ALMOST SURE ORBITS CLOSENESS

MAXIM KIRSEBOM, PHILIPP KUNDE, TOMAS PERSSON, AND MIKE TODD

Abstract. We consider the minimal distance between orbits of measure preserving dynamical systems. In the spirit of dynamical shrinking target problems we identify distance rates for which almost sure asymptotic closeness properties can be ensured. More precisely, we consider the set E_n of pairs of points whose orbits up to time n have minimal distance to each other less than the threshold r_n. We obtain bounds on the sequence $(r_n)_n$ to guarantee that $\limsup_n E_n$ and $\liminf_n E_n$ are sets of measure 0 or 1. Results for the measure 0 case are obtained in broad generality while the measure one case requires assumptions of exponential mixing for at least one of the systems. We also consider the analogous question of the minimal distance of points within a single orbit of one dimensional exponentially mixing dynamical systems.

1. Introduction

In a metric space (X, d), the problem of the shortest distance between two orbits of a dynamical system $T\colon X \to X$, with an ergodic measure μ, was introduced in [BLR]. That is, for $n \in \mathbb{N}$ and $x, y \in X$, they studied

$$\mathbb{M}_n(x,y) = \mathbb{M}_{T,n}(x,y) := \min_{0 \leqslant i,j < n} d(T^i(x), T^j(y)) \tag{1.1}$$

and showed that the decay of $\mathbb{M}_n$ depends on the correlation dimension.

The *lower correlation dimension* of μ is defined by

$$\underline{C}_\mu := \liminf_{r \to 0} \frac{\log \int \mu(B(x,r))\, d\mu(x)}{\log r},$$

and the *upper correlation dimension* $\overline{C}_\mu$ is analogously defined via the limsup. If these are equal, then this is C_μ, the *correlation dimension* of μ. This dimension plays an important role in the description of the fractal structure of invariant sets in dynamical systems and has been widely studied from different points of view: numerical estimates (e.g. [BB, BPTV, SR]), existence and relations with other fractal dimension (e.g. [BGT, P]) and relations with other dynamical quantities (e.g. [FV, M]).

In [BLR, Theorem 1], under the assumption $\underline{C}_\mu > 0$, a general upper bound for $\mathbb{M}_n$ was obtained:

Theorem 1.1. *For a dynamical system (X, T, μ), we have*

$$\limsup_n \frac{\log \mathbb{M}_{T,n}(x,y)}{-\log n} \leqslant \frac{2}{\underline{C}_\mu} \qquad (\mu \times \mu)\text{-a.e. } x, y.$$

Date: October 15, 2025.

2020 *Mathematics Subject Classification.* 37B20, 37A05, 37D05, 37E05.

We acknowledge financial support by the Hamburg–Lund Funding Program 2023–2025 which made several mutual research visits possible.

To replace the inequality above with equality, in [BLR, Theorems 3 and 6], the authors assumed that C_μ exists and proved

$$\liminf_n \frac{\log \mathbb{M}_{T,n}(x,y)}{-\log n} \geqslant \frac{2}{C_\mu} \qquad (\mu\times\mu)\text{-a.e. } x,y, \tag{1.2}$$

using some exponential mixing conditions on the system. This was shown in [RT] to be unnecessary in cases where there is an appropriate inducing scheme.

In the proofs of the above theorems, the approach was to find a sequences $(r_n)_n$ and show, using the Borel–Cantelli Lemma, that for almost every (x,y) for all large enough n either $\mathbb{M}_{T,n}(x,y) \leqslant r_n$, or $\mathbb{M}_{T,n}(x,y) \geqslant r_n$. The sequences were, for any $\varepsilon > 0$, of the form $r_n = \frac{1}{(n^2\log n)^{(\underline{C}_\mu-\varepsilon)}}$ in the former case and $r_n = \frac{1}{(n^2(\log n)^b)^{(\overline{C}_\mu+\varepsilon)}}$, for some $b < -4$, in the latter. Our aim in the current work is to refine these estimates and extend their applications.

Notation: we will sometimes use, for non-negative real sequences $(a_n)_n, (b_n)_n$, the notation $a_n \lesssim b_n$ to mean that there is some $C > 0$ such that $a_n \leqslant Cb_n$ for all n, similarly for $\gtrsim$. If $a_n \gtrsim b_n$ and $a_n \lesssim b_n$, we write $a_n \asymp b_n$.

2. Main results

We begin with the most general setup which consists of two probability preserving dynamical systems (T_1,μ_1) and (T_2,μ_2) on the same space X, i.e. $T_1, T_2\colon X \to X$. We let $(r_n)_n$ denote a sequence of positive numbers. Note that in this setting $\int \mu_1(B(y,r_n))\,d\mu_2(y) = \int \mu_2(B(y,r_n))\,d\mu_1(y)$. We define the minimum analogously to (1.1),

$$\mathbb{M}_n(x,y) = \mathbb{M}_{(T_1,T_2),n}(x,y) := \min_{0\leqslant i,j<n} d(T_1^i(x), T_2^j(y)),$$

and define

$$\begin{aligned} E_n = E_{n,r_n}^{T_1,T_2} : &= \left\{ (x,y)\in X\times X : \mathbb{M}_{(T_1,T_2),n}(x,y) < r_n \right\} \\ &= \left\{ (x,y)\in X\times X : d(T_1^i x, T_2^j y) < r_n \text{ for some } 0\leqslant i,j<n \right\}. \end{aligned}$$

Henceforth we abandon the use of $\mathbb{M}_n$ and instead use the latter form of the above equation. Ultimately one would like conditions on the choices for $(r_n)_n$ of the following form:

$$\text{Condition on } (r_n)_n \iff (\mu_1\times\mu_2)(\liminf_n E_{n,r_n}^{T_1,T_2}) = 1,$$

i.e., for almost every (x,y), for all large enough n there exist $0 \leqslant i,j < n$ such that $d(T_1^i x, T_2^j y) < r_n$; and

$$\text{Condition on } (r_n)_n \iff (\mu_1\times\mu_2)(\limsup_n E_{n,r_n}^{T_1,T_2}) = 1,$$

i.e., almost surely there are infinitely many n such that there exist $0 \leqslant i,j < n$ with $d(T_1^i x, T_2^j y) < r_n$.

Some results will be stated explicitely for the special case when $T_1 = T_2$ and $\mu_1 = \mu_2$ in which case we simply talk about the probability preserving dynamical system (X,T,μ) and write

$$E_n = E_{n,r_n}^T := \left\{ (x,y)\in X\times X : d(T^i x, T^j y) < r_n \text{ for some } 0\leqslant i,j<n \right\}.$$

In the special case of one dynamical system we will also consider the further specialized case of $x = y$, or in other words, the case of one single orbit's minimal internal distance. In this case we change notation and write

$$F_n = F_{n,r_n}^T := \{ x\in X : d(T^i x, T^j x) < r_n \text{ for some } 0\leqslant i<j<n \}.$$

The lim sup and lim inf results that we wish to obtain remain analogous for this set.

The main results, which are to follow, are initially split into the case of two distinct orbits and the case of one single orbit.

For two distinct orbits of two dynamical systems we obtain conditions on measure zero for both lim inf and lim sup sets in a very general setting. Specialising to two mixing systems we obtain conditions for the lim inf and lim sup sets to have measure one and simplify these conditions in a corollary when the two mixing systems are indeed identical. In the special case of one dynamical system which is chosen to be the doubling map we get particularly sharp results. For the lim sup set we obtain a dichotomy and for the lim inf set we obtain bounds sharp enough to produce a shrinking rate of the $(r_n)_n$ which gives measure zero for the lim inf set but measure one for the lim sup set. Our final result within the two distinct orbit case concerns the situation when one system is a rotation with a Diophantine condition and one system is mixing with respect to the Lebesgue measure. In this case we also obtain conditions for which the lim inf and lim sup sets have measure one.

The single orbit case is more difficult and we obtain conditions for measure zero and one of the lim inf and lim sup sets under somewhat stronger mixing assumptions.

2.1. Results for two distinct orbits.

Theorem 2.1. *Let (X, T_1, μ_1) and (X, T_2, μ_2) be two probability preserving dynamical systems and $(r_n)_n$ a sequence of positive numbers. Then*

(a) *If $n^2 \int \mu_1(B(y, r_n))\, d\mu_2(y) \to 0$, then*
$$(\mu_1 \times \mu_2)(\liminf_n E_{n,r_n}^{T_1,T_2}) = 0.$$
(b) *If $n \int \mu_1(B(y, r_n)) d\mu_2(y)$ is decreasing and $\sum_{n=1}^{\infty} n \int \mu_1(B(y, r_n))\, d\mu_2(y) < \infty$ then*
$$(\mu_1 \times \mu_2)(\limsup_n E_{n,r_n}^{T_1,T_2}) = 0.$$

We note that the result holds with the obvious simplifications when the two dynamical systems are identical. For part (b), the summability condition on the $(r_n)_n$ holds for example if $\int \mu_1(B(y, r_n))\, d\mu_2(y) \lesssim 1/(n^2(\log n)^{1+\varepsilon})$. Note that in the case of one dynamical system, criteria of this type appear in other results, such as [KKP, Theorem C], where $\sum_{n=1}^{\infty} \int \mu(B(y, r_n))\, d\mu(y) < \infty$ is assumed in order to get a result on returns of x to itself.

We will need the following definition of exponential mixing wrt. observables in the well-known function spaces BV and L^∞.

Definition 2.2. Let (X, T, μ) denote a measure preserving system where $X \subset \mathbb{R}$. If there are $C, \theta > 0$ such that for all $\psi \in BV$ and $\varphi \in L^\infty$,
$$\left| \int \psi \cdot \varphi \circ T^n \, d\mu - \int \psi \, d\mu \int \varphi \, d\mu \right| \leqslant C \|\psi\|_{BV} \|\varphi\|_{L^\infty} e^{-\theta n},$$
then we say that (X, T, μ) has *exponential mixing for BV against L^∞*.

In Section 2.3, some results will require L^∞ to be replaced by L^1 in this definition. Systems that are known to satisfy exponential mixing for BV against either L^1 or L^∞ include piecewise expanding interval maps with μ being a Gibbs measure and quadratic maps with Benedicks–Carleson parameter and μ being the absolutely continuous invariant measure ([LSV], [Y]).

The next definition concerns exponential mixing wrt. observables in L^∞ and the less-known function space V_α. This space along with its associated norm $|\cdot|_\alpha$ was introduced by Saussol in [S] where he defined and studied multidimensional piecewise expanding maps (see also [BLR, section 5]). The definition of these maps and of V_α is rather involved and extensive, hence we refrain from giving the details and refer the interested reader to the cited papers. For our purposes the intuition that the $|\cdot|_\alpha$-norm is an analogue of the BV-norm for higher dimensional maps suffices (in fact, here we only use that characteristic function on balls have norm bounded independently of the ball, as well as the maps $x \mapsto \mu(B(x,r)))$ along with the fact that Saussol proved that his multidimensional piecewise expanding maps defined on a compact subset of $\mathbb{R}^n$ satisfy the following mixing property.

Definition 2.3 (see [S] and [BLR])**.** Let (X,T,μ) denote a measure preserving system where $X \subset \mathbb{R}^n$. If there are $C,\theta > 0$ such that for all $\psi \in V_\alpha$ and $\varphi \in L^\infty$,

$$\left| \int \psi \cdot \varphi \circ T^n \, d\mu - \int \psi \, d\mu \int \varphi \, d\mu \right| \leqslant C|\psi|_\alpha \|\varphi\|_{L^\infty} e^{-\theta n},$$

then we say that (X,T,μ) has *exponential mixing for V_α against L^∞.*

Theorem 2.4. *Let (X,T_1,μ_1) and (X,T_2,μ_2) be two probability preserving systems which are exponentially mixing for either BV against L^∞ or V_α against L^∞. Let $(r_n)_n$ be a sequence of positive numbers.*

(a) *If $(r_n)_n$ is decreasing and for some $\varepsilon > 0$ we have*

$$\int \mu_1(B(y,r_n)) \, d\mu_2(y) \gtrsim \frac{(\log n)^3(\log\log n)^{1+\varepsilon}}{n^2}$$

and

$$\frac{\left(\int \mu_i(B(y,r_n)) \, d\mu_i(y)\right)^{\frac{1}{2}}}{\int \mu_1(B(y,r_n)) \, d\mu_2(y)} \lesssim \frac{n}{(\log n)^2(\log\log n)^{1+\varepsilon}}, \quad i = 1,2, \tag{2.1}$$

then

$$(\mu_1 \times \mu_2)\left(\liminf_n E_{n,r_n}^{T_1,T_2}\right) = 1.$$

(b) *If for some $h(n) \to \infty$ we have*

$$\int \mu_1(B(y,r_n)) \, d\mu_2(y) \geqslant \frac{(\log n)^2 h(n)}{n^2}$$

and

$$\frac{\left(\int \mu_i(B(y,r_n)) \, d\mu_i(y)\right)^{\frac{1}{2}}}{\int \mu_1(B(y,r_n)) \, d\mu_2(y)} \leqslant \frac{n}{(\log n)h(n)}, \quad i = 1,2, \tag{2.2}$$

then

$$(\mu_1 \times \mu_2)\left(\limsup_n E_{n,r_n}^{T_1,T_2}\right) = 1.$$

Remark 2.5. If μ_1 or μ_2 are Ahlfors regular, i.e., there exist $C,s > 0$ such that

$$\frac{1}{C} r^s \leqslant \mu(B(x,r)) \leqslant C r^s \text{ for all } x,$$

then all of our conditions on measures of balls, and their integrals, depend only on s and $(r_n)_n$.

We examine the conditions on $(r_n)_n$ in part (a) of Theorem 2.4 in two examples. Part (b) can be considered in the same way.

Example 2.6. If μ_1 and μ_2 are equivalent and there is a constant $c \geqslant 1$ such that

$$c^{-1} \leqslant \frac{d\mu_1}{d\mu_2} \leqslant c,$$

then for all large n it is in this case enough, since $\varepsilon > 0$ is arbitrary, to require in part (a) that

$$\int \mu_1(B(y,r_n)) \, d\mu_1(y) \geqslant \frac{(\log n)^4(\log\log n)^{2+\varepsilon}}{n^2},$$

and

$$\int \mu_2(B(y,r_n)) \, d\mu_2(y) \geqslant \frac{(\log n)^4(\log\log n)^{2+\varepsilon}}{n^2}.$$

Example 2.7. Suppose that (T_1, μ_1) is the doubling map with Lebesgue measure, and that (T_2, μ_2) is a quadratic map for a Benedicks–Carleson parameter, with the invariant measure which is absolutely continuous with respect to Lebesgue measure.

