

ON QUANTIZATION AND THE CLASSICAL VARIATIONAL PRINCIPLE FOR THE METRIC MEAN DIMENSION

MARIA CARVALHO AND GUSTAVO PESSIL

ABSTRACT. This paper investigates the relationship between quantization of measures and metric mean dimension of topological dynamical systems. We introduce the concept of mean quantization dimension for invariant probability measures and establish a classical variational principle: the metric mean dimension of a dynamical system is equal to the maximum mean quantization dimension among all invariant measures. This approach effectively characterizes the complexity of systems with infinite entropy by identifying a measure that captures information across all scales; and yields a fundamental property that allows for the exchange of limits and suprema in the Lindenstrauss-Tsukamoto variational principles, a feat that most known entropy-like maps fail to achieve due to convexity. Nevertheless, we show that the Katok and Shapira entropies do satisfy this property and, therefore, a classical variational principle for the metric mean dimension, for which maximizing measures always exist.

1. INTRODUCTION

Let (X, d) be a compact metric space and $T: X \rightarrow X$ be a homeomorphism. Denote by $\mathcal{P}(X)$ the set of Borel probability measures on X with the weak*-topology; by $\mathcal{P}_T(X)$ its subset of T -invariant measures; by $\mathcal{E}_T(X) \subset \mathcal{P}_T(X)$ the set of ergodic elements; and, given $\mu \in \mathcal{P}(X)$, $\text{supp}(\mu)$ stands for its support.

Measure-theoretic and topological entropy are classical invariants in the theory of dynamical systems and are related by the classical variational principle

$$h_{\text{top}}(T) = \sup_{\mu \in \mathcal{P}_T(X)} h_{\mu}(T) = \sup_{\mu \in \mathcal{E}_T(X)} h_{\mu}(T) \quad (1)$$

which sometimes provides a natural way to choose important elements from $\mathcal{P}_T(X)$. Entropy-dimensional concepts have been recently developed in order to estimate the complexity of infinite entropy systems from a better suited viewpoint. The *metric mean dimension*, introduced by E. Lindenstrauss and B. Weiss in [7], is both a dynamical analog of the box dimension and a dimensional analog of the topological entropy. In what follows, we denote the upper and lower metric mean dimensions by $\text{mdim}_{\text{M}}(X, d, T)$ and $\underline{\text{mdim}}_{\text{M}}(X, d, T)$, respectively, thus emphasizing their dependence on the fixed metric d on the space X where the dynamics T acts. These concepts vanish if the topological entropy of T is finite; if, otherwise, T has infinite entropy, they quantify the speed at which the entropy at scale ε approaches $+\infty$ as the scale goes to zero. The choice of the metric d has impact precisely on such convergence rate.

In [8], E. Lindenstrauss and M. Tsukamoto established a connection between the upper metric mean dimension and the rate distortion function \mathcal{R}_{μ} . More precisely, they showed that with the

Date: June 2, 2026.

2010 Mathematics Subject Classification. Primary: 37B99, 28D20, 37D35.

Key words and phrases. Quantization of measures; Metric mean dimension; Variational principle.

tame growth of covering numbers

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)}.$$

Meanwhile, in [5], Y. Gutman and A. Śpiewak proved a corresponding relation between the upper metric mean dimension and the Rényi entropy, namely

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\inf_{\text{diam}(P) \leq \varepsilon} h_{\mu}(P)}{\log(1/\varepsilon)}$$

where the infimum is taken over the finite Borel partitions P of X with diameter at most ε and $h_{\mu}(P)$ stands for the entropy of μ with respect to P . Moreover, Problem 3 in [5] asked whether the upper metric mean dimension could be expressed in terms of the Brin-Katok local entropy. An affirmative answer to this problem was given by R. Shi in [12]:

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\overline{h}_{\mu}^{BK}(\varepsilon)}{\log(1/\varepsilon)}.$$

Apart from the ones mentioned above, many other distinct notions of measure-theoretic metric mean dimension were shown to exist and be inspired by some version of entropy, such as Katok metric entropy (cf. [12]), Shapira entropy (cf. [12]) and Pfister-Sullivan entropy (cf. [18]), just to mention a few. A unified framework was provided by E. Chen, R. Yang and X. Zhou in [2], where they considered the following family of entropy-like functions of a measure μ at a given scale ε (see the precise definitions in [2] and Section 3):

$$\mathcal{D} = \left\{ \begin{array}{l} \underline{h}_{\mu}^K(\varepsilon, \delta), \overline{h}_{\mu}^K(\varepsilon, \delta), \underline{h}_{\mu}^K(\varepsilon), \overline{h}_{\mu}^K(\varepsilon), \overline{h}_{\mu}^{BK}(\varepsilon), h_{\mu}^S(\varepsilon, \delta), \\ PS_{\mu}(T, \varepsilon), \mathcal{R}_{\mu, L^{\infty}}(\varepsilon), \inf_{\text{diam}(P) \leq \varepsilon} h_{\mu}(P), \mathcal{P}_{\mu}(T, d, \varepsilon) \end{array} \right\}. \quad (2)$$

