



**HAL**  
open science

# TOPOLOGICAL MEAN DIMENSION OF INDUCED SYSTEMS

David Burguet, Ruxi Shi

► **To cite this version:**

David Burguet, Ruxi Shi. TOPOLOGICAL MEAN DIMENSION OF INDUCED SYSTEMS. 2022.  
hal-03907081

**HAL Id: hal-03907081**

**<https://hal.science/hal-03907081>**

Preprint submitted on 19 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# TOPOLOGICAL MEAN DIMENSION OF INDUCED SYSTEMS

DAVID BURGUET AND RUXI SHI

ABSTRACT. For a topological system with positive topological entropy, we show that the induced transformation on the set of probability measures endowed with the weak-\* topology has infinite topological mean dimension. We also estimate the rate of divergence of the entropy with respect to the Wasserstein distance when the scale goes to zero.

## 1. INTRODUCTION

Let  $(X, T)$  be a topological dynamical system, i.e.  $X$  is a compact metrizable space and  $T : X \rightarrow X$  is a continuous map. Let  $\mathcal{M}(X)$  be the space of Borel probability measures on  $X$  endowed with the weak-\* topology. Then the push-forward map  $T_* : \mu \mapsto \mu \circ T^{-1}$  is a continuous map on  $\mathcal{M}(X)$ . The topological system  $(\mathcal{M}(X), T_*)$  is called the *induced system* of  $(X, T)$  on probability measures.

W. Bauer and K. Sigmund [BS75] have investigated which dynamical properties of  $(\mathcal{M}(X), T_*)$  are inherited from  $(X, T)$ . In particular they observed that  $h_{top}(T_*)$  is infinite when  $h_{top}(T)$  is positive. Later B. Weiss and E. Glasner [GW95] showed that if  $(X, T)$  has zero topological entropy, then so does  $(\mathcal{M}(X), T_*)$ . Therefore we have the following equivalences :

$$h_{top}(T) > 0 \Leftrightarrow h_{top}(T_*) > 0 \Leftrightarrow h_{top}(T_*) = \infty.$$

M. Gromov [Gro99] has introduced a new topological invariant of dynamical systems called (*topological*) *mean dimension* as a dynamical analogue of topological covering dimension. In some sense the mean dimension  $\text{mdim}(T)$  estimates the number of parameters that we need to describe the system  $(X, T)$  per unity of time. A dynamical system of finite entropy or finite dimension always has zero mean dimension. The mean dimension of the shift on  $d$ -cubes  $([0, 1]^d)^{\mathbb{Z}}$  is equal to  $d$  ([LW00, Proposition 3.3]). Computing the mean dimension is in general difficult and it is only known for few systems ([Tsu19, BS21]). In the present paper we investigate the mean dimension of the induced system  $(\mathcal{M}(X), T_*)$ . Our main result states as below :

**Main Theorem.** *For any topological system  $(X, T)$  with positive topological entropy, the induced system  $(\mathcal{M}(X), T_*)$  has infinite topological mean dimension. Therefore,*

$$h_{top}(T) > 0 \Leftrightarrow \text{mdim}(T_*) > 0 \Leftrightarrow \text{mdim}(T_*) = \infty.$$

Given a distance  $d$  on  $X$ , Lindenstrauss and Weiss [LW00] has introduced the associated metric mean dimension  $\text{mdim}(X, d, T)$  as the dynamical quantity corresponding to the Minkowski dimension. They showed inter alia that the metric mean dimension is always larger than or equal to the associated mean topological dimension. It is conjectured that

the topological mean dimension is the infimum of the metric mean dimensions with respect to distances inducing the topology (this conjecture is known to hold true in some cases, see [LT19]). A dynamical system of finite entropy has zero metric mean dimension for any metric compatible with the topology ([LW00]). Together with the Main Theorem, we have the following corollary.

**Corollary.** *Given a topological system  $(X, T)$  we have the following equivalences for any metric  $D$  on  $\mathcal{M}(X)$  compatible with the weak- $*$  topology:*

$$h_{top}(T) > 0 \Leftrightarrow \text{mdim}(\mathcal{M}(X), D, T_*) > 0 \Leftrightarrow \text{mdim}(\mathcal{M}(X), D, T_*) = \infty.$$

For an integer  $a \geq 2$  we let  $T_a$  be the  $\times a$ -map on the circle. Kloeckner [Klo13] proved the following lower bound on the metric mean dimension of  $T_a$  for the  $p^{\text{th}}$ -Wasserstein metric:

$$\underline{\text{mdim}}(\mathcal{M}(X), W_p, (T_a)_*) \geq p(a - 1).$$

In fact, the metric mean dimension of the  $n$ -power of a system is always larger than or equal to  $n$  times the metric mean dimension of the system (see Inequality (2.5)). But the  $n$ -power of  $(T_a)_*$  is just  $(T_{a^n})_*$ , therefore without referring to our new results one can show directly from Kloeckner's theorem that the metric mean dimension of the induced map  $(T_a)_*$  with respect to the  $p^{\text{th}}$ -Wasserstein metric is in fact infinite:

$$\begin{aligned} \underline{\text{mdim}}(\mathcal{M}(X), W_p, (T_a)_*) &\geq \frac{1}{n} \cdot \underline{\text{mdim}}(\mathcal{M}(X), W_p, (T_a)_*^n), \\ &\geq \frac{1}{n} \cdot \underline{\text{mdim}}(\mathcal{M}(X), W_p, (T_{a^n})_*), \\ &\geq p \cdot \frac{a^n - 1}{n} \xrightarrow{n} \infty. \end{aligned}$$

Therefore, Kloeckner's theorem can be regarded as a special case of our Corollary.

For a topological system  $(X, T)$  with positive entropy, the Main Theorem implies the mean dimension <sup>1</sup> of  $T_*$  with the scale  $\epsilon$ , denoted by  $\text{mdim}_{W_1}(T_*, \epsilon)$ , goes to infinity as  $\epsilon$  tends to zero. The rate of divergence of  $\text{mdim}_{W_1}(T_*, \epsilon)$  may be precised as follows:

**Theorem 1.** *Suppose  $(X, T)$  is of positive entropy and Lipschitz. Then there exist  $C, \alpha > 0$  such that we have*

$$\forall 0 < \epsilon < 1, \quad \text{mdim}_{W_1}(T_*, \epsilon) \geq C\epsilon^{-\alpha}.$$

In general, we can use the modulus of continuity of  $T$  to estimate the rate of divergence of  $\text{mdim}(\mathcal{M}(X), T_*, W_1, \epsilon)$ , but for simplicity we only state the result for Lipschitz systems  $(X, T)$ .

The Corollary also implies that the mean dimension of  $T_*$  with respect to the Wasserstein metric  $W_1$  associated to some given compatible metric  $d$  on  $X$  is infinite, i.e.

$$\frac{h_{W_1}(T_*, \epsilon)}{\log \frac{1}{\epsilon}} \xrightarrow{\epsilon \rightarrow 0} +\infty.$$

---

<sup>1</sup>See the definition in Section 2.3.

We also investigate the rate of divergence of  $h_{W_1}(T_*, \epsilon)$  as  $\epsilon$  tends to zero for Lipschitz systems.

**Theorem 2.** *Suppose  $(X, T)$  is of positive entropy and Lipschitz. Then there exist  $C, \alpha > 0$  such that we have*

$$\forall 0 < \epsilon < 1, \quad h_{W_1}(T_*, \epsilon) \geq C\epsilon^{-\alpha}.$$

Under the same approach, our results (Main Theorem, Corollary, Theorem 1 and Theorem 2) can be generalized to the action of congruent monothilable amenable groups, like  $\mathbb{Z}^d, d \geq 2$ . It seems that we could not apply directly the same method to the action of amenable groups which is not congruent monothilable. We thus wonder whether our results still hold for general amenable actions.

**Organization of the paper.** In the next section we recall some basic facts on the mean dimension. In the third section we give a first direct proof of the Main theorem for full shifts on  $\{0, 1\}^{\mathbb{Z}}$ . This proof, which is independent from the proof in the general case, strongly uses the product structure on  $\{0, 1\}^{\mathbb{Z}}$ . The mean dimension is shown to be infinite by embedding the shift on cubes of arbitrarily large dimension in the induced system  $(\mathcal{M}(\{0, 1\}^{\mathbb{Z}}, \sigma_*))$ . Then we prove the main theorem in the general case. We work with the metric definition (2.2) of the topological mean dimension with respect to the Kantorovic-Rubinstein metric. We use the theory of independence developed by Weiss-Glasner [GW95], Kerr-Li [KL07] and others, which gives a *weak* product structure in any topological system with positive entropy. By embedding generalized cubes of dimension  $(k-1)n$  in the  $k^n$ -simplex we conclude the proof with a Lebesgue's like theorem for these cubes. In Section 6 we establish the rates of convergence for the mean dimension and entropy with respect to the Wasserstein distance given in Theorem 1 and Theorem 2. In the last section we study the mean dimension of the action induced on the hyperspace of compact subsets of  $X$ .

## 2. BACKGROUNDS

In the whole paper we let  $(X, T)$  for a topological dynamical system, i.e.  $X$  is compact metrizable and  $T : X \rightarrow X$  a continuous map. We denote by  $d$  any metric compatible with the topology on  $X$ .

