MINIMAL DISTANCE BETWEEN RANDOM ORBITS

SÉBASTIEN GOUËZEL, JÉRÔME ROUSSEAU, AND MANUEL STADLBAUER

ABSTRACT. We study the minimal distance between two orbit segments of length n, in a random dynamical system with sufficiently good mixing properties. This problem has already been solved in non-random dynamical system, and on average in random dynamical systems (the so-called annealed version of the problem): it is known that the asymptotic behavior for this question is given by a dimension-like quantity associated to the invariant measure, called its correlation dimension (or Rényi entropy). We study the analogous quenched question, and show that the asymptotic behavior is more involved: two correlation dimensions show up, giving rise to a non-smooth behavior of the associated asymptotic exponent.

1. INTRODUCTION

1.1. **Main results.** This article is devoted to the study of the minimal distance between pieces of orbits of length n, in a random dynamical system setting. By this, we mean the following standard setting. We start from an invertible, probability preserving dynamical system $(\Omega, \theta, \mathbb{P})$ on a compact metric space, and consider another metric space (X, d). For each $\omega \in \Omega$, let T_ω be a measurable map of X, such that the skew-product map $S : (\omega, x) \mapsto (\theta\omega, T_\omega x)$ is measurable and preserves a probability measure ν whose marginal on Ω is $\mathbb{P}$. The iterates of S are given by $S^n(\omega, x) = (\theta^n \omega, T_\omega^n x)$, where $T_\omega^n = T_{\theta^{n-1}\omega} \circ \cdots \circ T_\omega$ is a random composition of the T_ω's, where the randomness is dictated by the driving map θ. In this setting, the measure ν can be disintegrated above $\mathbb{P}$: there is a family of probability measures μ_ω, depending measurably on ω, such that for any bounded function f holds $\int f \, \mathrm{d}\nu = \int \left(\int f(x) \, \mathrm{d}\mu_\omega(x)\right) \mathrm{d}\mathbb{P}(\omega)$. We write informally $\nu = \mathbb{P} \otimes \mu_\omega$. Let $\mu = \int \mu_\omega \, \mathrm{d}\mathbb{P}(\omega)$ be the second marginal of ν. As ν is invariant under S, the measures μ_ω also satisfy an invariance property: $(T_\omega)_* \mu_\omega = \mu_{\theta\omega}$ for $\mathbb{P}$-a.e. ω.

We are interested in the minimal distance between two pieces of orbit of length n. In a classical dynamical system setting, this would amount to understanding the behavior of $\min_{i,j<n} d(T^i x, T^j y)$ for a typical pair (x, y). It has been shown in [3] that the rate of decay to zero of this quantity is related to a dimension-like quantity associated to the invariant measure, called its correlation dimension (or Rényi entropy for symbolic dynamical systems),

measuring the polynomial decay rate of the r-neighborhood of the diagonal in $X \times X$ in terms of r.

In the random dynamics situation, there are two possible interpretations for this question. One may consider typical pairs (ω, x) and (ω', y) and try to minimize $d(T_\omega^i x, T_{\omega'}^j y)$. This is the annealed question, where randomness is taken over the whole product space $\Omega \times X$. It has already been studied in [5], and the outcome is comparable to the situation of classical dynamical systems (the relevant quantity being the correlation dimension of the second marginal μ of ν). One may also consider a typical ω, and then for this fixed ω pick a typical pair (x, y) for μ_ω and try to minimize $d(T_\omega^i x, T_\omega^j y)$. This is the quenched variant of our main question, to which this article is devoted.

In many questions about random systems, the outcome in a quenched situation is similar to the outcome of the annealed situation, but harder to prove. This is not the case here: we observe a behavior which is genuinely different from the annealed case, with a phase transition: there are two competing phenomena to decide the decay rate of the minimal distance between orbits, one similar to the annealed situation and one that is specific to the quenched situation, and each of them can be prevalent in some situations.

We denote

$$\underline{D}_2^{\mathrm{an}} = \liminf_{r \to 0} \frac{\log \int \mu(B(x,r)) \, \mathrm{d}\mu(x)}{\log r}, \quad \overline{D}_2^{\mathrm{an}} = \limsup_{r \to 0} \frac{\log \int \mu(B(x,r)) \, \mathrm{d}\mu(x)}{\log r}$$

for the lower and upper correlation dimensions of the measure μ. The 2 in the notation comes from the fact that this in an L^2-like expression, which is easier to see in a symbolic setting as in Remark 1.3. These are annealed quantities, referring to the averaged measure $\mu = \int \mu_\omega \, \mathrm{d}\mathbb{P}(\omega)$, hence the superscript an. In the quenched version, one should rather compute the correlation dimension of each measure μ_ω and then average with respect to $\mathbb{P}$, giving rise to the following definitions:

$$\underline{D}_2^{\mathrm{qu}} = \liminf_{r \to 0} \frac{\log \int \mu_\omega(B(x,r)) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mathbb{P}(\omega)}{\log r},$$

$$\overline{D}_2^{\mathrm{qu}} = \limsup_{r \to 0} \frac{\log \int \mu_\omega(B(x,r)) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mathbb{P}(\omega)}{\log r}.$$

When the liminf and the limsup coincide, we denote the corresponding quantities by D_2^{an} and D_2^{qu}.

Our main theorem shows that the decay rate of the minimal distance between orbits, in the quenched situation, can be expressed in terms of D_2^{an} and D_2^{qu}. This result requires that the geometry of the space should be nice enough (spaces with bounded local complexity, see Definition 1.4 below – this is a very mild geometric condition on the space, satisfied for instance by shift spaces and Riemannian manifolds), that the measures μ_ω depend in a Lipschitz way on ω (see Definition 1.5) and that the system mixes quickly enough, both for the base map and the fiber maps (stretched exponential mixing, see Definitions 1.7 and 1.8). Finally, we also require that the map S is Lipschitz.

Theorem 1.1. *Let X be a compact metric space with bounded local complexity. Consider a random dynamical system $S : \Omega \times X \to \Omega \times X$ preserving a probability measure ν, for which D_2^{an} and D_2^{qu} are well defined, and for which $\omega \to \mu_\omega$ is Lipschitz. Assume that S is*

Lipschitz, has fiberwise stretched exponential mixing, and that the base map has stretched exponential 4-mixing. Then, for $\mathbb{P}$-almost every ω, for $\mu_\omega^{\otimes 2}$-almost every x, y, one has the convergence

$$(1.1) \qquad \frac{-\log\min_{i,j<n} d(T_\omega^i x, T_\omega^j y)}{\log n} \to \max\left\{\frac{2}{D_2^{\mathrm{an}}}, \frac{1}{D_2^{\mathrm{qu}}}\right\}.$$

This theorem should be compared with the corresponding statement in the annealed situation: under the same assumptions, for $\nu^{\otimes 2}$ almost-every pairs $(\omega, x), (\omega', y)$, one has

$$\frac{-\log\min_{i,j<n} d(T_\omega^i x, T_{\omega'}^j y)}{\log n} \to \frac{2}{D_2^{\mathrm{an}}},$$

by [5, Theorem 4.4].

Theorem 1.1 is a consequence of several statements on upper and lower bounds, which for some of them require weaker assumptions regarding mixing, and which can be expressed in terms of $\underline{D}_2^{\mathrm{an}}$, $\underline{D}_2^{\mathrm{qu}}$ and $\overline{D}_2^{\mathrm{an}}$, $\overline{D}_2^{\mathrm{qu}}$ respectively, without requiring that D_2^{an} and D_2^{qu} are well defined. These more precise versions are discussed in Paragraph 1.3, after the precise meaning of our assumptions is discussed in the next paragraph.

Let us stress that Theorem 1.1 applies to a large class of concrete uniformly and nonuniformly expanding random dynamical systems, see Paragraphs 2.2 and 2.3. Also, one can construct examples in which the maximum in the right hand side of (1.1) is realized either by the first or the second term. See in particular Paragraph 2.1 in which we exhibit a family of systems depending smoothly on a parameter for which there is a transition from the first behavior to the second behavior, in a non-smooth way, exhibiting a second-order phase transition for the minimal approximation rate of orbits in the quenched setting (while there is no such phase transition for the analogous annealed question).

When Ω is a point, the annealed and quenched correlation dimensions coincide, so the maximum is always realized by $2/D_2^{\mathrm{an}}$. This is also the case when we are close enough to a product situation (in which case all the μ_ω are close to μ), but $1/D_2^{\mathrm{qu}}$ may become dominant in more distorted situations, as testified in Paragraph 2.1.

The intuition as to which term is dominant is the following. If one considers i far away from j, then $\theta^i\omega$ and $\theta^j\omega$ are essentially independent, so $T_\omega^i x$ and $T_\omega^j y$ are essentially two independent points distributed according to μ, and one should get the same behavior as in the annealed situation. There are n^2 such pairs (i, j), and for each of them the probability that the points are close by is governed by the dimension D_2^{an}, hence an asymptotics $2/D_2^{\mathrm{an}}$. For $j = i$ on the other hand, the points $T_\omega^i x$ and $T_\omega^i y$ are independent points distributed according to the measure $\mu_{\theta^i\omega}$, so the correlation dimension of this measure should appear in the asymptotic. Since there are only n such pairs (i, i) (as opposed to n^2 before), we get an asymptotics $1/D_2^{\mathrm{qu}}$. The precise statements in Paragraph 1.3 will make this intuition precise, by showing that the on-diagonal and off-diagonal behaviors are genuinely different.

It is interesting to specialize Theorem 1.1 to the case of a deterministic dynamical system (taking Ω to be a point). Many of our assumptions become trivial in this situation. The statement becomes the following.

Theorem 1.2. *Let $T : X \to X$ be a Lipschitz map on a compact metric space with bounded local complexity, preserving a probability measure μ with a well-defined correlation dimension*

$D_2(\mu)$. *Assume that T mixes stretched exponentially. Then, for $\mu^{\otimes 2}$-almost all x, y,*

$$\frac{-\log \min_{i,j<n} d(T^i x, T^j y)}{\log n} \to \frac{2}{D_2(\mu)}.$$

This theorem is essentially proved in [3], although the assumptions there are phrased in a slightly different way.

Remark 1.3 (Longest common substring). When the (random) dynamical system is a (random) shift, i.e., $X = \mathcal{A}^{\mathbb{N}}$ for some alphabet $\mathcal{A}$ and $T = \sigma$ (or $T_\omega = \sigma$) with σ the left shift, it was observed in [3] that studying the minimal distance between orbits is equivalent to studying the length of the longest common substring between two sequences, that is:

$$-\log \min_{i,j<n} d(\sigma^i x, \sigma^j y) = \max\{m : \exists 0 \leqslant i, j < n \text{ s.t. } x_{i+k} = y_{j+k} \text{ for } k = 0, \ldots, m-1\}.$$

In this case, balls will correspond to cylinders and the correlation dimensions coincide with the annealed and quenched Rényi entropies

$$H_2^{\mathrm{an}} = \lim_{k \to \infty} \frac{\log \sum \mu(C_k)^2}{-k} \text{ and } H_2^{\mathrm{qu}} = \lim_{k \to \infty} \frac{\log \sum \int \mu_\omega(C_k)^2 d\mathbb{P}(\omega)}{-k},$$

where the sums are taken over all k-cylinders.

1.2. **The technical assumptions.** In this paragraph, we specify precisely the technical assumptions made in Theorem 1.1. The various assumptions will also be useful to highlight, in Paragraph 1.3, which statements require stronger or weaker assumptions.

A function $f : X \to \mathbb{R}$ is Lipschitz if it satisfies the inequality $|f(x) - f(y)| \leqslant Cd(x, y)$ for all x, y. The best such C is called the Lipschitz constant of f and denoted by $\mathrm{Lip}(f)$. We define the Lipschitz norm of f, denoted by $\|f\|_{\mathrm{Lip}}$, to be the sum of its sup norm and its Lipschitz constant. In this way, $\|fg\|_{\mathrm{Lip}} \leq \|f\|_{\mathrm{Lip}}\|g\|_{\mathrm{Lip}}$.

Here is our main geometric assumption on the spaces we consider.

Definition 1.4. *A compact metric space X has bounded local complexity if there exists a constant C_0 such that, for any small enough r, there exist a constant $k(r) < +\infty$ and points $x_1^{(r)}, \ldots, x_{k(r)}^{(r)}$ in X such that the space is covered by the balls $(B(x_p^{(r)}, r))_{1 \leq p \leq k(r)}$ and any point x belongs to at most C_0 balls $B(x_p^{(r)}, 4r)$.*

Basic examples are shift spaces on finitely many symbols: for these, one may take the balls $B(x_p^{(r)}, r)$ as the different cylinders of a given length N, and they are all disjoint. Compact Riemannian manifolds have also bounded local complexity: this follows from the fact that Euclidean spaces are, using finitely many charts and an approximation argument to reduce to this situation.

Definition 1.5. *Given a random dynamical system on $\Omega \times X$, the random fiber measures μ_ω depend on a Lipschitz way on ω if there exists $C_1 > 0$ such that, for any Lipschitz function $f : X \to \mathbb{R}$, for any ω, ω',*

$$\left| \int f \, d\mu_\omega - \int f \, d\mu_{\omega'} \right| \leqslant C_1 \|f\|_{\mathrm{Lip}} d(\omega, \omega').$$

Let us now turn to the various mixing conditions we need, for the base map or the fiber maps.

Definition 1.6. *The dynamical system $\theta : \Omega \to \Omega$ mixes stretched exponentially if there exist $c_2 > 0$ and $C_2 > 0$ such that, for any Lipschitz functions $f, g : \Omega \to \mathbb{R}$, for any $n \in \mathbb{N}$,*

$$\left| \int f \cdot g \circ \theta^n \, d\mathbb{P} - \left(\int f \, d\mathbb{P} \right) \left(\int g \, d\mathbb{P} \right) \right| \leqslant C_2 e^{-n^{c_2}} \|f\|_{\mathrm{Lip}} \|g\|_{\mathrm{Lip}}.$$

We will need a stronger property, ensuring that there is quantitative mixing for 4 functions instead of 2, if there is a large enough time gap between the second and third functions. This property, that we call stretched exponential 4-mixing, implies the usual stretched exponential mixing of Definition 1.6 (take $f_2 = 1$ and $g_1 = 1$)

Definition 1.7. *The dynamical system $\theta : \Omega \to \Omega$ has stretched exponential 4-mixing if there exist $c_2 > 0$ and $C_2 > 0$ such that, for any Lipschitz functions $f_1, f_2, g_1, g_2 : \Omega \to \mathbb{R}$, for any $n \in \mathbb{N}$, for any $a \leq b \leq c$ with $b - a \geq n$,*

$$(1.2) \quad \left| \int f_1 \cdot f_2 \circ \theta^a \cdot g_1 \circ \theta^b \cdot g_2 \circ \theta^c \, d\mathbb{P} - \left(\int f_1 \cdot f_2 \circ \theta^a \, d\mathbb{P} \right) \left(\int g_1 \cdot g_2 \circ \theta^{c-b} \, d\mathbb{P} \right) \right|$$
$$\leqslant C_2 e^{-n^{c_2}} \|f_1\|_{\mathrm{Lip}} \|f_2\|_{\mathrm{Lip}} \|g_1\|_{\mathrm{Lip}} \|g_2\|_{\mathrm{Lip}}.$$

Finally, we give a fiberwise mixing condition. In the case where Ω is a point (i.e., for a deterministic dynamical system), as in Theorem 1.2, this is the only nontrivial assumption.

