}\right)\subset Q_i\subset B\left(y_i,\frac{r_n}{2}\right).$$

Using this partition we infer that

$$\int_{X^{k-p}}\prod_{j=p+1}^{k-1}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j,r_n)}(x_l)\left(\int_{X^p}\prod_{j=1}^{p}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j,r_n)}(x_l)d\mu^p(1,p)\right)^2 d\mu^{k-p}(p+1,k)$$
$$\le\int_{X^{k-p}}\prod_{l=p+2}^{k}\mathbb{1}_{B(x_{p+1},r_n)}(x_l)\left(\int_{X^p}\prod_{j=1}^{p}\mathbb{1}_{B(x_j,r_n)}(x_{p+1})d\mu^p(1,p)\right)^2 d\mu^{k-p}(p+1,k)$$
$$=\int_X\mu\left(B(x_{p+1},r_n)\right)^{p+k-1}d\mu(x_{p+1})=\sum_i\int_{Q_i}\mu\left(B(x_{p+1},r_n)\right)^{p+k-1}d\mu(x_{p+1}).\qquad(31)$$

Now, for i fixed, there exist k_i elements $\{Q_{i,j}\}_{j=1}^{k_i}$ of the partition $\{Q_k\}_{k=1}^{\infty}$ such that $Q_{i,j}\cap B(y_i,2r_n)\neq\emptyset$ for $j=1,...,k_i$. Since the space is tight, there exists a constant K_0

depending only on N_0 such that $k_i \le K_0$ (see e.g. the proof of Theorem 4.1 in [28]). Defining $Q_{i,j} = \emptyset$ for $k_i < j \le K_0$ we infer that

$$\bigcup_{x_{p+1}\in Q_i} B(x_{p+1}, r_n) \subset B(y_i, 2r_n) \subset \bigcup_{j=1}^{K_0} Q_{i,j}. \tag{32}$$

From (32) we know that

$$\begin{aligned}\sum_i \int_{Q_i} \mu\left(B(x_{p+1}, r_n)\right)^{p+k-1} d\mu(x_{p+1}) &\le \sum_i \int_{Q_i} \left(\sum_{j=1}^{K_0} \mu\left(Q_{i,j}\right)\right)^{p+k-1} d\mu(x_{p+1}) \\ = \sum_i \mu(Q_i) \left(\sum_{j=1}^{K_0} \mu\left(Q_{i,j}\right)\right)^{p+k-1} &\le \sum_i \left(\sum_{j=1}^{K_0} \mu\left(Q_{i,j}\right)\right)^{p+k} \le K_0^{p+k-1} \sum_i \sum_{j=1}^{K_0} \mu\left(Q_{i,j}\right)^{p+k},\end{aligned}$$

where the last inequality is deduced from Jensen's inequality. Now, since the elements $Q_{i,j}$ cannot participate in more than K_0 different sums (one can see the arguments leading to (12) in [28]) and since $x \mapsto x^{k/(p+k)}$ is a countably subadditive function, we infer that

$$\begin{aligned}&\sum_i \int_{Q_i} \mu\left(B(x_{p+1}, r_n)\right)^{p+k-1} d\mu(x_{p+1}) \le K_0^{p+k} \sum_i \mu\left(Q_i\right)^{p+k} \\ &\le K_0^{p+k} \left(\sum_i \mu\left(Q_i\right)\right)^k\right)^{(p+k)/k} = K_0^{p+k} \left(\sum_i \int_{X^k} \prod_{l=1}^{k} \mathbb{1}_{Q_i}(x_l) d\mu^k(1,k)\right)^{(p+k)/k}. \end{aligned} \tag{33}$$

Note that for any $y \in Q_i$, we have $Q_i \subset B(y, r_n)$. Thus, if $\{x_1, \cdots, x_k\} \subset Q_i$, then we have $x_l \in B(x_j, r_n)$ for any $j, l = 1, \ldots, k$ and we conclude that

$$\sum_i \int_{X^k} \prod_{l=1}^{k} \mathbb{1}_{Q_i}(x_l) d\mu^k(1,k) \le \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^k(1,k). \tag{34}$$

Finally, (31), (33) and (34) give us

$$\begin{aligned}\int_{X^{k-p}} \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) \left(\int_{X^p} \prod_{j=1}^{p} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^p(1,p)\right)^2 d\mu^{k-p}(p+1,k) \\ \le K_0^{p+k} \left(\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^k(1,k)\right)^{(p+k)/k}.\end{aligned}$$

and the result follows with $K = K_0^{p+k}$. □

We are now ready to prove Theorem 2.6.

Proof of Theorem 2.6. Without loss of generality, we will assume in the proof that $\theta_n = e^{-n}$.
For $\varepsilon > 0$, let us define

$$k_n = \frac{1}{(k-1)\overline{D}_k(\mu) + \varepsilon}(k \log n + b \log\log n) \quad \text{and} \quad r_n = e^{-k_n}.$$

Using the same notation as in the proof of Theorem 2.2, we recall that

$$\mathbb{E}(S_n) = n^k \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j,r_n)}(x_l) d\mu^k(1,k). \tag{35}$$

To simplify our equations, from now on, we will denote by $B(x_j)$ the set $B(x_j, r_n)$.