**2.1. Dimension.** We first recall some standard concepts related to the dimension of the phase space  $X$ .

**2.1.1. Order of a cover and Lebesgue's Lemma.** We define the order  $\text{ord}(\mathcal{A})$  of a finite open cover  $\mathcal{A}$  of  $X$  as follows

$$\text{ord}(\mathcal{A}) := \sup_{x \in X} \sum_{A \in \mathcal{A}} 1_A(x) - 1,$$

H. Lebesgue proved in [Leb11] that cubes of different dimensions are not homeomorphic. We state below the key lemma of his approach.

**Lemma 3.** [Leb11] *Let  $\alpha$  be a finite open cover of the unit cube  $[0, 1]^n$ . Suppose that there is no element of  $\alpha$  meeting two opposite faces of  $[0, 1]^n$ . Then we have*

$$\text{ord}(\alpha) \geq n.$$

Let  $\Delta_k$  be the standard  $k$ -simplex,  $k \in \mathbb{N}^*$ , i.e.

$$\Delta_k := \left\{ (x_i)_i \in [0, 1]^k, \sum_i x_i = 1 \right\}.$$

To prove the Main Theorem, we will generalize Lebesgue's Lemma to  $n$ -direct products of the simplex  $\Delta_k$  (see Lemma 9) in place of  $[0, 1]^n$ .

2.1.2. *Topological dimension.* For two finite open covers  $\mathcal{A}$  and  $\mathcal{B}$  of  $X$ , we say that the cover  $\mathcal{B}$  is *finer* than the cover  $\mathcal{A}$ , and write  $\mathcal{B} \succ \mathcal{A}$ , if for every element of  $\mathcal{B}$ , one can find an element of  $\mathcal{A}$  which contains it. For a finite open cover  $\mathcal{A}$  of  $X$ , we define the quantity

$$D(\mathcal{A}) := \min_{\mathcal{B} \succ \mathcal{A}} \text{ord}(\mathcal{B}).$$

For finite open covers  $\mathcal{A}$  and  $\mathcal{B}$ , we set the joint  $\mathcal{A} \vee \mathcal{B} := \{U \cap V : U \in \mathcal{A}, V \in \mathcal{B}\}$ . It is easy to check that  $D(\mathcal{A} \vee \mathcal{B}) \leq D(\mathcal{A}) + D(\mathcal{B})$ . Clearly, if  $\mathcal{B} \succ \mathcal{A}$  then  $D(\mathcal{B}) \geq D(\mathcal{A})$ .

The (*topological*) *dimension* of  $X$  is defined by

$$\dim(X) := \sup_{\mathcal{A}} D(\mathcal{A}),$$

where  $\mathcal{A}$  runs over all finite open covers of  $X$ . For a non-empty compact  $X$ , the *stable topological dimension* of  $X$  is given by

$$\text{stabdim}(X) := \lim_{n \rightarrow \infty} \frac{\dim(X^n)}{n} = \inf_{n \rightarrow \infty} \frac{\dim(X^n)}{n},$$

where  $X^n$  denotes the  $n$ -direct product of  $X$ . The limit above exists by sub-additivity of the sequence  $\{\dim(X^n)\}_{n \geq 1}$ . Moreover, if  $X$  is finite dimensional, then we have either  $\text{stabdim}(X) = \dim(X)$  and  $\dim(X^2) = 2 \dim(X)$  (the space  $X$  is then said of *basic type*) or  $\text{stabdim}(X) = \dim(X) - 1$  and  $\dim(X^2) = 2 \dim(X) - 1$  (and  $X$  is said of *exceptional type*). M. Tsukamoto proved in [Tsu19] that the full shift on  $X^{\mathbb{N}}$  has topological mean dimension equal to  $\text{stabdim}(X)$  whenever  $X$  is finite dimensional.

2.1.3. *Metric  $\epsilon$ -dimension and width.* Recall  $d$  denotes a compatible distance on  $X$ . For a set  $Z$  and  $\epsilon > 0$ , a map  $f : X \rightarrow Z$  is called  *$(d, \epsilon)$ -injective* if  $\text{diam}(f^{-1}(z)) < \epsilon$  for all  $z \in Z$ . The *metric  $\epsilon$ -dimension*  $\dim_{\epsilon}(X, d)$  is defined by

$$\dim_{\epsilon}(X, d) = \inf_Y \dim(Y),$$

where  $Y$  runs over all compact metrizable spaces for which there exists a  $(d, \epsilon)$ -injective continuous map  $f : X \rightarrow Y$ .

The metric  $\epsilon$ -dimension of  $X$  is bounded from above by the topological dimension of  $X$  and is going non-increasingly to the topological dimension when  $\epsilon$  goes to zero.

A variant of the metric  $\epsilon$ -dimension introduced by Gromov [Gro99] is the  $\epsilon$ -width,  $\text{Widim}_\epsilon(X, d)$  which is the smallest integer  $n$  such that there exists an  $\epsilon$ -injective continuous map  $f : X \rightarrow P$  from  $X$  into some  $n$ -dimensional polyhedron  $P$ . We always have

$$\dim_\epsilon(X, d) \leq \text{Widim}_\epsilon(X, d) \leq 2 \dim_\epsilon(X, d) + 1.$$

Gromov showed that the unit ball  $B$  of a  $n$ -dimensional Banach  $(E, \|\cdot\|)$  space satisfies

$$\forall 1 > \epsilon > 0, \text{Widim}_\epsilon(B, \|\cdot\|) = n.$$

Motivated by a question raised by Gromov [Gro99, p.334], Gournay [Gou11] and Tsukamoto [Tsu08] obtained some estimates for  $\text{Widim}_\epsilon(B_p, \ell_q)$  with  $B_p$  being the unit ball in  $(\mathbb{R}^n, \ell_p)$ . Gromov also wondered about the width of the simplex  $\Delta_n$ .

**2.2. Entropy.** To fix the notations we recall here Bowen's definition of topological entropy. For any positive integer  $n$  we consider the  $n$ -dynamical distance  $d_n$  as follows,

$$\forall x, y \in X, d_n(x, y) = \max_{0 \leq i \leq n-1} d(T^i x, T^i y).$$

Let  $K \subset X$  and  $\epsilon > 0$ . A subset  $E$  of  $X$  is said to be  $(n, \epsilon)$ -separated if we have  $d_n(x, y) > \epsilon$  for any  $x \neq y \in E$ . Denote by  $s_n(d, T, K, \epsilon)$  the largest cardinality of any  $(n, \epsilon)$ -separated subset of  $K$ . Define

$$h_d(K, T, \epsilon) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log s_n(d, T, K, \epsilon).$$

We sometimes write  $h_d(T, \epsilon)$  when  $K = X$ . Observe that for all positive integer  $n$  and all  $\epsilon > 0$  we have the following estimate on the entropy of the  $n$ -power  $T^n$ :

$$(2.1) \quad h_d(T^n, \epsilon) \leq n \cdot h_d(T, \epsilon).$$

The topological entropy is the limit of  $h_d(T, \epsilon)$  when  $\epsilon$  goes to zero. This limit does not depend on the choice of the metric  $d$  (this is no more the case of the topological and metric mean dimensions).

The entropy is homogeneous, i.e.  $h_{top}(T^n) = n \cdot h_{top}(T)$  for all  $n \in \mathbb{N}$  and  $h_{top}(X, T) \geq h_{top}(Y, T)$  for any invariant closed subset  $Y$  of  $X$ . Also the entropy of the full shift over a finite alphabet  $\mathcal{A}$  is equal to  $\log \#\mathcal{A}$ . A topological system  $(X, T)$  is said to have a *horseshoe* when the full shift over  $\{0, 1\}^{\mathbb{N}}$  embeds in some power  $(X, T^m)$ . It follows from the above properties, that such systems have positive topological entropy. Many systems with positive topological entropy have a horseshoe, e.g.:

- coded subshifts,
- continuous interval and circle maps with positive entropy (by Misiurewicz horseshoe theorem [Mis79]),
- $C^{1+}$  surface diffeomorphisms (by Katok's horseshoe theorem [Kat80]),
- ...

But there are also systems with positive topological entropy without horseshoes, even uniquely ergodic minimal subshifts.

Inspired by the works on emergence introduced by P. Berger, Y. Ji, E. Chen and X. Zhou [YJ] defined a topological order which estimates the superexponential growth of orbits at arbitrarily small scaled. Let  $W_{d_n}$  be the 1-Wasserstein metric associated to  $d_n$ . They define the *entropy order*  $\mathcal{E}(T)$  of  $T$  by

$$\mathcal{E}(T) := \lim_{\epsilon \rightarrow 0} \limsup_n \frac{\log \log s_n(W_{d_n}, T_*, \mathcal{M}(X), \epsilon)}{n}.$$

They showed that the entropy order is always equal to the topological entropy by Jewett-Krieger theorem.