Definition 1.8. *The random dynamical system $S : \Omega \times X \to \Omega \times X$ mixes stretched exponentially along the fibers if there exist $c_3 > 0$ and $C_3 > 0$ such that, for any Lipschitz functions $f, g : X \to \mathbb{R}$, for any $n \in \mathbb{N}$, for any $\omega \in \Omega$*

$$\left| \int f \cdot g \circ T_\omega^n \, d\mu_\omega - \left(\int f \, d\mu_\omega \right) \left(\int g \, d\mu_{\theta^n \omega} \right) \right| \leqslant C_3 e^{-n^{c_3}} \|f\|_{\mathrm{Lip}} \|g\|_{\mathrm{Lip}}.$$

1.3. More fine-grained results. Let $m_n(\omega; x, y) = \min_{i,j<n} d(T_\omega^i x, T_\omega^j y)$ be the minimal distance between orbit segments of length n. For more precise results, we will need to split it further according to the allowed gap between i and j. Accordingly, let

$$(1.3) \qquad \qquad \alpha(n) = (\log n)^{C_4},$$

where C_4 is large enough (we will need $C_4 \geq \max(2/c_2, 2/c_3)$, where c_2 and c_3 are the rates of stretched exponential mixing along the basis and the fibers respectively). Let

$$m_n^0(\omega; x, y) = \min_{i<n} d(T_\omega^i x, T_\omega^i y),$$
$$m_n^{\leq}(\omega; x, y) = \min_{\substack{i,j<n \\ |j-i| \leq \alpha(n)}} d(T_\omega^i x, T_\omega^j y),$$
$$m_n^{>}(\omega; x, y) = \min_{\substack{i,j<n \\ |j-i| > \alpha(n)}} d(T_\omega^i x, T_\omega^j y),$$
$$m_n^{\gg}(\omega; x, y) = \min_{i<n/3, \, 2n/3 \leq j<n} d(T_\omega^i x, T_\omega^j y).$$

We start with the upper bounds for $-\log m_n$, i.e., with the lower bounds for m_n: we have to show that the orbits are never too close to each other. For this, we will split m_n as $\min(m_n^{\leq}, m_n^{>})$ and show separately that these two terms are almost surely not too small.

Proposition 1.9. *Assume that the space X has bounded local complexity. Then, for $\mathbb{P}$-almost every ω, for $\mu_\omega^{\otimes 2}$ every x, y, one has*

$$\limsup_{n\to\infty} \frac{-\log m_{\overline{n}}^{\le}(\omega; x, y)}{\log n} \le \frac{1}{\underline{D}_2^{\mathrm{qu}}}.$$

Proposition 1.10. *Assume that the space X has bounded local complexity, that the fiber measures μ_ω depend in a Lipschitz way on ω, and that θ mixes stretched exponentially. Then, for $\mathbb{P}$-almost every ω, for $\mu_\omega^{\otimes 2}$ every x, y, one has*

$$\limsup_{n\to\infty} \frac{-\log m_n^{>}(\omega; x, y)}{\log n} \le \frac{2}{\underline{D}_2^{\mathrm{an}}}.$$

Combining the two previous propositions, and since $m_n = \min(m_{\overline{n}}^{\le}, m_n^{>})$, one obtains almost surely

(1.4) $$\limsup_{n\to\infty} \frac{-\log m_n(\omega; x, y)}{\log n} \le \max\left\{ \frac{2}{\underline{D}_2^{\mathrm{an}}}, \frac{1}{\underline{D}_2^{\mathrm{qu}}} \right\},$$

proving the first (easy) half of Theorem 1.1.

Let us now deal with the lower bounds for $-\log m_n$, i.e., with the upper bounds for m_n: we have to show that there are some times at which the orbits are pretty close. We can select those times as we like. We will use either $i = j$ (given by m_n^0) or i and j very far apart, i.e., $i < n/3$ and $2n/3 \le j < n$ (given by $m_n^{\gg}$). In other words, we use the trivial inequality $m_n \le \min(m_n^0, m_n^{\gg})$, and we will get good upper bounds for these two terms.

Proposition 1.11. *Assume that the space X has bounded local complexity, that the fiber measures μ_ω depend in a Lipschitz way on ω, that θ mixes stretched exponentially and that S mixes stretched exponentially along the fibers. Then, for $\mathbb{P}$-almost every ω, for $\mu_\omega^{\otimes 2}$ every x, y, one has*

$$\liminf_{n\to\infty} \frac{-\log m_n^0(\omega; x, y)}{\log n} \ge \frac{1}{\overline{D}_2^{\mathrm{qu}}}.$$

Proposition 1.12. *Assume that the space X has bounded local complexity, that the fiber measures μ_ω depend in a Lipschitz way on ω, that θ has stretched exponential 4-mixing and that S mixes stretched exponentially along the fibers. Assume also that S is Lipschitz. Then, for $\mathbb{P}$-almost every ω, for $\mu_\omega^{\otimes 2}$ every x, y, one has*

$$\liminf_{n\to\infty} \frac{-\log m_n^{\gg}(\omega; x, y)}{\log n} \ge \frac{2}{\overline{D}_2^{\mathrm{an}}}.$$

Combining the two previous propositions, and since $m_n \le \min(m_n^0, m_n^{\gg})$, one obtains almost surely

$$\liminf_{n\to\infty} \frac{-\log m_n(\omega; x, y)}{\log n} \ge \max\left\{ \frac{2}{\overline{D}_2^{\mathrm{an}}}, \frac{1}{\overline{D}_2^{\mathrm{qu}}} \right\},$$

proving the second (harder) half of Theorem 1.1.

The proofs of all these theorems are given in Section 3, after several examples are discussed in Section 2. They go from the easiest one (Proposition 1.9) to the hardest one (Proposition 1.12). The results on the upper bounds (Propositions 1.9 and 1.10) are given in Paragraph 3.2. They are based on a first moment computation. The results on the lower

bounds (Propositions 1.11 and 1.12) are then established in Paragraph 3.3. They rely on technically more involved second moment estimates, that require stronger mixing conditions.

2. Examples

We now present an explicit example with a phase transition, as well as two classes of random dynamical systems which satisfy the hypothesis of Theorem 1.1. The first class is constructed from finitely many uniformly expanding maps whereas the second is given by a continuous family of non-uniformly expanding, local diffeomorphisms on a manifold.

2.1. **An explicit example with a phase transition.** We construct in this section a simple random Bernoulli shift where, depending on the sample measures, we will obtain a phase transition. As explained in Remark 1.3, in this example we will compute Rényi entropies which correspond to correlation dimensions on shift spaces.

Let (Ω, θ) be the full shift on the symbolic space $\Omega = \{A, B\}^{\mathbb{Z}}$ and let $\mathbb{P} = \mathcal{P}^{\mathbb{Z}}$ with $\mathcal{P}(A) = \mathcal{P}(B) = \frac{1}{2}$. We then consider the one-sided shift on $X = \{0, 1\}^{\mathbb{N}}$ as a random subshift by constructing a random Bernoulli measure as follows. Let $p_A, p_B \in (0, 1)$. The random Bernoulli measure $\{\mu_\omega : \omega \in \Omega\}$ is defined by

$$\mu_\omega([x_0, \ldots, x_n]) = \mu_{\omega_0}(x_0)\mu_{\omega_1}(x_1) \cdots \mu_{\omega_n}(x_n),$$

with $\mu_A(0) = p_A$ and $\mu_A(1) = 1 - p_A$ on the one hand, and $\mu_B(0) = p_B$ and $\mu_B(1) = 1 - p_B$ on the other hand.

To compute the Rényi entropy, observe that for a cylinder $C_n = [x_0, \ldots, x_{n-1}]$

$$\mu(C_n) = \int \mu_\omega([x_0, \ldots, x_{n-1}])d\mathbb{P}(\omega) = \int \prod_{i=0}^{n-1} \mu_{\omega_i}(x_i)d\mathbb{P}(\omega) = \prod_{i=0}^{n-1} \int \mu_{\omega_i}(x_i)d\mathbb{P}(\omega_i)$$

$$= \prod_{i=0}^{n-1} \left(\frac{1}{2}\mu_A(x_i) + \frac{1}{2}\mu_B(x_i) \right) = \frac{1}{2^n} (p_A + p_B)^{\#\{i:x_i=0\}} (2 - p_A - p_B)^{\#\{i:x_i=1\}}.$$

Thus, by the binomial identity,

$$\sum_{C_n} \mu(C_n)^2 = \frac{1}{2^{2n}} \left((p_A + p_B)^2 + (2 - p_A - p_B)^2 \right)^n = \frac{1}{2^n} \left((p_A + p_B)^2 - 2(p_A + p_B) + 2 \right)^n,$$

which implies that

$$H_2^{\mathrm{an}} = -\log\left(\frac{1}{2} \left((p_A + p_B)^2 - 2(p_A + p_B) + 2 \right) \right).$$

Moreover, it follows by the same arguments that

$$\int \mu_\omega(C_n)^2 d\mathbb{P}(\omega) = \prod_{i=0}^{n-1} \int \mu_{\omega_i}(x_i)^2 d\mathbb{P}(\omega_i) = \prod_{i=0}^{n-1} \left(\frac{1}{2}\mu_A(x_i)^2 + \frac{1}{2}\mu_B(x_i)^2 \right)$$

$$= \frac{1}{2^n} \left(p_A^2 + p_B^2 \right)^{\#\{i:x_i=0\}} \left((1 - p_A)^2 + (1 - p_B)^2 \right)^{\#\{i:x_i=1\}}.$$

Thus,

$$\sum_{C_n} \int \mu_\omega(C_n)^2 d\mathbb{P}(\omega) = \left(\frac{1}{2} \left(p_A^2 + p_B^2 + (1 - p_A)^2 + (1 - p_B)^2 \right) \right)^n$$

and

$$H_2^{\mathrm{qu}} = -\log\left(\frac{1}{2}\left(p_A^2 + p_B^2 + (1-p_A)^2 + (1-p_B)^2\right)\right)$$
$$= -\log\left(\left(p_A^2 - \tfrac{1}{2}\right)^2 + \left(p_B^2 - \tfrac{1}{2}\right)^2 + \tfrac{1}{2}\right).$$

Proposition 2.1 below shows that the hypothesis of Theorem 1.1 are satisfied. Hence, for $\mathbb{P}$-almost every ω, for $\mu_\omega \otimes \mu_\omega$-almost every (x, y),

$$\frac{-\log m_n(\omega; x, y)}{\log n} \xrightarrow[n \to \infty]{} \max\left\{\frac{2}{H_2^{\mathrm{an}}}, \frac{1}{H_2^{\mathrm{qu}}}\right\}$$
$$= \max\left\{\frac{2}{-\log\frac{1}{2}\left((p_A + p_B)^2 - 2(p_A + p_B) + 2\right)}, \frac{1}{-\log\left(\left(p_A^2 - \tfrac{1}{2}\right)^2 + \left(p_B^2 - \tfrac{1}{2}\right)^2 + \tfrac{1}{2}\right)}\right\}.$$

The behavior of this maximum depends on the values of p_A and p_B and some simple choices can give us distinctive behaviors. If p_A and p_B are close enough (for example if $|p_A - p_B| \leqslant \frac{1}{2}$) then $\max\{2/H_2^{\mathrm{an}}, 1/H_2^{\mathrm{qu}}\} = 2/H_2^{\mathrm{an}}$. In this case, we obtain the same behaviour as observed in [7, Example 2.1]. However, if p_A and p_B are sufficiently far from each other, there is a phase transition as the quenched parameter becomes dominant. In order to obtain a precise description of these domains, it suffices to determine the separating curve given by $H_2^{\mathrm{an}}(p_A, p_B) = 2H_2^{\mathrm{qu}}(p_A, p_B)$ (see the left hand side of Figure 2.1). The same argument then gives rise to the contours (or level sets) of the function $(p_A, p_B) \mapsto \max\{2/H_2^{\mathrm{an}}, 1/H_2^{\mathrm{qu}}\}$, which are displayed on the right hand side of Figure 2.1. Observe that the formulas for H_2^{an} and H_2^{qu} imply that these contours consist of straight lines of slope -1 and circle segments with center $(1/2, 1/2)$.
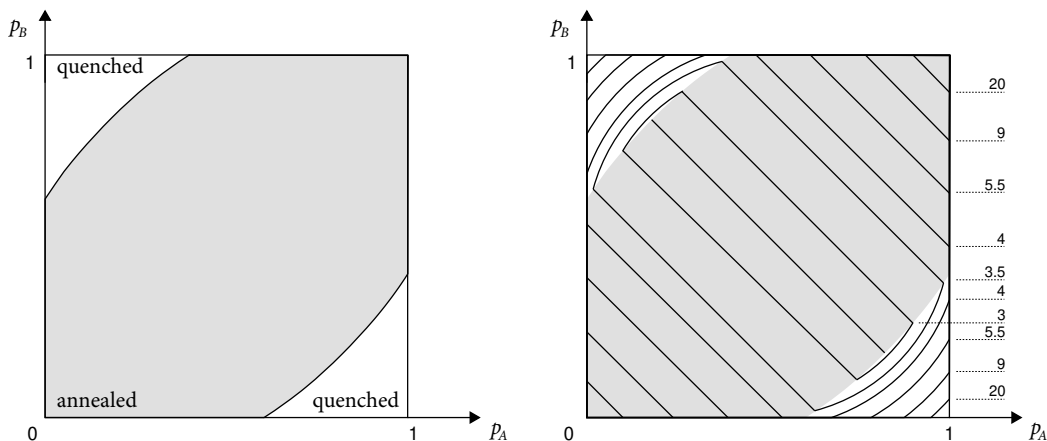

FIGURE 2.1. Regions with quenched and annealed behaviour

For example, if we choose $p_B = 1 - p_A$, and set $c_\pm := \frac{1}{2} \pm \frac{1}{2}\sqrt{2\sqrt{\sqrt{2}-1}-1}$, then

$$\lim_{n\to\infty} \frac{-\log m_n(\omega; x, y)}{\log n} = \max\left\{\frac{2}{H_2^{\mathrm{an}}}, \frac{1}{H_2^{\mathrm{qu}}}\right\} = \begin{cases} \frac{1}{H_2^{\mathrm{qu}}} & : \ 0 < p_A \leqslant c_- \\ \frac{2}{H_2^{\mathrm{an}}} & : c_- < p_A \leqslant c_+ \\ \frac{1}{H_2^{\mathrm{qu}}} & : c_+ < p_A \leqslant 1. \end{cases}$$

To illustrate this phase transition, the graph of $p_A \mapsto \max\{2/H_2^{\mathrm{an}}, 1/H_2^{\mathrm{qu}}\}$, provided that $p_A = 1 - p_B$, is presented in Figure 2.2.
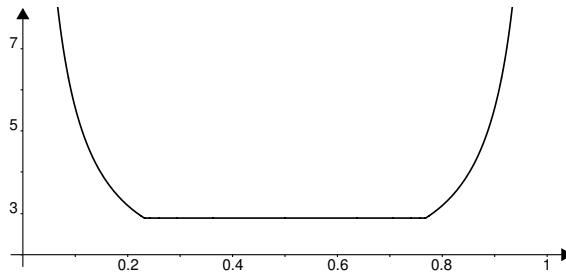

FIGURE 2.2. Graph of $-\lim_{n\to\infty} \frac{\log m_n(\omega; x, y)}{\log n}$

2.2. Finitely many Ruelle expanding maps.
In this paragraph, we describe a simple class of random dynamical systems to which Theorem 1.1 applies.