We have of course that $r \leqslant \mu_1(B(y,r)) \leqslant 2r$, so that

$$r \leqslant \int \mu_1(B(y,r))\, d\mu_1(y) \leqslant 2r$$

and

$$r \leqslant \int \mu_1(B(y,r))\, d\mu_2(y) \leqslant 2r.$$

To estimate $\int \mu_2(B(y,r))\, d\mu_2(y)$, we note that it is known that μ_2 has a density which is sum of a BV function, bounded away from zero, and countably many one-sided singularities of the form $1/\sqrt{x}$, see [Y, Theorem 1]. Because of this, we have $cr \leqslant \mu_2(B(y,r))$ and

$$cr \leqslant \int \mu_2(B(y,r))\, \mathrm{d}\mu_2(y).$$

To get an upper bound, it is enough to consider the case that the density of μ_2 is $h(x) = 1/\sqrt{x}$ on $[0,1]$. We then have that

$$\mu_2(B(y,r)) = \int_{y-r}^{y+r} \frac{1}{\sqrt{x}}\, dx \leqslant \frac{2r}{\sqrt{y-r}}$$

if $y - r \geqslant 0$ and

$$\mu_2(B(y,r)) = \int_0^{y+r} \frac{1}{\sqrt{x}}\, dx = 2\sqrt{y+r} \leqslant 2\sqrt{2}\sqrt{r}$$

if $y - r < 0$.

Hence,

$$\begin{aligned}
\int \mu_2(B(y,r))\, d\mu_2(y) &\leqslant \int_0^r 2\sqrt{2r}\frac{1}{\sqrt{y}}\, dy + \int_r^1 \frac{2r}{\sqrt{y-r}}\frac{1}{\sqrt{y}}\, dy \\
&\leqslant 4\sqrt{2r}\sqrt{r} + 2r\int_0^1 \frac{1}{\sqrt{y}\sqrt{y+r}}\, dy \\
&= 4\sqrt{2}r + 4r\log(1/\sqrt{r} + \sqrt{1+1/r}) \\
&\leqslant Cr|\log r|.
\end{aligned}$$

The conditions in Theorem 2.4 (a) on r_n are therefore the following:

$$\begin{aligned}
r_n &\geqslant \frac{(\log n)^3(\log\log n)^{1+\varepsilon}}{n^2}, \\
\frac{\sqrt{r_n}}{r_n} &\leqslant \frac{n}{(\log n)^2(\log\log n)^{1+\varepsilon}}, \\
\frac{\sqrt{r_n|\log r_n|}}{r_n} &\leqslant \frac{n}{(\log n)^2(\log\log n)^{1+\varepsilon}}.
\end{aligned}$$

This simplifies to

$$r_n \geqslant \frac{(\log n)^3(\log\log n)^{1+\varepsilon}}{n^2}, \quad \text{and} \quad \frac{r_n}{|\log r_n|} \geqslant \frac{(\log n)^4(\log\log n)^{2+2\varepsilon}}{n^2}.$$

Hence, it is enough to require that

$$r_n \geqslant \frac{(\log n)^5(\log\log n)^{2+\varepsilon}}{n^2},$$

for some $\varepsilon > 0$.

The proof of Theorem 2.4 will be given for the case of exponential mixing for BV against L^∞. The proof for V_α against L^∞ is essentially identical but would require us to give the precise definition of V_α and $|\cdot|_\alpha$.

As an immediate corollary of Theorem 2.4 we obtain the simpler statement for one system.

Corollary 2.8. *Let (X, T, μ) be a probability preserving dynamical system which is exponentially mixing for either BV against L^∞ or V_α against L^∞. Let $(r_n)_n$ be a sequence of positive numbers.*

(a) *If $(r_n)_n$ is decreasing and*
$$\int \mu(B(y, r_n))\, d\mu(y) \gtrsim \frac{(\log n)^4 (\log\log n)^{2+\varepsilon}}{n^2},$$
then
$$(\mu \times \mu)(\liminf_n E^T_{n,r_n}) = 1.$$

(b) *If for some function $h(n) \to \infty$*
$$\int \mu(B(y, r_n))\, d\mu(y) \geqslant \frac{(\log n)^2 h(n)}{n^2},$$
then
$$(\mu \times \mu)(\limsup_n E^T_{n,r_n}) = 1.$$

The doubling map case. Let $X = [0, 1]$, $T : X \to X$ be given by $Tx = 2x \mod 1$, let μ denote the Lebesgue measure on X and let d denote the Euclidian metric on X. We note that all results in this subsection also hold for the map $Tx = kx \mod 1$, $k \in \mathbb{N}$, $k \geqslant 2$.

For $\liminf_n E_n$ we get the following result.

Theorem 2.9. *Let $(r_n)_n$ be a sequence of positive numbers.*

(a) *If $n^2 r_n \to 0$ then*
$$(\mu \times \mu)\left(\liminf_n E^T_{n,r_n}\right) = 0.$$

(b) *If $(r_n)_n$ is decreasing and*
$$\sum_{n=0}^{\infty} \frac{1}{2^{2n} r_{2^n}} < \infty. \tag{2.3}$$
then
$$(\mu \times \mu)\left(\liminf_n E^T_{n,r_n}\right) = 1.$$

Remark 2.10. We note that by Cauchy condensation (see Lemma 3.1 below), we have
$$\sum_{n=1}^{\infty} \frac{1}{n^3 r_n} < \infty \quad \Leftrightarrow \quad \sum_{n=0}^{\infty} \frac{1}{2^{2n} r_{2^n}} < \infty,$$
provided $(n^3 r_n)_n$ is increasing. Condition (2.3) is for example satisfied for
$$r_n = \frac{\log n (\log\log n)^{1+\epsilon}}{n^2}$$
for any $\epsilon > 0$. In this case we see that
$$\sum_{n=0}^{\infty} \frac{1}{n^3 r_n} = \sum_{n=0}^{\infty} \frac{1}{n \log n (\log\log n)^{1+\epsilon}} < \infty.$$

We are able to prove the following dichotomy for $\limsup_n E_n$.

Theorem 2.11. *Let $(r_n)_n$ be a decreasing sequence of positive numbers s.t. also $(nr_n)_n$ is decreasing. Then*

$$(\mu \times \mu)(\limsup_n E^T_{n,r_n}) = \begin{cases} 0 & \text{if } \sum_{n=1}^\infty nr_n < \infty, \\ 1 & \text{if } \sum_{n=1}^\infty nr_n = \infty. \end{cases}$$

Remark 2.12. Actually, $(nr_n)_n$ being decreasing is only required for the measure 0 case. For the measure 1 case it is sufficient to assume that $(r_n)_n$ is decreasing.

Remark 2.13. We note that Theorem 2.9 and Theorem 2.11 together show that radii sequences with $0 = (\mu \times \mu)(\liminf_n E_n) < (\mu \times \mu)(\limsup_n E_n) = 1$ exist. Take for example

$$r_n = \frac{1}{n^2 \log n}.$$

Then $n^2 r_n = \frac{1}{\log n} \to 0$ and $\sum_{n=1}^\infty nr_n = \sum_{n=1}^\infty \frac{1}{n \log n} = \infty$ which means that the conditions for both results are satisfied.

2.2. Results when one system is a rotation.

Theorem 2.14. *Suppose that X is the circle and both μ_1 and μ_2 are Lebesgue measure. Moreover, suppose that (X, T_1, μ_1) is exponentially mixing for BV against L^∞ and that (X, T_2, μ_2) is a rotation by an angle α. Let $(r_n)_n$ be a sequence of positive numbers.*

(a) *If $n^2 r_n \to 0$, then $(\mu_1 \times \mu_2)(\liminf_{n\to\infty} E^{T_1,T_2}_{n,r_n}) = 0$.*
(b) *If $(nr_n)_n$ is decreasing and $\sum_{n=1}^\infty nr_n < \infty$, then $(\mu_1 \times \mu_2)(\limsup_n E^{T_1,T_2}_{n,r_n}) = 0$.*
(c) *Suppose some $\varepsilon > 0$ that α satisfies*
$$|q\alpha - p| \geqslant c(\alpha)(\log q)^2 \cdot (\log\log(q))^{1+\varepsilon}/q^2$$
for all sufficiently large $q \in \mathbb{N}$. If $(r_n)_n$ is decreasing and satisfies $r_n \gtrsim \frac{(\log n)^2 (\log\log n)^{1+\delta}}{n^2}$ with $0 < \delta < \varepsilon$, then $(\mu_1 \times \mu_2)(\liminf_n E^{T_1,T_2}_{n,r_n}) = 1$.
(d) *Suppose α satisfies $|q\alpha - p| \geqslant c(\alpha) \log q \cdot \varphi(q)/q^2$ for all sufficiently large $q \in \mathbb{N}$, where $\varphi(q) \to \infty$ as $q \to \infty$. If $(r_n)_n$ satisfies $r_n \geqslant \frac{\log n \cdot h(n)}{n^2}$ for some $h(n)$ with $h(n) \to \infty$ and $h(n)/\varphi(n) \to 0$ as $n \to \infty$, then $(\mu_1 \times \mu_2)(\limsup_n E^{T_1,T_2}_{n,r_n}) = 1$.*

We emphasize that the Diophantine condition imposed on the angle α in parts (c) and (d) is satisfied for almost every α. This follows from Khinchin's Theorem on metric Diophantine approximation which implies that if $\sum_q \psi(q)$ converges, then a.e. number α satisfies $|q\alpha - p| \geqslant \psi(q)$ for all sufficiently large $q \in \mathbb{N}$. For example, our Diophantine conditions are satisfied for all numbers of Diophantine exponent strictly less than 1. Here we recall that α has Diophantine exponent σ if there exists a constant $C > 0$ such that $|q\alpha - p| \geqslant \frac{C}{q^{1+\sigma}}$ for all $q \in \mathbb{N}$.

2.3. Results for a single orbit.

Recall the notation

$$F_n = F^T_{n,r_n} = \{\, x : d(T^i x, T^j x) < r_n \text{ for some } 0 \leqslant i < j < n \,\}.$$

We will need the following definition.

Definition 2.15. Let (X, T, μ) denote a measure preserving system where $X \subset \mathbb{R}$. If there exist $C', \theta' > 0$ such that for all $\psi_1, \psi_2 \in BV$, $\varphi_1, \varphi_2 \in L^\infty$, and $0 \leqslant a < b \leqslant c$,

$$\left| \int \psi_1 \cdot \psi_2 \circ T^a \cdot \varphi_1 \circ T^b \cdot \varphi_2 \circ T^c \, d\mu - \int \psi_1 \cdot \psi_2 \circ T^a \, d\mu \int \varphi_1 \cdot \varphi_2 \circ T^{c-b} \, d\mu \right| \leqslant C'(\psi_1, \psi_2, \varphi_1, \varphi_2) e^{-\theta'(b-a)},$$

then we say that (X, T, μ) has *exponential 4-mixing for BV against L^∞*.

Note that by setting $\psi_1 = \varphi_2 = \mathrm{Id}$ we get 2-mixing for BV against L^∞ with the same constants C' and θ' as in the 4-mixing estimate. The 4-mixing property is known to hold for Gibbs–Markov interval maps, see [Z, Lemma 4.16]. In our case each of the observables will be characteristic functions on an interval, and we will assume that C' can be taken independently of the intervals. See [Z, Lemma 4.16] for a case of this.

Remark 2.16. The uniformity of C' is automatic for intervals. By Collet [C] (using Banach–Steinhaus), we have $C'(\psi_1, \psi_2, \varphi_1, \varphi_2) = C\|\psi_1\|_{BV}\|\psi_2\|_{BV}\|\varphi_1\|_\infty\|\varphi_2\|_\infty$, and for intervals all these norms are uniformly bounded.

Let $A(r, n) := \{ x : d(x, T^n x) < r \}$. We will need to assume the following short return time estimate which states that there exist C and $s > 0$ such that

$$\mu(A(r, n)) \leqslant Cr^s. \tag{2.4}$$

There are several known estimates of this type in the literature, see for instance Holland, Nicol and Török [HNT, Lemma 3.4], Kirsebom, Kunde and Persson [KKP, Section 6.2] and Holland, Kirsebom, Kunde and Persson [HKKP, Lemma 13.7].

Theorem 2.17. *Let (X, T, μ) be probability preserving dynamical system with an interval $X \subset \mathbb{R}$. Let $(r_n)_n$ be a sequence of positive numbers.*

(a) *Assume (X, T, μ) is exponentially mixing for BV against L^1, and that the short return time estimate (2.4) holds with constants $C, s > 0$. If $r_n \leqslant 1/((n \log n)^{\frac{1}{s}} h(n))$ and $\int \mu(B(x, r_n))\, d\mu(x) \leqslant \frac{1}{n^2 h(n)}$ for some function $h(n) \to \infty$ when $n \to \infty$, then $\mu(\liminf_n F^T_{n,r_n}) = 0$.*

(b) *Assume (X, T, μ) is exponentially mixing for BV against L^1, that the short return time estimate (2.4) holds with constants $C, s > 0$ and that $r_n \lesssim n^{-1/s}(\log n)^{-2/s-\varepsilon}$ for some ε. If $n \int \mu(B(x, r_n))\, d\mu(x)$ is decreasing and $\sum_{n=1}^\infty n \int \mu(B(x, r_n))\, \mathrm{d}\mu(x) < \infty$, then $\mu(\limsup_n F^T_{n,r_n}) = 0$.*

(c) *Assume (X, T, μ) is exponentially 4-mixing for BV against L^∞. Let $(r_n)_n$ be decreasing such that $r_n \geqslant n^{-\beta}$ for some $\beta > 0$ and $\int \mu(B(y, r_n))\, d\mu(y) \geqslant \frac{(\log n)^4 (\log \log n)^{2+\varepsilon}}{n^2}$, then $\mu(\liminf_n F^T_{n,4r_n}) = 1$.*

(d) *Assume (X, T, μ) is exponentially 4-mixing for BV against L^∞, $r_n \geqslant n^{-\beta}$ for some $\beta > 0$ and $\int \mu(B(y, r_n))\, d\mu(y) \geqslant \frac{(\log n)^2 h(n)}{n^2}$, for some function $h(n) \to \infty$ as $n \to \infty$, then $\mu(\limsup_n F^T_{n,4r_n}) = 1$.*

Remark 2.18. Theorem 2.17 is stated for $X \subset \mathbb{R}$ being an interval. We could have stated it more generally for a compact subset $X \subset \mathbb{R}^n$ by changing the definition of 4-mixing analogue to Definition 2.3, i.e. by exchanging the BV-norm with the $|\cdot|_\alpha$-norm. However, beyond interval maps there are, to our knowledge, no known examples of systems satisfying either this generalized 4-mixing for V_α or the short return time estimate of (2.4). For this reason we opted for the simpler formulation.