Using (30) and Chebyshev's inequality, we obtain

$$\mu^k \left(m_n(x_1,\dots,x_k) \geq r_n\right) \leq \mu^k \left(S_n(x_1,\dots,x_k) = 0\right) \leq \frac{\operatorname{var}(S_n)}{\mathbb{E}(S_n)^2}. \tag{36}$$

Thus, we need to control the variance of S_n. First of all, we have

$$\operatorname{var}(S_n) = \sum_{\substack{i_1,\dots,i_k=0,\dots,n-1\\ i'_1,\dots,i'_k=0,\dots,n-1}} \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l) \mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) d\mu^k(1,k) - \mathbb{E}(S_n)^2.$$

Let $g = g(n) = \log(n^\gamma)$ where $\gamma > 0$ will be defined later.

We will split the last sum depending on the relative position of i_l and i'_l. Without loss of generality we can always suppose $i_l > i'_l$.

We first consider the case $i_1 - i'_1 > g, \dots, i_k - i'_k > g$.

Since $i_1 - i'_1 > g$ then by (H1) and (H2),

$$\begin{aligned}
&\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l) \mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) d\mu^k(1,k) \qquad (37)\\
= &\int_{X^{k-1}} \prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l) \mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) \times \\
&\quad \times \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(T^{i_1-i'_1}x_1)}(T^{i_l}x_l) \mathbb{1}_{B(x_1)}(T^{i'_l}x_l) d\mu(x_1)\right] d\mu^{k-1}(2,k) \\
\leq &\int_{X^{k-1}} \prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l) \mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) \\
&\quad \times \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i_l}x_l) d\mu(x_1)\right] \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i'_l}x_l) d\mu(x_1)\right] d\mu^{k-1}(2,k) + c^2 r_n^{-2\xi}\theta_g \\
=: &\; I + c^2 r_n^{-2\xi}\theta_g.
\end{aligned}$$

Now we use that $i_2 - i_2' > g$ and the same ideas to find that

$$\begin{aligned}
I \quad = \quad & \int_{X^{k-2}} \prod_{j=3}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) \\
& \times \int_X \left[\prod_{l=3}^{k} \mathbb{1}_{B(T^{i_2-i'_2}x_2)}(T^{i_l}x_l) \int_X \mathbb{1}_{B(x_1)}(T^{i_2-i'_2}x_2) \prod_{l=3}^{k} \mathbb{1}_{B(x_1)}(T^{i_l}x_l) d\mu(x_1) \right] \qquad (38) \\
& \quad \times \left[\prod_{l=3}^{k} \mathbb{1}_{B(x_2)}(T^{i'_l}x_l) \int_X \mathbb{1}_{B(x_1)}(x_2) \prod_{l=3}^{k} \mathbb{1}_{B(x_1)}(T^{i'_l}x_l) d\mu(x_1) \right] d\mu(x_2) d\mu^{k-2}(3,k) \\
& \leq \int_{X^{k-2}} \prod_{j=3}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) \\
& \times \int_{X^2} \mathbb{1}_{B(x_1)}(x_2) \prod_{l=3}^{k} \mathbb{1}_{B(x_1)}(T^{i_l}x_l)\mathbb{1}_{B(x_2)}(T^{i_l}x_l) d\mu^2(x_1,x_2) \\
& \times \int_{X^2} \mathbb{1}_{B(x_1)}(x_2) \prod_{l=3}^{k} \mathbb{1}_{B(x_1)}(T^{i'_l}x_l)\mathbb{1}_{B(x_2)}(T^{i'_l}x_l) d\mu^2(x_1,x_2) d\mu^{k-2}((3,k)) + c^2 r_n^{-2\xi}\theta_g.
\end{aligned}$$

Applying this argument again we will have on the p-th step ($i_p - i_p' > g$)

$$\begin{aligned}
& \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) d\mu^k(1,k) \qquad (39) \\
& \leq \int_{X^{k-p}} \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) \\
& \quad \times \int_{X^p} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l}x_l) \right] d\mu^p(1,p) \\
& \quad \times \int_{X^p} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i'_l}x_l) \right] d\mu^p(1,p) d\mu^{k-p}(p+1,k) \\
& \quad + c^2 p r_n^{-2\xi}\theta_g =: II + c^2 p r_n^{-2\xi}\theta_g.
\end{aligned}$$

Therefore, when $i_1 - i_1' > g$, $i_2 - i_2' > g, \ldots,$ $i_k - i_k' > g$ we have

$$\begin{aligned}
& \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l) d\mu^k(1,k) \\
& \qquad \leq c^2 k r_n^{-2\xi}\theta_g + \left(\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) d\mu^k(1,k) \right)^2. \qquad (40)
\end{aligned}$$

Now, if $i_1 - i_1' > g$, $i_2 - i_2' \leq g, \ldots,$ $i_k - i_k' \leq g$, we first proceed as in (37) and then, to estimate the term I, we use Hölder's inequality to find that

$$\begin{aligned} I = & \int_{X^{k-1}} \prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l} x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i_l} x_l) d\mu(x_1) \right] \\ & \times \prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j'} x_j)}(T^{i_l'} x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i_l'} x_l) d\mu(x_1) \right] d\mu^{k-1}(2,k) \\ \leq & \left(\int_{X^{k-1}} \left(\prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l} x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i_l} x_l) d\mu(x_1) \right] \right)^2 d\mu^{k-1}(2,k) \right)^{1/2} \\ & \times \left(\int_{X^{k-1}} \left(\prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l'} x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(T^{i_l'} x_l) d\mu(x_1) \right] \right)^2 d\mu^{k-1}(2,k) \right)^{1/2} . \end{aligned}$$