We begin with the description of the fibre maps. Let (X, d) be a compact metric space of bounded local complexity and recall that a Lipschitz-continuous and surjective map $T : X \to X$ is *Ruelle expanding* if there exist $a > 0$ and $\lambda \in (0, 1)$, such that for any $x, y, \tilde{x} \in X$ with $d(x, y) < a$ and $T(\tilde{x}) = x$, there exists a unique $\tilde{y} \in X$ with $T(\tilde{y}) = y$ and $d(\tilde{x}, \tilde{y}) < a$. Moreover, $d(\tilde{x}, \tilde{y}) \leqslant \lambda d(x, y)$. This class of maps was introduced by Ruelle in [8] and contains subshifts of finite type as well as uniformly expanding maps on manifolds.

Now assume that $T_1, \ldots, T_k : X \to X$ are Ruelle expanding maps which are jointly mixing in the following sense. For any pair of nonempty open sets $U, V \subset X$ we require that there exists $m \in \mathbb{N}$ such that $(T_{i_n} \circ \cdots T_{i_1})^{-1}(U) \cap V \neq \emptyset$ for each choice $i_1, \ldots i_n \in \{1, \ldots k\}$ and $n > m$. In order to construct the skew product, assume that $\Omega \subseteq \{1, \ldots k\}^{\mathbb{Z}}$ is a topologically mixing subshift of finite type and let

$$S : \Omega \times X \to \Omega \times X, ((\omega_i), x) \mapsto (\sigma((\omega_i)), T_{\omega_0}(x)),$$

where σ refers to the left shift. Here, it is worth noting that the choice of Ω as a shift space is natural in the setting of finitely many maps.

It remains to construct measures $\mathbb{P}$ and $\{\mu_\omega : \omega \in \Omega\}$ which satisfy the Lipschitz and mixing conditions in Definitions 1.5, 1.8 and 1.7. In order to do so, we first define a metric on Ω by $d_s(\omega, \tilde{\omega}) := s^{\min\{|k| : \omega_k \neq \tilde{\omega}_k\}}$ for some fixed $s \in (0, 1)$. Secondly, we fix Lipschitz continuous functions $\psi : \Omega \to \mathbb{R}$ and $\varphi_i : X \to \mathbb{R}$ for $i = 1, \ldots k$ and assume that $\mathbb{P}$ is the

unique Gibbs measure associated to ψ (see [4]). For the construction of μ_ω, we proceed as follows. It is well known that the operator defined by

$$(2.1) \qquad \mathcal{L}_i(f)(x) := \sum_{T_i(y)=x} e^{\varphi_i(y)} f(y)$$

acts on the space of Lipschitz functions. Furthermore, as shown in [10, Prop. 6.3], there exists $a > 0$ such that for any $\omega \in \Omega$, there exists a probability measure μ_ω such that for any Lipschitz continuous function f and $m, n > 0$,

$$(2.2) \qquad \left\| \frac{\mathcal{L}_{\omega_m} \cdots \mathcal{L}_{\omega_0}(f\, \mathcal{L}_{\omega_{-1}} \cdots \mathcal{L}_{\omega_{-n}}(\mathbf{1}))}{\mathcal{L}_{\omega_m} \cdots \mathcal{L}_{\omega_{-n}}(\mathbf{1})} - \int f\, \mathrm{d}\mu_\omega \right\|_\infty \leqslant Ce^{-a\min\{m,n\}} \operatorname{Lip}(f),$$

where $\operatorname{Lip}(f)$ refers to the best Lipschitz constant of f.

Proposition 2.1. *Assume that $T_1, \ldots T_k$ are jointly mixing Ruelle expanding maps of the compact metric space of bounded local complexity X and that (Ω, σ) is a two-sided, topologically mixing subshift of finite type. Furthermore, assume that $\psi : \Omega \to \mathbb{R}$ and $\varphi_i : X \to \mathbb{R}$ are Lipschitz continuous. Then the conclusions of Theorem 1.1 hold with respect to the equilibrium state $\mathbb{P}$ of ψ and $\{\mu_\omega : \omega \in \Omega\}$ as defined in (2.2).*

Proof. Observe that it follows from (2.2) that $\omega \mapsto \mu_\omega$ is Lipschitz continuous with respect to d_t, for $t := \max\{s, e^{-a}\}$. For ease of notation, set $\mathcal{L}_{\omega_n \cdots \omega_0} := \mathcal{L}_{\omega_n} \cdots \mathcal{L}_{\omega_0}$. Fix $x_0 \in X$. For $f, g : X \to \mathbb{R}$ Lipschitz continuous and $k > 0$, it follows from (2.2) that

$$\int fg \circ T_\omega^k\, \mathrm{d}\mu_\omega - \int f\, \mathrm{d}\mu_\omega \int g\, \mathrm{d}\mu_{\theta^k \omega}$$

$$= \lim_{n \to \infty} \frac{\mathcal{L}_{\omega_n \ldots \omega_0}(fg \circ T_\omega^k \mathcal{L}_{\omega_{-1} \ldots \omega_{-n}}(\mathbf{1}))(x_0)}{\mathcal{L}_{\omega_n \ldots \omega_{-n}}(\mathbf{1})(x_0)} - \int f\, \mathrm{d}\mu_\omega \int g\, \mathrm{d}\mu_{\theta^k \omega}$$

$$= \lim_{n \to \infty} \frac{\mathcal{L}_{\omega_n \cdots \omega_k}\left(g\mathcal{L}_{\omega_{k-1} \cdots \omega_{-n}}(\mathbf{1})\left(\frac{\mathcal{L}_{\omega_{k-1} \cdots \omega_0}(f\mathcal{L}_{\omega_{-1} \cdots \omega_{-n}}(\mathbf{1}))}{\mathcal{L}_{\omega_{k-1} \cdots \omega_{-n}}(\mathbf{1})} - \int f\, \mathrm{d}\mu_\omega\right)\right)(x_0)}{\mathcal{L}_{\omega_n \cdots \omega_{-n}}(\mathbf{1})(x_0)}$$

$$\leqslant Ce^{-a(k-1)} \operatorname{Lip}(f) \int |g|\, \mathrm{d}\mu_{\theta^k \omega} \leqslant Ce^{-a(k-1)} \|f\|_{\operatorname{Lip}} \|g\|_{\operatorname{Lip}}$$

as $\|\cdot\|_{\operatorname{Lip}} = \|\cdot\|_\infty + \operatorname{Lip}(\cdot)$. Furthermore, by considering $f = \mathbf{1}$, it follows from the above calculation that $\mu_\omega \circ T_\omega^{-1} = \mu_{\theta\omega}$.

The fact that $\mathbb{P}$ is exponential 4-mixing is standard. Let us nevertheless explain the proof quickly, using the non-invertible, canonical factor (Ω_+, θ_+) of (Ω, θ), where $\Omega_+ := \{(\omega_i : i = 0, 1, \ldots) : (\omega_i) \in \Omega\}$, θ_+ is the one-sided shift, d_t^+ is the usual shift metric with respect to the parameter t and $\pi : \Omega \to \Omega_+$ the canonical projection. It is now crucial to recall some results from [4]. Firstly, we may assume without loss of generality that $\psi = \psi_+ \circ \pi$ for a Lipschitz function $\psi_+ : \Omega_+ \to \mathbb{R}$. The regularity of ψ_+ then implies that the operator defined by

$$P(f)(\omega) := \sum_{\theta_+ \tilde{\omega} = \omega} e^{\psi_+(\tilde{\omega})} f(\tilde{\omega})$$

acts on the space of Lipschitz continuous functions on Ω_+. Secondly, by adding a coboundary we may then assume that $P(\mathbf{1}) = \mathbf{1}$ and $\|P^n(f) - \int f\, \mathrm{d}\mathbb{P}_+\|_{\operatorname{Lip}} \ll \lambda^n \operatorname{Lip}(f)$ for $\mathbb{P}_+ := \mathbb{P} \circ \pi^{-1}$ and some $\lambda \in (0, 1)$.

Now fix $m \in \mathbb{N}$ and choose for a given function $f : \Omega \to \mathbb{R}$ a function $f^* : \Omega \to \mathbb{R}$ which on each ball of radius t^{m+1} is constant and equal to some value of f there. If f is Lipschitz continuous, then $\|f - f^*\|_\infty \leqslant \mathrm{Lip}(f)t^m$ and there exists $f^+ : \Omega_+ \to \mathbb{R}$ such that $f^+ \circ \pi = f^* \circ \theta^m$ and $\mathrm{Lip}(f^+) \leqslant t^{-m-1}\mathrm{Lip}(f)$.

Now assume that f_1, f_2, g_1, g_2 are Lipschitz continuous and that $0 \leqslant a \leqslant b \leqslant c$, with $b - a \geqslant n \geqslant 0$. Letting $m = \varepsilon n$ for some positive ε to be fixed below, we choose functions $f_1^*, f_2^*, g_1^*, g_2^*$ as above. To check the 4-mixing property (1.2), we may replace f_i and g_i with f_i^* and g_i^* respectively, as this introduces an exponentially small error.

Then the quantity to be estimated is

$$
\left| \int (f_1^* f_2^* \circ \theta^a g_1^* \circ \theta^b g_2^* \circ \theta^c) \circ \theta^m \, d\mathbb{P} - \int f_1^* f_2^* \circ \theta^a \, d\mathbb{P} \int g_1^* g_2^* \circ \theta^{c-b} \, d\mathbb{P} \right|
$$

$$
= \left| \int f_1^+ f_2^+ \circ \theta_+^a g_1^+ \circ \theta_+^b g_2^+ \circ \theta_+^c \, d\mathbb{P}_+ - \int f_1^+ f_2^+ \circ \theta_+^a \, d\mathbb{P}_+ \int g_1^+ g_2^+ \circ \theta_+^{c-b} \, d\mathbb{P}_+ \right|
$$

$$
= \left| \int P^{b-a}(f_2^+ P^a(f_1^+))g_1^+ g_2^+ \circ \theta_+^{c-b} \, d\mathbb{P}_+ - \int f_2^+ P^a(f_1^+) \, d\mathbb{P}_+ \int g_1^+ g_2^+ \circ \theta_+^{c-b} \, d\mathbb{P}_+ \right|
$$

$$
\ll \lambda^{b-a} \mathrm{Lip}(f_2^+ P^a(f_1^+)) \int |g_1^+ g_2^+ \circ \theta_+^{c-b}| \, d\mathbb{P}_+
$$

$$
\ll \lambda^{b-a} t^{-2m} \|f_1\|_{\mathrm{Lip}} \|f_2\|_{\mathrm{Lip}} \|g_1\|_\infty \|g_2\|_\infty.
$$

The prefactor is bounded by $\lambda^n t^{-2\varepsilon n}$. If ε is small enough, it is exponentially small as desired. $\qquad\square$

Remark 2.2. If $T_1, \ldots, T_k$ are Ruelle expanding maps defined on a connected and compact Riemannian manifold X, then the semigroup generated by these maps is always jointly mixing (see [10, Prop. 3.3]). Furthermore, as Riemannian manifolds are always of bounded local complexity, the conclusions of Proposition 2.1 hold without these two hypotheses.

2.3. **Non-uniformly expanding local diffeomorphisms.** We now give an example in which the fiber maps are nonuniformly expanding: contrary to Paragraph 2.2, there may be some region where the maps are contracting, but the expansion is still winning on average. We adapt the setting in [9]. In contrast to the situation in there, we have to assume that the base transformation has the following mixing property.

(H0) Assume that (Ω, d) is a compact metric space, that $\theta : \Omega \to \Omega$ is a bi-Lipschitz homeomorphism and that $\mathbb{P}$ is a θ-invariant probability measure with stretched exponential 4-mixing.

Furthermore, let X be a compact connected Riemannian manifold and let $\{T_\omega\}_{\omega \in \Omega}$ be a family of C^1-maps on X with the following properties.

(H1) For each ω, the map T_ω is a surjective, local diffeomorphism.
(H2) There exists $\delta > 0$ such that for every $(\omega, x) \in X$, there exists an open neighborhood U_x^ω of x with $T_\omega|_{U_x^\omega} : U_x^\omega \to B(T_\omega(x), \delta)$ invertible,
(H3) There exists $C > 0$ such that $\|DT_\omega(x)\| \leqslant C$ for all $\omega \in \Omega$ and all $x \in X$.

Observe that the connectedness and compactness of X imply that the degree $\deg(T_\omega)$ of T_ω, i.e., the number of preimages of T_ω is finite and constant for each ω. Moreover, (H3) ensures that this degree is uniformly bounded.

We also assume the following geometric conditions. There are random variables $\sigma_\omega > 1$, $L_\omega \geqslant 0$ and $0 \leqslant p_\omega, q_\omega < \deg(f_\omega)$ with $L := \sup L_\omega < \infty$ such that for each ω,

(H4) there exists a covering $\mathcal{P}_\omega = \{P_1^\omega, \ldots, P_{p_\omega}^\omega, \ldots, P_{p_\omega + q_\omega}^\omega\}$ of X_ω such that every $T_\omega|_{P_i}$ is injective, $\|DT_\omega(x)^{-1}\| \leqslant \sigma_\omega^{-1} < 1$ for each $x \in P_1^\omega \cup \cdots \cup P_{p_\omega}^\omega$, and $\|DT_\omega(x)^{-1}\| \leqslant L_\omega$ for every $x \in X$,

(H5)

$$\sup\left\{\log \frac{\sigma_\omega^{-1} p_\omega + L_\omega q_\omega}{\deg(f_\omega)} : \omega \in \Omega\right\} < 0.$$

Finally, assume that $\{\varphi_\omega\}$ is a family of real valued functions in $C^1(X)$, referred to as potentials, such that $\sup\{\|D\varphi_\omega\|_\infty : \omega \in \Omega\} < \infty$ and for all ω

(H6) $\displaystyle\sup_{x \in X_\omega} \varphi_\omega - \inf_{x \in X_\omega} \varphi_\omega + \log(1 + \|D\varphi_\omega\|_\infty \operatorname{diam} X) < -\log \frac{\sigma_\omega^{-1} p_\omega + L_\omega q_\omega}{\deg(f_\omega)}$.

We now motivate conditions (H4–6). (H4) states that regions of contraction and expansion may coexist whereas (H5) implies that expansion dominates the contraction. Finally, (H6) can be seen as an upper bound of the global and local oscillation by the combinatorial expansion in (H4). However, in contrast to the bounds in average in [9], we have to ask for uniform bounds in (H5) and (H6) due to the uniform hypothesis in Theorem 1.1.