Examples of systems for which Theorem 2.17 holds include the doubling map (and more generally $Tx = kx \mod 1$ for some $k \in \mathbb{N}$, $k \geqslant 2$), piecewise expanding interval maps where μ is absolutely continuous wrt. Lebesgue, and the Gauss map with the Gauss measure.

The proof of Theorem 2.17(c) is similar to the proof of [Z, (4.2) in Theorem 4.8]: note that many of the ideas there were developed for a different case in [GRS].

3. Two mixing systems proofs

We include a useful version of Cauchy Condensation, see for example the proof of [HKKP, Proposition 10.1]:

Lemma 3.1. *Suppose* $(c_n)_n$ *has* $c_n \geqslant c_{n+1} > 0$ *and* $a > 1$. *Then*

$$\sum_{n=1}^{\infty} c_n < \infty \qquad \iff \qquad \sum_{k=1}^{\infty} a^k c_{\lfloor a^k \rfloor} < \infty.$$

We also require the following results.

Lemma 3.2. *Suppose that* (X, μ_1), (X, μ_2) *are probability spaces with* $X \subset \mathbb{R}^n$. *Then there exists* $K > 0$ *such that if* $r > 0$ *is sufficiently small then*

$$\mu_i(B(y, r)) \leqslant K \left(\int \mu_i(B(x, r))\, d\mu_i(x) \right)^{\frac{1}{2}} \quad \text{for } i = 1, 2.$$

Hence also

$$\int \mu_i(B(x, r))^2\, d\mu_j(x) \leqslant K \left(\int \mu_i(B(x, r))\, d\mu_i(x) \right)^{\frac{1}{2}} \int \mu_i(B(x, r))\, d\mu_j(x).$$

for $(i, j) = (1, 2), (2, 1)$.

Note that when $\mu_1 = \mu_2$, the second statement is the conclusion of [BLR, Lemma 14].

Proof. Since $X \subset \mathbb{R}^n$, we can cover $B(y, r)$ by K_0 balls of radius $r/2$. Let $B(z, r/2)$ be the ball of largest μ_i-measure, then $\mu_i(B(z, r/2)) \geqslant \mu_i(B(y, r))/K_0$. Then

$$\begin{aligned} \int \mu_i(B(x, r))\, d\mu_i(x) &\geqslant \int_{B(z,r/2)} \mu_i(B(x, r))\, d\mu_i(x) \geqslant \int_{B(z,r/2)} \mu_i(B(z, r/2))\, d\mu_i(x) \\ &= \mu_i(B(z, r/2))^2 \geqslant \frac{1}{K_0^2} \mu_i(B(y, r))^2, \end{aligned}$$

so setting $K = K_0$ the first statement of the lemma is proved. The second statement is then immediate. □

Lemma 3.3. *Let* (X, μ) *denote a probability space where* $X \subset \mathbb{R}$. *For any* $r > 0$, $\psi_r \colon X \to \mathbb{R}$ *given by* $y \mapsto \mu_1(B(y, r))$ *is BV with total variation bounded above by 2.*

Proof. We have that $\mu(B(y, r)) = \mu((-\infty, y + r)) - \mu((-\infty, y - r])$. Hence, the function ψ_r is a difference of two increasing functions, both increasing from 0 to 1. It follows immediately that the total variation of ψ_r is at most $1 + 1 = 2$. □

Remark 3.4. An analogue of this lemma in higher dimension for $|\cdot|_\alpha$ also holds. Indeed, a stronger statement is proved in [BLR, Section 5].

Proof of Theorem 2.1(a). Given a sequence $(r_n)_n$, let

$$S_n(x, y) = \sum_{0 \leqslant i,j < n} \mathbb{1}_{B(T_2^j y, r_n)}(T_1^i x). \tag{3.1}$$

Note that $S_n(x, y) = 0$ implies $(x, y) \notin E_n$. We argue that the result follows if $\mathbb{E}(S_n) \to 0$. Indeed, since $S_n \geqslant 0$, there exists a subsequence $(n_k)_k$ s.t. $S_{n_k}(x, y) \to 0$ a.s. Since $S_n(x, y)$ is integer-valued this means that a.s. $S_{n_k}(x, y) = 0$ for all sufficiently large k. In particular, this means that $(x, y) \notin \liminf_n E_n$. Hence we compute the expectation of S_n. By the T_1-invariance of μ_1 and T_2-invariance of μ_2,

$$\begin{aligned} \mathbb{E}(S_n) &= \sum_{0 \leqslant i,j < n} \iint \mathbb{1}_{B(T_2^j y, r_n)}(T_1^i x)\, d\mu_1(x) d\mu_2(y) \\ &= \sum_{0 \leqslant i,j < n} \int \mu_1(B(y, r_n))\, d\mu_2(y) = n^2 \int \mu_1(B(y, r_n))\, d\mu_2(y) \to 0. \end{aligned}$$

This concludes the proof. □

Proof of Theorem 2.1(b). Given a sequence $(r_n)_n$ which decreases such that also $n\int \mu_1(B(y,r_n))\,d\mu_2(y)$ is decreasing, let

$$\tilde{S}_n(x,y) := \sum_{i,j\in[0,2^{n+1})} \mathbb{1}_{B(T_2^j y, r_{2^n})}(T_1^i x),$$
$$\tilde{S}(x,y) := \sum_{n=0}^{\infty} \tilde{S}_n(x,y).$$

We argue that the result follows if $\mathbb{E}(\tilde{S}) < \infty$. Indeed, then $\tilde{S}(x,y) < \infty$ a.s., which in turn means that a.s. $\tilde{S}_n(x,y) = 0$ for all but finitely many $n \in \mathbb{N}$ since the $\tilde{S}_n$ are integer-valued. The definition of the $\tilde{S}_n$ along with the assumption $r_n \geqslant r_{n+1}$ means that $\tilde{S}_n(x,y) = 0$ implies $(x,y) \notin E_{2^n+l}$ for all $l \in \{0,1,\ldots,2^{n+1}-2^n\}$. So if a.s. $\tilde{S}_n(x,y) = 0$ for all sufficiently large n, then a.s. $(x,y) \notin \limsup_n E_n$.

Hence we compute the expectation of $\tilde{S}$. Since $\tilde{S}_n \geqslant 0$ we have $\mathbb{E}(\tilde{S}) = \sum_{n=1}^{\infty} \mathbb{E}(\tilde{S}_n)$. By the T_1-invariance of μ_1 and T_2-invariance of μ_2,

$$\begin{aligned}
\mathbb{E}(\tilde{S}) &= \sum_{n=0}^{\infty} \sum_{i,j\in[0,2^{n+1})} \int\int \mathbb{1}_{B(T_2^j y, r_{2^n})}(T_1^i x)\,d\mu_1(x)d\mu_2(y) \\
&= \sum_{n=0}^{\infty} \sum_{i,j\in[0,2^{n+1})} \int \mu_1(B(T_2^j y, r_{2^n}))\,d\mu_2(y) \\
&= \sum_{n=0}^{\infty} 2^{2(n+1)} \int \mu_1(B(y, r_{2^n}))\,d\mu_2(y).
\end{aligned}$$

Hence the assumption that $\sum_{n=0}^{\infty} 2^{2n}\int \mu_1(B(y,r_{2^n}))\,d\mu_2(y) < \infty$ implies that $\mathbb{E}(\tilde{S})$ is bounded. Since $n\int \mu_1(B(y,r_n))\,d\mu_2(y)$ is assumed decreasing, Lemma 3.1 gives that this is equivalent to the assumption

$$\sum_{n=0}^{\infty} n \int \mu_1(B(y,r_n))\,d\mu_2(y) < \infty.$$

This concludes the proof. □

Proof of Theorem 2.4(a). Suppose, without loss of generality, that the exponential mixing constants are the same for both systems.

Given a decreasing sequence $(r_n)_n$, for $x,y \in X$, define

$$\hat{S}_n(x,y) := \sum_{i,j\in[0,2^n)} \mathbb{1}_{B(T_2^j y, r_{2^{n+1}})}(T_1^i x).$$

The motivation for defining $\hat{S}_n$ along the subsequence 2^n will become clear later in the proof. Note that $\hat{S}_n$ is constructed so that if for some $n \in \mathbb{N}$ we have $\hat{S}_n(x,y) \geqslant 1$, then since $(r_n)_n$ is decreasing, also $\sum_{i,j\in[0,l)} \mathbb{1}_{B(T_2^j y, r_l)}(T_1^i x) \geqslant 1$ for all $l \in [2^n, 2^{n+1}]$. Hence if $\hat{S}_n(x,y) \geqslant 1$ for all large n, then $(x,y) \in \liminf_n E_n$ and if $\hat{S}_n(x,y) \geqslant 1$ for all large n is true a.s., then $(\mu_1 \times \mu_2)(\liminf_n E_n) = 1$. We argue that the result follows if $\mathbb{E}(\hat{S}_n) \to \infty$ and

$$\sum_{n=1}^{\infty} \mathbb{E}\left(\frac{\hat{S}_n}{\mathbb{E}(\hat{S}_n)} - 1\right)^2 < \infty. \tag{3.2}$$

The argument is standard and well-known but we include it for completeness. By the Markov inequality, since $\left(\frac{\hat{S}_n}{\mathbb{E}(\hat{S}_n)} - 1\right)^2 \geqslant 0$, we get that for any $\delta > 0$,

$$(\mu_1 \times \mu_2)\left(\left(\frac{\hat{S}_n(x,y)}{\mathbb{E}(\hat{S}_n)} - 1\right)^2 \geqslant \delta\right) \leqslant \frac{1}{\delta}\mathbb{E}\left(\frac{\hat{S}_n}{\mathbb{E}(\hat{S}_n)} - 1\right)^2.$$

By (3.2), the right hand side is summable and hence the left hand side is as well By the Borel–Cantelli Lemma, along with the fact that $\delta > 0$ is arbitrary, we conclude that a.s. $\left|\frac{\hat{S}_n(x,y)}{\mathbb{E}(\hat{S}_n)} - 1\right| \geqslant \delta$ is true for at most finitely many $n \in \mathbb{N}$. In other words, a.s. for all n sufficiently large we have $\frac{\hat{S}_n(x,y)}{\mathbb{E}(\hat{S}_n)} \in (1-\delta, 1+\delta)$. Since $\mathbb{E}(\hat{S}_n) \to \infty$ also a.s. $\hat{S}_n(x,y) \to \infty$ and in particular a.s. $\hat{S}_n(x,y) \geqslant 1$ for all n sufficiently large which concludes the argument.

We now prove $\mathbb{E}(\hat{S}_n) \to \infty$ and (3.2). Using T_1-invariance of μ_1 and T_2-invariance of μ_2 we first compute

$$\mathbb{E}(\hat{S}_n) = 2^{2n}\int \mu_1(B(y, r_{2^{n+1}}))\, d\mu_2(y) \to \infty.$$

by the assumption on the lower bound of shrinking rate of the integral.

Next, note that the summands in (3.2) are equal to $\frac{\mathbb{E}(\hat{S}_n^2) - \mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2}$. Since we already have $\mathbb{E}(\hat{S}_n)$ we proceed to estimate $\mathbb{E}(\hat{S}_n^2)$. Because

$$\mathbb{E}(\hat{S}_n^2) = \sum_{i_1,i_2,j_1,j_2 \in [0,2^n)} \int\int \mathbb{1}_{B(T_2^{j_1}y, r_{2^{n+1}})}(T_1^{i_1}x)\cdot \mathbb{1}_{B(T_2^{j_2}y, r_{2^{n+1}})}(T_1^{i_2}x)\, d\mu_1(x) d\mu_2(y),$$

it suffices to assume that $i_1 < i_2$, $j_1 < j_2$ since for the three other combinations of inequalities, the upcoming estimates will give the same bound. For this case of indices we then split the sum into the four pairs arising from the cases where for some $c > 0$ (to be chosen later) $i_2 - i_1 \leqslant cn$, $i_2 - i_1 > cn$, $j_2 - j_1 \leqslant cn$, $j_2 - j_1 > cn$.

For $i_2 - i_1 > cn$ and $j_2 - j_1 > cn$, i.e. *the totally separated case*,

$$\begin{aligned}
&\sum_{\substack{i_1,i_2 \in [0,2^n) \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in [0,2^n) \\ j_2 - j_1 > cn}} \int\int \mathbb{1}_{B(T_2^{j_1}y, r_{2^{n+1}})}(T_1^{i_1}x) \mathbb{1}_{B(T_2^{j_2}y, r_{2^{n+1}})}(T_1^{i_2}x)\, d\mu_1(x) d\mu_2(y) \\
&= \sum_{\substack{i_1,i_2 \in [0,2^n) \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in [0,2^n) \\ j_2 - j_1 > cn}} \int\int \mathbb{1}_{B(T_2^{j_1}y, r_{2^{n+1}})}(x) \mathbb{1}_{B(T_2^{j_2}y, r_{2^{n+1}})}(T_1^{i_2 - i_1}x)\, d\mu_1(x) d\mu_2(y) \\
&\leqslant \sum_{\substack{i_1,i_2 \in [0,2^n) \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in [0,2^n) \\ j_2 - j_1 > cn}} \left(\int \mu_1(B(y, r_{2^{n+1}}))\mu_1(B(T_2^{j_2 - j_1}y, r_{2^{n+1}}))\, d\mu_2(y) + \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \left. + 2Ce^{-n\theta c}\right) \\
&\leqslant \sum_{\substack{i_1,i_2 \in [0,2^n) \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in [0,2^n) \\ j_2 - j_1 > cn}} \left(\left(\int \mu_1(B(y, r_{2^{n+1}}))\, d\mu_2(y)\right)^2 + 4Ce^{-n\theta c}\right) \\
&\leqslant 2^{4n}\left(\left(\int \mu_1(B(y, r_{2^{n+1}}))\, d\mu_2(y)\right)^2 + 4Ce^{-n\theta c}\right) = \mathbb{E}(\hat{S}_n)^2 + 2^{4n+2}Ce^{-n\theta c},
\end{aligned}$$

where the first two inequalities follow from the mixing properties of the two systems, and in the second we use Lemma 3.3 or Remark 3.4.