Finally we use the invariance of μ to conclude that

$$\begin{aligned} I \leq & \int_{X^{k-1}} \left(\prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(x_l) d\mu(x_1) \right] \right)^2 d\mu^{k-1}(2,k) \\ = & \int_{X^{k-1}} \prod_{j=2}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) \left[\int_X \prod_{l=2}^{k} \mathbb{1}_{B(x_1)}(x_l) d\mu(x_1) \right]^2 d\mu^{k-1}(2,k). \end{aligned} \tag{41}$$

In the case $i_1 - i_1' > g, \ldots, i_p - i_p' > g$ and $i_{p+1} - i_{p+1}' \leq g, \ldots,$ $i_k - i_k' \leq g$ we proceed as in (37)- (39) and then we use Holder's inequality to infer that

$$\begin{aligned} II = & \int_{X^{k-p}} f(p+1,k) g(p+1,k) d\mu^{k-p}(p+1,k) \\ \leq & \left(\int_{X^{k-p}} f^2(p+1,k) d\mu^{k-p}(p+1,k) \right)^{1/2} \left(\int_{X^{k-p}} g^2(p+1,k) d\mu^{k-p}(p+1,k) \right)^{1/2}, \end{aligned}$$

where $f(p+1,k)$ denotes the function $f(x_{p+1}, \ldots, x_k)$ defined as

$$f(p+1,k) = \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l} x_l) \int_{X^p} \prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l} x_l) d\mu^p(1,p),$$

and analogously for $g(p+1,k) = g(x_{p+1}, \ldots, x_k)$

$$g(p+1,k) = \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j'} x_j)}(T^{i_l'} x_l) \int_{X^p} \prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l'} x_l) d\mu^p(1,p).$$

Then we use the invariance of μ to infer that

$$II \leq \int_{X^{k-p}} \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) \left(\int_{X^p} \prod_{j=1}^{p} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) d\mu^p(1,p) \right)^2 d\mu^{k-p}(p+1,k). \tag{42}$$

Finally we observe that if $i_1 - i'_1 \leq g, \ldots, i_k - i'_k \leq g$ then

$$\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_j}x_j)}(T^{i'_l}x_l)d\mu^k(1,k) \leq \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j}x_j)}(T^{i_l}x_l)d\mu^k(1,k)$$

$$= \int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l)d\mu^k(1,k) = n^{-k}\mathbb{E}(S_n). \tag{43}$$

One can notice that all the other cases can be treated by symmetry.

Thus from (36) and (40)- (43) we conclude that

$$\mu^k\left(m_n(x_1,\ldots,x_k) \geq r_n\right) \leq \frac{1}{\mathbb{E}(S_n)^2}\Bigg[n^{2k}c^2kr_n^{-2\xi}\theta_g + \sum_{p=1}^{k-1}\binom{k}{p}n^{2p+k-p}g^{k-p} \times$$

$$\times \int_{X^{k-p}} \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l)\left(\int_{X^p}\prod_{j=1}^{p}\prod_{l=j+1}^{k}\mathbb{1}_{B(x_j)}(x_l)d\mu^p\right)^2 d\mu^{k-p}(p+1,k)$$

$$+\sum_{p=1}^{k-1}\binom{k}{p}n^{2p+k-p}g^{k-p}c^2pr_n^{-2\xi}\theta_g + g^k\mathbb{E}(S_n)\Bigg].$$

Thus, by Lemma 7.3, we deduce

$$\mu^k\left(m_n(x_1,\ldots,x_k) \geq r_n\right) \leq \frac{1}{\mathbb{E}(S_n)^2}\Bigg[n^{2k}c^2kr_n^{-2\xi}\theta_g + \sum_{p=1}^{k-1}\binom{k}{p}n^{p+k}g^{k-p}c^2pr_n^{-2\xi}\theta_g$$

$$+\sum_{p=1}^{k-1}\binom{k}{p}n^{p+k}g^{k-p}\left(n^{-k}\mathbb{E}(S_n)\right)^{(p+k)/k} + g^k\mathbb{E}(S_n)\Bigg]$$

$$= \theta_g\frac{n^{2k}c^2kr_n^{-2\xi} + \sum_{p=1}^{k-1}\binom{k}{p}n^{p+k}g^{k-p}c^2pr_n^{-2\xi}}{\mathbb{E}(S_n)^2} + \sum_{p=1}^{k-1}\frac{\binom{k}{p}g^{k-p}}{\mathbb{E}(S_n)^{2-(p+k)/k}} + \frac{g^k}{\mathbb{E}(S_n)}.$$