Furthermore, in order to guarantee the continuous variation with respect to ω, we assume that T and φ vary Lipschitz continuously with respect to ω, that is we require that

(H7) there exists $C_T > 0$ with $d(T_\omega(x), T_{\omega'}(x)) \leqslant C_T d(\omega, \omega')$ for all $\omega, \omega' \in \Omega$ and all $x \in X$,

(H8) there exists $C_\varphi > 0$ with $|\varphi_\omega(x) - \varphi_{\omega'}(x)| \leqslant C_\varphi d(\omega, \omega')$ for all $\omega, \omega' \in \Omega$ and all $x \in X$.

The relevant measures $\{\mu_\omega\}$ are now given by application of Theorem A in [9]: denoting by $\mathcal{L}_\omega$ the transfer operator associated to T_ω and φ_ω as in (2.1), there exist families of positive constants $\{\lambda_\omega\}$, of differentiable functions $\{h_\omega\}$ and probability measures ν_ω such that $\mathcal{L}_\omega(h_\omega) = \lambda_\omega h_{\theta\omega}$ and $\mathcal{L}_\omega^*(\nu_{\theta\omega}) = \lambda_\omega \nu_\omega$. Furthermore, for $d\mu_\omega := h_\omega \, d\nu_\omega$, it follows that $\mu_\omega = \mu_{\theta\omega} \circ T_\omega^{-1}$ and that a fibered exponential decay of correlation for C^1-observables holds (Corollary 1 in [9]). Finally, we would like to remark that the standard examples for this class are random Manneville-Pomeau maps or sufficiently random perturbations of non-uniformly expanding maps with respect to a potential sufficiently close to zero (see Examples 3.1 and 3.2 in [9] for more details).

Proposition 2.3. *Assume that (H0–H8) hold. Then there exists $\alpha > 0$ such that the assumptions of Theorem 1.1 hold with respect to $\mathbb{P}$, the metric d_α defined by $d_\alpha(\omega, \omega') := d(\omega, \omega')^\alpha$ on Ω and $\{\mu_\omega : \omega \in \Omega\}$ as defined above.*

Proof. Before starting with the proof, we remark that we will write $a \ll b$ whenever there exists $C > 0$, depending exclusively on (H0–H8), such that $a \leqslant Cb$.

In order to verify the assumptions of Theorem 1.1, we have to modify the construction in [9] slightly. In order to do so, recall that it is shown there that there exists a positive and almost surely finite random variable κ_ω such that the family of cones

(2.3) $$\Lambda_\omega := \left\{g_\omega \in C^1(X) : g_\omega > 0 \text{ and } \|Dg_\omega\|_\infty \leqslant \kappa_\omega \inf_{x \in X} g_\omega(x)\right\}$$

satisfies $\mathcal{L}_\omega(\Lambda_\omega) \subset \Lambda_{\theta\omega}$ a.s. However, as (H6) provides a uniform bound, it follows from the construction (see Formula 5.2 in [9]) that κ_ω is in fact uniformly bounded. Hence (see [9, Claim 1]), there exist $C > 0$ and $\vartheta \in (0,1)$ such that, for every $m, n \geqslant 1$ and every $\varphi \in \Lambda_{\theta^{-(n+m)}(\omega)}$ and $\psi \in \Lambda_{\theta^{-n}\omega}$,

$$(2.4) \qquad \Theta_\omega\Big(\mathcal{L}^{n+m}_{\theta^{-(n+m)}\omega}\varphi, \mathcal{L}^n_{\theta^{-n}\omega}\psi\Big) \leqslant C\,\vartheta^n\,\Theta_{\theta^{-n}\omega}\big(\mathcal{L}^m_{\theta^{-(m+n)}\omega}\varphi, \psi\big).$$

In here, Θ_ω refers to the Hilbert metric on Λ_ω and $\mathcal{L}^n_\omega$ stands for $\mathcal{L}_{\theta^{n-1}\omega}\cdots\mathcal{L}_\omega$ which we now analyze briefly. Firstly, by dividing the defining relation by g_ω in (2.3), one obtains that $\|D\log g_\omega\|_\infty \leqslant \kappa_\omega$ and, in particular, that $q_\omega(x)/q_\omega(y) \leqslant \kappa_\omega \operatorname{diam}(X)$ for all $x, y \in X$. Secondly, if $\Theta_\omega(f,g) < \varepsilon$ for some $\varepsilon > 0$, then there are $t' \geqslant t > 0$ with $\log t' - \log t < \varepsilon$ such that $0 \leqslant \kappa_\omega \inf(f - tg)$ and $0 \leqslant \kappa_\omega \inf(t'g - f)$. Hence, $t \leqslant f/g \leqslant t'$ and $0 \leqslant f/g - t < \varepsilon$.

Construction of ν_ω. Assume that $f \in \Lambda_\omega$. Then (2.4) implies, for $\varepsilon_n := C\vartheta^n\Theta_\omega(f,\mathbf{1})$, that $\Theta_{\theta^n\omega}(\mathcal{L}^n_\omega(f), \mathcal{L}^n_\omega(\mathbf{1})) \leqslant \varepsilon_n$. With $r_n := \inf_x \mathcal{L}^n_\omega(f)(x)/\mathcal{L}^n_\omega(\mathbf{1})(x)$, one then obtains from the above that $r_n \leqslant \mathcal{L}^n_\omega(f)/\mathcal{L}^n_\omega(\mathbf{1}) \leqslant r_n e^{\varepsilon_n}$. Hence, as

$$r_{m+n} \leqslant \frac{\mathcal{L}^m_\omega(\mathcal{L}^n_{\theta^m\omega}(f))}{\mathcal{L}^{m+n}_\omega(\mathbf{1})} \leqslant r_n e^{\varepsilon_n} \text{ and } r_{m+n}e^{\varepsilon_{m+n}} \geqslant \frac{\mathcal{L}^m_\omega(\mathcal{L}^n_{\theta^m\omega}(f))}{\mathcal{L}^{m+n}_\omega(\mathbf{1})} \geqslant r_n,$$

$(\log r_n)$ is a Cauchy sequence and, in particular, $\lim_n r_n \in (0,\infty)$. Moreover, as $\Theta_\omega(f,\mathbf{1})$ is uniformly bounded, the function $\mathcal{L}^n_\omega(f)/\mathcal{L}^n_\omega(\mathbf{1})$ converges uniformly to a constant $\nu_\omega(f)$. Moreover, this constant satisfies

$$(2.5) \qquad \left\|\log\nu_\omega(f) - \log\frac{\mathcal{L}^n_\omega(f)}{\mathcal{L}^n_\omega(\mathbf{1})}\right\|_\infty \ll \vartheta^n \quad \forall f \in \Lambda_\omega, n \in \mathbb{N}, \omega \in \Omega.$$

We now extend the domain of ν_ω to Lipschitz functions, using the following standard fact: There exists $C > 0$ such that for any $\varepsilon > 0$ and Lipschitz function $f : X \to \mathbb{R}$, there is $f^* \in C^1(X)$ with $\|f - f^*\|_\infty \leqslant \varepsilon$ and $\|Df^*\| \leqslant C\operatorname{Lip}(f)$. Let us recall how this fact is proved, using *mollifiers* as follows. Assume that $\ell : \mathbb{R}^d \to [0,\infty)$ is a function in C^1 such that ℓ is supported on $\{x : \|x\| \leqslant 1\}$ and $\int \ell \,\mathrm{d}x = 1$. For $\ell_\varepsilon(x) := \varepsilon^{-d}\ell(x/\varepsilon)$, the convolution $g * \ell_\varepsilon$ with a Lipschitz function g is in $C^1(X)$ and $\|g - g*\ell_\varepsilon\|_\infty \leqslant 2\operatorname{Lip}(g)\varepsilon$. If, in addition, $\ell(x) = j(\|x\|)$ for some $j : [0,1] \to \mathbb{R}$ (i.e., ℓ is constant on spheres), a straightforward calculation shows that $\|D(g*\ell_\varepsilon)\| \ll \operatorname{Lip}(g)$ where the implicit constant in '$\ll$' only depends on ℓ. Finally, by employing an argument based on a partition of unity, one obtains that the same holds for the Riemannian manifold X.

In order to employ (2.5), note that $f^* + c \in \Lambda_\omega$ for $c = \|Df^*\|_\infty/\kappa_\omega - \inf f^*$. Hence,

$$\left\|\nu_\omega(f^*) - \frac{\mathcal{L}^n_\omega(f)}{\mathcal{L}^n_\omega(\mathbf{1})}\right\|_\infty - \varepsilon \leqslant \left\|\nu_\omega(f^*+c) - \frac{\mathcal{L}^n_\omega(f^*+c)}{\mathcal{L}^n_\omega(\mathbf{1})}\right\|_\infty$$
$$\ll \vartheta^n\nu_\omega(f^*+c) \ll \vartheta^n(\|Df^*\|_\infty/\kappa_\omega + \|f^* - \inf f^*\|_\infty)$$
$$\ll \vartheta^n\operatorname{Lip}(f)\big(\kappa_\omega^{-1} + \operatorname{diam}(X)\big) \ll \vartheta^n\operatorname{Lip}(f).$$

As $\varepsilon > 0$ is arbitrary, the definition of ν_ω extends to Lipschitz continuous functions and

$$(2.6) \qquad \|\mathcal{L}^n_\omega(f)/\mathcal{L}^n_\omega(\mathbf{1}) - \nu_\omega(f)\|_\infty \ll \vartheta^n\operatorname{Lip}(f).$$

Construction of h_ω. Let $\lambda^n_{\theta^{-n}\omega} := \nu_\omega(\mathcal{L}^n_{\theta^{-n}\omega}(\mathbf{1}))$ and $h^n_\omega := \mathcal{L}^n_{\theta^{-n}\omega}(\mathbf{1})/\lambda^n_{\theta^{-n}\omega}$. As Θ_ω is a projective metric, (2.4) implies that $(\mathcal{L}^n_{\theta^{-n}\omega}(\mathbf{1}))$ and (h_n) are Cauchy sequences with respect to Θ_ω. Hence, for $r_{n,k} := \inf_x h^{n+k}_\omega(x)/h^n_\omega(x)$, it follows in analogy to the construction of ν_ω that

$$r_{n,k} \leqslant h^{n+k}_\omega/h^n_\omega \leqslant r_{n,k}e^{\varepsilon_n}.$$

Moreover, by multiplying with h^n_ω and integrating with respect to ν_ω, it follows that $|\log r_{n,k}| \leqslant \varepsilon_n$. Hence, $h_\omega = \lim h^n_\omega$ exists and satisfies

$$(2.7) \qquad\qquad \|\log h^n_\omega - \log h_\omega\|_\infty \ll \vartheta^n.$$

Exponential mixing. We now show that $\mathrm{d}\mu_\omega = h_\omega\,\mathrm{d}\nu_\omega$ mixes exponentially along fibers with respect to Lipschitz functions (in [9], it is only shown for functions in $C^1(X)$). However, it follows from (2.6) and (2.7) that, for f Lipschitz and some uniform $C > 0$,

$$\left\|\log\nu_\omega(fh_\omega) - \log\frac{\mathcal{L}^n_\omega(fh_\omega)}{\lambda^n_\omega h_{\theta^n\omega}}\right\|_\infty \leqslant \left\|\log\nu_\omega(fh_\omega) - \log\frac{\mathcal{L}^n_\omega(fh_\omega)}{\mathcal{L}^n_\omega(\mathbf{1})}\right\|_\infty + \left\|\log\frac{\lambda^n_\omega h_{\theta^n\omega}}{\mathcal{L}^n_\omega(\mathbf{1})}\right\|_\infty$$

$$\leqslant C\vartheta^n + \|\log h_{\theta^n\omega} - \mathcal{L}^n_\omega(\mathbf{1})/\lambda^n_\omega\|_\infty \ll \vartheta^n.$$

Observe that $\tilde{\mathcal{L}}_\omega : f \mapsto \mathcal{L}^n_\omega(fh_\omega)/(\lambda^n_\omega h_{\theta^n\omega})$ is the transfer operator of T_ω with respect to μ_ω. Hence, as $\mathrm{Lip}(h_\omega)$ is uniformly bounded, it follows that $\|\tilde{\mathcal{L}}^n_\omega(f) - \mu_\omega(f)\|_\infty \ll \vartheta^n \mathrm{Lip}(f)$. The exponential mixing along fibers follows from this.

Hölder continuity of μ_ω. Note that Kantorovich's duality for the Wasserstein metric W on probability measures implies that the condition in Definition 1.5 is equivalent to Lipschitz continuity with respect to W. So assume that f is a Lipschitz function on X with $\mathrm{Lip}(f) \leqslant 1$ and $\inf_{x \in X_\omega} f(x) = 0$. Then $\|f\|_\infty \leqslant \mathrm{Lip}(f)\,\mathrm{diam}(X)$ and

$$\int f\,\mathrm{d}(\mu_\omega - \mu_{\omega'}) = \int f(h_\omega - h_{\omega'})\,\mathrm{d}\nu_\omega + \int fh_{\omega'}\,\mathrm{d}(\nu_\omega - \nu_{\omega'})$$

$$\leqslant \|h_\omega - h_{\omega'}\|_\infty \int f\,\mathrm{d}\nu_\omega + \mathrm{Lip}(fh_{\omega'})W(\nu_\omega, \nu_{\omega'})$$

$$\leqslant \mathrm{diam}(X)\|h_\omega - h_{\omega'}\|_\infty + (\mathrm{Lip}(h_{\omega'})\,\mathrm{diam}(X) + \|h_{\omega'}\|_\infty)W(\nu_\omega, \nu_{\omega'})$$

$$\ll \|h_\omega - h_{\omega'}\|_\infty + \|h_{\omega'}\|_{\mathrm{Lip}}W(\nu_\omega, \nu_{\omega'}).$$

As $h_{\omega'} \in \Lambda_{\omega'}$ and $\int h_{\omega'}\,\mathrm{d}\nu_{\omega'} = 1$, the Lipschitz norm $\|h_\omega\|_{\mathrm{Lip}}$ is uniformly bounded. So it remains to show that $\omega \to h_\omega$ is Hölder with respect to the sup-norm and that ν_ω is Hölder with respect to W, which essentially is a corollary of the above and the following estimate.

We recall that constants δ, L and C_T have been introduced respectively in (H2), (H4) and (H7).

Lemma 2.4. *There exists $A \geqslant 1$ such that for $n \in \mathbb{N}$, $\omega, \omega' \in \Omega$ with $A^n d(\omega, \omega') < 1$ (or $d(\theta^n\omega, \theta^n\omega')A^n < 1$, respectively) and $f \in \Lambda_\omega \cap \Lambda_{\omega'}$,*

$$\|\log\mathcal{L}^n_\omega(f) - \log\mathcal{L}^n_{\omega'}(f)\|_\infty \ll \begin{cases} A^n d(\omega, \omega') & \text{if } d(\omega, \omega')A^n < 1, \\ A^n d(\theta^n\omega, \theta^n\omega') & \text{if } d(\theta^n\omega, \theta^n\omega')A^n < 1. \end{cases}$$

Proof. The estimate relies on the control of the distances between the preimages of T_ω^k and $T_{\omega'}^k$, where $k = 1, \ldots, n$ and ω and ω' are sufficiently close.