**2.3. Mean dimension.** Let  $(X, T)$  be a topological dynamical system, i.e.  $X$  is a compact metrizable space and  $T : X \rightarrow X$  is a continuous map. The *mean dimension* of  $(X, T)$  is defined by

$$\text{mdim}(X, T) = \sup_{\alpha} \lim_{n \rightarrow \infty} \frac{D(\bigvee_{i=0}^{n-1} T^{-i} \alpha)}{n},$$

where  $\alpha$  runs over all finite open covers of  $X$ . The existence of the limit follows from the sub-additivity of the sequence  $(D(\bigvee_{i=0}^{n-1} T^{-i} \alpha))_{n \geq 1}$ . We put also  $\text{mdim}(X, T, \alpha) = \lim_{n \rightarrow \infty} \frac{D(\bigvee_{i=0}^{n-1} T^{-i} \alpha)}{n}$ . When there is no ambiguity we write also  $\text{mdim}(T)$  for  $\text{mdim}(X, T)$ .

We mention some basic properties of mean dimension. We refer to the book [Coo05] for the proofs and further properties.

- For a metric  $d$  on  $X$  compatible with the topology, we have

$$(2.2) \quad \text{mdim}(X, T) = \lim_{\epsilon \rightarrow 0} \text{mdim}_d(T, \epsilon),$$

where we let  $\text{mdim}_d(T, \epsilon) = \lim_{n \rightarrow \infty} \frac{\dim_{\epsilon}(X, d_n)}{n}$ . Observe the limit in  $\epsilon$  in the definition (2.2) of  $\text{mdim}(X, T)$  is also a supremum over  $\epsilon$ .

- If  $(Y, T)$  is a subsystem of  $(X, T)$ , i.e.  $Y$  is a closed  $T$ -invariant subset of  $X$ , then

$$(2.3) \quad \text{mdim}(Y, T) \leq \text{mdim}(X, T).$$

- For  $n \in \mathbb{N}$ ,

$$(2.4) \quad \text{mdim}(X, T^n) = n \cdot \text{mdim}(X, T).$$

- For dynamical systems  $(X_i, T_i)$ ,  $1 \leq i \leq n$ , we have

$$\text{mdim}(X_1 \times X_2 \times \cdots \times X_n, T_1 \times T_2 \times \cdots \times T_n) \leq \sum_{i=1}^n \text{mdim}(X_i, T_i).$$

The *upper metric mean dimension* of the system  $(X, d, T)$  is defined by

$$\overline{\text{mdim}}(X, d, T) = \limsup_{\epsilon \rightarrow 0} \frac{h_d(T, \epsilon)}{\log \frac{1}{\epsilon}}.$$

Similarly, the *lower metric mean dimension* is defined by

$$\underline{\text{mdim}}(X, d, T) = \liminf_{\epsilon \rightarrow 0} \frac{h_d(T, \epsilon)}{\log \frac{1}{\epsilon}}.$$

If the upper and lower metric mean dimensions coincide, then we call their common value the metric mean dimension of  $(X, d, T)$  and denote it by  $\text{mdim}(X, d, T)$ . Unlike the topological entropy, the metric mean dimension depends on the metric  $d$ . An important property of metric mean dimension is

$$\text{mdim}(X, T) \leq \underline{\text{mdim}}(X, d, T),$$

for any metric  $d$  [LW00].

It follows from the definitions and Inequality (2.1) that:

$$(2.5) \quad \forall n \in \mathbb{N}, \quad \underline{\text{mdim}}(X, d, T^{(n)}) \leq n \cdot \underline{\text{mdim}}(X, d, T).$$

**2.4. Embedding in induced system.** Let  $(X, T)$  be a dynamical system. For any  $n \in \mathbb{N}^*$  we denote by  $(X^{(n)}, T^{(n)})$  the  $n$ -direct product of  $(X, T)$ . These products embed in  $(\mathcal{M}(X), T_*)$ . Indeed we may find  $n$  positive integers  $k_1 \ll \dots \ll k_n$  such that for any two subfamilies  $I \neq J$  of  $\{1, \dots, n\}$  we have  $\sum_{i \in I} k_i \neq \sum_{j \in J} k_j$ . Then we define a dynamical embedding  $\pi : (X^{(n)}, T^{(n)}) \rightarrow (\mathcal{M}(X), T_*)$  as follows

$$\forall (x_1, \dots, x_n) \in X^{(n)}, \quad \pi(x_1, \dots, x_n) = \frac{\sum_{i=1, \dots, n} k_i \delta_{x_i}}{\sum_{i=1, \dots, n} k_i},$$

where  $\delta_x$  is the Dirac measure at  $x$ .

The topological entropy of  $T^{(n)}$  is  $n \cdot h_{\text{top}}(T)$ . The topological mean dimension satisfies the same property as proved recently in [JQ21] :

$$\forall n \in \mathbb{N}, \quad \text{mdim}(X^{(n)}, T^{(n)}) = n \cdot \text{mdim}(X, T).$$

Therefore as the topological entropy (resp. topological mean dimension) is a topological invariant which is smaller for subsystems, we get  $h_{\text{top}}(T_*) \geq n \cdot h_{\text{top}}(T)$  (resp.  $\text{mdim}(T_*) \geq n \cdot \text{mdim}(T)$ ) for all  $n$ . In particular if  $T$  has positive topological entropy (resp. mean dimension) then  $T_*$  has infinite topological entropy (resp. mean dimension). For the topological entropy this property was first noticed in [Sig78].

### 3. A FIRST DIRECT PROOF FOR SYSTEMS WITH A HORSESHOE

We first prove the Main Theorem for systems with a horseshoe. By (2.4) and (2.3) it is enough to show it for a full shift on a finite alphabet  $\mathcal{A}$ . Let  $\sigma$  be the shift on  $\mathcal{A}^{\mathbb{N}}$ , i.e.  $(x_n)_{n \in \mathbb{N}} \mapsto (x_{n+1})_{n \in \mathbb{N}}$ . For a finite word  $a = (a_0, \dots, a_{n-1}) \in \mathcal{A}^n$  with some integer  $n$ , we denote by  $[a]$  the cylinder with preword  $a$ , i.e.

$$[a] := \{x \in \mathcal{A}^{\mathbb{N}} : x_i = a_i, \forall 0 \leq i \leq n-1\}.$$

**Theorem 4.** For  $\sharp \mathcal{A} \geq 2$ , we have  $\text{mdim}(\mathcal{M}(\mathcal{A}^{\mathbb{N}}), \sigma_*) = +\infty$ .



Using the product structure of  $\mathcal{A}^n$ , we will embed the full shift over a simplex of arbitrarily large dimension in  $(\mathcal{M}(\mathcal{A}^{\mathbb{N}}), \sigma_*)$ .

*Proof.* Fix  $m \in \mathbb{N}^*$ . Let  $\mathcal{A}^m$  be the set of words of length  $m$ . Recall that  $\Delta_{\sharp\mathcal{A}^m}$  denotes the  $\sharp\mathcal{A}^m$ -simplex, which can be seen as the set of probability measures on  $\mathcal{A}^m$ :

$$\Delta_{\sharp\mathcal{A}^m} := \left\{ (x_i)_{i \in \mathcal{A}^m} \in [0, 1]^{\mathcal{A}^m} : \sum_{i \in \mathcal{A}^m} x_i = 1 \right\}.$$

In this case we have  $\text{stabdim}(\Delta_{\sharp\mathcal{A}^m}) = \dim(\Delta_{\sharp\mathcal{A}^m}) = \sharp\mathcal{A}^m - 1$ . For any  $a \in \Delta_{\sharp\mathcal{A}^m}$  we let  $\mu_a$  be the probability measure on  $\sharp\mathcal{A}^m$  with probability vector  $a = (a^j)_{j \in \mathcal{A}^m}$ , i.e.

$$\mu_a(\{j\}) = a^j \text{ for all } j \in \mathcal{A}^m.$$

We identify  $(\mathcal{A}^m)^{\mathbb{N}}$  with  $\mathcal{A}^{\mathbb{N}}$  by the natural homeomorphism  $\pi : (\mathcal{A}^m)^{\mathbb{N}} \rightarrow \mathcal{A}^{\mathbb{N}}$ , which associates to  $(x_j^i)_{i \in \{0, \dots, m-1\}, j \in \mathbb{N}}$  the sequence  $(y_k)_{k \in \mathbb{N}}$  defined by  $y_{jm+i} = x_j^i$  for any  $j \in \mathbb{N}$  and  $i \in \{0, \dots, m-1\}$ . Then define  $f_m : \Delta_{\sharp\mathcal{A}^m}^{\mathbb{N}} \rightarrow \mathcal{M}(\mathcal{A}^{\mathbb{N}})$  as follows:

$$\forall a = (a_n)_{n \in \mathbb{N}} \in \Delta_{\sharp\mathcal{A}^m}^{\mathbb{N}}, \quad f_m(a) = \pi_*(\mu_{a_1} \times \mu_{a_2} \times \dots),$$

where  $\mu_{a_1} \times \mu_{a_2} \times \dots$  denotes the direct product measure on  $(\mathcal{A}^m)^{\mathbb{N}}$ . Clearly  $f_m$  is injective, continuous and for all  $a = (a_n)_{n \in \mathbb{N}} \in (\mathcal{A}^m)^{\mathbb{N}}$  we have with  $\sigma, \sigma'$  and  $\sigma''$  being respectively the shifts on  $\mathcal{A}^{\mathbb{N}}, (\mathcal{A}^m)^{\mathbb{N}}$  and  $\Delta_{\sharp\mathcal{A}^m}^{\mathbb{N}}$ :

$$\begin{aligned} f_m \circ \sigma''(a) &= \pi_*(\mu_{a_2} \times \mu_{a_3} \times \dots), \\ &= \pi_* \circ \sigma'_*(\mu_{a_1} \times \mu_{a_2} \times \dots), \\ &= \sigma_*^m \circ \pi_*(\mu_{a_1} \times \mu_{a_2} \times \dots), \\ &= \sigma_*^m \circ f_m(a). \end{aligned}$$