For $i_2 - i_1 \leqslant cn$ and $j_2 - j_1 \leqslant cn$, i.e. *the totally non-separated case*,

$$\begin{aligned}\sum_{\substack{i_1,i_2\in[0,2^n)\\ 0\leqslant i_2-i_1\leqslant cn}} \sum_{\substack{j_1,j_2\in[0,2^n)\\ 0\leqslant j_2-j_1\leqslant cn}} \int\int \mathbb{1}_{B(T_2^{j_1}y,r_{2^n+1})}(T_1^{i_1}x)\cdot \mathbb{1}_{B(T_2^{j_2}y,r_{2^n+1})}(T_1^{i_2}x)\,d\mu_1(x)d\mu_2(y) \\ \leqslant \sum_{\substack{i_1,i_2\in[0,2^n)\\ 0\leqslant i_2-i_1\leqslant cn}} \sum_{\substack{j_1,j_2\in[0,2^n)\\ 0\leqslant j_2-j_1\leqslant cn}} \int \mu_1(B(T_2^{j_1}y,r_{2^{n+1}}))\,d\mu_2(y) \\ = (cn)^2 2^{2n}\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y) = (cn)^2\mathbb{E}(\hat{S}_n).\end{aligned}$$

For $i_2 - i_1 > cn$ and $j_2 - j_1 \leqslant cn$, i.e. *a half-separated case* (we can swap T_1 and T_2 to get the other half-separated case),

$$\begin{aligned}&\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1>cn}} \sum_{\substack{j_1,j_2\in[0,2^n)\\ 0\leqslant j_2-j_1\leqslant cn}} \int\int \mathbb{1}_{B(T_2^{j_1}y,r_{2^n+1})}(T_1^{i_1}x)\cdot \mathbb{1}_{B(T_2^{j_2}y,r_{2^n+1})}(T_1^{i_2}x)\,d\mu_1(x)d\mu_2(y) \\ &= \sum_{\substack{i_1,i_2,j_1,j_2\in[0,2^n)\\ i_2-i_1>cn\\ 0\leqslant j_2-j_1\leqslant cn}} \int\int \mathbb{1}_{B(T_2^{j_1}y,r_{2^n+1})}(x)\cdot \mathbb{1}_{B(T_2^{j_2}y,r_{2^n+1})}(T_1^{i_2-i_1}x)\,d\mu_1(x)d\mu_2(y) \\ &\leqslant \sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1>cn}} \sum_{\substack{j_1,j_2\in[0,2^n)\\ 0\leqslant j_2-j_1\leqslant cn}} \Bigg(\int \mu_1(B(T_2^{j_1}y,r_{2^{n+1}}))\mu_1(B(T_2^{j_2}y,r_{2^{n+1}}))\,d\mu_2(y)+ \\ &\qquad\qquad + 2Ce^{-n\theta c}\Bigg) \\ &\leqslant \sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1>cn}} \sum_{\substack{j_1,j_2\in[0,2^n)\\ 0\leqslant j_2-j_1\leqslant cn}} \left(\int \mu_1(B(y,r_{2^{n+1}}))^2\,d\mu_2(y) + 2Ce^{-n\theta c}\right) \\ &\leqslant cn2^{3n}\left(\int \mu_1(B(y,r_{2^{n+1}}))^2\,d\mu_2(y) + 2Ce^{-n\theta c}\right),\end{aligned}$$

where in the penultimate line we use the Hölder inequality and T_2-invariance of μ_2.

As mentioned earlier, we aim to show that $\mathbb{E}(\hat{S}_n^2) - \mathbb{E}(\hat{S}_n)^2$, when divided by $\mathbb{E}(\hat{S}_n)^2$ are summable. Note that the definition of $\hat{S}_n$ along the subsequence 2^n becomes essential at this stage in the proof. Collecting our estimates from above we see that if we ignore multiplicative constants, the $\frac{\mathbb{E}(\hat{S}_n^2)-\mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2}$ are bounded by

$$\frac{2^{4n}e^{-n\theta c}}{\mathbb{E}(\hat{S}_n)^2} + \frac{n^2\mathbb{E}(\hat{S}_n)}{\mathbb{E}(\hat{S}_n)^2} + \frac{n2^{3n}\int \mu_1(B(y,r_{2^{n+1}}))^2\,d\mu_2(y)}{\mathbb{E}(\hat{S}_n)^2} + \frac{n2^{3n}e^{-n\theta c}}{\mathbb{E}(\hat{S}_n)^2}.$$

The fourth summand is dominated by the first so it suffices to demonstrate that the first three are summable. For the first summand we have,

$$\frac{2^{4n}e^{-n\theta c}}{\mathbb{E}(\hat{S}_n)^2} = \frac{e^{-n\theta c}}{\left(\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y)\right)^2} \leqslant \frac{2^{4n+4}e^{-n\theta c}}{(n^3(\log\log 2+\log n)^{1+\varepsilon}(\log 2)^3)^2}.$$

by the assumption that $\int \mu_1(B(y,r_n))\,d\mu_2(y) \geqslant \frac{(\log n)^3(\log\log n)^{1+\varepsilon}}{n^2}$. So choosing $c \geqslant 4\log 2/\theta$, this is summable.

Using the same assumption on the integral we get for the second summand that

$$\frac{n^2\mathbb{E}(\hat{S}_n)}{\mathbb{E}(\hat{S}_n)^2} = \frac{n^2}{2^{2n}\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y)} \leqslant \frac{1}{n(\log\log 2+\log n)^{1+\varepsilon}(\log 2)^3},$$

which is summable.

For the third summand we apply Lemma 3.2 and the assumption (2.1) for $i = 1$ to obtain

$$\begin{aligned}
&\frac{n2^{3n}\int \mu_1(B(y,r_{2^{n+1}}))^2\,d\mu_2(y)}{\mathbb{E}(\hat{S}_n)^2}\\
&\leqslant \frac{n2^{3n}\big(\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_1(y)\big)^{\frac{1}{2}}\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y)}{2^{4n}\left(\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y)\right)^2}\\
&= \frac{n\big(\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_1(y)\big)^{\frac{1}{2}}}{2^n\int \mu_1(B(y,r_{2^{n+1}}))\,d\mu_2(y)}.\\
&\leqslant \frac{2}{n(\log\log 2+\log n)^{1+\varepsilon}(\log 2)^2}
\end{aligned}$$

which is also summable.

As mentioned above, the other half-separated case is obtained by switching T_1 and T_2 and applying (2.1) for $i = 2$ instead of $i = 1$. This concludes the proof of part (a). $\square$

Proof of Theorem 2.4(b). Given a sequence $(r_n)_n$, let again

$$S_n(x,y) = \sum_{0\leqslant i,j<n} \mathbb{1}_{B(T_2^j y, r_n)}(T_1^i x).$$

Note that by the assumption on the shrinking rate of the integral,

$$\mathbb{E}(S_n) = n^2\int \mu_1(B(y,r_n))\,d\mu_2(y) \geqslant (\log n)^2 h(n) \to \infty.$$

We argue that the result follows if $\mathbb{E}\Big(\frac{S_n}{\mathbb{E}(S_n)}-1\Big)^2 \to 0$. Indeed, then there exists a subsequence $(n_k)_k$ s.t. $\Big(\frac{S_{n_k}(x,y)}{\mathbb{E}(S_{n_k})}-1\Big)^2 \to 0 \Rightarrow \frac{S_{n_k}(x,y)}{\mathbb{E}(S_{n_k})} \to 1$ a.s. Since the denominator goes to infinity, so must $S_{n_k}(x,y)$ for $k\to\infty$ a.s. This implies that $(x,y)\in\limsup_n E_n$ a.s.

To estimate the quantity $\mathbb{E}\Big(\frac{S_n}{\mathbb{E}(S_n)}-1\Big)^2$ we follow the steps of part (a). The difference to part (a) is limited to the fact that we sum over the indices in $[0,n)$ instead of $[0,2^n)$ and the threshold for indices to be "separated" is given by $c\log n$ instead of cn. Repeating the same calculations as in part (a), we obtain for part (b) the estimate (again ignoring multiplicative constants)

$$\frac{n^4e^{-c\theta\log n}}{\mathbb{E}(S_n)^2}+\frac{(\log n)^2\mathbb{E}(S_n)}{\mathbb{E}(S_n)^2}+\frac{(\log n)n^3\int \mu_1(B(y,r_n))^2\,d\mu_2(y)}{\mathbb{E}(S_n)^2}+\frac{(\log n)n^3e^{-c\theta\log n}}{\mathbb{E}(S_n)^2}$$

on the quantity of interest. Again, the fourth summand is dominated by the first so it suffices to demonstrate that the first three summands vanish.

For the first summand we get,

$$\frac{n^4e^{-c\theta\log n}}{\mathbb{E}(S_n)^2} = \frac{1}{n^{c\theta}\big(\int \mu_1(B(y,r_n))\,d\mu_2(y)\big)^2} \leqslant \frac{n^4}{n^{c\theta}(\log n)^4(h(n))^2}$$

by the assumption that $\int \mu_1(B(y,r_n))\,d\mu_2(y) \geqslant \frac{(\log n)^2h(n)}{n^2}$. So choosing $c\geqslant\frac{4}{\theta}$, this goes to zero.

Using the same assumption on the integral we get for the second summand that

$$\frac{(\log n)^2\mathbb{E}(S_n)}{\mathbb{E}(S_n)^2} = \frac{(\log n)^2}{n^2\int \mu_1(B(y,r_n))\,d\mu_2(y)} \leqslant \frac{1}{h(n)},$$

which goes to zero.

For the third summand we apply Lemma 3.2 and the assumption (2.2) to obtain

$$\frac{(\log n)n^3 \int \mu_1(B(y,r_n))^2\, d\mu_2(y)}{\mathbb{E}(S_n)^2} \leqslant \frac{(\log n)n^3 \left(\int \mu_1(B(y,r_n))^2\, d\mu_1(y)\right)^{\frac{1}{2}}}{n^4 \int \mu_1(B(y,r_n))^2\, d\mu_2(y)} \leqslant \frac{1}{h(n)},$$

which goes to zero.

This concludes the proof of part (b). □

3.1. **Proofs for the doubling map.** For the proofs relating to the doubling map we will use the notation $\mu^2 := (\mu \times \mu)$ to simplify expressions.

Proof of Theorem 2.11. The zero measure case follows from Theorem 2.1(b). We now consider the measure 1 case.

Clearly,

$$\begin{aligned} G_n &:= \bigcup_{0 \leqslant i \leqslant j < n} \{(x,y) : d(T^i x, T^j y) \leqslant r_n\} \\ &\subset \bigcup_{0 \leqslant i,j < n} \{(x,y) : d(T^i x, T^j y) \leqslant r_n\} = E_n, \end{aligned}$$

implying that $\limsup_n G_n \subset \limsup_n E_n$. In the following we will construct a subset of $\limsup_n G_n$ which is easier to work with. Consider now

$$H_n := \bigcup_{0 \leqslant i \leqslant j < n} \{(x,y) : d(T^i x, T^j y) \leqslant r_j\}.$$

Write

$$\begin{aligned} C_{i,j} &:= \{(x,y) : d(T^i x, T^j y) \leqslant r_j\} \\ B_{i,j,n} &:= \{(x,y) : d(T^i x, T^j y) \leqslant r_n\} \end{aligned}$$

such that

$$\begin{aligned} H_n &= \bigcup_{0 \leqslant i \leqslant j < n} C_{i,j} \\ G_n &= \bigcup_{0 \leqslant i \leqslant j < n} B_{i,j,n}. \end{aligned}$$

For a given $j \leqslant n$, $C_{i,j} = B_{i,j,j} \subset G_j$. Hence

$$H_n \subset \bigcup_{j=1}^{n} G_j.$$

We next want to restructure $J_n := \{(i,j) : 1 \leqslant i \leqslant j \leqslant n\}$ from being an array to a sequence. We do this by introducing the ordering $\preceq$ whereby $(i_1, j_1) \preceq (i_2, j_2)$ if $j_2 > j_1$ or if $j_2 = j_1$ and $i_2 \geqslant i_1$. Using this ordering we may reenumerate the $C_{i,j}$'s chronologically from 1 up to $|J_n| = \frac{n(n+1)}{2}$, i.e. for each $(i,j) \in J_n$, $C_{i,j} =: A_l$ for some $l \in \{1, \ldots, \frac{n(n+1)}{2}\}$. Note that this restructuring may equally well be done for the infinite array $J := \{(i,j) : 1 \leqslant i \leqslant j\}$ giving rise to an infinite sequence. This restructuring allows us to consider the lim sup set of the $C_{i,j} = A_l$'s, i.e.

$$\limsup_l A_l = \bigcap_{l=1}^{\infty} \bigcup_{k=l}^{\infty} A_k.$$

Suppose $(x,y) \in \limsup_l A_l$. This means that $(x,y) \in A_l$ for infinitely many l, or similarly, $(x,y) \in C_{i,j}$ for infinitely many distinct pairs $(i,j) \in J$. In particular, since there are for each j only finitely many options for i, there exists pairs $(i,j) \in J$ with arbitrarily large value of j for which $(x,y) \in C_{i,j}$. Since $C_{i,j} \subset G_j$ we have that $(x,y) \in G_j$ for infinitely many j. In other words, $(x,y) \in \limsup_n G_n$.

We conclude that

$$\limsup_l A_l \subset \limsup_n G_n \subset \limsup_n E_n$$

and consequently

$$\mu^2(\limsup_l A_l) \leqslant \mu^2\left(\limsup_n E_n\right).$$

We will prove that $\mu^2(\limsup_l A_l) = 1$, thereby concluding $\mu^2(\limsup_n E_n) = 1$.