So assume that $x, y, y' \in X$ and $\omega, \omega' \in \Omega$ with $T_\omega(x) = y$, $d(y, y') < \delta$, and $d(\omega, \omega') < C_T^{-1}\delta$. As $d(y, y') < \delta$, $z := (T_\omega|_{U_x^\omega})^{-1}(y')$ is well defined and $d(x, z) \leqslant Ld(y, y')$ by (H4). It now follows from (H7) that $d(T_\omega(z), T_{\omega'}(z)) < \delta$. Hence, $x' := (T_{\omega'}|_{U_z^{\omega'}})^{-1}(y')$ is also well defined and

$$
(2.8) \quad
\begin{aligned}
d(x, x') &\leqslant d(x, z) + d(z, x') \leqslant Ld(T_\omega x, T_\omega z) + Md(T_{\omega'}z, T_{\omega'}x') \\
&= Ld(y, y') + Ld(T_{\omega'}(z), T_\omega z) \leqslant Ld(y, y') + LC_T d(\omega', \omega).
\end{aligned}
$$

Note that (H3) implies that $T_{\omega'}|_{B(x, C^{-1}\delta)}$ is injective. In particular, there is at most one $x' \in X$ with $T_{\omega'}(x') = y'$ and $d(x, x') < C^{-1}\delta$. Hence, if $d(x, x') < C^{-1}\delta$, then x' is unique.

Now assume that $n \in \mathbb{N}$ and $x, y \in X$ with $T_\omega^n(x) = y$ are given. By iterating the above construction of preimages, one then obtains a unique $x' \in X$ with $T_{\omega'}^n(x') = y$ whose ω'-orbit stays close to the ω-orbit of x until time n, provided that $d(\omega, \omega')$ is sufficiently small. Namely, it easily follows by induction that this happens whenever $d(\theta^k\omega, \theta^k\omega') < (CC_T)^{-1}\delta(2L)^{-k-1}$ for $k = 0, \ldots n - 1$. As $d(\theta^k\omega, \theta^k\omega') \leqslant \mathrm{Lip}(\theta)^k d(\omega, \omega')$, there exists $A \geqslant 1$ such that ω and ω' are sufficiently close if $d(\omega, \omega') < A^{-n}$.

Furthermore, if $d(\omega, \omega') < A^{-n}$, it follows from (H8), (H3) and again by induction from (2.8) and eventually enlarging A that

$$
(2.9)
\begin{aligned}
\left| \sum_{k=0}^{n-1} \varphi_{\theta^k\omega}(T_\omega^k(x)) - \varphi_{\theta^k\omega'}(T_{\omega'}^k(x')) \right| &\leqslant \sum_{k=0}^{n-1} C_\varphi d(\theta^k\omega, \theta^k\omega') + \sum_{k=0}^{n-1} Cd(T_\omega^k(x), T_{\omega'}^k(x')) \\
&\leqslant C_\varphi \sum_{k=0}^{n-1} \mathrm{Lip}(\theta)^k d(\omega, \omega') + CC_T \sum_{k=0}^{n-1} \sum_{\ell=k}^{n-1} L^{n-\ell} d(\theta^\ell\omega, \theta^\ell\omega') \\
&\ll A^n d(\omega, \omega').
\end{aligned}
$$

As there is a one-to-one relation between the preimages of T_ω^n and $T_{\omega'}^n$, it follows from (2.9) by a standard argument, for $f \in \Lambda_\omega \cap \Lambda_{\omega'}$, that

$$
\begin{aligned}
|\mathcal{L}_\omega^n(f)(x) - \mathcal{L}_{\omega'}^n(f)(x)| &\ll A^n d(\omega, \omega')(\mathcal{L}_\omega^n(f)(x) + \mathrm{Lip}(\log f)\mathcal{L}_{\omega'}^n(f)(x)) \\
&\ll A^n d(\omega, \omega')(\mathcal{L}_\omega^n(f)(x) + \mathcal{L}_{\omega'}^n(f)(x)),
\end{aligned}
$$

where we have used that $\mathrm{Lip}(\log f)$ is uniformly bounded as $f \in \Lambda_{\omega'}$. Furthermore, observe that (2.9), $d(\omega, \omega')A^n < 1$ and the uniform bound on $\mathrm{Lip}(\log f)$ imply that $\mathcal{L}_\omega^n(f)(x) \ll \mathcal{L}_{\omega'}^n(f)(x) \ll \mathcal{L}_\omega^n(f)(x)$. Hence, the first assertion of the lemma follows by dividing the last estimate by $\mathcal{L}_\omega^n(f)(x)$. The second part of Lemma 2.4 follows by precisely the same arguments. $\qquad\square$

By combining Lemma 2.4 with estimate (2.5), one obtains for ω, ω' sufficiently close, $x \in X$ and $f \in \Lambda_\omega \cap \Lambda_{\omega'}$ and $\nu_\omega^n(f) := \mathcal{L}_\omega^n(f)(x)/\mathcal{L}_\omega^n(\mathbf{1})(x)$ that

$$
(2.10)
\begin{aligned}
|\log \nu_\omega(f) - \log \nu_{\omega'}(f)| &\leqslant |\log \nu_\omega(f) - \log \nu_\omega^n(f)| + |\log \mathcal{L}_\omega^n(f)(x) - \log \mathcal{L}_{\omega'}^n(f)(x)| \\
&\quad + |\log \mathcal{L}_\omega^n(\mathbf{1})(x) - \log \mathcal{L}_{\omega'}^n(\mathbf{1})(x)| + |\log \nu_{\omega'}^n(f) - \log \nu_{\omega'}(f)| \\
&\ll \vartheta^n + A^n d(\omega, \omega').
\end{aligned}
$$

For t given by $d(\omega, \omega') = A^{-t}$, $n := \lfloor t \log A / (\log A - \log \vartheta) \rfloor$, it then follows that

$$|\log \nu_\omega(f) - \log \nu_{\omega'}(f)| \ll d(\omega, \omega')^{\frac{-\log \vartheta}{\log A - \log \vartheta}} = d(\omega, \omega')^\alpha,$$

where $\alpha := -\log \vartheta / (\log A - \log \vartheta)$. By repeating the approximation argument in (2.6), one obtains that $|\nu_\omega(f) - \nu_{\omega'}(f)| \ll \mathrm{Lip}(f) d(\omega, \omega')^\alpha$ and by Kantorovich's duality that $W(\nu_\omega, \nu_{\omega'}) \ll d(\omega, \omega')^\alpha$. With respect to h_ω, estimate (2.7), Lemma 2.4, estimate (2.10), a further application of Lemma 2.4 and (2.7) (in this order) imply that

$$|\log h_\omega(x) - \log h_{\omega'}(x)| \leqslant \left| \log \frac{h_\omega(x)}{h_\omega^n(x)} \right| + \left| \log \frac{\mathcal{L}_{\theta^{-n}\omega}^n(\mathbf{1})(x)}{\mathcal{L}_{\theta^{-n}\omega'}^n(\mathbf{1})(x)} \right| + \left| \log \frac{\nu_\omega(\mathcal{L}_{\theta^{-n}\omega}^n(\mathbf{1}))}{\nu_{\omega'}(\mathcal{L}_{\theta^{-n}\omega}^n(\mathbf{1}))} \right|$$

$$+ \left| \log \frac{\nu_{\omega'}(\mathcal{L}_{\theta^{-n}\omega}^n(\mathbf{1}))}{\nu_{\omega'}(\mathcal{L}_{\theta^{-n}\omega'}^n(\mathbf{1}))} \right| + \left| \log \frac{h_{\omega'}(x)}{h_{\omega'}^n(x)} \right|$$

$$\ll \vartheta^n + A^n d(\omega, \omega') + d(\omega, \omega')^\alpha + A^n d(\omega, \omega') + \vartheta^n.$$

With respect to the same choice of t and n as above, it follows that $\|h_\omega - h_{\omega'}\|_\infty \ll d(\omega, \omega')^\alpha$. This concludes the proof of Proposition 2.3, except for the fact that θ is exponentially 4-mixing with respect to Hölder functions, which is proved in the next lemma. $\square$

Lemma 2.5. *Let $\theta : \Omega \to \Omega$ be a map on a compact metric space which is exponentially 4-mixing for Lipschitz functions. Then it is also exponentially 4-mixing with respect to Hölder-continuous functions of any given positive exponent.*

This fact is not specific to exponential 4-mixing: it works for any kind of mixing rate. It follows from a standard interpolation argument, approximating Hölder-continuous functions with Lipschitz ones. This is a standard fact on Riemannian manifolds using mollifiers as we discussed above in the construction of ν_ω, but it works in any metric space as we explain now.

Lemma 2.6. *Let (Ω, d) be a metric space, and $\alpha > 0$, $\kappa > 0$. For any α-Hölder-continuous $f : \Omega \to \mathbb{R}$, there exists a Lipschitz function $f^* : \Omega \to \mathbb{R}$ with $\|f - f^*\|_\infty \leqslant 2\mathrm{Hol}_\alpha(f)\kappa^\alpha$ and $\mathrm{Lip}(f^*) \leqslant \kappa^{-(1-\alpha)}\mathrm{Hol}_\alpha(f)$, where*

$$\mathrm{Hol}_\alpha(f) = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)^\alpha}$$

is the best α-Hölder constant of f.

Proof. Let A be a maximal κ-separated set in Ω. Let also $M = \mathrm{Hol}_\alpha(f)\kappa^{-(1-\alpha)}$. The restriction of f to A satisfies, for $x \neq y$, the inequality

$$|f(x) - f(y)| \leqslant \mathrm{Hol}_\alpha(f) d(x, y)^\alpha \leqslant M d(x, y),$$

as $d(x, y) \geqslant \kappa$. Define f^* on Ω by

$$f^*(x) = \inf_{y \in A} f(y) + M d(x, y).$$

The previous inequality ensures that this function coincides with f on A, and that it is M-Lipschitz globally. Let us check that $\|f - f^*\|_\infty \leqslant 2\mathrm{Hol}_\alpha(f)\kappa^\alpha$. Take $x \in \Omega$. By maximality of A, there exists $y \in A$ with $d(x, y) \leqslant \kappa$. Then, as $f(y) = f^*(y)$, we get

$$|f(x) - f^*(x)| \leqslant |f(x) - f(y)| + |f^*(x) - f^*(y)| \leqslant \mathrm{Hol}_\alpha(f)\kappa^\alpha + \mathrm{Lip}(f^*)\kappa \leqslant 2\mathrm{Hol}_\alpha(f)\kappa^\alpha. \ \square$$

Proof of Lemma 2.5. Start from α-Hölder continuous functions f_1, f_2, g_1, g_2 and $a \leqslant b \leqslant c$ for which one wants to prove (1.2) with the Lipschitz norm replaced by the Hölder norm. Let $\kappa = e^{-\varepsilon n}$ with ε suitable small, and apply Lemma 2.5 to get new functions $f_1^*, f_2^*, g_1^*, g_2^*$. Replacing each f_i, g_i with its starred version in (1.2) introduces an error controlled by $e^{-\alpha\varepsilon n}$, and therefore exponentially small. Thanks to (1.2) for Lipschitz functions, the remaining difference for the starred functions is bounded by

$$Ce^{-cn}\|f_1^*\|_{\mathrm{Lip}}\|f_2^*\|_{\mathrm{Lip}}\|g_1^*\|_{\mathrm{Lip}}\|g_2^*\|_{\mathrm{Lip}} \leqslant Ce^{-cn}e^{4(1-\alpha)\varepsilon n}\|f_1^*\|_{\mathrm{Hol}_\alpha}\|f_2^*\|_{\mathrm{Hol}_\alpha}\|g_1^*\|_{\mathrm{Hol}_\alpha}\|g_2^*\|_{\mathrm{Hol}_\alpha},$$

which is exponentially small if ε was chosen small enough at the beginning of the argument. $\qquad\square$

Remark 2.7. With respect to the proof of Proposition 2.3, we would like to remark that probably all of the arguments are well known but that they had to be adapted to our situation in order to prove Hölder continuity of $\omega \mapsto \mu_\omega$ and mixing with respect to Lipschitz functions (instead of functions in $C^1(X)$) along the fibers. In here, it turned out to be advantageous to first construct the family of conformal measures and thereafter the invariant functions for the family of transfer operators as the regularity of $\omega \mapsto h_\omega$ and $\omega \mapsto \mu_\omega$ are consequences of the regularity of $\omega \mapsto \nu_\omega$. Moreover, as the arguments essentially depend on the existence of the cone field, they probably can be easily adapted to other settings. It is also worth noting that we did not make use of all features of the cone field as we never touched the convergence of the logarithmic derivatives, which is provided by the definition of the cones.

Furthermore, we also would like to draw attention to [1, 2], where the authors studied similar cones adapted to random interval transformations. In there, the cones are defined through the BV-norm instead of the $C^1(X)$ norm. In particular, provided that $\log \mathcal{L}^n$ is sufficiently regular (see Lemma 2.4), the above proof is applicable in verbatim.

We also would like to point out that the Hölder continuity of ν_ω with respect to ω was obtained in [6] in a uniformly expanding setting for the more general setting of fibered systems.

3. Proofs

3.1. **Preliminaries.** Given a space with bounded local complexity, consider for each r a finite sequence $x_1^{(r)}, \ldots, x_{k(r)}^{(r)}$ of points as in the definition, such that the space is covered by the balls $B(x_p^{(r)}, r)$ and such that no point is in more than C_0 balls $B(x_p^{(r)}, 4r)$. Fix also functions $\rho_p^{(r)}$ supported around $x_p^{(r)}$, which are equal to 1 on $B(x_p^{(r)}, 2r)$, to 0 outside of $B(x_p^{(r)}, 4r)$, and take values in $[0,1]$. For instance, one can take $\rho_p^{(r)}(x) = \rho(d(x, x_p)/r))$ where $\rho : \mathbb{R} \to \mathbb{R}$ is equal to 1 on $(-\infty, 2]$, to 0 on $[4, \infty)$ and affine in between. With this specific choice, one has a Lipschitz control

$$(3.1) \qquad\qquad \|\rho_p^{(r)}\|_{\mathrm{Lip}} \leqslant \frac{1}{r}$$

that will prove useful later.

The main point of these definitions is that one can approximate $(x, y) \mapsto 1_{d(x,y) \leqslant r}$ by a sum of functions $\rho_p^{(r)}$, to which we will be able to apply mixing arguments:

Lemma 3.1. *For any $x, y \in X$,*

$$(3.2) \qquad 1_{d(x,y)\leq r} \leqslant \sum_{p=1}^{k(r)} \rho_p^{(r)}(x)\rho_p^{(r)}(y) \leqslant C_0 1_{d(x,y)\leq 8r}.$$

Proof. Assume $d(x,y) \leq r$. There is some $p = p(x,r)$ such that $x \in B(x_p^{(r)}, r)$, as these balls cover the space. Then $y \in B(x,r) \subseteq B(x_p^{(r)}, 2r)$. Therefore, $\rho_p^{(r)}(x) = \rho_p^{(r)}(y) = 1$, proving the left inequality.