By definitions of r_n, k_n, (35) and Lemma 7.1, we observe that for n large enough we have

$$\mathbb{E}(S_n) \geq (\log n)^{-b}.$$

Since $g = \log(n^\gamma)$ we have for γ large enough that

$$\theta_g\frac{n^{2k}c^2kr_n^{-2\xi} + \sum_{p=1}^{k-1}\binom{k}{p}n^{p+k}g^{k-p}c^2pr_n^{-2\xi}}{\mathbb{E}(S_n)^2} = \mathcal{O}\left(\frac{1}{\log n}\right).$$

Then, we can choose $b \ll -1$ such that

$$\sum_{p=1}^{k-1}\frac{\binom{k}{p}g^{k-p}}{\mathbb{E}(S_n)^{2-(p+k)/k}} + \frac{g^k}{\mathbb{E}(S_n)} \leq \sum_{p=1}^{k-1}\frac{\binom{k}{p}g^{k-p}}{(\log n)^{-b(k-p)/k}} + \frac{g^k}{(\log n)^{-b}} = \mathcal{O}\left(\frac{1}{\log n}\right),$$

and we have

$$\mu^k\left(m_n(x_1,\ldots,x_k) \geq r_n\right) = \mathcal{O}\left(\frac{1}{\log n}\right). \tag{44}$$

To conclude the proof it suffices to take a subsequence n_ℓ and use Borel-Cantelli Lemma as in the proof of (5). □

In order to simplify the proof of Theorem 2.7, we state and prove the following technical lemma:

Lemma 7.4. *Let φ be given by*

$$\varphi(x_p) = \int_{X^{p-1}} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(x_j)}(x_l) \right] d\mu^{p-1}(1, p-1).$$

and suppose that (HA) is satisfied. Then there exist $0 < r_0 < 1$, $c > 0$ and $\zeta \geq 0$ such that for every $p \in \{2, \ldots, k\}$, for μ^{k-p}-almost every $x_{p+1}, \ldots, x_k \in X$ and for any $0 < r < r_0$, the function φ belongs to $\mathcal{H}^\alpha(X, \mathbb{R})$ and

$$||\varphi||_{\mathcal{H}^\alpha} \leq c r^{-\zeta}.$$

Proof. Let $0 < r < r_0$ and $x, y \in X$. We have

$$\begin{aligned}
|\varphi(x) - \varphi(y)| &= \left| \int_{X^{p-1}} \prod_{j=1}^{p-1} \prod_{\substack{l=j+1 \\ l \neq p}}^{k} \mathbb{1}_{B(x_j)}(x_l) \left[\prod_{j=1}^{p-1} \mathbb{1}_{B(x_j)}(x) - \prod_{j=1}^{p-1} \mathbb{1}_{B(x_j)}(y) \right] d\mu^{p-1}(1, p-1) \right| \\
&\leq \int_{X^{p-1}} \left| \prod_{j=1}^{p-1} \mathbb{1}_{B(x)}(x_j) - \prod_{j=1}^{p-1} \mathbb{1}_{B(y)}(x_j) \right| d\mu^{p-1}(1, p-1) \\
&\leq \sum_{l=0}^{p-2} \int_{X^{p-1}} \left| \prod_{j=1}^{l} \mathbb{1}_{B(y)}(x_j) \prod_{j=l+1}^{p-1} \mathbb{1}_{B(x)}(x_j) - \prod_{j=1}^{l+1} \mathbb{1}_{B(y)}(x_j) \prod_{j=l+2}^{p-1} \mathbb{1}_{B(x)}(x_j) \right| d\mu^{p-1}(1, p-1) \\
&\leq \sum_{l=0}^{p-2} \int_{X} \left| \mathbb{1}_{B(x)}(x_{l+1}) - \mathbb{1}_{B(y)}(x_{l+1}) \right| d\mu(x_{l+1}) = (p-1) \int_X \left| \mathbb{1}_{B(x)}(z) - \mathbb{1}_{B(y)}(z) \right| d\mu(z).
\end{aligned}$$

If $d(x, y) \geq r$ then

$$|\varphi(x) - \varphi(y)| \leq 2(p-1) \leq \frac{2(p-1)}{r} d(x, y). \tag{45}$$

If otherwise $d(x, y) < r$ then from (HA) we conclude that

$$\begin{aligned}
|\varphi(x) - \varphi(y)| \quad &\leq \quad (p-1) \int_X \left| \mathbb{1}_{B(x)}(z) - \mathbb{1}_{B(y)}(z) \right| d\mu(z) \qquad (46) \\
&\leq (p-1) \left[\mu\left(B(x,r) \backslash \left(B(x,r) \cap B(y,r)\right)\right) + \mu\left(B(y,r) \backslash \left(B(x,r) \cap B(y,r)\right)\right) \right] \\
&\leq (p-1) \mu\left(B(x, r + d(x,y)) \backslash B(x, r - d(x,y))\right) \leq (p-1) r^{-\xi} d(x,y)^\beta,
\end{aligned}$$

and the lemma follows from inequalities (45) and (46).

□

Proof of Theorem 2.7. When our Banach space $\mathcal{C}$ is the space of Hölder functions, (H2) cannot be satisfied since characteristic functions are not continuous. Thus, we need to adapt the proof of Theorem 2.6 to this setting, approximating characteristic functions by Lipschitz functions, following the construction of the proof of Lemma 9 in [46]. We will only prove the key part here, which is obtaining the equivalent of our inequality (39).