Assume that

$$\sum_{n=1}^{\infty} n r_n = \infty. \tag{3.3}$$

Extending $\mathbb{1}_{B(0,r_j)}$ to a periodic function on $\mathbb{R}$, for any given $l \in \mathbb{N}$ we have

$$\mu^2(A_l) = \mu^2(C_{i,j}) = \int \mathbb{1}_{C_{i,j}} d\mu^2 = \iint \mathbb{1}_{B(0,r_j)}(2^i x - 2^j y)\, dx\, dy$$

for some $(i,j) \in J$. We may write the periodic extension of $\mathbb{1}_{B(0,r_j)}$ via its Fourier series,

$$\mathbb{1}_{B(0,r_j)}(z) = \sum_{k\in\mathbb{Z}} c_{j,k} e^{2\pi i k z}$$

which means that

$$\mu^2(A_l) = \sum_{k\in\mathbb{Z}} c_{j,k} \iint e^{2\pi i k(2^i x - 2^j y)}\, dx\, dy = c_{j,0} = \int \mathbb{1}_{B(0,r_j)} d\mu^2 = 2r_j$$

since the integrals are zero unless $k = 0$ in which case it is 1. From this we can compute the partial sums of the $\mu^2(A_l)$, namely, we write any given $M \in \mathbb{N}$ as $M = k_n + a$ for some $n \in \mathbb{N}$ where $k_n = \sum_{i=1}^n i$ and $1 \leqslant a < n+1$, thereby obtaining

$$\begin{aligned}\sum_{l=1}^{M} \mu^2(A_l) &= \sum_{l=1}^{k_n} \mu^2(A_l) + \sum_{l=k_n+1}^{M} \mu^2(A_l) = \sum_{1\leqslant i\leqslant j\leqslant n} \mu^2(C_{i,j}) + \sum_{i=1}^{a} \mu^2(C_{i,n+1}) \\ &= \sum_{j=1}^{n} j 2r_j + a 2r_{n+1} \xrightarrow{n\to\infty} \infty\end{aligned}$$

where the divergence holds by assumption (3.3). Clearly $M \to \infty$ for $n \to \infty$ implying the divergence of the first sum as well. We aim to employ the following consequence of the Erdős–Rényi version of the Borel–Cantelli lemma [ER] .

Lemma 3.5. *Let $(Y, \mathcal{B}, \nu)$ denote a probability space. If $A_l \in \mathcal{B}$ are sets such that*

$$\sum_{l=1}^{\infty} \nu(A_l) = \infty \tag{3.4}$$

and

$$\liminf_{N\to\infty} \frac{\displaystyle\sum_{1\leqslant l_1 < l_2 \leqslant N} \big(\nu(A_{l_1} \cap A_{l_2}) - \nu(A_{l_1})\nu(A_{l_2})\big)}{\left(\displaystyle\sum_{l=1}^{N} \nu(A_l)\right)^2} \leqslant 0, \tag{3.5}$$

then $\nu(\limsup_l A_l) = 1$.

Since we already showed that property (3.4) is satisfied, we are left with showing that (3.5) also holds. For this purpose we again invoke Fourier series. In the following, sets A_{l_1}, A_{l_2} will be identified with the sets C_{i_1,j_1}, C_{i_2,j_2} respectively when indexed via an array. Since we sum over $l_1 < l_2$ in condition (3.5) we may assume without loss of generality that $j_2 \geqslant j_1$. We have that

$$\begin{aligned}\mu^2(A_{l_1} \cap A_{l_2}) &= \iint \mathbb{1}_{B(0,r_{j_1})}(2^{i_1}x - 2^{j_1}y)\mathbb{1}_{B(0,r_{j_2})}(2^{i_2}x - 2^{j_2}y)\,dx\,dy \\ &= \sum_{k_1,k_2\in\mathbb{Z}} c_{j_1,k_1}c_{j_2,k_2}\iint e^{2\pi i\left(k_1(2^{i_1}x-2^{j_1}y)+k_2(2^{i_2}x-2^{j_2}y)\right)}\,dx\,dy \\ &= \sum_{k_1,k_2\in\mathbb{Z}} c_{j_1,k_1}c_{j_2,k_2}\int e^{2\pi i(k_1 2^{i_1}+k_2 2^{i_2})x}\,dx\int e^{-2\pi i(k_1 2^{j_1}+k_2 2^{j_2})y}\,dy.\end{aligned}$$

We see that the integrals are only non-zero if

$$\begin{cases} k_1 2^{i_1} + k_2 2^{i_2} &= 0 \\ k_1 2^{j_1} + k_2 2^{j_2} &= 0 \end{cases} \quad\Rightarrow\quad \begin{cases} k_1 &= -k_2 2^{i_2-i_1} \\ k_1 &= -k_2 2^{j_2-j_1}. \end{cases}$$

The last two equations can only be true if $p := j_2 - j_1 = i_2 - i_1$ or $k_1 = k_2 = 0$. Suppose that $j_2 - j_1 = i_2 - i_1$. In this case the sum over $k_1 \in \mathbb{Z}$ can be replaced by summing over the values $-k_2 2^p$. Furthermore, the central coefficients satisfy $c_{j_1,0}c_{j_2,0} = \mu(A_{l_1})\mu(A_{l_2})$ which we may subtract to the left hand side to obtain

$$\begin{aligned}\mu^2(A_{l_1} \cap A_{l_2}) - \mu^2(A_{l_1})\mu^2(A_{l_2}) &= \sum_{k_2\in\mathbb{Z}\setminus\{0\}} c_{j_1,-k_2 2^p}c_{j_2,k_2} \\ &\leqslant \sum_{k_2\in\mathbb{Z}\setminus\{0\}} |c_{j_1,-k_2 2^p}c_{j_2,k_2}|.\end{aligned}$$

We have two upper bounds for the coefficients, namely the inverse of the index over which the coefficient is summed as well as the integral of the function that the Fourier series represents. More precisely,

$$|c_{j_1,-k_2 2^p}c_{j_2,k_2}| \leqslant \begin{cases} \frac{1}{|k_2|2^p}\frac{1}{|k_2|} = \frac{1}{k_2^2 2^p} \\ \int \mathbb{1}_{B(0,r_{j_1})}\,d\mu^2 \int \mathbb{1}_{B(0,r_{j_2})}\,d\mu^2 = 4r_{j_1}r_{j_2} \end{cases}.$$

To optimize our upper bound we compute

$$4r_{j_1}r_{j_2} \leqslant \frac{1}{k_2^2 2^p} \quad\Rightarrow\quad k_2 \leqslant (4r_{j_1}r_{j_2})^{-\frac{1}{2}}2^{-\frac{p}{2}}$$

and split up our sum accordingly, that is, for $S_{j_1,j_2} := (4r_{j_1}r_{j_2})^{-\frac{1}{2}}2^{-\frac{p}{2}}$ we have

$$\begin{aligned}\sum_{k_2\in\mathbb{Z}\setminus\{0\}} |c_{j_1,-k_2 2^p}c_{j_2,k_2}| &= \sum_{k_2=1}^{\lfloor S_{j_1,j_2}\rfloor} |c_{j_1,-k_2 2^p}c_{j_2,k_2}| + \sum_{k_2=S_{j_1,j_2}}^{\infty} |c_{j_1,-k_2 2^p}c_{j_2,k_2}| \\ &\leqslant 2\left(\sum_{k_2=1}^{\lfloor S_{j_1,j_2}\rfloor} 4r_{j_1}r_{j_2} + \sum_{k_2=S_{j_1,j_2}}^{\infty}\frac{1}{k_2^2 2^p}\right) \\ &\leqslant 8r_{j_1}r_{j_2}(4r_{j_1}r_{j_2})^{-\frac{1}{2}}2^{-\frac{p}{2}} + \frac{2}{S_{j_1,j_2}2^p} \\ &\ll (r_{j_1}r_{j_2})^{\frac{1}{2}}2^{-\frac{p}{2}}.\end{aligned}$$

In total we conclude that

$$\mu^2(A_{l_1} \cap A_{l_2}) - \mu^2(A_{l_1})\mu^2(A_{l_2}) \begin{cases} = 0 & \text{if } j_2 - j_1 \neq i_2 - i_1, \\ \ll (r_{j_1}r_{j_2})^{\frac{1}{2}}2^{-\frac{j_2-j_1}{2}} & \text{if } j_2 - j_1 = i_2 - i_1. \end{cases}$$

With this estimate in hand we are ready to prove property (3.5) of Lemma 3.5. We have

$$\sum_{1\leqslant l_1<l_2\leqslant N}\left(\mu^2(A_{l_1} \cap A_{l_2}) - \mu^2(A_{l_1})\mu^2(A_{l_2})\right) \ll \sum_{1\leqslant l_1<l_2\leqslant N}(r_{j_1}r_{j_2})^{\frac{1}{2}}2^{-\frac{j_2-j_1}{2}}$$

$$\leqslant \sum_{1 \leqslant l_1 \leqslant l_2 \leqslant N} r_{j_1} 2^{-\frac{j_2 - j_1}{2}}$$

since $r_j \geqslant r_{j+1}$ by assumption. Recall that the sets A_{l_1}, A_{l_2} correspond to the sets C_{i_1,j_1}, C_{i_2,j_2}, hence we need to rewrite the sum over $1 \leqslant l_1 \leqslant l_2 \leqslant N$ as a sum over the appropriate index set of the i_1, i_2, j_1, j_2. Since the condition in (3.5) is given via the lim inf, it is sufficient if we can prove that the fraction vanishes along a subsequence. We will do this for the subsequence $N = k_n$ where $k_n = \sum_{i=1}^n i$. Define the sets

$$I_n := \left\{ (i_1, i_2, j_1, j_2) \in \mathbb{N}^4 : \begin{array}{ll} i_1 \leqslant j_1 \leqslant n & j_2 - j_1 = i_2 - i_1 \\ i_2 \leqslant j_2 \leqslant n & j_1 \leqslant j_2 \end{array} \right\}.$$

So we have

$$\sum_{1 \leqslant l_1 \leqslant l_2 \leqslant N} r_{j_1} 2^{-\frac{j_2 - j_1}{2}} = \sum_{I_n} r_{j_1} 2^{-\frac{j_2 - j_1}{2}}.$$

First we note that the sum on the right hand side is independent of i_1, i_2. For fixed $j_1 \leqslant j_2 \leqslant n$ there are exactly j_1 pairs (i_1, i_2) such that $(i_1, i_2, j_1, j_2) \in I_n$, namely

$$\begin{gathered} (1, 1 + (j_2 - j_1)) \\ (2, 2 + (j_2 - j_1)) \\ \vdots \\ (j_1, j_1 + (j_2 - j_1)). \end{gathered}$$

Hence we may write

$$\begin{aligned} \sum_{I_n} r_{j_1} 2^{-\frac{j_2 - j_1}{2}} &= \sum_{1 \leqslant j_1 \leqslant j_2 \leqslant n} j_1 r_{j_1} 2^{-\frac{j_2 - j_1}{2}} = \sum_{j_1=1}^{n} j_1 r_{j_1} \sum_{j_2 = j_1}^{n} 2^{-\frac{j_2 - j_1}{2}} \\ &\leqslant \sum_{j_1=1}^{n} j_1 r_{j_1} \sum_{m=0}^{\infty} 2^{-\frac{m}{2}} = C_1 \sum_{j_1=1}^{n} j_1 r_{j_1}. \end{aligned}$$

This gives (3.5) since then

$$\begin{aligned} \lim_{n \to \infty} \frac{\sum\limits_{1 \leqslant l_1 < l_2 \leqslant k_n} \left(\mu^2(A_{l_1} \cap A_{l_2}) - \mu^2(A_{l_1}) \mu^2(A_{l_2}) \right)}{\left(\sum\limits_{l=1}^{N} \mu^2(A_l) \right)^2} &\leqslant \lim_{n \to \infty} \frac{C_1 \sum_{j=1}^n j r_j}{(\sum_{j=1}^n j r_j)^2} \\ &= \lim_{n \to \infty} \frac{C_1}{\sum_{j=1}^n j r_j} = 0 \end{aligned}$$

since $\sum_{j=1}^n j r_j \to \infty$ for $n \to \infty$ by assumption. We conclude that under this assumption $\mu^2(\limsup_n E_n) = 1$. $\square$

Proof of Theorem 2.9. Part (a) is an immediate consequence of Theorem 2.1(a).