Conversely, in the sum $\sum_p \rho_p^{(r)}(x)\rho_p^{(r)}(y)$, a term can only be nonzero if $d(x, x_p) \leq 4r$. There are at most C_0 such values of p, by definition of bounded local complexity. For each such p, the factor $\rho_p^{(r)}(y)$ can only be nonzero if $d(y, x_p^{(r)}) \leqslant 4r$, which implies $d(x,y) \leqslant 8r$. This proves the right inequality. $\qquad\square$

This lemma makes it possible to express the different correlation dimensions in terms of the discretization $\rho_p^{(r)}$:

Lemma 3.2. *One has*

$$\limsup_{r\to 0} \frac{\log\left(\sum_{p=1}^{k(r)}\left(\int \rho_p^{(r)}\,\mathrm{d}\mu\right)^2\right)}{\log r} = \overline{D}_2^{\mathrm{an}},$$

and

$$\limsup_{r\to 0} \frac{\log\left(\sum_{p=1}^{k(r)}\int\left(\int \rho_p^{(r)}\,\mathrm{d}\mu_\omega\right)^2\mathrm{d}\mathbb{P}(\omega)\right)}{\log r} = \overline{D}_2^{\mathrm{qu}},$$

Using liminfs instead, similar equations hold for $\underline{D}_2^{\mathrm{an}}$ and $\underline{D}_2^{\mathrm{qu}}$.

Proof. We claim that, for any probability measure η on X,

$$\int \eta(B(x,r))\,\mathrm{d}\eta(x) \leqslant \sum_{p=1}^{k(r)}\left(\int \rho_p^{(r)}\,\mathrm{d}\eta\right)^2 \leqslant C_0 \int \eta(B(x,8r))\,\mathrm{d}\eta(x).$$

This follows from integrating the inequalities in Lemma 3.1 with respect to $\eta^{\otimes 2}$.

Applying these inequalities to the measures μ or μ_ω, the lemma follows readily from the definitions of the correlation dimensions. $\qquad\square$

Lemma 3.3. *Let $k = k(r)$ be as in Definition 1.4. Then, for small enough r, one has $k(r) \leqslant r^{-C_0'}$ for $C_0' = 4\log C_0$.*

Proof. Let us show that

$$(3.3) \qquad k(r) \leqslant C_0 k(2r).$$

Since the balls $(B(x_q^{(2r)}))_{q\leqslant k(2r)}$ cover the space, any point $x_p^{(r)}$ belongs to one of these balls, for some $q = q(p)$. This defines a map $\{1,\ldots,k(r)\} \to \{1,\ldots,k(2r)\}$. Moreover, each q has at most C_0 preimages, by definition of the bounded local complexity. This shows (3.3).

We deduce that $k(r) \leqslant C_0^n k(2^n r)$. Take n so that $2^n r$ is of the order of magnitude of 1 (so that $k(2^n r)$ is bounded), e.g., $n = -\lfloor \log r / \log 2 \rfloor$. This gives

$$k(r) \leqslant C C_0^{-\log r / \log 2} = C r^{-\log C_0 / \log 2}.$$

As $4 \log C_0 > \log C_0 / \log 2$, the conclusion follows. $\square$

All the forthcoming proofs will be based on the same scheme. We will fix a small enough r, and use the functions $\rho_p^{(r)}$, omitting the superscript $^{(r)}$ for readability. We will need to compute the first or second moment of some functions, to apply Markov inequalities to estimate ultimately the probability that m_n is smaller or larger than r, and then conclude with a Borel-Cantelli argument. Depending on the precise quantity to be estimated, we will need stronger or weaker mixing conditions. The following notations will be used throughout:

- We define a function $R_p : \Omega \to \mathbb{R}$ by

$$(3.4) \qquad R_p(\omega) = \int \rho_p \, d\mu_\omega.$$

  As ρ_p is Lipschitz with $\|\rho_p\|_{\mathrm{Lip}} \leqslant 1/r$ by (3.1), the assumption that $\omega \mapsto \mu_\omega$ is Lipschitz ensures that R_p is also Lipschitz, with

$$(3.5) \qquad \|R_p\|_{\mathrm{Lip}} \leqslant C_2/r.$$

  With this notation, the second part of Lemma 3.2 can be reformulated as:

$$(3.6) \qquad \limsup_{r \to 0} \frac{\log\left(\sum_p \int R_p^2 \, d\mathbb{P} \right)}{\log r} = \overline{D}_2^{\mathrm{qu}},$$

  and similarly for the liminf.
- Both the base map θ and the fiber map T_ω are mixing stretched exponentially, with respective constants (C_2, c_2) and (C_3, c_3). We define a function

$$\varphi(n) = C_\varphi \exp(-n^{c_\varphi}),$$

  with $C_\varphi = \max(C_2, C_3)$ and $c_\varphi = \min(c_2, c_3)$, that bounds from above the mixing rate of both maps. Its main property is that, with our choice of the gap function α in (1.3), one has

$$(3.7) \qquad \varphi(\alpha(n)) \leqslant \frac{C}{n^{\log n}}.$$

  In particular, $\varphi(\alpha(n))$ tends to zero faster than any polynomial, which is the property we will use below.

3.2. **Upper bounds.** In this paragraph, we prove Propositions 1.9 and 1.10, giving upper bounds for $-\log m_n^{\leqslant}$ and $-\log m_n^{>}$, i.e., showing that respectively on-diagonal and off-diagonal distances along orbits cannot be too small. The probability that they are too small will be estimated thanks to a first moment computation, and we will conclude with a Borel-Cantelli argument. These arguments are pretty soft, as testified by the fact that the assumptions in these propositions are much milder than for the corresponding lower bounds.

Proof of Proposition 1.9 on on-diagonal upper bounds. First of all, we observe that if $\underline{D}_2^{\mathrm{qu}} = 0$, the theorem is empty. Assume now that $\underline{D}_2^{\mathrm{qu}} > 0$. Let $\varepsilon < \underline{D}_2^{\mathrm{qu}}$. Fix a small enough r.

Define a function

$$S_n^{\leq}(\omega; x, y) = \sum_{\substack{i,j < n \\ |j-i| \leqslant \alpha(n)}} \sum_p \rho_p(T_\omega^i x) \rho_p(T_\omega^j y),$$

where $\rho_p = \rho_p^{(r)}$ is the discretization at scale r.

If $m_n^{\leq}(\omega; x, y) \leq r$, then there are two indices i, j such that $d(T_\omega^i x, T_\omega^j y) \leq r$. Then $S_n^{\leq}(\omega; x, y) \geq 1$ by Lemma 3.1. With Markov's inequality, we obtain

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^{\leq}(\omega; x, y) \leq r\}$$
$$\leq \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : S_n^{\leq}(\omega; x, y) \geq 1\} \leq \mathbb{E}(S_n^{\leq}).$$

Let us now compute this expectation. We have

$$\mathbb{E}(S_n^{\leq}) = \sum_p \sum_{|j-i| \leq \alpha(n)} \int \rho_p(T_\omega^i x) \rho_p(T_\omega^j y) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega)$$
$$= \sum_p \sum_{|j-i| \leq \alpha(n)} \int \left(\int \rho_p \, \mathrm{d}\mu_{\theta^i \omega} \right) \left(\int \rho_p \, \mathrm{d}\mu_{\theta^j \omega} \right) \, \mathrm{d}\mathbb{P}(\omega),$$

since $T_\omega^i x$ is distributed according to $\mu_{\theta^i \omega}$ when x is distributed according to μ_ω, by equivariance. We recall the notation R_p introduced in (3.4). Applying the Cauchy-Schwarz inequality and then using the invariance of $\mathbb{P}$ under θ, we obtain

$$\mathbb{E}(S_n^{\leq}) \leq \sum_p \sum_{|j-i| \leq \alpha(n)} \left[\int R_p(\theta^i \omega)^2 \, \mathrm{d}\mathbb{P}(\omega) \right]^{1/2} \left[\int R_p(\theta^j \omega)^2 \, \mathrm{d}\mathbb{P}(\omega) \right]^{1/2}$$
$$= \sum_p \sum_{|j-i| \leq \alpha(n)} \int R_p(\omega)^2 \, \mathrm{d}\mathbb{P}(\omega).$$

The sum over i, j reduces to $2n \cdot \alpha(n)$. Thanks to (3.6), if r is small enough one has $\sum_p \int R_p(\omega)^2 \, \mathrm{d}\mathbb{P}(\omega) \leqslant r^{\underline{D}_2^{\mathrm{qu}} - \varepsilon}$. We have obtained

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^{\leq}(\omega; x, y) \leq r\} \leq 2n\alpha(n) r^{\underline{D}_2^{\mathrm{qu}} - \varepsilon}.$$

Take now $r = r_n = 1/n^{(1+2\varepsilon)/(\underline{D}_2^{\mathrm{qu}} - \varepsilon)}$. As $\alpha(n) = O(n^\varepsilon)$, the previous bound gives

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^{\leq}(\omega; x, y) \leq r_n\} = O(1/n^\varepsilon).$$

Choose a subsequence $n_s = \lfloor s^{2/\varepsilon} \rfloor$. Along this subsequence, the error probability is $O(1/s^2)$, which is summable. By Borel-Cantelli, for almost every (ω, x, y), one has eventually $m_{n_s}^{\leq}(\omega; x, y) > r_{n_s}$, and therefore

$$\frac{-\log m_{n_s}^{\leq}(\omega; x, y)}{\log n_s} \leq \frac{1 + 2\varepsilon}{\underline{D}_2^{\mathrm{qu}} - \varepsilon}.$$

As $-\log m_{\overline{n}}^{\leq}(\omega; x, y)$ is a non-decreasing function of n and $\log n_{s+1}/\log n_s \to 1$, this inequality along the subsequence n_s passes to the whole sequence. We obtain almost surely

$$\limsup \frac{-\log m_{\overline{n}}^{\leq}(\omega; x, y)}{\log n} \leq \frac{1+2\varepsilon}{\underline{D}_2^{\mathrm{qu}}-\varepsilon}.$$

As ε is arbitrary, this proves the result. $\qquad\square$

Proof of Proposition 1.10 on off-diagonal upper bounds. We follow the same strategy as in the previous proof. The result is obvious if $\underline{D}_2^{\mathrm{an}} = 0$. Assume that $\underline{D}_2^{\mathrm{an}} > 0$. Let $\varepsilon < \underline{D}_2^{\mathrm{an}}$. Fix a small enough r.

Define a function

$$S_n^{>}(\omega; x, y) = \sum_{\substack{i,j<n \\ |j-i|>\alpha(n)}} \sum_p \rho_p(T_\omega^i x)\rho_p(T_\omega^j y).$$

As above, we have

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega\{(\omega, x, y) \ : m_n^{>}(\omega; x, y) \leq r\}$$
$$\leq \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega\{(\omega, x, y) \ : \ S_n^{>}(\omega; x, y) \geq 1\} \leq \mathbb{E}(S_n^{>}).$$

Moreover,

$$\mathbb{E}(S_n^{>}) = \sum_{\substack{i,j<n \\ |j-i|>\alpha(n)}} \sum_p \int R_p(\theta^i \omega) R_p(\theta^j \omega) \, \mathrm{d}\mathbb{P}(\omega),$$

as in the proof of Proposition 1.9.

As the speed of mixing of θ is at least φ by definition, we have

$$\int R_p(\theta^i \omega) R_p(\theta^j \omega) \, \mathrm{d}\mathbb{P}(\omega) = \left(\int R_p(\omega) \, \mathrm{d}\mathbb{P}(\omega)\right)^2 + O(\varphi(\alpha(n))r^{-2}),$$

as R_p has a Lipschitz norm bounded by C_2/r (see (3.5)), and i and j are separated by at least $\alpha(n)$.

Summing over i, j (there are less than n^2 of them) and then p (there are $k(r)$ of them), we get

$$\mathbb{E}(S_n^{>}) \leqslant n^2 \sum_p \left(\int R_p(\omega) \, \mathrm{d}\mathbb{P}(\omega)\right)^2 + Cn^2 k(r)\varphi(\alpha(n))r^{-2}.$$

Note that $\int R_p(\omega) \, \mathrm{d}\mathbb{P}(\omega) = \int \rho_p \, \mathrm{d}\mu$. By Lemma 3.2, we have for small enough r

$$\sum_p \left(\int \rho_p \, \mathrm{d}\mu\right)^2 \leqslant r^{\underline{D}_2^{\mathrm{an}}-\varepsilon}.$$

Moreover, $k(r) \leqslant r^{-C_0'}$ by Lemma 3.3. Finally,

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega\{(\omega, x, y) \ : \ m_n^{>}(\omega; x, y) \leq r\} \leq n^2 r^{\underline{D}_2^{\mathrm{an}}-\varepsilon} + Cn^2 r^{-(C_0'+2)}\varphi(\alpha(n)).$$

Take now $r = r_n = 1/n^{(2+\varepsilon)/(\underline{D}_2^{\mathrm{an}}-\varepsilon)}$. As $\varphi(\alpha(n)) \leqslant C/n^{\log n}$ by (3.7), this gives

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega\{(\omega, x, y) \ : \ m_n^{>}(\omega; x, y) \leq r_n\} \leq 1/n^\varepsilon + C' n^{C(\varepsilon)}/n^{\log n}.$$

For large enough n, this is bounded by $2/n^\varepsilon$.

Choose a subsequence $n_s = \lfloor s^{2/\varepsilon} \rfloor$. Along this subsequence, the error probability is $O(1/s^2)$, which is summable. By Borel-Cantelli, for almost every (ω, x, y), one has eventually $m_{n_s}^{>}(\omega; x, y) > r_{n_s}$, and therefore

$$\frac{-\log m_{n_s}^{>}(\omega; x, y)}{\log n_s} \leq \frac{2 + \varepsilon}{\underline{D}_2^{\mathrm{an}} - \varepsilon}.$$

Contrary to the proof of Proposition 1.9, we can not argue from there by monotonicity, as $-\log m_n^{>}$ is not a non-decreasing function of n (we are considering more times, but the constraint $|j - i| > \alpha(n)$ becomes stronger when n increases). What is true, though, is that for all $n \in [n_s, n_{s+1}]$, one has

$$-\log m_n^{>}(\omega; x, y) \leq -\log \min_{\substack{i,j < n_{s+1} \\ |j-i| > \alpha(n_s)}} d(T_\omega^i x, T_\omega^j y) =: -\log m'_{n_{s+1}}.$$

One can show exactly as above that, almost surely, eventually,

$$\frac{-\log m'_{n_s}(\omega; x, y)}{\log n_s} \leq \frac{2 + \varepsilon}{\underline{D}_2^{\mathrm{an}} - \varepsilon}.$$

With the previous equation, this inequality passes to the whole sequence $m_n^{>}$. We obtain almost surely

$$\limsup \frac{-\log m_n^{>}(\omega; x, y)}{\log n} \leq \frac{2 + \varepsilon}{\underline{D}_2^{\mathrm{an}} - \varepsilon}.$$

As ε is arbitrary, this proves the result. $\qquad\square$

Remark 3.4. Stretched exponential mixing is not essential for the proof of Proposition 1.10: using a gap size $\alpha(n) = n^\delta$ for an arbitrarily small δ instead of $\alpha(n) = (\log n)^{C_4}$ as we did, the proof goes through if the mixing speed of θ is faster than any polynomial. Therefore, (1.4) holds under this assumption. The stretched exponential mixing is however necessary in our argument for the lower bound for m_n (see the proof of Proposition 1.12).