To do so we fix $q \in \{1, \dots, p\}$ and consider the the following term:

$$
\begin{aligned}
I := \int_X \Big[\prod_{l=q+1}^{k} & \mathbb{1}_{B(T^{i_q}x_q)}(T^{i_l}x_l)\mathbb{1}_{B(T^{i'_q}x_q)}(T^{i'_l}x_l) \\
& \times \int_{X^{q-1}} \left[\prod_{j=1}^{q-2} \prod_{l=j+1}^{q-1} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{q-1} \prod_{l=q}^{k} \mathbb{1}_{B(x_j)}(T^{i_l}x_l) \right] d\mu^{q-1}(1, q-1) \\
& \times \int_{X^{q-1}} \left[\prod_{j=1}^{q-2} \prod_{l=j+1}^{q-1} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{q-1} \prod_{l=q}^{k} \mathbb{1}_{B(x_j)}(T^{i'_l}x_l) \right] d\mu^{q-1}(1, q-1) \Big] d\mu(x_q).
\end{aligned}
$$

We assume $|i_q - i'_q| > g$. Let $\rho > 0$ (to be choosen later). Let $\eta_{r_n} : [0, \infty) \to \mathbb{R}$ be the $\frac{1}{\rho r_n}$-Lipschitz function such that $\mathbb{1}_{[0,r_n]} \le \eta_{r_n} \le \mathbb{1}_{[0,(1+\rho)r_n]}$ and set

$$
\varphi_{x_{q+1},\dots,x_k,r_n}(x) = \prod_{l=q+1}^{k} \eta_{r_n}(d(x, x_l)). \tag{47}
$$

We observe that $\varphi_{x_{q+1},\dots,x_k,r_n}$ is $\frac{k-q}{\rho r_n}$-Lipschitz. Moreover, we have

$$
\prod_{l=q+1}^{k} \mathbb{1}_{B(x,r_n)}(T^{i_l}x_l) \le \varphi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(x) \le \prod_{l=q+1}^{k} \mathbb{1}_{B(x,(1+\rho)r_n)}(T^{i_l}x_l). \tag{48}
$$

Now we define the following auxiliary function

$$
\begin{aligned}
\Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r}(x_q) &= \varphi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r}(x_q) \times \\
&\times \int_{X^{q-1}} \left[\prod_{j=1}^{q-2} \prod_{l=j+1}^{q} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{q-1} \prod_{l=q+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l}x_l) \right] d\mu^{q-1}(1, q-1).
\end{aligned} \tag{49}
$$

From Lemma 7.4, we observe that for μ^{k-q+1}-almost every $x_{q+1}, \dots, x_k \in X$ and for any $0 < r < r_0$ the function Φ belongs to $\mathcal{H}^\alpha(X, \mathbb{R})$ and

$$
||\Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r}||_{\mathcal{H}^\alpha} \le cr^{-\zeta} + (k-q)(\rho r)^{-1} \le cr^{-\zeta} + k(\rho r)^{-1}.
$$

Using (H1), (48) and (49) we deduce that

$$
\begin{aligned}
I \le\ & \int_X \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(T^{i_q}x_q)\Phi_{T^{i'_{q+1}}x_{q+1},\dots,T^{i'_k}x_k,r_n}(T^{i'_q}x_q)d\mu(x_q) \\
\le\ & \int_X \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(x_q)d\mu(x_q) \int_X \Phi_{T^{i'_{q+1}}x_{q+1},\dots,T^{i'_k}x_k,r_n}(x_q)d\mu(x_q) \\
+\ & \theta_g \left\| \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n} \right\|_{\mathcal{H}^\alpha} \left\| \Phi_{T^{i'_{q+1}}x_{q+1},\dots,T^{i'_k}x_k,r_n} \right\|_{\mathcal{H}^\alpha} \\
\le\ & \int_X \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(x_q)d\mu(x_q) \int_X \Phi_{T^{i'_{q+1}}x_{q+1},\dots,T^{i'_k}x_k,r_n}(x_q)d\mu(x_q) \\
& + (cr_n^{-\zeta} + k(\rho r_n)^{-1})^2\theta_g.
\end{aligned} \tag{50}
$$

At this point we observe that the inequality (50) together with (48) will not be sufficient to obtain our equivalent of (39) since the radius of the balls will be $(1+\rho)r_n$ (instead of r_n).

To overcome this problem we use (HA). For simplicity, we use the following notation:

$$g(x_q)=\int_{X^{q-1}}\left[\prod_{j=1}^{q-2}\prod_{l=j+1}^{q}\mathbb{1}_{B(x_j)}(x_l)\right]\left[\prod_{j=1}^{q-1}\prod_{l=q+1}^{k}\mathbb{1}_{B(x_j)}(T^{i_l}x_l)\right]d\mu^{q-1}(1,q-1).$$