In other terms,  $f_m : (\Delta_{\sharp\mathcal{A}^m}^{\mathbb{N}}, \sigma'') \rightarrow (\mathcal{M}(\mathcal{A}^{\mathbb{N}}), \sigma_*^m)$  is a dynamical embedding. In particular we get for all positive integers  $m$ :

$$\begin{aligned} \text{mdim}(\sigma_*) &= \frac{\text{mdim}(\sigma_*^m)}{m}, \\ &\geq \frac{\text{mdim}(\sigma'')}{m}, \\ &\geq \frac{\text{stabdim}(\Delta_{\sharp\mathcal{A}^m})}{m}. \\ &\geq \frac{\sharp\mathcal{A}^m - 1}{m} \xrightarrow{m \rightarrow \infty} +\infty. \end{aligned}$$

□

From Theorem 4, the conclusion of the Main Theorem holds for any topological system  $(X, T)$  with a horseshoe. For a positive integer  $a$  we recall  $T_a$  denotes the  $\times a$ -map on the circle. For  $a > 1$  we have  $h_{\text{top}}(T_a) = \log a > 0$  and  $T_a$  has then a horseshoe. In particular the topological mean dimension of  $T_a$  is infinite. Therefore we recover Kloeckner's result:

$$\underline{\text{mdim}}(\mathcal{M}(X), W_1, (T_a)_*) \geq \text{mdim}(\mathcal{M}(X), (T_a)_*) = +\infty.$$

4. TOOLS FOR THE PROOF IN THE GENERAL CASE

We need some preparation before proving the Main Theorem for a general topological system with positive entropy.

**4.1. Wasserstein distance.** To estimate the topological mean dimension we use the metric definition (2.2) of the topological mean dimension for the Wasserstein distance  $W_1$  on  $\mathcal{M}(X)$  (which is compatible with the weak-\* topology). The  $p^{\text{th}}$ -Wasserstein distance between two probability measures  $\mu$  and  $\nu$  in  $\mathcal{M}(X)$  is defined as

$$W_p(\mu, \nu) = \left( \inf_{\gamma \in \Gamma(\mu, \nu)} \int d(x, y)^p d\gamma(x, y) \right)^{1/p},$$

where  $\Gamma(\mu, \nu)$  denotes the collection of all measures on  $X \times X$  with marginals  $\mu$  and  $\nu$  on the first and second factors respectively. In the present paper we only make use of  $W = W_1$ . The Kantorovich-Rubinstein dual representation of  $W$  is given by

$$\forall \mu, \nu \in \mathcal{M}(X), \quad W(\mu, \nu) = \sup_f \int f (d\mu - d\nu),$$

where the supremum holds over all 1-Lipschitz functions  $f : X \rightarrow \mathbb{R}$ .

For a metric space  $(Y, \rho)$  we let  $\text{diam}_\rho(Y)$  be the diameter of  $Y$  with respect to the distance  $\rho$ . From the Kantorovich-Rubinstein dual representation of  $W$  we get easily

$$\text{diam}_W(\mathcal{M}(X)) \leq \text{diam}_d(X).$$

We also recall below a useful property of  $W$ , which follows from the Lipschitz property of the distance.

**Lemma 5.** *Let  $\mu$  and  $\nu$  be two measures in  $\mathcal{M}(X)$  respectively supported on two compact sets  $S$  and  $S'$ . Then*

$$W(\mu, \nu) \geq \mu(S \setminus S')d(S \setminus S', S').$$

*Proof.* Write  $\mu$  as  $\mu = \lambda\mu' + (1 - \lambda)\mu''$  with  $\mu', \mu'' \in \mathcal{M}(X)$  and  $\lambda \in [0, 1]$  satisfying  $\mu'(S \setminus S') = 1$  and  $\mu''(S') = 1$ . Note that we have  $\lambda = \mu(S \setminus S')$ . Then, as  $d(\cdot, S')$  is 1-Lipschitz, we get:

$$\begin{aligned} W(\mu, \nu) &\geq \int d(\cdot, S') (d\mu - d\nu), \\ &\geq \lambda \int d(\cdot, S') d\mu', \text{ because } \nu \text{ and } \mu'' \text{ are supported on } S', \\ &\geq \lambda d(S \setminus S', S') = \mu(S \setminus S')d(S \setminus S', S'). \end{aligned}$$

□

**4.2. Independence set.** For a subset  $E$  of  $\mathbb{N}$  we let  $E_n = E \cap [1, n]$  for all  $n \in \mathbb{N}$ . Then we let  $\underline{d}(E)$  be the *lower asymptotic density* of  $E$  given by

$$\underline{d}(E) = \liminf_n \frac{\#E_n}{n}.$$

B. Weiss and E. Glasner [GW95], then D. Kerr and J. Li showed positive entropy implies some weak product structure of dynamical system.

**Lemma 6** ([KL07] Proposition 3.9 (2)). *Let  $(X, T)$  be a topological system with positive topological entropy. There exist non empty open subsets  $U_0, U_1$  of  $X$  with  $\overline{U_0} \cap \overline{U_1} = \emptyset$  and  $I \subset \mathbb{N}$  with  $\underline{d}(I) > 0$  such that for any finite subset  $J$  of  $I$  and for any  $\zeta : J \rightarrow \{0, 1\}$  we have*

$$\bigcap_{j \in J} T^{-j} U_{\zeta(j)} \neq \emptyset.$$

Conversely the existence of such a pair  $U_0$  and  $U_1$  clearly implies that  $(X, T)$  has positive entropy. The set  $I$  is called an *independence set* and  $(U_0, U_1)$  is an *IE-pair*. Existence of IE-pairs follows from a combinatorial result due to N. Sauer [Sau72] and S. Shelah [She72]. This result was also a key point in the "non-ergodic" proof of  $[h_{top}(T) = 0] \Rightarrow [h_{top}(T_*) = 0]$  by B. Weiss and E. Glasner [GW95].

Given  $(X, T)$  with  $h_{top}(T) > 0$  we fix from now the IE-pair  $(U_0, U_1)$  and the independence set  $I$  as in Lemma 6. For all  $m \in \mathbb{N}$  we let

$$q_m = \left\lfloor \frac{m \cdot \underline{d}(I)}{2} \right\rfloor$$

where  $\lfloor x \rfloor$  is the largest integer  $n \leq x$  and we consider

$$I^m = \{k \in \mathbb{N}, \#[km, (k+1)m] \cap I > q_m\}.$$

As in the proof of Theorem 4 we will bound from below the mean dimension of  $T_*^m$  to show the Main Theorem and then take the limit in  $m$ . We need the following lower bound on the lower asymptotic density of  $I_m$ .

**Lemma 7.** *For any  $m \geq 1$ , we have*

$$\underline{d}(I^m) \geq \frac{\underline{d}(I)}{2}.$$

*Proof.* For  $n \in \mathbb{N}$ , we have

$$\#I_{mn} \leq m\#I_n^m + q_m(n - \#I_n^m),$$

therefore by taking the limit in  $n$  we get:

$$\begin{aligned} \underline{d}(I) &\leq \frac{m - q_m}{m} \underline{d}(I^m) + \frac{q_m}{m}, \\ &\leq \underline{d}(I^m) + \frac{\underline{d}(I)}{2}. \end{aligned}$$

□

**4.3. Simplex and generalized cube.** Recall that  $\Delta_k$  is the standard  $k$ -simplex of dimension  $k - 1$ . A  $\ell$ -face of the simplex  $\Delta_k$  is

$$\{(x_i)_{i \in I} \in [0, 1]^I : \sum_{i \in I} x_i = 1\}$$

for some  $I \subset \{1, 2, \dots, k\}$  with  $\sharp I = \ell$ . Clearly, any  $\ell$ -face is affinely homeomorphic to  $\Delta_\ell$ . The *opposite face*  $\bar{F}$  of a face  $F = \{(x_i)_{i \in I} \in [0, 1]^I : \sum_{i \in I} x_i = 1\}$  is

$$\{(x_i)_{i \in I^c} \in [0, 1]^{I^c} : \sum_{i \in I^c} x_i = 1\}$$

where  $I^c = \{1, 2, \dots, k\} \setminus I$ . The opposite face of a  $(k - 1)$ -face  $F$  is just the vertex of the simplex which does not belong to  $F$ . Notice also that for  $m (\leq k)$  distinct  $(k - 1)$ -faces  $(F_j)_{1 \leq j \leq m}$ , the intersection  $\cap_{1 \leq j \leq m} F^j$  is a  $(k - m + 1)$ -face.