Proof of part (b). For a decreasing sequence $(r_n)_n$, we again let

$$\hat{S}_n(x, y) = \sum_{i,j \in [0, 2^n)} \mathbb{1}_{B(T^j y, r_{2^{n+1}})}(T^i x).$$

The structure of the proof is the same as that of the proof of Theorem 2.4(a). As explained in detail there, the result follows if we can prove that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}(\hat{S}_n^2) - \mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2} < \infty.$$

To this end, we first compute the expectation of $\hat{S}_n$. By T-invariance of μ we get

$$\begin{aligned}\mathbb{E}(\hat{S}_n) &= \sum_{i,j\in[0,2^n)} \int \mu(B(T^j y, r_{2^{n+1}}))\,d\mu(y)\\ &= \sum_{i,j\in[0,2^n)} 2r_{2^{n+1}} = 2^{2n+1} r_{2^{n+1}}. \qquad (3.6)\end{aligned}$$

Note that condition (2.3) implies $\mathbb{E}(\hat{S}_n) \to \infty$. We are left with computing $\mathbb{E}(\hat{S}_n^2)$. We have

$$\begin{aligned}\hat{S}_n(x,y)^2 &= \sum_{i_1,i_2,j_1,j_2\in[0,2^n)} \mathbb{1}_{B(T^{j_1}y, r_{2^{n+1}})}(T^{i_1}x)\mathbb{1}_{B(T^{j_2}y, r_{2^{n+1}})}(T^{i_2}x)\\ &= \sum_{i_1,i_2,j_1,j_2\in[0,2^n)} \mathbb{1}_{B(0, r_{2^{n+1}})}(T^{i_1}x - T^{j_1}y)\mathbb{1}_{B(0, r_{2^{n+1}})}(T^{i_2}x - T^{j_2}y)\\ &= \sum_{i_1,i_2,j_1,j_2\in[0,2^n)} \mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_1}x - 2^{j_1}y)\mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_2}x - 2^{j_2}y).\end{aligned}$$

Taking the expectation then gives

$$\mathbb{E}(\hat{S}_n^2) = \sum_{\substack{i_1,i_2\in[0,2^n)\\ j_1,j_2\in[0,2^n)}} \iint \mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_1}x - 2^{j_1}y)\mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_2}x - 2^{j_2}y)\,dx\,dy.$$

We now proceed as in the proof of Theorem 2.11 by writing $\mathbb{1}_{B(0,r_{2^{n+1}})}$ via its Fourier series. For given $i_1, i_2, j_1, j_2 \in [1, 2^n]$ we obtain

$$\begin{aligned}&\iint \mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_1}x - 2^{j_1}y)\mathbb{1}_{B(0, r_{2^{n+1}})}(2^{i_2}x - 2^{j_2}y)\,dx\,dy\\ &\quad= \sum_{k_1,k_2\in\mathbb{Z}} c_{j_1,k_1}c_{j_2,k_2} \iint e^{2\pi i\left(k_1(2^{i_1}x - 2^{j_1}y) + k_2(2^{i_2}x - 2^{j_2}y)\right)}\,dx\,dy\\ &\quad= \sum_{k_1,k_2\in\mathbb{Z}} c_{j_1,k_1}c_{j_2,k_2} \int e^{2\pi i(k_1 2^{i_1} + k_2 2^{i_2})x}\,dx \int e^{2\pi i(k_1 2^{j_1} + k_2 2^{j_2})y}\,dy.\end{aligned}$$

As in the proof of Theorem 2.11 we see that the integrals are only non-zero if

$$\begin{cases} k_1 2^{i_1} + k_2 2^{i_2} = 0 \\ k_1 2^{j_1} + k_2 2^{j_2} = 0 \end{cases} \Rightarrow \begin{cases} k_1 = -k_2 2^{i_2 - i_1} \\ k_1 = -k_2 2^{j_2 - j_1}, \end{cases}$$

that is, when $k_1 = k_2 = 0$ or when $p := j_2 - j_1 = i_2 - i_1$. Suppose $j_2 - j_1 = i_2 - i_1$. In this proof we need to consider the two cases $p \geqslant 0$ and $p < 0$. For $p \geqslant 0$ we use the equation above for k_1, i.e. we sum over $k_1 = -k_2 2^p$. For $p < 0$ we instead replace the sum over k_2 by $k_2 = -k_1 2^{-p} = -k_1 2^{|p|}$. However, referring again to the proof of Theorem 2.11, we see that the final estimate of the Fourier coefficients does not depend on k_1 or k_2, hence we can use the same bound. In total we get

$$\begin{aligned}&\sum_{k_1,k_2\in\mathbb{Z}} c_{j_1,k_1}c_{j_2,k_2} \int e^{2\pi i(k_1 2^{i_1} + k_2 2^{i_2})x}\,dx \int e^{2\pi i(k_1 2^{j_1} + k_2 2^{j_2})y}\,dy\\ &\qquad\qquad \leqslant c_{j_1,0}c_{j_2,0} + \begin{cases} 0 & \text{if } j_2 - j_1 \neq i_2 - i_1, \\ r_{2^{n+1}} 2^{-\frac{|j_2 - j_1|}{2}} & \text{if } j_2 - j_1 = i_2 - i_1. \end{cases}\end{aligned}$$

Define the index set

$$\tilde{I}_n := \left\{(i_1, i_2, j_1, j_2) \in \mathbb{N}^4 : i_1, i_2, j_1, j_2 \leqslant n\ ,\ j_2 - j_1 = i_2 - i_1\right\}.$$

The desired expectation can now be estimated through the expression

$$\mathbb{E}(\hat{S}_n)^2 \leqslant \sum_{i_1,i_2,j_1,j_2\in[0,2^n)} 4r_{2^{n+1}}^2 + \sum_{i_1,i_2,j_1,j_2\in\tilde{I}_{2^n}} r_{2^{n+1}} 2^{-\frac{|j_2 - j_1|}{2}}$$

$$= 2^{4n+2} r_{2^{n+1}}^2 + r_{2^{n+1}} \sum_{i_1,i_2,j_1,j_2 \in \tilde{I}_{2^n}} 2^{-\frac{|j_2-j_1|}{2}}.$$

We focus on the remaining sum. Since it does not depend on i_1, i_2 we consider the number of pairs (i_1, i_2) for a given pair (j_1, j_2), for which $j_2 - j_1 = i_2 - i_1$. Certainly this can be bounded from above by 2^n. This means that

$$\begin{aligned}
\mathbb{E}(\hat{S}_n)^2 &\leqslant 2^{4n+2} r_{2^{n+1}}^2 + 2^n r_{2^{n+1}} \sum_{j_1,j_2 \in [0,2^n)} 2^{-\frac{|j_2-j_1|}{2}} \\
&\leqslant 2^{4n+2} r_{2^{n+1}}^2 + 2^n r_{2^{n+1}} 2 \sum_{j_1=0}^{2^n-1} \sum_{j_2=j_1}^{2^n} 2^{-\frac{j_2-j_1}{2}} \\
&= 2^{4n+2} r_{2^{n+1}}^2 + 2^n r_{2^{n+1}} 2 \sum_{j_1=0}^{2^n-1} \sum_{m=0}^{\infty} 2^{-\frac{m}{2}} \\
&\leqslant 2^{4n+2} r_{2^{n+1}}^2 + 2^n r_{2^{n+1}} 2 \sum_{j_1=0}^{2^n-1} C_1 \\
&\leqslant 2^{4n+2} r_{2^{n+1}}^2 + C_1 2^{2n+1} r_{2^{n+1}},
\end{aligned}$$

for some $C_1 > 0$.

Using this estimate along with (3.6) we can now compute

$$\begin{aligned}
\frac{\mathbb{E}(\hat{S}_n^2) - \mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2} &\leqslant \frac{2^{4n+2} r_{2^{n+1}}^2 + C_1 2^{2n+1} r_{2^{n+1}} - (2^{2n+1} r_{2^{n+1}})^2}{(2^{2n+1} r_{2^{n+1}})^2} \qquad (3.7)\\
&= \frac{C_1}{2^{2n+1} r_{2^{n+1}}}.
\end{aligned}$$

Hence

$$\sum_{n=0}^{\infty} \frac{C_1}{2^{2n+1} r_{2^{n+1}}} = C_1 \sum_{n=1}^{\infty} \frac{1}{2^{2n} r_{2^n}} < \infty$$

by (2.3). We conclude that $\mu^2(\liminf_n E_n) = 1$. □

4. Proofs for Rotations

Parts (a) and (b) of Theorem 2.14 follow immediately from Theorem 2.1 exploiting that μ_1 and μ_2 are Lebesgue measure. The proof of parts (c) and (d), given below, is similar to that of Theorem 2.4.

Proof of Theorem 2.14(c). Given a decreasing sequence $(r_n)_n$, set again

$$\hat{S}_n(x,y) := \sum_{i,j \in [0,2^n)} \mathbb{1}_{B(T_2^j y, r_{2^{n+1}})}(T_1^i x).$$

As argued in Theorem 2.4(a), it is sufficient to show that $\sum_{n=1}^{\infty} \frac{\mathbb{E}(\hat{S}_n^2) - \mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2} < \infty$. As in the doubling map case, we can write $\mathbb{E}(\hat{S}_n) = 2^{2n+1} r_{2^{n+1}}$.

We begin by estimating $\mathbb{E}(\hat{S}_n^2)$ which we may also write as

$$\sum_{i_1,i_2,j_1,j_2 \in [0,2^n)} \iint \mathbb{1}_{B(T_2^{j_1} y, r_{2^{n+1}})}(T_1^{i_1} x) \cdot \mathbb{1}_{B(T_2^{j_2} y, r_{2^{n+1}})}(T_1^{i_2} x)\, d\mu_1(x) d\mu_2(y).$$

Since only one the systems is mixing we distinguish just between two cases, namely the separated case $i_2 - i_1 > cn$ and the non-separated case $i_2 - i_1 \leqslant cn$ with

$c > 4\log(2)/\theta$. For the separated case we use exponential mixing of T_1 with respect to Lebesgue measure μ_1 to obtain

$$\begin{aligned}&\sum_{j_1,j_2\in[0,2^n)}\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1>cn}}\iint \mathbb{1}_{B(T_2^{j_1}y,r_{2^{n+1}})}(T_1^{i_1}x)\cdot\mathbb{1}_{B(T_2^{j_2}y,r_{2^{n+1}})}(T_1^{i_2}x)\ d\mu_1(x)d\mu_2(y)\\ &\leqslant \sum_{j_1,j_2\in[0,2^n)}\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1>cn}}\int\Big(\mu_1(B(T_2^{j_1}y,r_{2^{n+1}})\mu_1(B(T_2^{j_2}y,r_{2^{n+1}})+\\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + 2C\mathrm{e}^{-n\theta c}\Big)d\mu_2(y)\\ &\leqslant 2^{4n}\cdot\big(4r_{2^{n+1}}^2+2C\mathrm{e}^{-n\theta c}\big).\end{aligned}$$

To investigate the non-separated case $i_2 - i_1 \leqslant cn$ we write

$$\begin{aligned}&\sum_{j_1,j_2\in[0,2^n)}\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}}\iint \mathbb{1}_{B(T_2^{j_1}y,r_{2^{n+1}})}(T_1^{i_1}x)\mathbb{1}_{B(T_2^{j_2}y,r_{2^{n+1}})}(T_1^{i_2}x)\ d\mu_1(x)d\mu_2(y)\\ &=\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}}\sum_{j_1,j_2\in[0,2^n)}\iint \mathbb{1}_{B(T_1^{i_1}x,r_{2^{n+1}})}(y)\cdot\\ &\qquad\qquad\qquad\cdot\mathbb{1}_{B(T_1^{i_2}x,r_{2^{n+1}})}(T_2^{j_2-j_1}y)\ d\mu_2(y)d\mu_1(x)\end{aligned}$$

and we note that it is enough to prove part (c) under the assumption that

$$r_n = \frac{(\log n)^2(\log\log n)^{1+\varepsilon}}{n^2}.$$

This will allow the application of the following claim.

Claim 1. *Suppose α admits a function $\Psi_\alpha\colon \mathbb{N}\to\mathbb{R}$ such that $\|j\alpha\| \geqslant \Psi_\alpha(n)$ for all $j \in [1,n]$ and that $r_n = o(\Psi_\alpha(n))$.*

Then, for all large enough n, if y and z are two given points, there is at most one $j \in [1,n]$ such that $T_2^j y \in B(z,r_n)$.

Proof. The claim follows if we can show for all large enough n that it is not possible for distinct $j_1, j_2 \in [0,n)$ to have $d(T_2^{j_1}y, T_2^{j_2}y) < 2r_n$. So it is sufficient to show that $d(T_2^j 0, 0) \geqslant \Psi_\alpha(n)$ for all $j \in [1,n]$. Another way of writing this is $\|j\alpha\| \geqslant \Psi_\alpha(n)$ for $j \in [1,n]$, which is guaranteed by the Diophantine assumption on α. □

Set $\Psi_\alpha(n) = c(\alpha)(\log n)^2(\log\log(n))^{1+2\varepsilon}/n^2$. By the assumptions in part (c), we see that $r_n = o(\Psi_\alpha(n))$ and since $\Psi_\alpha(n)$ is decreasing also $\|j\alpha\| \geqslant \Psi_\alpha(n)$ for all $j \in [1,n]$. Thus $\Psi_\alpha(n)$ satisfies the requirements of the claim.

By the claim, for each each fixed x, y, i_1, i_2 and j_1 there can be at most one $j_2 = j_2(x,y,i_1,i_2,j_1) \in [0,2^n)$ such that $\mathbb{1}_{B(T_1^{i_1}x,r_{2^{n+1}})}(y)\cdot\mathbb{1}_{B(T_1^{i_2}x,r_{2^{n+1}})}(T_2^{j_2-j_1}y) \neq 0$. We then write $\tilde{j} = \tilde{j}(x,y,i_1,i_2,j_1) := j_2 - j_1$ and our sum from above as

$$\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}}\sum_{j_1\in[0,2^n)}\iint \mathbb{1}_{B(T_1^{i_1}x,r_{2^{n+1}})}(y)\cdot\mathbb{1}_{B(T_1^{i_2}x,r_{2^{n+1}})}(T_2^{\tilde{j}}y)\ d\mu_2(y)d\mu_1(x)$$

noting that $\tilde{j}$ may not exist in which case the integrad above is 0. All together we get

$$\sum_{j_1,j_2\in[0,2^n)}\sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}}\iint \mathbb{1}_{B(T_2^{j_1}y,r_{2^{n+1}})}(T_1^{i_1}x)\cdot\mathbb{1}_{B(T_2^{j_2}y,r_{2^{n+1}})}(T_1^{i_2}x)\ d\mu_1(x)d\mu_2(y)$$

$$= \sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}} \sum_{j_1\in[0,2^n)} \iint \mathbb{1}_{B(y,r_{2^{n+1}})}(T_1^{i_1}x)\cdot \mathbb{1}_{B(T_2^{\tilde{j}}y,r_{2^{n+1}})}(T_1^{i_2}x)\, d\mu_1(x)d\mu_2(y)$$

$$\leqslant \sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}} \sum_{j_1\in[0,2^n)} \iint \mathbb{1}_{B(y,r_{2^{n+1}})}(T_1^{i_1}x)\, d\mu_1(x)d\mu_2(y)$$

$$= \sum_{\substack{i_1,i_2\in[0,2^n)\\ i_2-i_1\leqslant cn}} \sum_{j_1\in[0,2^n)} \int \mu_1(B(y,r_{2^{n+1}}))\, d\mu_2(y) \leqslant cn2^{2n}2r_{2^{n+1}}.$$

So if $r_n \geqslant \frac{(\log n)^2(\log\log n)^{1+\varepsilon}}{n^2}$, then

$$\frac{\mathbb{E}(\hat{S}_n^2)-\mathbb{E}(\hat{S}_n)^2}{\mathbb{E}(\hat{S}_n)^2} \leqslant \frac{2^{4n+1}C\mathrm{e}^{-n\theta c}+cn2^{2n}2r_{2^{n+1}}}{2^{4n+2}r_{2^{n+1}}^2} \tag{4.1}$$

is summable since $c > 4\log(2)/\theta$. □

Proof of Theorem 2.14(d). Analogously to the proof of Theorem 2.4 (b), it is enough to observe that (4.1) with $\hat{S}_n$ replaced by S_n from (3.1) goes to zero under the assumption $r_n = \log(n)h(n)/n^2$ with $h(n)\to\infty$. Furthermore, the Diophantine condition on α ensures $r_n = o(\Psi_\alpha(n))$ and, hence, the applicability of Claim 1 in this case. □

5. Proofs for one orbit

Suppose $X \subset \mathbb{R}^d$ is compact and $r > 0$. Then we can cover X by balls $\{B(x_p,r)\}_{p=1}^{k(r)}$ where $k(r)\asymp r^{-d}$ and there exists a $C_0\in\mathbb{N}$ such that each $x\in X$ belongs to at most C_0 elements of $\{B(x_p,2r)\}_{p=1}^{k(r)}$.[1]

We will write $\mathbb{1}_{p,r} := \mathbb{1}_{B(x_p,2r)}$. For a proof of the following easy fact, see the ideas in [GRS, Lemma 12].