Thus from (48) we know that

$$\begin{aligned}
&\int_X \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(x_q)d\mu(x_q)=\int_X \varphi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r}(x_q)g(x_p)d\mu(x_q) \qquad (51)\\
&\le \int_X g(x_q)\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q,(1+\rho)r_n)}(T^{i_l}x_l)d\mu(x_q)=\int_X g(x_q)\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q,r_n)}(T^{i_l}x_l)d\mu(x_q)\\
&+\int_X g(x_q)\left(\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q,(1+\rho)r_n)}(T^{i_l}x_l)-\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q,r_n)}(T^{i_l}x_l)\right)d\mu(x_q)\\
&\le \int_X g(x_q)\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q,r_n)}(T^{i_l}x_l)d\mu(x_q)+\mu\left(\bigcap_{l=q+1}^{k}B(T^{i_l}x_l,(1+\rho)r_n)\setminus\bigcap_{l=q+1}^{k}B(T^{i_l}x_l,r_n)\right).
\end{aligned}$$

Using (HA) we conclude that

$$\begin{aligned}
&\mu\left(\bigcap_{l=q+1}^{k}B(T^{i_l}x_l,(1+\rho)r_n)\setminus\bigcap_{l=q+1}^{k}B(T^{i_l}x_l,r_n)\right)\le\mu\left(\bigcup_{l=q+1}^{k}B(T^{i_l}x_l,(1+\rho)r_n)\setminus B(T^{i_l}x_l,r_n)\right)\\
&\le\sum_{l=q+1}^{k}\mu\left(B(T^{i_l}x_l,(1+\rho)r_n)\setminus B(T^{i_l}x_l,r_n)\right)\le(k-q)r_n^{-\xi}\rho^{\beta}. \qquad (52)
\end{aligned}$$

Then from (51) and (52) and taking ρ small enough we infer that

$$\begin{aligned}
&\int_X \Phi_{T^{i_{q+1}}x_{q+1},\dots,T^{i_k}x_k,r_n}(x_q)d\mu(x_q)\int_X \Phi_{T^{i'_{q+1}}x_{q+1},\dots,T^{i'_k}x_k,r_n}(x_q)d\mu(x_q)\\
\le\ &\int_X\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q)}(T^{i_l}x_l)\int_{X^{q-1}}\prod_{j=1}^{q-2}\prod_{l=j+1}^{q}\mathbb{1}_{B(x_j)}(x_l)\prod_{j=1}^{q-1}\prod_{l=q+1}^{k}\mathbb{1}_{B(x_j)}(T^{i_l}x_l)d\mu^{q-1}(1,q-1)d\mu(x_q)\\
&\times\int_X\prod_{l=q+1}^{k}\mathbb{1}_{B(x_q)}(T^{i'_l}x_l)\int_{X^{q-1}}\prod_{j=1}^{q-2}\prod_{l=j+1}^{q}\mathbb{1}_{B(x_j)}(x_l)\prod_{j=1}^{q-1}\prod_{l=q+1}^{k}\mathbb{1}_{B(x_j)}(T^{i'_l}x_l)d\mu^{q-1}(1,q-1)d\mu(x_q)\\
&+3(k-q)r_n^{-\xi}\rho^{\beta}\\
=\ &\int_{X^q}\left[\prod_{j=1}^{q-2}\prod_{l=j+1}^{q}\mathbb{1}_{B(x_j)}(x_l)\right]\left[\prod_{j=1}^{q}\prod_{l=q+1}^{k}\mathbb{1}_{B(x_j)}(T^{i_l}x_l)\right]d\mu^q(1,q)\\
&\times\int_{X^q}\left[\prod_{j=1}^{q-2}\prod_{l=j+1}^{q}\mathbb{1}_{B(x_j)}(x_l)\right]\left[\prod_{j=1}^{q}\prod_{l=q+1}^{k}\mathbb{1}_{B(x_j)}(T^{i'_l}x_l)\right]d\mu^q(1,q)+3(k-q)r_n^{-\xi}\rho^{\beta}. \qquad (53)
\end{aligned}$$

Thus, since $q \leq k$, (53) together with (50) gives us

$$\begin{aligned} I \leq & \int_{X^q} \left[\prod_{j=1}^{q-2} \prod_{l=j+1}^{q} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{q} \prod_{l=q+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l} x_l) \right] d\mu^q(1,q) \\ & \times \int_{X^q} \left[\prod_{j=1}^{q-2} \prod_{l=j+1}^{q} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{q} \prod_{l=q+1}^{k} \mathbb{1}_{B(x_j)}(T^{i'_l} x_l) \right] d\mu^q(1,q) \\ & + 3kr_n^{-\xi}\rho^\beta + (cr_n^{-\zeta} + k(\rho r_n)^{-1})^2 \theta_g. \end{aligned}$$

Repeating this process for each $q \in \{1, \dots, p\}$ we obtain our equivalent of (39)

$$\int_{X^k} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l} x_l) \mathbb{1}_{B(T^{i'_j} x_j)}(T^{i'_l} x_l) d\mu^k(1,k)$$

$$\leq 3pkr_n^{-\xi}\rho^\beta + p(cr_n^{-\zeta} + k(\rho r_n)^{-1})^2 \theta_g + \int_{X^{k-p}} \prod_{j=p+1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{B(T^{i_j} x_j)}(T^{i_l} x_l) \mathbb{1}_{B(T^{i'_j} x_j)}(T^{i'_l} x_l)$$

$$\times \int_{X^p} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i_l} x_l) \right] d\mu^p(1,p)$$

$$\times \int_{X^p} \left[\prod_{j=1}^{p-1} \prod_{l=j+1}^{p} \mathbb{1}_{B(x_j)}(x_l) \right] \left[\prod_{j=1}^{p} \prod_{l=p+1}^{k} \mathbb{1}_{B(x_j)}(T^{i'_l} x_l) \right] d\mu^p(1,p) d\mu^{k-p}(p+1k).$$

Thus, the rest of the proof follows exactly as in Theorem 2.6 where at the end, one must choose $\rho = n^{-\delta}$ with δ large enough so that (44) holds.