3.3. **Lower bounds.** In this paragraph, we prove Propositions 1.11 and 1.12, giving lower bounds for $-\log m_n^0$ and $-\log m_n^{\gg}$, i.e., showing that respectively on-diagonal and off-diagonal distances along orbits can not be too large. The probability that they are too large will be estimated thanks to a second moment computation: we will control the average of a suitable function, and its variance to show that this function can not deviate much from its average. We will then conclude with a Borel-Cantelli argument. These arguments are more technical than the ones for the corresponding upper bounds, as we need estimates on moments of order two instead of one (and therefore stronger mixing assumptions to be able to deal with the more involved terms that show up).

Proof of Proposition 1.11 on on-diagonal lower bounds. Let $\varepsilon > 0$. Fix a small enough r.
Define a function
$$S_n^0(\omega; x, y) = \sum_{i < n} \sum_p \rho_p(T_\omega^i x) \rho_p(T_\omega^i y),$$

where $\rho_p = \rho_p^{(r)}$ is the discretization at scale r.

If $m_n^0(\omega; x, y) > 8r$, then for all i one has $d(T_\omega^i x, T_\omega^i y) > 8r$. Then $S_n^0(\omega; x, y) = 0$ by Lemma 3.1. We obtain

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^0(\omega; x, y) > 8r\}$$
$$\leq \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : S_n^0(\omega; x, y) = 0\}.$$

When $S_n^0 = 0$, then $(S_n^0 - \mathbb{E}(S_n^0))^2 / \mathbb{E}(S_n^0)^2 = 1$. With Markov inequality, this gives

$$(3.8) \qquad \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^0(\omega; x, y) > 8r\} \leqslant \frac{\mathrm{var}(S_n^0)}{\mathbb{E}(S_n^0)^2} = \frac{\mathbb{E}((S_n^0)^2) - \mathbb{E}(S_n^0)^2}{\mathbb{E}(S_n^0)^2}.$$

We have

$$\mathbb{E}(S_n^0) = \sum_p \sum_i \int \rho_p(T_\omega^i x) \rho_p(T_\omega^i y) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega)$$
$$= \sum_p \sum_i \int R_p(\theta^i \omega) \cdot R_p(\theta^i \omega) \, \mathrm{d}\mathbb{P}(\omega)$$
$$= n \sum_p \int R_p(\omega)^2 \, \mathrm{d}\mathbb{P}(\omega),$$

by θ-invariance of $\mathbb{P}$. Therefore, its asymptotic behavior is described by (3.6).

Let us now estimate $\mathbb{E}((S_n^0)^2)$. Expanding the square, we get

$$\mathbb{E}((S_n^0)^2) = \sum_{i,i'} \sum_{p,q} \int \rho_p(T_\omega^i x) \rho_p(T_\omega^i y) \rho_q(T_\omega^{i'} x) \rho_q(T_\omega^{i'} y) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega).$$

We split the sum according to whether $|i' - i| \leq \alpha(n)$ or $|i' - i| > \alpha(n)$. In the former case, we will use a crude upper bound, and in the latter we will use mixing. Let us fix i, i'.

We have $\sum_q \rho_q(T_\omega^{i'} x) \rho_q(T_\omega^{i'} y) \leqslant C_0$, by (3.2). Therefore, we get a bound

$$C_0 \sum_p \int \rho_p(T_\omega^i x) \rho_p(T_\omega^i y) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega) = C_0 \sum_p \int R_p(\omega)^2 \, \mathrm{d}\mathbb{P}(\omega).$$

We will only use this crude bound when $|i' - i| \leq \alpha(n)$. Therefore, we will get $2n\alpha(n)$ of them.

Assume now $|i' - i| > \alpha(n)$. Then

$$\int \rho_p(T_\omega^i x) \rho_q(T_\omega^{i'} x) \, \mathrm{d}\mu_\omega(x) = \left(\int \rho_p \, \mathrm{d}\mu_{\theta^i \omega} \right) \left(\int \rho_q \, \mathrm{d}\mu_{\theta^{i'} \omega} \right) + O(\varphi(\alpha(n)) \|\rho_p\|_{\mathrm{Lip}} \|\rho_q\|_{\mathrm{Lip}}),$$

by fiberwise mixing (Definition 1.8). As $\|\rho_p\|_{\mathrm{Lip}} \leqslant 1/r$, the contribution of the error terms to $\mathbb{E}((S_n^0)^2)$ is bounded by

$$(3.9) \qquad C \sum_{i,i'} \sum_{p,q} \varphi(\alpha(n)) r^{-2} \leqslant C n^2 r^{-2C_0' - 2} \varphi(\alpha(n)),$$

as there are n possible values of i, i', and $k(r) \leq r^{-C_0'}$ possible values of p, q.

Let us use the same estimate for the integral with respect to $\mathrm{d}\mu_\omega(y)$. The remaining term is

$$\sum_{p,q} \sum_{|i'-i|>\alpha(n)} \int R_p(\theta^i \omega)^2 R_q(\theta^{i'} \omega)^2 \, \mathrm{d}\mathbb{P}(\omega).$$

The function R_p is Lipschitz, with $\|R_p\|_{\mathrm{Lip}} \leqslant C_1/r$, see (3.5). As R_p is bounded by 1, it follows that R_p^2 is also Lipschitz, with $\|R_p^2\|_{\mathrm{Lip}} \leqslant 2C_1/r$. One may therefore use the mixing of the base transformation θ, to obtain

$$\int R_p(\theta^i \omega)^2 R_q(\theta^{i'} \omega)^2 \, \mathrm{d}\mathbb{P}(\omega) = \left(\int R_p^2 \, \mathrm{d}\mathbb{P} \right) \left(\int R_q^2 \, \mathrm{d}\mathbb{P} \right) + O(\varphi(\alpha(n)) r^{-2}).$$

The error terms add up exactly as in (3.9).

Adding up all the terms, we are left with

$$\mathbb{E}((S_n^0)^2) \leqslant n^2 \sum_{p,q} \left(\int R_p^2 \, \mathrm{d}\mathbb{P} \right) \left(\int R_q^2 \, \mathrm{d}\mathbb{P} \right) + Cn^2 r^{-2C_0'-2} \varphi(\alpha(n)) + Cn\alpha(n) \left[\sum_p \int R_p^2 \, \mathrm{d}\mathbb{P} \right]$$

$$= \mathbb{E}(S_n^0)^2 + Cn^2 r^{-2C_0'-2} \varphi(\alpha(n)) + C\alpha(n) \mathbb{E}(S_n^0).$$

Together with (3.8), this yields

$$(3.10) \qquad \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) \,:\, m_n^0(\omega; x, y) > 8r\} \leqslant \frac{Cn^2 r^{-2C_0'-2} \varphi(\alpha(n))}{\mathbb{E}(S_n^0)^2} + \frac{C\alpha(n)}{\mathbb{E}(S_n^0)}.$$

For small enough r, the limit (3.5) ensures that

$$\sum_p \int R_p^2 \, \mathrm{d}\mathbb{P} \geqslant r^{\overline{D_2^{\mathrm{qu}}}+\varepsilon}.$$

Take $r = r_n = 1/n^{(1-2\varepsilon)/(\overline{D_2^{\mathrm{qu}}}+\varepsilon)}$. For this value of r, we get

$$\mathbb{E}(S_n^0) = n \sum_p \int R_p^2 \, \mathrm{d}\mathbb{P} \geqslant n \cdot n^{-(1-2\varepsilon)} = n^{2\varepsilon}.$$

As $\varphi(\alpha(n)) = O(1/n^{\log n})$ by (3.7), the first term in (3.10) decays faster than any polynomial. Moreover, as $\alpha(n) = (\log n)^{C_4}$, the second term decays at least like $1/n^\varepsilon$. For large enough n, we obtain

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) \,:\, m_n^0(\omega; x, y) > 8r_n\} \leqslant \frac{1}{n^\varepsilon}.$$

Choose a subsequence $n_s = \lfloor s^{2/\varepsilon} \rfloor$. Along this subsequence, the error probability is $O(1/s^2)$, which is summable. By Borel-Cantelli, for almost every (ω, x, y), one has eventually $m_{n_s}^0(\omega; x, y) \leq 8r_{n_s}$, and therefore

$$\liminf \frac{-\log m_{n_s}^0(\omega; x, y)}{\log n_s} \geq \frac{1-2\varepsilon}{\overline{D_2^{\mathrm{qu}}}+\varepsilon}.$$

By monotonicity of m_n^0, this behavior passes to the whole sequence. As ε is arbitrary, this concludes the proof of the proposition. $\square$

Remark 3.5. As in Remark 3.4, a mixing rate faster than any polynomial would be enough for the proof of Proposition 1.11.

Proof of Proposition 1.12 on off-diagonal lower bounds. Fix a small enough r. In this proof, we will consider indices i or i' that will always be restricted to the range $[0, n/3)$, and indices j or j' that will always be restricted to the range $[2n/3, n)$. Instead of always specifying this, we will use the notation $\sum'$ to enforce these restrictions implicitly. In this proof, an error term (depending on n and r) will be called *admissible* if it is bounded by $r^{-C}u(n)$, for some C and some function u that tends to zero faster than any n^{-D}. A generic admissible error term will be denoted by a_n, keeping r implicit.

Define a function

$$S_n^{\gg}(\omega; x, y) = \sum_{i,j}' \sum_p \rho_p(T_\omega^i x)\rho_p(T_\omega^j y).$$

As in (3.8), we have

$$(3.11) \quad \mathbb{P} \otimes \mu_\omega \otimes \mu_\omega\{(\omega, x, y) \,:\, m_n^{\gg}(\omega; x, y) > 8r\} \leqslant \frac{\mathrm{var}(S_n^{\gg})}{\mathbb{E}(S_n^{\gg})^2} = \frac{\mathbb{E}((S_n^{\gg})^2) - \mathbb{E}(S_n^{\gg})^2}{\mathbb{E}(S_n^{\gg})^2}.$$

Let us first estimate $\mathbb{E}(S_n^{\gg})$. We have

$$\mathbb{E}(S_n^{\gg}) = \sum_{i,j}' \sum_p \int \left(\int \rho_p(T_\omega^i x) \, \mathrm{d}\mu_\omega(x) \right) \left(\int \rho_p(T_\omega^j y) \, \mathrm{d}\mu_\omega(y) \right) \mathrm{d}\mathbb{P}(\omega)$$

$$= \sum_{i,j}' \sum_p \int R_p(\theta^i \omega) R_p(\theta^j \omega) \, \mathrm{d}\mathbb{P}(\omega).$$

By (3.5), the function R_p is Lipschitz, with $\|R_p\|_{\mathrm{Lip}} \leqslant C_2/r$. The mixing of θ gives then

$$\int R_p(\theta^i \omega) R_p(\theta^j \omega) \, \mathrm{d}\mathbb{P}(\omega) = \left(\int R_p \, \mathrm{d}\mathbb{P} \right)^2 + O(\varphi(n/3)r^{-2}),$$

as the gap between i and j is at least $n/3$. Summing over n, and writing

$$(3.12) \qquad A_n = \sum_p \left(\int \rho_p \, \mathrm{d}\mu \right)^2 = \sum_p \left(\int R_p \, \mathrm{d}\mathbb{P} \right)^2, \quad B_n = (n/3)^2 A_n,$$

we get

$$\mathbb{E}(S_n^{\gg}) = (n/3)^2 A_n + O(n^2 \varphi(n/3)r^{-2}) = B_n + O(a_n),$$

where we recall that a_n is a generic admissible error term. Note that B_n also depends on r, but we keep this implicit in the notation. Also, A_n only depends on r and not on n, but since in the end we want to let r depend on n we keep the index n.

Let us now estimate $\mathbb{E}((S_n^{\gg})^2)$, up to an admissible error term. We expand the square, getting 6 indices i, i', j, j', p, q.

$$\mathbb{E}((S_n^{\gg})^2) = \sum_{i,i',j,j'}' \sum_{p,q} \int \rho_p(T_\omega^i x)\rho_q(T_\omega^{i'} x)\rho_p(T_\omega^j y)\rho_q(T_\omega^{j'} y) \, \mathrm{d}\mu_\omega(x) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega).$$

We will split this sum depending on whether i and i' are close, i.e., $|i' - i| \leq \alpha(n)$, or whether they are far, and similarly for j and j'.

Assume first that i and i' are far. Then

$$(3.13) \qquad \int \rho_p(T_\omega^i x)\rho_q(T_\omega^{i'} x) \, \mathrm{d}\mu_\omega(x) = R_p(\theta^i \omega)R_q(\theta^{i'} \omega) + O(\varphi(\alpha(n))r^{-2}),$$

thanks to the fiberwise mixing (Definition 1.8). When adding these error terms over all possible i, i', j, j', p, q, one gets an error at most

$$C\varphi(\alpha(n))r^{-2} \cdot n^4 k(r)^2 \leqslant Cn^4 r^{-2C_0'-2}\varphi(\alpha(n)),$$

which is admissible as $\varphi(\alpha(n)) \leqslant C/n^{\log n}$.

In the same way, when j and j' are far, we can replace $\int \rho_p(T_\omega^j y)\rho_q(T_\omega^{j'} y)\,\mathrm{d}\mu_\omega(y)$ with $R_p(\theta^j \omega)R_q(\theta^{j'}\omega)$, up to an admissible error.

Case 1: when (i, i') are far away and (j, j') are far away.
Up to an admissible error, the contribution of these terms to $\mathbb{E}((S_n^\gg)^2)$ is bounded by

$$\sideset{}{'}\sum_{|i'-i|>\alpha(n),|j'-j|>\alpha(n)} \sum_{p,q} \int R_p(\theta^i \omega)R_q(\theta^{i'}\omega)R_p(\theta^j \omega)R_q(\theta^{j'}\omega)\,\mathrm{d}\mathbb{P}(\omega).$$

There is a gap of $n/3$ between $\max\{i, i'\}$ and $\min\{j, j'\}$. Therefore, using 4-mixing, we can replace the above integral with

$$\left(\int R_p(\theta^i \omega)R_q(\theta^{i'}\omega)\,\mathrm{d}\mathbb{P}(\omega)\right)\left(\int R_p(\theta^j \omega)R_q(\theta^{j'}\omega)\,\mathrm{d}\mathbb{P}(\omega)\right)$$

up to an error of $C\varphi(n/3)r^{-4}$. The sum of these errors over p, q, i, i', j, j' is admissible. Then, using 2-mixing, one can replace $\int R_p(\theta^i \omega)R_q(\theta^{i'}\omega)\,\mathrm{d}\mathbb{P}(\omega)$ with $\int R_p \cdot \int R_q$, up to an error which is again admissible as the gap between i and i' is at least $\alpha(n)$. The same goes for $\int R_p(\theta^j \omega)R_q(\theta^{j'}\omega)\,\mathrm{d}\mathbb{P}(\omega)$.

Finally, up to an admissible error, the contribution of this case to $\mathbb{E}((S_n^\gg)^2)$ is

$$\sideset{}{'}\sum_{|i'-i|>\alpha(n),|j'-j|>\alpha(n)} \sum_{p,q} \left(\int R_p\,\mathrm{d}\mathbb{P}\right)^2\left(\int R_q\,\mathrm{d}\mathbb{P}\right)^2 \leqslant \left((n/3)^2 A_n\right)^2 = B_n^2,$$

as $\int R_p\,\mathrm{d}\mathbb{P} = \int \rho_p\,\mathrm{d}\mu$ by definition of μ. Recall that, with A_n and B_n defined as in (3.12), then B_n is the dominant term in the expansion of $\mathbb{E}(S_n^\gg)$.