Lemma 5.1. *Let $X\subset\mathbb{R}^d$ be compact. For all $x,y\in X$ we have,*

$$\mathbb{1}_{B(x,r)}(y) \leqslant \sum_{p=1}^{k(r)} \mathbb{1}_{p,r}(x)\mathbb{1}_{p,r}(y) \leqslant C_0\mathbb{1}_{B(x,4r)}(y).$$

Consequently, for a probability measure μ, we have,

$$\int \mu(B(x,r))\, d\mu(x) \leqslant \sum_{p=1}^{k(r)} \mu(B(x_p,2r))^2 \leqslant C_0\int \mu(B(x,4r))\, d\mu(x).$$

Recall the notation $A(r,n) = \{x : d(x,T^nx) < r\}$ and the short return time estimate (2.4),

$$\mu(A(r,n)) \leqslant Cr^s. \tag{5.1}$$

From this point on, let $X\subset\mathbb{R}$ denote an interval. A lemma by Kirsebom, Kunde and Persson [KKP, Lemma 3.1], implies that under exponential mixing for BV against L^1 observables,

$$\mu(A(r,n)) \leqslant \int \mu(B(x,r))\, \mathrm{d}\mu(x) + C_1e^{-\theta n}. \tag{5.2}$$

for some constant $C_1 > 0$.

[1]For general metric spaces this property is known as *bounded local complexity*.

Proof of Theorem 2.17(a). Given a sequence $(r_n)_n$, let

$$Q_n(x) := \sum_{0 \leqslant i < j < n} \mathbb{1}_{B(T^i x, r_n)}(T^j x).$$

As argued in the proof of Theorem 2.1(a), the result follows if $\mathbb{E}(Q_n) \to 0$. We have that

$$\mathbb{E}(Q_n) = \sum_{0 \leqslant i < j < n} \int \mathbb{1}_{B(T^i x, r_n)}(T^j x)\, d\mu(x) = \sum_{0 \leqslant i < j < n} \mu(A(r_n, j-i)).$$

Using (5.2) when $j - i > 3\theta^{-1} \log n$, and (5.1) otherwise, we obtain

$$\begin{aligned}\sum_{0 \leqslant i < j < n} \mu(A(r_n, j-i)) &= \sum_{\substack{0 \leqslant i < j < n \\ j-i \leqslant 3\theta^{-1} \log n}} \mu(A(r_n, j-i)) + \sum_{\substack{0 \leqslant i < j < n \\ j-i > 3\theta^{-1} \log n}} \mu(A(r_n, j-i)) \\ &\leqslant Cn(\log n) r_n^s + n^2 \int \mu(B(x, r_n))\, d\mu(x) + \frac{C_1}{n}.\end{aligned}$$

Hence $\mathbb{E}(Q_n) \to 0$ as $n \to \infty$. $\square$

Proof of Theorem 2.17(b). Given a decreasing sequence $(r_n)_n$, let

$$\begin{aligned}\tilde{Q}_n(x) &:= \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j} \mathbb{1}_{B(T^i x, r_{2^n})}(T^j x). \\ \tilde{Q}(x) &:= \sum_{n=0}^{\infty} \tilde{Q}_n(x).\end{aligned}$$

As argued in the proof of Theorem 2.1(b), the result follows if $\mathbb{E}(\tilde{Q}) < \infty$.

We have

$$\begin{aligned}\mathbb{E}(\tilde{Q}) &= \sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j} \int \mathbb{1}_{B(T^i x, r_{2^n})}(T^j x)\, \mathrm{d}\mu \\ &= \sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j} \mu(A(r_{2^n}, j-i)).\end{aligned}$$

As in part (a), we combine (5.2) with (5.1) in order to estimate $\mathbb{E}(\tilde{Q})$. Take $c > 2/\theta$. We split the above sum as

$$\begin{aligned}\mathbb{E}(\tilde{Q}) \leqslant \sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j - c \log j} \mu(A(r_{2^n}, j-i)) \\ + \sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{j - c \log j \leqslant i < j} \mu(A(r_{2^n}, j-i)).\end{aligned} \tag{5.3}$$

For the first part we use (5.2) and obtain that

$$\begin{aligned}\sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j - c \log j} \mu(A(r_{2^n}, j-i)) \\ \leqslant \sum_{n=0}^{\infty} \sum_{2^n < j < 2^{n+1}} \sum_{0 \leqslant i < j - c \log j} \left(\int \mu(B(x, r_{2^n}))\, \mathrm{d}\mu(x) + C_1 e^{-\theta(j-i)} \right).\end{aligned}$$

Then

$$\sum_{n=0}^{\infty} \sum_{2^n \leqslant j < 2^{n+1}} \sum_{0 \leqslant i < j - c \log j} \int \mu(B(x, r_{2^n}))\, \mathrm{d}\mu(x) \leqslant \sum_{n=0}^{\infty} 2^{2n} \int \mu(B(x, r_{2^n}))\, \mathrm{d}\mu(x),$$

which, since $n\int \mu(B(x,r_n))\,d\mu(x)$ is assumed decreasing, is finite by Cauchy condensation, Lemma 3.1. For the part containing the remainder term $C_1e^{-\theta(j-i)}$ we estimate that

$$\sum_{n=0}^{\infty}\sum_{2^n\leqslant j<2^{n+1}}\sum_{0\leqslant i<j-c\log j} C_1e^{-\theta(j-i)} \leqslant \sum_{n=0}^{\infty}\sum_{2^n<j<2^{n+1}} C_1\frac{1}{j^2} \leqslant \sum_{n=0}^{\infty} C_1 2^{-n},$$

which is also finite.

We now turn to the second part of (5.3). Here we use (5.1) to get

$$\sum_{n=0}^{\infty}\sum_{2^n\leqslant j<2^{n+1}}\sum_{j-c\log j\leqslant i<j} \mu(A(r_{2^n},j-i)) \leqslant \sum_{n=0}^{\infty}\sum_{2^n\leqslant j<2^{n+1}} c(\log j)Cr_{2^n}^s.$$

Using that $r_n\leqslant n^{-1/s}(\log n)^{-2/s-\varepsilon}$, we see that this is finite as well.

These estimates together show that $\mathbb{E}(\tilde{Q})<\infty$ and hence that $\mu(\limsup_n F_n)=0$. $\square$

Proof of Theorem 2.17(c). Let $\gamma\in(0,1/2)$ and define for a decreasing sequence $(r_n)_n$,

$$\hat{Q}_n(x) := \sum_{\substack{i\in[0,\gamma 2^n)\\ j\in[(1-\gamma)2^n,2^n)}}\sum_{p=1}^{k(r_{2^{n+1}})} \mathbb{1}_{p,r_{2^{n+1}}}(T^ix)\mathbb{1}_{p,r_{2^{n+1}}}(T^jx).$$

For notational simplicity, set $L_n^\gamma := [0,\gamma 2^n)$, $R_n^\gamma := [(1-\gamma)2^n,2^n)$ and $\bar{r} := r_{2^{n+1}}$. Suppose $\hat{Q}_n(x)\geqslant 1$ for some $n\in\mathbb{N}$. This means that there exists $(i,j)\in L_n^\gamma\times R_n^\gamma$ and an x_p with $\mathbb{1}_{p,\bar{r}}(T^ix)\mathbb{1}_{p,\bar{r}}(T^jx)\geqslant 1$ which in turn implies that $d(T^ix,T^jx)\leqslant 4\bar{r}$. Since $(r_n)_n$ is decreasing, we then also have that $\hat{Q}_l(x)\geqslant 1$ for all $l\in[2^n,2^{n+1}]$. Hence if $\hat{Q}_n(x)\geqslant 1$ for all sufficiently large n a.s., then $x\in\liminf_n F^T_{n,4r_n}$ a.s. As argued in the proof of Theorem 2.4 (a), this will follow if $\sum_{n=1}^{\infty}\frac{\mathbb{E}(\hat{Q}_n^2)-\mathbb{E}(\hat{Q}_n)^2}{\mathbb{E}(\hat{Q}_n)^2}<\infty$.

First we compute $\mathbb{E}(\hat{Q}_n)$.

$$\begin{aligned}
\mathbb{E}(\hat{Q}_n) &= \sum_{i\in L_n^\gamma,j\in R_n^\gamma}\sum_{p=1}^{k(\bar{r})}\int \mathbb{1}_{p,\bar{r}}(T^ix)\mathbb{1}_{p,\bar{r}}(T^jx)\,d\mu(x)\\
&= \sum_{i\in L_n^\gamma,j\in R_n^\gamma}\sum_{p=1}^{k(\bar{r})}\left(\left(\int \mathbb{1}_{p,\bar{r}}(x)\,d\mu(x)\right)^2+\mathrm{Err}(n)\right)\\
&= \gamma^2 2^{2n}\bar{r}^{-d}\mathrm{Err}(n)+\gamma^2 2^{2n}\sum_{p=1}^{k(\bar{r})}\mu(B(x_p,2\bar{r}))^2.
\end{aligned}$$

where $\mathrm{Err}(n)$ denotes the error which we can bound by $C'e^{-2^n(1-2\gamma)\theta'}$ using the exponential mixing estimate. Since $r_n\geqslant n^{-\beta}$ we see that the first term in the estimate above vanishes. For the second term we get from Lemma 5.1 that

$$\int \mu(B(x,\bar{r}))\,d\mu(x) \leqslant \sum_{p=1}^{k(\bar{r})}\mu(B(x_p,2\bar{r}))^2, \tag{5.4}$$

so by the lower bound assumption on $\int\mu(B(x,\bar{r}))\,d\mu(x)$, we have that

$$2^{2n}\sum_{p=1}^{k(\bar{r})}\mu(B(x_p,2\bar{r}))^2\to\infty \quad \text{as} \quad n\to\infty.$$

Now we have

$$\mathbb{E}(\hat{Q}_n^2) = \sum_{\substack{i_1 \in L_n^\gamma \\ j_1 \in R_n^\gamma}} \sum_{\substack{i_2 \in L_n^\gamma \\ j_2 \in R_n^\gamma}} \sum_{p,q=1}^{k(\bar{r})} \int \mathbb{1}_{p,\bar{r}}(T^{i_1}x)\mathbb{1}_{p,\bar{r}}(T^{j_1}x)\mathbb{1}_{q,\bar{r}}(T^{i_2}x)\mathbb{1}_{q,\bar{r}}(T^{j_2}x)\,d\mu(x).$$

We split the sum as in [Z, Section 4, Lower bound]. We start with the totally separated case, were we suppose that $|i_2 - i_1| > cn$ and $|j_2 - j_1| > cn$, where we will choose $c > 0$ later: it is sufficient to assume that $i_2 - i_1 > cn$ and $j_2 - j_1 > cn$ since the other combinations lead to the same estimates.

$$\begin{aligned}
&\sum_{\substack{i_1,i_2 \in L_n^\gamma \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ j_2 - j_1 > cn}} \sum_{p,q} \int \mathbb{1}_{p,\bar{r}}(T^{i_1}x)\mathbb{1}_{p,\bar{r}}(T^{j_1}x)\mathbb{1}_{q,\bar{r}}(T^{i_2}x)\mathbb{1}_{q,\bar{r}}(T^{j_2}x)\,d\mu(x) \\
&= \sum_{\substack{i_1,i_2 \in L_n^\gamma \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ j_2 - j_1 > cn}} \sum_{p,q} \int \mathbb{1}_{p,\bar{r}}(x)\mathbb{1}_{q,\bar{r}}(T^{i_2-i_1}x)\mathbb{1}_{p,\bar{r}}(T^{j_1-i_1}x)\mathbb{1}_{q,\bar{r}}(T^{j_2-i_1}x)\,d\mu(x) \\
&\leqslant \sum_{\substack{i_1,i_2 \in L_n^\gamma \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ j_2 - j_1 > cn}} \sum_{p,q} \Bigg(\int \mathbb{1}_{p,\bar{r}}(x)\mathbb{1}_{q,\bar{r}}(T^{i_2-i_1}x)\,d\mu(x) \int \mathbb{1}_{p,\bar{r}}(x)\mathbb{1}_{q,\bar{r}}(T^{j_2-j_1}x)\,d\mu(x) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + C'e^{-(j_1-i_2)\theta'} \Bigg) \\
&\leqslant \sum_{\substack{i_1,i_2 \in L_n^\gamma \\ i_2 - i_1 > cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ j_2 - j_1 > cn}} \sum_{p,q} \left[\left(\mu(B(x_p, 2\bar{r}))\mu(B(x_q, 2\bar{r})) + C'e^{-n\theta' c} \right)^2 + C'e^{-2^n(1-2\gamma)\theta'} \right] \\
&\leqslant \gamma^4 2^{4n} \sum_{p,q} \left[\left(\mu(B(x_p, 2\bar{r}))\mu(B(x_q, 2\bar{r})) + C'e^{-n\theta' c} \right)^2 + C'e^{-2^n(1-2\gamma)\theta'} \right].
\end{aligned}$$

where the third line is by 4-mixing. Since we assume that $r_n \geqslant n^{-\beta}$ and $k(r) \asymp r^{-d}$, up to constants this differs from $\mathbb{E}(\hat{Q}_n)^2$ by at most

$$\begin{aligned}
\gamma^4 2^{4n} \sum_{p,q} (e^{-n\theta' c} + e^{-2^n(1-2\gamma)\theta'}) &\leqslant \gamma^4 2^{4n} k(\bar{r})^2 (e^{-n\theta' c} + e^{-2^n(1-2\gamma)\theta'}) \\
&\leqslant \gamma^4 2^{4n} 2^{n2\beta d} (e^{-n\theta' c} + e^{-2^n(1-2\gamma)\theta'}).
\end{aligned}$$

So if $c > 2(2 + \beta d)\log 2/\theta'$, then this is summable, without any conditions coming from dividing by $\mathbb{E}(\hat{Q}_n)^2$.