We denote by  $v(\Delta_k)$  the set of vertices of  $\Delta_k$  and by  $*$  the center of  $\Delta_k$ , i.e. the point with coordinates  $(1/k, \dots, 1/k)$ . Note that  $\Delta_1$  is a compact interval. The polyhedron given by the  $n$ -product  $\Delta_k^n$  will be called a *generalized cube*. Let  $F$  be a face of  $\Delta_k$ . For  $i \in \{1, \dots, n\}$  we let  $F_i$  be the face of  $\Delta_k^n$  given by

$$F_i = \underbrace{\Delta_k \times \dots \times \Delta_k}_{i-1} \times F \times \underbrace{\Delta_k \times \dots \times \Delta_k}_{n-i}.$$

The *boundary* of  $\Delta_k^n$ , denoted by  $\partial \Delta_k^n$ , is the union of  $F_i$  for all  $(k - 1)$ -faces  $F$  and  $1 \leq i \leq n$ . The *interior* of  $\Delta_k^n$  is  $\Delta_k^n \setminus \partial \Delta_k^n$ .

We introduce an adapted notion of separation of faces for generalized cubes.

**Definition 8.** A cover  $\alpha$  of  $\Delta_k^n$  is said to be **separating** if for any  $i \in \{1, \dots, n\}$  and for any subfamily  $(U_j)_{j \in J}$  of  $\alpha$  with  $(k - 1)$ -faces  $(F^j)_{j \in J}$  of  $\Delta_k$  satisfying  $F_i^j \cap U_j \neq \emptyset$  for all  $j \in J$ , the intersection set  $\cap_{j \in J} U_j$  and the opposite face  $(\overline{\cap_{j \in J} F^j})_i$  are disjoint.

**4.4. Generalized Lebesgue's lemma.** We need a Lebesgue's like Lemma for generalized cubes.

**Lemma 9** (Generalized Lebesgue's lemma). *Any separating open cover  $\alpha$  of  $\Delta_k^n$  has order larger than or equal to  $nk$ .*

*Proof.* We argue by contradiction by considering a separating open cover  $\alpha$  of  $\Delta_k^n$  with order less than  $nk$ . Let  $(f_U)_{U \in \alpha}$  be a partition of unity associated to  $\alpha$ , i.e.  $f_U : \Delta_k^n \rightarrow [0, 1]$  is a continuous function supported on  $U$  for any  $U \in \alpha$  and  $\sum_{U \in \alpha} f_U = 1$ . We define a function  $\phi^\alpha = (\phi_1^\alpha, \dots, \phi_n^\alpha) : \alpha \rightarrow \Delta_k^n$  as follows: for any  $U \in \alpha$  and any  $i \in \{1, \dots, n\}$  we let

$$\phi_i^\alpha(U) := \begin{cases} *, & \text{if } U \cap F_i = \emptyset \text{ for any } (k - 1)\text{-face } F, \\ \underbrace{\{0\} \times \dots \times \{0\}}_{i-1} \times \bar{F} \times \underbrace{\{0\} \times \dots \times \{0\}}_{n-i}, & \text{for some } (k - 1)\text{-face } F \text{ with } F_i \cap U \neq \emptyset. \end{cases}$$

Note that  $\phi^\alpha$  is not uniquely defined in this way as  $U$  may intersect different faces of  $F_i$ , but it does not affect the proof. We define a function  $g : \Delta_k^n \circlearrowleft$  by

$$\forall x \in \Delta_k^n, \quad g(x) = \sum_{U \in \alpha} \phi^\alpha(U) f_U(x).$$

Since the order of  $\alpha$  is less than  $nk$  the function  $g$  takes values in a finite union of  $(nk - 1)$ -dimensional convex subsets of  $\Delta_k^n$ . Therefore there is a point  $\bar{x}$  in the interior of  $\Delta_k^n$  which does not lie in the image of  $g$ . Let  $r : \Delta_k^n \setminus \{\bar{x}\} \rightarrow \partial\Delta_k^n$  be a retraction map on the boundary  $\partial\Delta_k^n$  of  $\Delta_k^n$ .

Notice that if  $g(x)$  does not lie in the interior of  $\Delta_k^n$  then there is  $i \in \{1, \dots, n\}$  and a family  $\mathcal{F}$  of  $(k - 1)$ -faces such that :

- for each element  $U$  of  $\alpha$  containing  $x$  there is a  $(k - 1)$ -faces  $F \in \mathcal{F}$  with  $U \cap F_i \neq \emptyset$ ,
- $g(x)$  belongs to  $(\overline{\cap_{F \in \mathcal{F}} F})_i$ .

As the cover  $\alpha$  is assumed to be separating, we have in particular  $g(x) \neq x$  for any  $x \in \partial\Delta_k^n$ . Therefore the restriction of  $g$  to the boundary of  $\Delta_k^n$  has no fixed point. Finally the map  $r \circ g : \Delta_k^n \circlearrowleft$  has also no fixed point. This contradicts Brower's fixed point theorem.  $\square$

**4.5. A multi-affine embedding of the generalized cube  $\Delta_k^n$  in the simplex  $\Delta_{k^n}$ .** To simplify the notations we let  $\llbracket n \rrbracket := \{1, \dots, n\}$  for any positive integer  $n$ . The generalized cube  $\Delta_k^n$  may be embedded in the simplex  $\Delta_{k^n}$  by the following map:

$$(4.1) \quad \begin{aligned} \Theta : \Delta_k^n &\rightarrow \Delta_{k^n}, \\ (t_m)_{m \in \llbracket n \rrbracket} &\mapsto \sum_{\mathbf{i}=(i_m)_{m \in \llbracket n \rrbracket}} t_{\mathbf{i}} e_{\mathbf{i}}, \end{aligned}$$

where  $e_{\mathbf{i}}$  is a standard basis of  $\mathbb{R}^{k^n}$  and  $t_{\mathbf{i}} = \prod_{m \in \llbracket n \rrbracket} t_{m, i_m}$  with  $t_m = (t_{m, i})_{i \in \llbracket k \rrbracket} \in \Delta_k$  for any  $m \in \llbracket n \rrbracket$ .

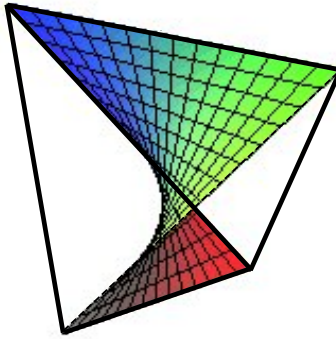


FIGURE 1. The hyperbolic paraboloid inside the simplex  $\Delta_4$ .

Clearly  $\Theta$  is continuous and injective. It is not hard to check that the embedding  $\Theta$  is not affine, but *multi-affine*, i.e. it is affine separately in each variable  $t_m$ . In particular the image  $S$  of  $\Delta_k^n$  under  $\Theta$  is not convex, but its faces may still be joined by segment lines inside  $S$ . For example if one takes  $n = k = 2$ , then the image of  $\Theta$  is a hyperbolic paraboloid, which is a doubly ruled two-dimension surface in  $\Delta_4$  (see Figure 1).

## 5. PROOF OF MAIN THEOREM

In this section, we prove our main theorem. To this end, we shall prove the mean dimension of the induced map  $T_*^m$  grows exponentially fast in  $m$ , whenever  $(X, T)$  has positive topological entropy. Recall that  $W$  denotes the 1-Wasserstein distance. We let  $W_n^m$  be the  $n$ -dynamical 1-Wasserstein distance for  $T_*^m$ , i.e.

$$\forall \mu, \nu \in \mathcal{M}(X), W_n^m(\mu, \nu) = \max_{0 \leq k < n} W(T_*^{km} \mu, T_*^{km} \nu).$$

Remark that  $W_n^1$  may differ from the Wasserstein distance  $W_{d_n}$  associated to the dynamical distance  $d_n$ . In general we only have  $W_n^1 \leq W_{d_n}$  (see Proposition 2.1 in [YJ]).

Suppose  $(X, T)$  is a dynamical system with positive topological entropy. Let  $(U_0, U_1)$  be the IE-pair and  $I$  be the independence set of  $(X, T)$  as in Lemma 6. We state below the main proposition towards the Main Theorem.

**Proposition 10.** *For any  $m \geq 1$  there is some  $\epsilon_m > 0$  such that*

$$(5.1) \quad \dim_{\epsilon_m}(\mathcal{M}(X), W_n^m) \geq 2^{q_m} \# I_n^m.$$

We first conclude the proof of the Main Theorem assuming Proposition 10.