Case 2: when (i, i') are far away and (j, j') are close (or conversely).
By symmetry, we assume that (i, i') are far away and (j, j') are close. In this case, after doing the substitution (3.13), we should study

$$(3.14) \qquad \int R_p(\theta^i \omega)R_q(\theta^{i'}\omega)\left(\int \rho_p(T_\omega^j y)\rho_q(T_\omega^{j'} y)\,\mathrm{d}\mu_\omega(y)\right)\mathrm{d}\mathbb{P}(\omega).$$

Assume for instance $j \leq j'$. Then, changing variables with $y' = T_\omega^j y$, and writing $\omega' = \theta^j \omega$, the inner integral becomes

$$F(\omega') = \int \rho_p(y')\rho_q(T_{\omega'}^{j'-j}y')\,\mathrm{d}\mu_{\omega'}(y').$$

We claim that this function F is Lipschitz continuous, with

$$(3.15) \qquad \|F\|_{\mathrm{Lip}} \leqslant CC_S^{\alpha(n)}r^{-2},$$

where $C_S \geq 1$ is a Lipschitz constant for S. To prove this, let us compute

$$(3.16) \quad F(\omega_1) - F(\omega_2) = \left[\int \rho_p \cdot (\rho_q \circ T^{j'-j}_{\omega_1} - \rho_q \circ T^{j'-j}_{\omega_2}) \, \mathrm{d}\mu_{\omega_1} \right]$$
$$+ \left[\int \rho_p \cdot \rho_q \circ T^{j'-j}_{\omega_2} \, \mathrm{d}\mu_{\omega_1} - \int \rho_p \cdot \rho_q \circ T^{j'-j}_{\omega_2} \, \mathrm{d}\mu_{\omega_2} \right].$$

For any y, we have

$$\rho_q \circ T^{j'-j}_{\omega_1}(y) - \rho_q \circ T^{j'-j}_{\omega_2}(y) = \rho_q \circ \pi_2(S^{j'-j}(\omega_1, y)) - \rho_q \circ \pi_2(S^{j'-j}(\omega_2, y)),$$

where $\pi_2 : \Omega \times X \to X$ is the second projection. As S is Lipschitz with Lipschitz constant C_S, we have $d(S^{j'-j}(\omega_1, y), S^{j'-j}(\omega_2, y)) \leqslant C_S^{j'-j} d(\omega_1, \omega_2)$. Therefore, the term on the first line of (3.16) is bounded by $C_S^{\alpha(n)} r^{-1} d(\omega_1, \omega_2)$, as ρ_q has Lipschitz constant at most r^{-1} and $j' - j \leq \alpha(n)$. For the term on the second line, we note that $\rho_p \cdot \rho_q \circ T^{j'-j}_{\omega_2}$ is Lipschitz, with Lipschitz constant at most $r^{-2} C_S^{\alpha(n)}$, by the same argument. Since $\omega' \mapsto \mu_{\omega'}$ is Lipschitz (Definition 1.5), it follows that this term is bounded by $C r^{-2} C_S^{\alpha(n)} d(\omega_1, \omega_2)$. This completes the proof of (3.15).

We can write (3.14) as $\int R_p(\theta^i \omega) R_q(\theta^{i'} \omega) F(\theta^j \omega) \, \mathrm{d}\mathbb{P}(\omega)$. Using 3-mixing (which follows from 4-mixing by taking the last function equal to 1), this is equal to

$$\left(\int R_p(\theta^i \omega) R_q(\theta^{i'} \omega) \, \mathrm{d}\mathbb{P}(\omega) \right) \left(\int F \, \mathrm{d}\mathbb{P} \right) + O(\|R_p\|_{\mathrm{Lip}} \|R_q\|_{\mathrm{Lip}} \|F\|_{\mathrm{Lip}} \varphi(n/3)).$$

The error terms add up to at most $n^4 r^{-4-2C'_0} C_S^{\alpha(n)} \varphi(n/3)$, thanks to (3.15). As $\varphi(n/3) \leqslant C e^{-(n/3)^c}$ for some $c > 0$ while $\alpha(n) = (\log n)^{C_4}$, this error term is again admissible.

We can also replace $\int R_p(\theta^i \omega) R_q(\theta^{i'} \omega) \, \mathrm{d}\mathbb{P}(\omega)$ with $\left(\int R_p \right) \left(\int R_q \right)$ up to an admissible error term, thanks to the mixing of θ and since i and i' are far apart. Finally, up to an admissible error term, the contribution of these terms to $\mathbb{E}((S_n^{\gg})^2)$ is at most

$$\sum_{p,q} \left(\int R_p \, \mathrm{d}\mathbb{P} \right) \cdot \left(\int R_q \, \mathrm{d}\mathbb{P} \right) \cdot \left(\int \rho_p(y) \rho_q(T^{j'-j}_\omega y) \, \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega) \right)$$

$$= \int \left(\sum_p \left[\int R_p \, \mathrm{d}\mathbb{P} \right] \rho_p(y) \right) \left(\sum_q \left[\int R_q \, \mathrm{d}\mathbb{P} \right] \rho_q(T^{j'-j}_\omega y) \right) \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega)$$

$$\leqslant \left(\int \left(\sum_p \left[\int R_p \, \mathrm{d}\mathbb{P} \right] \rho_p(y) \right)^2 \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega) \right)^{1/2}$$

$$\times \left(\int \left(\sum_q \left[\int R_q \, \mathrm{d}\mathbb{P} \right] \rho_q(T^{j'-j}_\omega y) \right)^2 \mathrm{d}\mu_\omega(y) \, \mathrm{d}\mathbb{P}(\omega) \right)^{1/2},$$

by Cauchy-Schwartz. By invariance of the measure $\mathbb{P} \otimes \mu_\omega$ under S, the two factors in the last product coincide, eliminating the square roots.

For any y, there are at most C_0 nonzero terms in the sum $\sum_p \left[\int R_p \, \mathrm{d}\mathbb{P} \right] \rho_p(y)$, by bounded local complexity. We can therefore use the convexity inequality $(a_1 + \ldots + a_{C_0})^2 \leq C_0(a_1^2 + $

$\ldots + a_{C_0}^2$) to bound the square of the sum by the sum of the squares. We get a bound

$$C_0 \int \sum_p \left[\int R_p \, d\mathbb{P} \right]^2 \rho_p(y)^2 \, d\mu_\omega(y) \, d\mathbb{P}(\omega).$$

Bounding ρ_p^2 by ρ_p, and since $\int \rho_p(y) \, d\mu_\omega(y) \, d\mathbb{P}(\omega) = \int R_p \, d\mathbb{P}$, we are left with a bound

$$C_0 \sum_p \left[\int R_p \, d\mathbb{P} \right]^3.$$

By concavity of the function $x \mapsto x^{2/3}$, one has the inequality $(\sum a_i)^{2/3} \leq \sum a_i^{2/3}$, and therefore $\sum a_i \leq \left(\sum a_i^{2/3} \right)^{3/2}$. Applying this inequality to the previous equation, one obtains a bound

$$C_0 \left(\sum_p \left[\int R_p \, d\mathbb{P} \right]^2 \right)^{3/2} = C_0 A_n^{3/2}.$$

Summing over the possible values of i, i', j (there are $(n/3)^3$ of them) and then j' (there are at most $2\alpha(n)$ of them, as $|j' - j| \leqslant \alpha(n)$), we get that the total contribution of this case is bounded by

$$C\alpha(n)n^3 A_n^{3/2} = C\alpha(n)B_n^{3/2},$$

up to an admissible error term.

Case 3: when both (i, i') and (j, j') are close.
In this case, we can use rough estimates. We have

$$\sum_{p,q} \int \rho_p(T_\omega^i x)\rho_q(T_\omega^{i'} x)\rho_p(T_\omega^j y)\rho_q(T_\omega^{j'} y) \, d\mu_\omega(x) \, d\mu_\omega(y) \, d\mathbb{P}(\omega)$$

$$\leq C_0 \sum_p \int \rho_p(T_\omega^i x)\rho_p(T_\omega^j y) \, d\mu_\omega(x) \, d\mu_\omega(y) \, d\mathbb{P}(\omega),$$

as $\sum_q \rho_q(z)\rho_q(z') \leq C_0$ for any z, z', by (3.2). This is equal to $C_0 \int R_p(\theta^i \omega) R_p(\theta^j \omega) \, d\mathbb{P}(\omega)$, which coincides with $C_0 \left(\int R_p \, d\mathbb{P} \right)^2$ by mixing, up to an error term which is admissible, as above. Finally, the total contribution of this case is, up to an admissible error term, bounded by

$$Cn^2\alpha(n)^2 \sum_p \left(\int R_p \, d\mathbb{P} \right)^2 = C\alpha(n)^2 B_n.$$

Conclusion.
Combining all three cases, we get

$$\mathbb{E}((S_n^{\gg})^2) \leqslant B_n^2 + C\alpha(n)B_n^{3/2} + C\alpha(n)^2 B_n + a_n,$$

where a_n is an admissible error term.

Let $\varepsilon > 0$. For small enough r, the convergence (3.6) ensures that, for large n,

$$A_n = \sum_p \left(\int R_p \, d\mathbb{P} \right)^2 \geqslant r^{\overline{D}_2^{\mathrm{an}}+\varepsilon}.$$

Take $r = r_n = 1/n^{(2-3\varepsilon)/(\overline{D}_2^{\mathrm{an}}+\varepsilon)}$. Then $n^2 A_n \geqslant n^{3\varepsilon}$. Note that $\mathbb{E}(S_n^{\gg}) = B_n + a_n$ where a_n is an admissible error term. We get $\mathbb{E}(S_n^{\gg})^2 = B_n^2 + a_n^2 + 2B_n a_n$, where both a_n^2 and $2B_n a_n$ are again admissible error terms, and may thus be written as a_n'. In particular, $\mathbb{E}(S_n^{\gg})^2$ is asymptotic to B_n^2, and is larger than $B_n^2/2$ for large n. With (3.11), we get

$$\mathbb{P} \otimes \mu_\omega \otimes \mu_\omega \{(\omega, x, y) : m_n^{\gg}(\omega; x, y) > 8r\} \leqslant \frac{\mathbb{E}((S_n^{\gg})^2) - \mathbb{E}(S_n^{\gg})^2}{\mathbb{E}(S_n^{\gg})^2}$$

$$\leqslant \frac{B_n^2 + C\alpha(n)B_n^{3/2} + C\alpha(n)^2 B_n + a_n - (B_n + a_n)^2}{B_n^2/2}$$

$$= \frac{B_n^2 + C\alpha(n)B_n^{3/2} + C\alpha(n)^2 B_n + a_n - B_n^2 + a_n'}{B_n^2/2}$$

$$\leqslant C\frac{\alpha(n)}{B_n^{1/2}} + C\frac{\alpha(n)^2}{B_n} + C\frac{a_n + a_n'}{B_n^2}.$$

As $\alpha(n)$ grows more slowly than any polynomial while $B_n \geqslant n^{3\varepsilon}$, it follows that the above probability is $O(n^{-\varepsilon})$.

Choose a subsequence $n_s = \lfloor s^{2/\varepsilon} \rfloor$. Along this subsequence, the error probability is $O(1/s^2)$, which is summable. By Borel-Cantelli, for almost every (ω, x, y), one has eventually $m_{n_s}^{\gg}(\omega; x, y) \leq 8r_{n_s}$, and therefore

$$\liminf \frac{-\log m_{n_s}^{\gg}(\omega; x, y)}{\log n_s} \geq \frac{2 - 3\varepsilon}{\overline{D}_2^{\mathrm{an}} + \varepsilon}.$$

One can not conclude directly from this, as the sequence $m_n^{\gg}$ is not monotone. What is true, though, is that for all $n \in [n_s, n_{s+1}]$, one has

$$-\log m_n^{\gg}(\omega; x, y) \geq -\log \min_{i < n_s/3, \ 2n_{s+1}/3 \leq j < n_s} d(T_\omega^i x, T_\omega^j y) =: -\log m_{n_s}'.$$

One can show exactly as above that, almost surely,

$$\liminf \frac{-\log m_{n_s}'(\omega; x, y)}{\log n_s} \geq \frac{2 - 3\varepsilon}{\overline{D}_2^{\mathrm{an}} + \varepsilon}.$$

With the previous equation, this inequality passes to the whole sequence $m_n^{\gg}$. We obtain almost surely

$$\liminf \frac{-\log m_n^{\gg}(\omega; x, y)}{\log n} \geq \frac{2 - 3\varepsilon}{\overline{D}_2^{\mathrm{an}} + \varepsilon}.$$

As ε is arbitrary, this proves the result. $\qquad\square$

References

[1] J. Atnip, G. Froyland, C. González-Tokman and S. Vaienti. *Thermodynamic formalism for random weighted covering systems.* Comm. Math. Phys. **386** (2021), no. 2, 819–902. Cited page 17.

[2] J. Atnip, G. Froyland, C. González-Tokman and S. Vaienti. *Equilibrium states for non-transitive random open and closed dynamical systems. Arxiv* 2107.03776 (2021). Cited page 17.

[3] 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. Cited pages 1 and 4.

[4] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* 1975. Lecture Notes in Mathematics, Vol. 470. Cited page 10.

[5] A. Coutinho, R. Lambert and J. Rousseau, *Matching strings in encoded sequences*, Bernoulli **26** (2020), no. 3, 2021–2050. Cited pages 2 and 3.

[6] M. Denker and M. Gordin. *Gibbs measures for fibred systems.* Adv. Math. **148** (1999), no. 2,161–192. Cited page 17.

[7] J. Rousseau, *Longest common substring for random subshifts of finite type*, Ann. Inst. Henri Poincaré Probab. Stat **57** (2021), no. 3, 1768–1785 Cited page 8.

[8] D. Ruelle. *The thermodynamic formalism for expanding maps.* Comm. Math. Phys. **125** (1989), no. 2, 239–262. Cited page 9.

[9] M. Stadlbauer, S. Suzuki and P. Varandas. *Thermodynamic formalism for random non-uniformly expanding maps.* Comm. Math. Phys. **385** (2021), no. 1, 369–427. Cited pages 11, 12, 13, and 14.

[10] M. Stadlbauer, P. Varandas and Xu. Zhang. *Quenched and annealed equilibrium states for random Ruelle expanding maps and applications.* Ergod. Theory Dyn. Syst. (2022), doi 10.1017/etds.2022.60. Cited pages 10 and 11.

Sébastien Gouëzel, IRMAR, CNRS UMR 6625, Université de Rennes 1, 35042 Rennes, France
Email address: `sebastien.gouezel@univ-rennes1.fr`

Jérôme Rousseau, CREC,Académie Militaire de St Cyr Coëtquidan, 56381 GUER Cedex, France

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

IRMAR, CNRS UMR 6625, Université de Rennes 1, 35042 Rennes, France
Email address: `jerome.rousseau@ufba.br`

Manuel Stadlbauer, Instituto de Matemática, Universidade Federal do Rio de Janeiro, Avenida Athos da Silveira Ramos 149, 21941-909 Rio de Janeiro (RJ), Brazil
Email address: `manuel@im.ufrj.br`