□

Proof of Theorem 3.2. The proof follows the lines of the proof of Theorems 2.2 and 2.7, replacing S_n by

$$S_n^f(x_1, \dots, x_k) = \sum_{i_1, \dots, i_k = 0}^{n-1} \prod_{j=1}^{k-1} \prod_{l=j+1}^{k} \mathbb{1}_{f^{-1}B(f(T^{i_j} x_j), r_n)}(T^{i_l} x_l).$$

Another modification worth mentioning is that in (47), $\varphi_{x_{q+1}, \dots, x_k, r_n}(x)$ must be replaced by

$$\varphi^f_{x_{q+1}, \dots, x_k, r_n}(x) = \prod_{l=q+1}^{k} \eta_{r_n}(d(f(x), f(x_l))),$$

which is a $\frac{L(k-q)}{\rho r_n}$-Lipschitz function (if f is L-Lipschitz). □

Acknowledgements. The authors would like to thank Benoit Saussol for useful discussions. This work was partially supported by CNPq, by FCT project PTDC/MAT-PUR/28177/2017, with national funds, and by CMUP (UIDB/00144/2020), which is funded by FCT with national (MCTES) and European structural funds through the programs FEDER, under the partnership agreement PT2020.

References

[1] M. Abadi and L. Cardeno, *Renyi entropies and large deviations for the first-match function,* IEEE Trans. Inf. Theory(61), 4 (2015), 1629–1639.

[2] M. Abadi and R. Lambert, *From the divergence between two measures to the shortest path between two observables,* Ergod. Theory Dyn., **39** (2019), no. 7, 1729–1744.

[3] M. Alber, J. Peinke, *Improved multifractal box-counting algorithm, virtual phase transitions, and negative dimensions*, Phys. Rev. E **57** (1998) 5489–5493.

[4] J. F. Alves, J. M. Freitas, S. Luzzatto, and S. Vaienti, *From rates of mixing to recurrence times via large deviations*, Adv. Math., **228** (2011), no. 2, 1203–1236.

[5] R. Arratia, P. Morris and M. Waterman, *Stochastic scrabble: Large deviations for sequences with scores,* J. Appl. Prob. **25** no 1 (1988), 106-119.

[6] R. Arratia and M. Waterman, *An Erdös–Rényi law with shifts*, Adv. Math., **55** (1985), 13–23.

[7] R. Badii, G. Broggi, *Measurement of the dimension spectrum $f(\alpha)$: Fixed-mass approach*, Phys. Lett. A. **131** (1988) 339–343.

[8] R. Badii, A. Politi, *Statistical description of chaotic attractors - the dimension function*, J. Stat. Phys. **40** (1985) 725–750.

[9] V. Baladi, *"Positive Transfer Operators and Decay of Correlations,"* Advances Series in Nonlinear Dynamics, World Scientific Publishing Co. Inc., River Edge, NJ, **16**, (2000).

[10] J.-M. Barbaroux, F. Germinet and S. Tcheremchantsev, *Generalized fractal dimensions: equivalences and basic properties*, J. Math. Pures Appl. **80**, 10 (2001), 977–1012.

[11] V. Barros, L. Liao and J. Rousseau, *On the shortest distance between orbits and the longest common substring problem,* Adv. Math., **334** (2019), 311-339.

[12] D. Bessis, G. Paladin, G. Turchetti and S. Vaienti, *Generalized dimensions, entropies, and Liapunov exponents from the pressure function for strange sets*, J. Statist. Phys. **51** (1988), no. 1-2, 109–134.

[13] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Math, Vol. **470**, Springer, Berlin, 1975.

[14] R. Bradley, *Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions*, Probability Surveys Vol. **2** (2005) 107–144.

[15] T. Caby, D. Faranda, G. Mantica, S. Vaienti and P. Yiou, *Generalized dimensions, large deviations and the distribution of rare events*, Phys. D **400** (2019), 132143.

[16] P. Collet, *Statistics of closest return for some non-uniformly hyperbolic systems*, Ergodic Theory Dynam. Systems, **21** (2001), 401–420.

[17] A. Coutinho, R. Lambert and J. Rousseau, *Matching strings in encoded sequences*, Bernoulli **26** (2020), no. 3, 2021–2050.

[18] A. Coutinho, J. Rousseau, B. Saussol, *Large deviation for return times*, Nonlinearity **31** (11) (2018) 5162–5179.

[19] A.H. Fan, K.S. Lau and H. Rao, *Relationships between different dimensions of a measure*, Monatsh. Math., **135** (2002), no. 3, 191–201.

[20] S. Fan, L. Liao and Y. Qiu, *Stationary determinantal processes: ψ-mixing property and L^q-dimensions*, preprint, arXiv:1911.04718.