*Proof of the Main Theorem.* By taking  $n \rightarrow +\infty$  in (5.1) we get the following inequality according to Lemma 7:

$$\lim_{n \rightarrow +\infty} \frac{\dim_{\epsilon_m}(\mathcal{M}(X), W_n^m)}{n} \geq 2^{q_m-1} \cdot \underline{d}(I).$$

Then by (2.2) we have

$$\begin{aligned} \text{mdim}(\mathcal{M}(X), T_*^m) &\geq \lim_{n \rightarrow \infty} \frac{\dim_{\epsilon_m}(\mathcal{M}(X), W_n^m)}{n}, \\ &\geq 2^{q_m-1} \cdot \underline{d}(I). \end{aligned}$$

Finally by using the formula (2.4) of the mean dimension of powers, we conclude that (recall that  $q_m$  is equal to  $\lfloor \frac{m \cdot \underline{d}(I)}{2} \rfloor$ ):

$$\begin{aligned} \text{mdim}(\mathcal{M}(X), T_*) &= \frac{\text{mdim}(\mathcal{M}(X), T_*^m)}{m}, \\ &\geq \frac{\text{mdim}_W(T_*^m, \epsilon_m)}{m}, \\ &\geq \lim_{n \rightarrow \infty} \frac{\dim_{\epsilon_m}(\mathcal{M}(X), W_n^m)}{n}, \\ &\geq \frac{2^{q_m-1}}{m} \cdot \underline{d}(I) \xrightarrow{m \rightarrow +\infty} +\infty. \end{aligned}$$

This completes the proof.  $\square$

The end of this section is devoted to the proof of Proposition 10. Fix  $m \geq 1$ . For any  $k \in I^m$  we let  $E_k$  be the  $q_m$  first integers in  $[km, (k+1)m] \cap I$  and we consider a bijection

$$\psi_k : \llbracket 2^{q_m} \rrbracket \rightarrow \{0, 1\}^{E_k}.$$

Let  $n \geq 1$ . Since  $I$  is the independent set of  $(X, T)$ , we can pick, for each sequence  $\mathbf{i} = (i_k)_{k \in I_n^m} \in \llbracket 2^{q_m} \rrbracket^{I_n^m}$ , a point  $x_{\mathbf{i}}$  with

$$x_{\mathbf{i}} \in \bigcap_{\substack{k \in I_n^m, \\ \ell \in E_k}} T^{-\ell} U_{a_{\ell}^k} \text{ where } \psi_k(i_k) = (a_{\ell}^k)_{\ell \in E_k}.$$

Let us denote by  $\delta_x$  the Dirac measure at  $x \in X$ . Let  $K_n$  inside  $\mathcal{M}(X)$  be the  $(2^{q_m})^{\sharp I_n^m}$ -simplex given by the convex hull of all  $\delta_{x_{\mathbf{i}}}$  for  $\mathbf{i} \in \llbracket 2^{q_m} \rrbracket^{I_n^m}$ . To prove Proposition 5.1 it is enough to show that

$$\dim_{\epsilon_m}(K_n, T_*^m) \geq 2^{q_m} \sharp I_n^m.$$

We consider the affine homeomorphism

$$(5.2) \quad \begin{aligned} \Psi : \Delta_{2^{q_m} \sharp I_n^m} &\rightarrow K_n, \\ (t_{\mathbf{i}})_{\mathbf{i} \in \llbracket 2^{q_m} \rrbracket^{I_n^m}} &\mapsto \sum_{\mathbf{i}} t_{\mathbf{i}} \delta_{x_{\mathbf{i}}}. \end{aligned}$$

We let  $\Xi = \Psi \circ \Theta$  with  $\Theta$  being the multi-affine embedding of  $\Delta_{2^{q_m}}^{\sharp I_n^m}$  in  $\Delta_{2^{q_m} \sharp I_n^m}$  defined in Subsection 4.5. Notice that  $\Xi$  is also a multi-affine embedding. We denote the image of  $\Xi$  by  $L_n$ . For  $E \subset \mathcal{M}(X)$ , we put  $S_E := \bigcup_{\mu \in E} \text{supp}(\mu)$ . Note that  $S_{K_n}$  consists of finitely many points  $x_{\mathbf{i}}$  for  $\mathbf{i} \in \llbracket 2^{q_m} \rrbracket^{I_n^m}$ . We have the following decomposition of the measure in  $L_n$ .

**Lemma 11.** *Let  $1 \leq \ell \leq k-1$ . Let  $F$  be a  $\ell$ -face of  $\Delta_{2^{q_m}}$  and  $\bar{F}$  be its opposite  $(2^{q_m} - \ell)$ -face. Let  $i \in I_n^m$ . For any  $\mu \in L_n$ , there is  $\lambda \in [0, 1]$ ,  $\mu' \in \Xi(F_i)$  and  $\mu'' \in \Xi(\bar{F}_i)$  with*

$$\mu = \lambda \mu' + (1 - \lambda) \mu'',$$

where  $\mu(S_{\Xi(F_i)}) = \lambda$  and  $\mu(S_{\Xi(\bar{F}_i)}) = 1 - \lambda$ .

*Proof.* Let  $F$  be a  $\ell$ -face of  $\Delta_{2^{q_m}}$  and  $i \in I_n^m$ . Suppose  $\mu = \Xi(t)$  for some  $t \in \Delta_{2^{q_m}}^{I_n^m}$ . Since the simplex  $\Delta_{2^{q_m}}$  is the convex hull of  $F$  and its opposite  $(2^{q_m} - \ell)$ -face  $\overline{F}$ , the generalized cube  $\Delta_{2^{q_m}}^{I_n^m}$  is the convex hull of  $F_i$  and  $\overline{F}_i$ . Then we write  $t = \lambda t' + (1 - \lambda)t''$  for some  $\lambda \in [0, 1]$ ,  $t' \in F_i$  and  $t'' \in \overline{F}_i$ . As  $\Xi$  is affine in the  $i^{\text{th}}$  coordinate, we conclude that

$$\mu = \lambda \Xi(t') + (1 - \lambda) \Xi(t'').$$

□

In the following we identify  $L_n$  and the generalized cube  $\Delta_{2^{q_m}}^{I_n^m}$  via the embedding  $\Xi$ . In particular we will talk about the faces of  $L_n$  as the images by  $\Xi$  of the faces of  $\Delta_{2^{q_m}}^{I_n^m}$ . For  $m \in \mathbb{N}$ , by continuity of  $T$ , there is  $\gamma_m > 0$  such that

$$[d(x, y) < \gamma_m] \Rightarrow [d(T^k x, T^k y) < d(U_0, U_1), \forall k = 0, \dots, m - 1].$$

**Lemma 12.** *Let  $\mu \in L_n$  and  $i \in I_n^m$ . For any face  $F$  of  $\Delta_{2^{q_m}}$  we have*

(1)

$$W_n^m(\mu, F_i) \geq \mu(S_{\overline{F}_i})\gamma_m,$$

*in particular,*

$$W_n^m(F_i, \overline{F}_i) \geq \gamma_m.$$

(2)

$$W_n^m(\mu, F_i) \leq \text{diam}_d(X)\mu(S_{\overline{F}_i}).$$

*Proof.* We write  $\mu$  as

$$\mu = \beta \mu' + (1 - \beta) \mu''$$

with  $\mu'' \in F_i$ ,  $\mu' \in \overline{F}_i$  and  $\beta = \mu(S_{\overline{F}_i})$  as in Lemma 11.

(1) If  $x_{\mathbf{i}}$  and  $x_{\mathbf{i}'}$  belong respectively to  $S_{\overline{F}_i}$  and  $S_{F_i}$ , then there exists  $0 \leq k < m$  with  $T^{im+k}x_{\mathbf{i}} \in U_j$  and  $T^{im+k}x_{\mathbf{i}'} \in U_{j'}$ ,  $j \neq j' \in \{0, 1\}$ . By definition of  $\gamma_m$  we have

$$d(T^{im}x_{\mathbf{i}}, T^{im}x_{\mathbf{i}'}) \geq \gamma_m.$$

It follows that

$$d(T^{im}S_{\overline{F}_i}, T^{im}S_{F_i}) \geq \gamma_m.$$

Consequently by Lemma 5 we have for all  $\nu \in F_i$  :

$$\begin{aligned} W_n^m(\mu, \nu) &\geq W(T^{im}\mu, T^{im}\nu), \\ &\geq \beta d(T^{im}S_{\overline{F}_i}, T^{im}S_{F_i}), \\ &\geq \beta \gamma_m. \end{aligned}$$

Therefore  $W_n^m(\mu, F_i) \geq \mu(S_{\overline{F}_i})\gamma_m$ . Finally, notice that if  $\mu \in \overline{F}_i$ , then  $\beta = 1$ . This implies that  $W_n^m(F_i, \overline{F}_i) \geq \gamma_m$ .