Now consider the totally non-separated case, where it is sufficient to consider the case $0 \leqslant i_2 - i_1 \leqslant cn$ and $0 \leqslant j_2 - j_1 \leqslant cn$. Since any $x \in X$ is in at most C_0 of the covering sets $\{B(x_p, 2r)\}_{p=1}^{k(r)}$, we get $\sum_{q=1}^{k(\bar{r})} \mathbb{1}_{q,\bar{r}}(x) \cdot \mathbb{1}_{q,\bar{r}}(y) \leqslant C_0$. Using this we have,

$$\begin{aligned}
&\sum_{\substack{i_1,i_2 \in L_n^\gamma \\ 0 \leqslant i_2 - i_1 \leqslant cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ 0 \leqslant j_2 - j_1 \leqslant cn}} \sum_{p,q} \int \mathbb{1}_{p,\bar{r}}(T^{i_1}x)\mathbb{1}_{p,\bar{r}}(T^{j_1}x)\mathbb{1}_{q,\bar{r}}(T^{i_2}x)\mathbb{1}_{q,\bar{r}}(T^{j_2}x)\,d\mu(x) \\
&\leqslant C_0 \sum_{\substack{i_1,i_2 \in L_n^\gamma \\ 0 \leqslant i_2 - i_1 \leqslant cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ 0 \leqslant j_2 - j_1 \leqslant cn}} \sum_{p} \int \mathbb{1}_{p,\bar{r}}(T^{i_1}x)\mathbb{1}_{p,\bar{r}}(T^{j_1}x)\,d\mu(x) \\
&\leqslant C_0 \sum_{\substack{i_1,i_2 \in L_n^\gamma \\ 0 \leqslant i_2 - i_1 \leqslant cn}} \sum_{\substack{j_1,j_2 \in R_n^\gamma \\ 0 \leqslant j_2 - j_1 \leqslant cn}} \sum_{p} (\mu(B(x_p, 2\bar{r}))^2 + C'e^{-2^n(1-2\gamma)\theta'}) \\
&\leqslant C_0 \gamma^2 2^{2n} c^2 n^2 \sum_{p} (\mu(B(x_p, 2\bar{r}))^2 + C'e^{-2^n(1-2\gamma)\theta'}).
\end{aligned}$$

So the important term to estimate here is $2^{2n}n^2\sum_p \mu(B(x_p,2\bar r))^2 \lesssim n^2\mathbb{E}(\hat Q_n)$, which when divided by $\mathbb{E}(\hat Q_n)^2$ is $n^2/\mathbb{E}(\hat Q_n)$. By using (5.4) this is seen to be summable if we assume $\int \mu(B(x,r_n))\,d\mu(x) \geqslant \frac{(\log n)^3(\log\log n)^{1+\varepsilon}}{n^2}$. Note that this is a weaker requirement than the assumption made in the theorem.

Now for the half-separated case. Let us suppose that $0\leqslant i_2-i_1\leqslant cn$ and $j_2-j_1\geqslant cn$. Then

$$\begin{aligned}
&\sum_{\substack{i_1,i_2\in L_n^\gamma\\ 0\leqslant i_2-i_1\leqslant cn}}\ \sum_{\substack{j_1,j_2\in R_n^\gamma\\ j_2-j_1>cn}}\ \sum_{p,q}\int \mathbb{1}_{p,\bar r}(T^{i_1}x)\mathbb{1}_{p,\bar r}(T^{j_1}x)\mathbb{1}_{q,\bar r}(T^{i_2}x)\mathbb{1}_{q,\bar r}(T^{j_2}x)\,d\mu(x)\\
&=\sum_{\substack{i_1,i_2\in L_n^\gamma\\ 0\leqslant i_2-i_1\leqslant cn}}\ \sum_{\substack{j_1,j_2\in R_n^\gamma\\ j_2-j_1>cn}}\ \sum_{p,q}\int \mathbb{1}_{p,\bar r}(x)\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\mathbb{1}_{p,\bar r}(T^{j_1-i_1}x)\mathbb{1}_{q,\bar r}(T^{j_2-i_1}x)\,d\mu(x)\\
&\leqslant\sum_{\substack{i_1,i_2\in L_n^\gamma\\ 0\leqslant i_2-i_1\leqslant cn}}\ \sum_{\substack{j_1,j_2\in R_n^\gamma\\ j_2-j_1>cn}}\ \sum_{p,q}\Bigg(\int \mathbb{1}_{p,\bar r}(x)\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\,d\mu(x)\\
&\qquad\qquad\cdot\int \mathbb{1}_{p,\bar r}(x)\mathbb{1}_{q,\bar r}(T^{j_2-j_1}x)\,d\mu(x)+C'e^{-(j_1-i_2)\theta'}\Bigg)\\
&\leqslant\sum_{\substack{i_1,i_2\in L_n^\gamma\\ 0\leqslant i_2-i_1\leqslant cn}}\ \sum_{\substack{j_1,j_2\in R_n^\gamma\\ j_2-j_1>cn}}\ \sum_{p,q}\Bigg(\int \mathbb{1}_{p,\bar r}(x)\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\,d\mu(x)\\
&\qquad\qquad\cdot\left(\mu(B(x_p,2\bar r))\mu(B(x_q,2\bar r))+C'e^{-n\theta' c}\right)+C'e^{-2^n(1-2\gamma)\theta'}\Bigg),
\end{aligned}$$

where the third line is by 4-mixing. Note that our choice of c again makes the part with the terms $C'e^{-n\theta' c}$ and $C'e^{-2^n(1-2\gamma)\theta'}$ summable over n. Now by the Cauchy–Schwarz Inequality,

$$\begin{aligned}
&\sum_{p,q}\int \mathbb{1}_{p,\bar r}(x)\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\,d\mu(x)\mu(B(x_p,2\bar r))\mu(B(x_q,2\bar r))\\
&=\int\sum_p \mu(B(x_p,2\bar r))\mathbb{1}_{p,\bar r}(x)\sum_q \mu(B(x_q,2\bar r))\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\,d\mu(x)\\
&\leqslant\left(\int\Big(\sum_p \mu(B(x_p,2\bar r))\mathbb{1}_{p,\bar r}(x)\Big)^2 d\mu(x)\right)^{\frac12}\\
&\qquad\qquad\cdot\left(\int\Big(\sum_q \mu(B(x_q,2\bar r))\mathbb{1}_{q,\bar r}(T^{i_2-i_1}x)\Big)^2 d\mu(x)\right)^{\frac12}\\
&=\int\Big(\sum_p \mu(B(x_p,2\bar r))\mathbb{1}_{p,\bar r}(x)\Big)^2 d\mu(x).
\end{aligned}$$

Recall that for each $x\in X$, the sum $\sum_p \mu(B(x_p,2\bar r))\mathbb{1}_{p,\bar r}(x)$ contains at most C_0 non-zero terms. For these non-zero terms we apply the following consequence of Jensen's inequality

$$(a_1+\cdots+a_m)^2\leqslant m(a_1^2+\cdots+a_m^2),$$

for numbers $a_1,\ldots,a_m\geqslant 0$ along with the fact that $\mathbb{1}_{p,\bar r}^2=\mathbb{1}_{p,\bar r}$ to get the bound $C_0\sum_p \mu(B(x_p,2\bar r))^2\mathbb{1}_{p,\bar r}(x)$. Hence, by superadditivity of convex functions, the above sum is bounded by,

$$\int C_0\sum_p \mu(B(x_p,2\bar r))^2\mathbb{1}_{p,\bar r}(x)\,d\mu(x)=C_0\sum_p \mu(B(x_p,2\bar r))^3$$

$$\leqslant C_0 \left(\sum_p \mu(B(x_p, 2\bar{r}))^2 \right)^{\frac{3}{2}}.$$

So the entire sum we must estimate is

$$2^{3n} cn \left(\sum_p \mu(B(x_p, 2\bar{r}))^2 \right)^{\frac{3}{2}} \lesssim n\mathbb{E}(\hat{Q}_n)^{\frac{3}{2}},$$

which when divided by $\mathbb{E}(\hat{Q}_n)^2$ can be estimated as $n/\mathbb{E}(\hat{Q}_n)^{\frac{1}{2}}$. Hence choosing $\int \mu(B(x, r_n))\, d\mu(x) \gtrsim \frac{(\log n)^4 (\log\log n)^{2+\varepsilon}}{n^2}$, this is bounded, up to constants, by

$$\frac{n}{(n^4 (\log n + \log\log 2)^{2+\varepsilon} (\log 2)^4)^{\frac{1}{2}}},$$

which is summable. □

Proof of Theorem 2.17(d). The proof follows the same principle as the proof of Theorem 2.4(b). We define

$$\bar{Q}_n(x) := \sum_{\substack{i \in [0, \gamma n) \\ j \in [(1-\gamma)n, n)}} \sum_{p=1}^{k(r_n)} \mathbb{1}_{p, r_n}(T^i x) \mathbb{1}_{p, r_n}(T^j x).$$

and observe that it is sufficient to prove that $\frac{\mathbb{E}(\bar{Q}_n^2) - \mathbb{E}(\bar{Q}_n)^2}{\mathbb{E}(\bar{Q}_n)^2}$ goes to 0. Again, all estimates and computations of part (c) are repeated with 2^n replaced by n and the separation gap cn replaced by $c \log n$. Using the corresponding assumptions on r_n imposed in part (d) of the theorem one easily reaches the conclusion that the aforementioned quantity vanishes. □

Remark 5.2. In part (c) and (d) the assumptions on $(r_n)_n$ have some flexibility in the following sense. The condition that $r_n \geqslant n^{-\beta}$ can be relaxed to $(r_n)_n$ decreasing at most subexponentially at the cost of increasing the power of $\log n$ by an arbitrarily small amount in the lower bound on the shrinking rate of $\int \mu(B(x, r_n))\, d\mu(x)$. In the proof, this would be reflected by replacing our time gap cn with $n^{1+\iota}$ for $\iota > 0$ in part (c) and $c \log n$ by $(\log n)^{1+\iota}$ for $\iota > 0$.

References

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

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

[BLR] V. Barros, L. Liao, J. Rousseau, *On the shortest distance between orbits and the longest common substring problem,* Adv. Math. **344** (2019) 311–339.

[BPTV] D. Bessis, G. Paladin, G. Turchetti and S. Vaienti, *Generalized dimensions, entropies, and Liapunov exponents from the pressure function for strange sets*, J. Statist. Phys. **51** (1988) 109–134.

[C] P. Collet, *A remark about uniform de-correlation prefactors,* unpublished manuscript.

[ER] P. Erdős, A. Rényi, *On Cantor's series with convergent $\sum \frac{1}{q_n}$*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math. **2** (1959) 93–109.

[FLR] A.-H. Fan, K.-S. Lau, H. Rao, *Relationships between different dimensions of a measure,* Monatsh. Math. **135** (2002) 191–201.

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

[GRS] S. Gouëzel, J. Rousseau, M. Stadlbauer, *Minimal distance between random orbits*, Probab. Theory Related Fields. **189** (2024) 811–847.

[HKKP] M. Holland, M. Kirsebom, P. Kunde, T. Persson, *Dichotomy results for eventually always hitting time statistics and almost sure growth of extremes,* Trans. Amer. Math. Soc. **377** (2024) 3927–3982.

[HNT] M. Holland, M. Nicol, A. Török, *Extreme value theory for non-uniformly expanding dynamical systems*, Trans. Amer. Math. Soc. **364** (2012) 661–688.

[KKP] M. Kirsebom, P. Kunde, T. Persson, *On shrinking targets and self-returning points,* Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **24** (2023) 1499–1535.

[LSV] C. Liverani, B. Saussol and S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems*, Ergodic Theory Dynam. Systems **18** (1998) 1399–1420.

[M] G. Mantica, *The global statistics of return times: Return time dimensions versus generalized measure dimensions*, J. Stat. Phys. **138** (2010) 701–727.

[P] Y. Pesin, *On rigorous mathematical definitions of correlation dimension and generalized spectrum for dimensions,* J. Statist. Phys. **71** (1993) 529–547.

[RT] J. Rousseau, M. Todd, *Orbits closeness for slowly mixing dynamical systems,* Ergodic Theory Dynam. Systems **44** (2024) 1192–1208.

[S] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps,* Isr. J. Math **116** (2000) 223–248.

[SR] J. C. Sprott, G. Rowlands, *Improved correlation dimension calculation*, Int. J. Bifurcation Chaos **11** (2001) 1865–1880.

[Y] L.-S. Young, *Decay of correlations for certain quadratic maps,* Comm. Math. Phys. **146** (1992) 123–138.

[Z] B. Zhao, *Closest distance between iterates of typical points*, Discrete Contin. Dyn. Syst. **44** (2024) 2252–2279.

Maxim Kirsebom, University of Bremen, Department 3 – Mathematics, Institute for Dynamical Systems, Bibliothekstr. 5, 28359 Bremen, Germany
Email address: kirsebom@uni-bremen.de

Philipp Kunde, Department of Mathematics, Oregon State University, Kidder Hall, Corvallis, OR 97331, USA
Email address: kundep@oregonstate.edu
URL: https://sites.google.com/view/pkunde

Tomas Persson, Centre for Mathematical Sciences, Lund University, Box 118, 221 00 Lund, Sweden
Email address: tomasp@gmx.com
URL: https://tomaspersson.ouvaton.org

Mike Todd, Mathematical Institute, University of St Andrews, North Haugh, St Andrews, KY16 9SS, Scotland
Email address: m.todd@st-andrews.ac.uk
URL: https://mtoddm.github.io/