[21] D. Faranda, A.C. Moreira Freitas, J. Milhazes Freitas, M. Holland, T. Kuna, V. Lucarini, M. Nicol, M. Todd, S. Vaienti, *Extremes and Recurrence in Dynamical Systems*, Wiley, New York, 2016.

[22] D. Faranda and S. Vaienti, *Correlation dimension and phase space contraction via extreme value theory*, Chaos, **28** (2018), 041103.

[23] A. C. M. Freitas, J. M. Freitas and M. Todd, *Hitting time statistics and extreme value theory.* Probab. Theory Related Fields, **147** (2010), 675–710.

[24] A. C. M. Freitas, J. M. Freitas and M. Todd, *The extremal index, hitting time statistics and periodicity.* Adv. Math. **231** (2012), 2626–2665.

[25] P. Grassberger, *Generalized dimension of strange attractors*, Phys. Lett. A **97** (1983) 227–230.

[26] P. Grassberger and I. Procaccia, *Measuring the strangeness of strange attractors*, Physica D, **9** (1–2) (1983), 189–208.

[27] P. Grassberger and I. Procaccia, *Characterization of strange attractors*, Phys. Rev. Lett., **50** (1983), 346–349.

[28] M. Guysinsky and S. Yaskolko, *Coincidence of various dimensions associated with metrics and measures on metric spaces*, Discrete Contin. Dyn. Syst., **3** (1997) 591–603.
[29] N. Haydn, J. Luevano, G. Mantica, S. Vaienti, *Multifractal properties of return time statistics*, Phys. Rev. Lett. **88** (2003) 224502.
[30] N. Haydn and S. Vaienti, *The Rényi entropy function and the large deviation of short return times*, Ergodic Theory Dynam. Systems, **30** (2010), no. 1, 159–179.
[31] F. Hofbauer, G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z., **180** (1982), no. 1, 119–140.
[32] S. Joly, G. Le Calvé, *Three-way distances*, J. Classification **12** (1995), no. 2, 191–205.
[33] S. Karlin, F. Ost, *Counts of long aligned word matches among random letter sequences*, Adv. in Appl. Probab. **19** (1987), no. 2, 293–351.
[34] S. Karlin, F. Ost, *Maximal length of common words among random letter sequences*, Ann. Probab. **16** (1988) 535–563.
[35] J. Li and X. Yang, *On longest matching consecutive subsequence*, Int. J. Number Theory, **15** (2019), no. 8, 1745–1758.
[36] T. Luczak and W. Szpankowski, *A suboptimal lossy data compression based on approximate pattern matching*, IEEE Trans. Inform. Theory 43 (1997), 1439–1451.
[37] G. Mantica, *The global statistics of return times: Return time dimensions versus generalized measure dimensions*, J. Stat. Phys. **138** (2010) 701–727.
[38] W. Parry, *On the β-expansions of real numbers*, Acta Math. Acad. Sci. Hungar., **11** (1960) 401–416.
[39] R. Pastor-Satorras, R.H. Riedi, *Numerical estimates of the generalized dimensions of the Hénon attractor for negative q*, J. Phys. A: Math. Gen. **29** (1996) L391–L398.
[40] Y. Pesin, *Dimension Theory in Dynamical Systems*, University of Chicago Press, Chicago, 1997.
[41] Y. Pesin and H. Weiss, *A multifractal analysis of equilibrium measures for conformal expanding maps and Moran-like geometric constructions*, J. Stat. Phys., **86** (1997), 233–275.
[42] W. Philipp, *Some metric theorems in number theory*, Pacific J. Math., **20** (1967) 109–127.
[43] W. Philipp, *Some metrical theorems in number theory II*, Duke Math. J., **37** (1970) 447–458.
[44] J. Rousseau, *Longest common substring for random subshifts of finite type,* to appear in Ann. Inst. Henri Poincar Probab. Stat, arXiv: 1905.08131.
[45] J. Rousseau, *Hitting time statistics for observations of dynamical systems,* Nonlinearity **27** (2014), 2377–2392.
[46] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps,* Israel J. Math., **116** (2000), 223–248.
[47] B. Saussol, *An introduction to quantitative Poincaré recurrence in dynamical systems,* Rev. Math. Phys., **21** (2009) 949–979.
[48] P. Walters, *Ruelle's operator theorem and g-measures*, Trans. Am. Math. Soc., **214**, (1975), 375–387.
[49] M. J. Warrens, *n-way metrics*, J. Classification **27** (2010), no. 2, 173–190.
[50] M. Waterman, *Introduction to Computational Biology: Maps, Sequences and Genomes*,Chapman and Hall, London (1995).

Vanessa Barros, UniLaSalle - Ecole des Métiers de l'Environnement, Campus de Ker Lann, 12 Avenue Robert Schuman, 35170 Bruz, France

Jérôme Rousseau, Departamento de Matemática, Faculdade de Ciências da Universidade do Porto,, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Vanessa Barros and Jérôme Rousseau, Departamento de Matemática, Universidade Federal da Bahia, Av. Ademar de Barros s/n, 40170-110 Salvador, Brazil

E-mail address: vanessa.barros@unilasalle.fr

URL: https://sites.google.com/site/vbarrosoliveira/home

E-mail address: jerome.rousseau@ufba.br

URL: http://www.sd.mat.ufba.br/~jerome.rousseau