(2) Since  $\mu'' \in F_i$ , we have

$$\begin{aligned} W_n^m(\mu, F_i) &\leq W_n^m(\mu, \mu''), \\ &\leq \beta \cdot W_n^m(\mu', \mu''), \\ &\leq \beta \cdot \text{diam}_d(X). \end{aligned}$$

□

Now we let

$$(5.3) \quad \epsilon_m = \frac{\gamma_m^2}{\text{diam}_d(X) 2^{q_m+1}}.$$

**Lemma 13.** *Let  $i \in I_n^m$ . Then for any  $\mu \in L_n$  and any family  $\mathcal{F}$  of  $(2^{q_m} - 1)$ -faces of  $\Delta_{2^{q_m}}$  with  $W_n^m(\mu, F_i) < \epsilon_m$  for all  $F \in \mathcal{F}$ , we have*

$$W_n^m(\mu, \bigcap_{F \in \mathcal{F}} F_i) \leq \frac{\gamma_m}{2}.$$

*Proof.* By Lemma 12 (1), we have

$$\forall F \in \mathcal{F}, \mu(S_{\overline{F}_i}) \leq \frac{\epsilon_m}{\gamma_m}.$$

Then according to Lemma 12 (2), we have

$$\begin{aligned} W_n^m(\mu, \bigcap_{F \in \mathcal{F}} F_i) &\leq \text{diam}_d(X) \mu \left( S_{\left( \overline{\bigcap_{F \in \mathcal{F}} F} \right)_i} \right), \\ &\leq \text{diam}_d(X) \sum_{F \in \mathcal{F}} \mu(S_{\overline{F}_i}), \text{ as } S_{\left( \overline{\bigcap_{F \in \mathcal{F}} F} \right)_i} \subset \bigcup_{F \in \mathcal{F}} S_{\overline{F}_i}, \\ &\leq \text{diam}_d(X) \frac{\epsilon_m}{\gamma_m} \#\mathcal{F}, \\ &\leq \text{diam}_d(X) \frac{\epsilon_m}{\gamma_m} 2^{q_m} = \frac{\gamma_m}{2}. \end{aligned}$$

□

**Lemma 14.** *Any cover  $\alpha$  of  $L_n$  with  $W_n^m$ -diameter less than  $\epsilon_m$  is separating.*

*Proof.* Fix  $i \in I_n^m$ . According to Lemma 13, if  $(U_j)_{j \in J}$  is a subfamily of  $\alpha$  with  $\bigcap_{j \in J} U_j \neq \emptyset$  and  $F_i^j \cap U_j \neq \emptyset$ ,  $j \in J$ , for some  $(2^{q_m} - 1)$ -faces  $F^j$  of  $\Delta_{2^{q_m}}$ , then  $\bigcap_{j \in J} U_j$  lies in the  $\gamma_m/2$ -neighborhood of  $\bigcap_{j \in J} F_i^j$ . Noting that  $\overline{\bigcap_j F^j}$  is the opposite face of  $\bigcap_j F^j$  we have in the other hand by Lemma 12 (1):

$$W_n^m \left( \bigcap_j F_i^j, \left( \overline{\bigcap_j F^j} \right)_i \right) \geq \gamma_m.$$

Therefore  $\bigcap_j U_j$  and  $\left( \overline{\bigcap_j F^j} \right)_i$  are disjoint.

□

We can now conclude the proof of Proposition 10.

*Proof.* Let  $m, n \geq 1$ . Let  $\epsilon_m$  be defined as (5.3). Suppose  $f : L_n \rightarrow Z$  is a  $(W_n^m, \epsilon_m)$ -injective map with  $Z$  a compact metrizable space of finite dimension. Then there is an open cover  $\alpha$  of  $L_n$  with diameter less than  $\epsilon_m$  and with order less than or equal to  $\dim(Z)$ . According to Lemma 14, the cover  $\alpha$  of  $L_n$  is separating. From Lemma 9 the order of  $\alpha$  is larger than  $2^{qm} \#I_n^m$ . Then we conclude that

$$\begin{aligned} \dim_{\epsilon_m}(\mathcal{M}(X), W_n^m) &\geq \dim_{\epsilon_m}(L_n, W_n^m), \\ &\geq 2^{qm} \#I_n^m. \end{aligned}$$

This completes the proof of Proposition 10, therefore of the Main Theorem.  $\square$

*Proof of Theorem 1.* Suppose  $T$  is  $K$ -Lipschitz, so that we can take

$$\gamma_m = K^{-m} \cdot \frac{d(U_0, U_1)}{2}.$$

Then we have

$$\epsilon_m = c_1 e^{-mc_2},$$

where  $c_1 = \frac{d(U_0, U_1)^2}{4 \text{diam}_d(X)}$  and  $c_2 = 2 \log K + \frac{d(I) \log 2}{2}$ . It follows that

$$\begin{aligned} \text{mdim}_W(T_*, \epsilon_m) &\geq \frac{1}{m} \text{mdim}_W(T_*^m, \epsilon_m), \\ &\geq \lim_{n \rightarrow +\infty} \frac{\dim_{\epsilon_m}(\mathcal{M}(X), W_n^m)}{n}, \\ &\geq \lim_n \frac{2^{qm} \#I_n^m}{n}, \\ &\geq 2^{m \frac{d(I)}{2} - 2} \underline{d}(I). \end{aligned}$$

Since  $\epsilon_{m+1}/\epsilon_m$  is bounded, we can find  $C > 0$  such that with  $\alpha = \frac{d(I) \log 2}{2c_2}$  it holds that :

$$\forall 0 < \epsilon < 1, \quad \text{mdim}_W(T_*, \epsilon) \geq C \epsilon^{-\alpha}.$$

$\square$

## 6. RATE OF DIVERGENCE OF $h_W(T_*, \epsilon)$

In this section, we discuss about the precise rate of divergence of  $h_W(T_*, \epsilon)$ . We first give a rough upper bound on the ratio  $h_W(T_*, \epsilon)$ . For  $\epsilon > 0$  and a metric space  $(Z, \rho)$ , let  $s(Z, \rho, \epsilon)$  be the smallest cardinality of a  $\epsilon$ -covering set in  $Z$  with respect to  $\rho$ . The (upper) Minkowski dimension  $\dim_d(Z)$  of  $(Z, d)$  is defined as  $\dim_d(Z) = \limsup_{\epsilon \rightarrow 0} \frac{\log s(Z, \rho, \epsilon)}{\log \frac{1}{\epsilon}}$ .

F. Bolley, A. Guillin, and C. Villani proved the following estimate on  $s(\mathcal{M}(X), W, \epsilon)$ :

**Theorem 15.** (*Theorem A1 in [BGV07]*) For  $\epsilon > 0$  small enough, we have:

$$s(\mathcal{M}(X), W, \epsilon) \leq \left( \frac{1}{\epsilon} \right)^{s(X, d, \epsilon/2)}.$$

**Lemma 16.** *For any  $\alpha > \dim_d(Z)$ , there is  $C > 0$  such that*

$$\forall 0 < \epsilon < 1, \quad h_W(T_*, \epsilon) \leq C\epsilon^{-\alpha}.$$

*Proof.* Obviously, we have  $h_W(T_*, \epsilon) \leq \log s(\mathcal{M}(X), W, \epsilon/2)$  for any  $\epsilon > 0$ . By Theorem 15, we get

$$h_W(T_*, \epsilon) \leq s(X, d, \epsilon/4) \log \frac{2}{\epsilon} \leq C\epsilon^{-\alpha},$$

for some constant  $C$ . □

Moverover, when  $T$  is  $K$ -Lipschitz, we can refine the inequality in Lemma 16. Indeed one easily checks in this case that  $T_*$  is  $K$ -Lipschitz with respect to the Wasserstein distance. Then we have

$$h_W(T_*, \epsilon) \leq \sup_{\mu \in \mathcal{M}(X)} s(B_W(\mu, K\epsilon), W, \epsilon),$$

where  $B_W(\mu, K\epsilon)$  denotes the  $W$ -ball of radius  $K\epsilon$  centered at  $\mu$ . However we do not know any better estimate of  $\sup_{\mu \in \mathcal{M}(X)} s(B_W(\mu, K\epsilon), W, \epsilon)$  than the one given in Theorem 15.

To estimate  $h_W(T_*, \epsilon)$  from below, we first prove a technical lemma.

**Lemma 17.** *Let  $(X, T)$  be a dynamical system of positive entropy. With the notations of Subsection 4.2 we have*

$$s_n \left( W, T_*^m, \mathcal{M}(X), \frac{\gamma_m}{2^{q_m}} \right) \geq 2^{2^{q_m} \sharp I_n^m}.$$

*Proof.* Define

$$H := \{(t_k)_{k \in \llbracket 2^{q_m} \rrbracket} : \exists \ell \geq 1, \exists 1 \leq i_1 < i_2 < \dots < i_\ell \leq 2^{q_m}, \text{ s.t. } t_{i_1} = t_{i_2} = \dots = t_{i_\ell} = 1/\ell\},$$

which is a subset of  $\Delta_{2^{q_m}}$ . Moreover, note that  $\sharp H = 2^{2^{q_m}}$ . Let  $H_n$  be the image of  $H^{I_n^m}$  via  $\Xi$ . Clearly,  $\sharp H_n = 2^{2^{q_m} \sharp I_n^m}$ . We will show that  $H_n$  is a  $(n, \frac{\gamma_m}{2^{q_m}})$ -separated set of  $(\mathcal{M}(X), W, T_*^m)$ . Let  $\mu \neq \nu \in H_n$ . Then there exists  $i \in \llbracket I_n^m \rrbracket$  and a face  $F$  of  $\Delta_{2^{q_m}}$  such that either  $\mu(S_{F_i}^-) \geq 2^{-q_m}$ ,  $\nu(S_{F_i}) = 1$  or  $\nu(S_{F_i}^-) \geq 2^{-q_m}$ ,  $\mu(S_{F_i}) = 1$ . Without loss of generality, we assume that it is the former case. It follows by Lemma 5 that

$$\begin{aligned} W_n^m(\mu, \nu) &\geq W(T_*^{im} \mu, T_*^{im} \nu), \\ &\geq \mu(S_{F_i}^-) d(T_*^{im} S_{F_i}^-, T_*^{im} S_{F_i}), \\ &\geq 2^{-q_m} \gamma_m. \end{aligned}$$

It means that  $\mu$  and  $\nu$  are  $(n, \frac{\gamma_m}{2^{q_m}})$ -separated. This completes the proof. □

Now we are in a position to prove Theorem 2.

*Proof of Theorem 2.* Suppose  $T$  is of positive entropy and  $K$ -Lipschitz. Let  $U_0$  and  $U_1$  be an IE-pair and let  $I$  be the associated independence set. Then we pick

$$\gamma_m = K^{-m} \cdot \frac{d(U_0, U_1)}{2}.$$

Thus we have

$$\frac{\gamma_m}{2^{q_m}} \geq c_1 e^{-mc_2},$$

where  $c_1 = \frac{d(U_0, U_1)}{2}$  and  $c_2 = \log K + \frac{d(I) \log 2}{2}$ . It follows from Lemma 17 that

$$\begin{aligned} h_W \left( T_*^m, \frac{\gamma_m}{2^{q_m}} \right) &\geq \frac{1}{m} h_W \left( T_*^m, \frac{\gamma_m}{2^{q_m}} \right), \\ &\geq \frac{c_2 \underline{d}(I)}{\log(1/\epsilon_m)} \left( \frac{c_1}{\gamma_m/2^{q_m}} \right)^{-\frac{d(I) \log 2}{2c_2}}. \end{aligned}$$

Since  $\frac{\gamma_{m+1}/2^{q_{m+1}}}{\gamma_m/2^{q_m}}$  is bounded, we conclude that for any  $\alpha < \frac{d(I) \log 2}{2 \log K + d(I) \log 2}$ , there is  $C > 0$  such that

$$\forall 0 < \epsilon < 1, \quad h_W(T_*, \epsilon) \geq C \epsilon^{-\alpha}.$$

□

**Remark 18.** *In the case that  $(X, T)$  is of positive entropy and non-Lipschitz, we may have  $\lim_{\epsilon \rightarrow 0} \frac{\log h_W(T_*, \epsilon)}{\log \frac{1}{\epsilon}} = 0$ , e.g. when  $X$  has zero topological dimension by Lemma 16. As an explicit example, one can consider the full shift on  $K^{\mathbb{Z}}$  with  $K$  being an infinite zero-dimensional compact space endowed with a metric of zero Minkowski dimension.*

When  $(X, T)$  is Lipschitz with positive entropy, the lower Minkowski dimension of  $(X, T, d)$  is positive (see e.g. [Mil]). Then by combining Lemma 16 and Theorem 2 we get:

**Corollary 19.** *Suppose  $(X, T)$  is of positive entropy and Lipschitz with  $X$  a compact metric space with finite Minkowski dimension. Then there exists  $\alpha > 1$  such that we have*

$$\forall 0 < \epsilon < 1, \quad \alpha^{-1} \leq \frac{\log h_W(T_*, \epsilon)}{\log \frac{1}{\epsilon}} \leq \alpha.$$

## 7. MEAN DIMENSION OF THE DYNAMICS ON THE HYPERSPACE

For a compact metrizable space we let  $2^X$  be the hyperspace of  $X$ , i.e. the set of closed non empty subsets of  $X$  endowed with the Hausdorff distance. For a continuous map  $T : X \rightarrow X$  we let  $T_{\mathcal{K}} : 2^X \rightarrow 2^X$  be the dynamical system induced by  $(X, T)$  on  $2^X$ . These induced systems have also been studied in [BS75] and [GW95]. In particular it still holds true that

$$[h_{top}(T) > 0] \Rightarrow [h_{top}(T_{\mathcal{K}}) = \infty],$$

but there are zero entropy system  $(X, T)$  with  $h_{top}(T_{\mathcal{K}}) > 0$ . One may wonder about the relations between the topological entropy of  $T$  and the mean dimension (either metric or topological) of  $T_{\mathcal{K}}$ . We present below two examples.

**7.1. Example of  $h_{top}(T) = 0$  and  $\text{mdim}(T_{\mathcal{K}}) > 0$ .** Let  $K$  be a compact metric space of finite dimension. Let  $T : \cup_{\mathbb{Z} \cup \{\infty\}} K \rightarrow \cup_{\mathbb{Z} \cup \{\infty\}} K$  to be  $x_n \mapsto x_{n+1}$  for  $n \in \mathbb{Z}$  and  $x_{\infty} \mapsto x_{\infty}$  where  $\mathbb{Z} \cup \{\infty\}$  is seen as the one-point compactification of the integers. Then it is clear that  $h_{top}(T) = 0$ . On the other hand, it is not hard to verify that the hyperspace system contains the full shift  $(K^{\mathbb{Z}}, \sigma)$ . It follows that

$$\text{mdim}(2^{\cup_{\mathbb{Z} \cup \{\infty\}} K}, T_{\mathcal{K}}) \geq \text{stabdim}(K).$$

In particular, we get  $\text{mdim}(2^{\cup_{\mathbb{Z} \cup \{\infty\}} K}, T_{\mathcal{K}}) = \infty$  when  $K = [0, 1]^{\mathbb{N}}$ .

**7.2. Example of  $h_{top}(T) > 0$  and  $\text{mdim}(T_{\mathcal{K}}) = 0$ .** Let  $X$  be a zero dimensional compact metric space. As well known,  $2^X$  is a zero dimensional compact metric space. Thus for any continuous transformation  $T : X \rightarrow X$ , we have  $\text{mdim}(2^X, T_{\mathcal{K}}) = 0$ . In particular, for a full shift  $(\mathcal{A}^{\mathbb{Z}}, \sigma)$  with  $1 < \#\mathcal{A} < \infty$ , we have  $h_{top}(\sigma) = \log \#\mathcal{A} > 0$  and  $\text{mdim}(\sigma_{\mathcal{K}}) = 0$ .

## REFERENCES

- [BGV07] François Bolley, Arnaud Guillin, and Cédric Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields*, 137(3-4):541–593, 2007.
- [BS75] Walter Bauer and Karl Sigmund. Topological dynamics of transformations induced on the space of probability measures. *Monatsh. Math.*, 79:81–92, 1975.
- [BS21] David Burgnet and Ruxi Shi. Mean dimension of continuous cellular automata. *arXiv preprint arXiv:2105.09708, Isr. J. Math., to appear*, 2021.
- [Coo05] Michel Coornaert. *Dimension topologique et systèmes dynamiques*, volume 14 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2005.
- [Gou11] Antoine Gournay. Widths of  $\ell^p$  balls. *Houston J. Math.*, 37(4):1227–1248, 2011.
- [Gro99] Misha Gromov. Topological invariants of dynamical systems and spaces of holomorphic maps. I. *Math. Phys. Anal. Geom.*, 2(4):323–415, 1999.
- [GW95] Eli Glasner and Benjamin Weiss. Quasi-factors of zero-entropy systems. *Journal of the American Mathematical Society*, 8(3):665–686, 1995.
- [JQ21] Lei Jin and Yixiao Qiao. Mean dimension of product spaces: a fundamental formula, 2021.
- [Kat80] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Inst. Hautes Études Sci. Publ. Math.*, (51):137–173, 1980.
- [KL07] David Kerr and Hanfeng Li. Independence in topological and  $C^*$ -dynamics. *Math. Ann.*, 338(4):869–926, 2007.
- [Klo13] Benoît Kloeckner. Optimal transport and dynamics of expanding circle maps acting on measures. *Ergodic Theory Dynam. Systems*, 33(2):529–548, 2013.
- [Leb11] Henri Lebesgue. Sur la non-applicabilité de deux domaines appartenant respectivement à des espaces à  $n$  et  $n + p$  dimensions. *Math. Ann.*, 70(2):166–168, 1911.
- [LT19] Elon Lindenstrauss and Masaki Tsukamoto. Double variational principle for mean dimension. *Geometric and Functional Analysis*, pages 1–62, 2019.
- [LW00] Elon Lindenstrauss and Benjamin Weiss. Mean topological dimension. *Israel J. Math.*, 115:1–24, 2000.
- [Mil] John Milnor. Introductory dynamics lectures. <http://www.math.stonybrook.edu/jack/DYNOTES/>.
- [Mis79] Michał Misiurewicz. Horseshoes for mappings of the interval. *Bull. Acad. Polon. Sci. Sér. Sci. Math.*, 27(2):167–169, 1979.
- [Sau72] N. Sauer. On the density of families of sets. *J. Combinatorial Theory Ser. A*, 13:145–147, 1972.
- [She72] Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41:247–261, 1972.

- [Sig78] Karl Sigmund. Affine transformations on the space of probability measures. In *Dynamical systems, Vol. III—Warsaw*, pages 415–427. Astérisque, No. 51. 1978.
- [Tsu08] Masaki Tsukamoto. Macroscopic dimension of the  $l^p$ -ball with respect to the  $l^q$ -norm. *J. Math. Kyoto Univ.*, 48(2):445–454, 2008.
- [Tsu19] Masaki Tsukamoto. Mean dimension of full shifts. *Israel Journal of Mathematics*, 230(1):183–193, 2019.
- [YJ] Xiaoyao Zhou Yong Ji, Ercai Chen. Entropy and emergence of topological dynamical systems, <https://arxiv.org/abs/2005.01548>.

SORBONNE UNIVERSITE, LPSM, 75005 PARIS, FRANCE  
*Email address:* david.burguet@upmc.fr

SORBONNE UNIVERSITE, LPSM, 75005 PARIS, FRANCE  
*Email address:* ruxi.shi@upmc.fr