The authors show that, for any ergodic probability measure $\mu \in \mathcal{E}_T(X)$, the limits

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \quad \text{and} \quad \liminf_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \quad (3)$$

do not depend on the choice of $F \in \mathcal{D}$ (so they coincide with the upper and lower L^{∞} -rate distortion dimensions, respectively) and that, for every $F \in \mathcal{D}$,

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}. \quad (4)$$

The corresponding equalities also hold for the lower limits. We stress that, in equality (4), maximization over the space of probability measures is done for a fixed scale, which only afterwards is made to go to zero. A recurring question in the literature asks whether we may exchange the order of \limsup_{ε} and \sup_{μ} in (4). An affirmative answer to this question would suggest natural definitions of measure-theoretic metric mean dimension, namely

$$\overline{\text{mdim}}_{\text{M}}(X, d, T, F)(\mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

and provide a classical variational principle for the metric mean dimension in terms of F ; moreover, if \sup_{μ} could be replaced by \max_{μ} , this principle would single out a measure which captures the complexities of the system on all scales ε . It is known that this order can be changed under

the marker property (cf. [9]), but there are relevant systems that do not satisfy this condition, as shown in [13, 15]. The following examples shed some light on this problem.

Example 1.1. [8, Section VIII] Consider the set $A = \{0\} \cup \{1/n : n \in \mathbb{N}\}$, $X = A^{\mathbb{Z}}$ endowed with the metric

$$d(x, y) = \sum_{n \in \mathbb{Z}} 2^{-|n|} |x_n - y_n|$$

and $T: X \rightarrow X$ the standard shift. Then $\overline{\text{mdim}}_{\text{M}}(X, d, T) = 1/2$ (cf. [17]) and

$$\forall \mu \in \mathcal{P}_T(X) \quad \lim_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)} = 0.$$

So,

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) > \sup_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)}.$$

Example 1.2. [14, Theorem 1.4] In [14] Tsukamoto constructs a homeomorphism T on a compact metric space (X, d) with the following properties:

- For any ergodic measure $\nu \in \mathcal{P}_T(X)$ we have $\limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\nu}(\varepsilon)}{\log(1/\varepsilon)} = 0$.
- There exists $\eta \in \mathcal{P}_T(X)$ such that $\limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\eta}(\varepsilon)}{\log(1/\varepsilon)} = +\infty$.

Consequently, one has $\overline{\text{mdim}}_{\text{M}}(X, d, T) = +\infty$ and

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) > \sup_{\mu \in \mathcal{E}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)}.$$

Yet, it is still true for this example that

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \sup_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)}. \quad (5)$$

Why do they behave differently with respect to the change in the order in which \limsup_{ε} and $\sup_{\mu \in \mathcal{P}_T(X)}$ appear? It turns out that a major role in this issue is played by the behavior of the measure-theoretic ε -entropy under ergodic decompositions. On the one hand, if the upper metric mean dimension is finite, then the upper rate distortion dimension

$$\overline{R}(X, d, T, \mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\mathcal{R}_{\mu}(\varepsilon)}{\log(1/\varepsilon)}$$

is a convex function of $\mu \in \mathcal{P}_T(X)$ with respect to its ergodic decomposition (cf. [14, Theorem 1.1]). Thus, in Example 1.1, the change in order between \limsup_{ε} and $\sup_{\mu \in \mathcal{P}_T(X)}$ produces a strict inequality. On the other hand, the system described in Example 1.2 has $\overline{\text{mdim}}_{\text{M}}(X, d, T) = +\infty$ and the convexity no longer holds. This is the reason why (5) is possible.

The purpose of this paper is to explain that we can understand this issue more clearly by comparing the maps $F \in \mathcal{D}$ with a new concept that we introduce in the next section.

2. MAIN RESULTS

After recalling the problem of quantization of measures, we motivate the questions we address and describe our main contributions.

2.1. Quantization. Let (X, d) be a compact metric space. Quantization of measures, as addressed in [4], consists in approximating, at arbitrarily small resolutions, a given measure by another with finite support. Let $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ stand for either a Wasserstein distance or the Lévy-Prokhorov distance on the space of Borel probability measures $\mathcal{P}(X)$ (see the precise definitions in Subsection 3.5).

Definition 2.1. *Given $\mu \in \mathcal{P}(X)$ and $\varepsilon > 0$, the quantization number $Q_{\mu, D}(\varepsilon)$ of μ with respect to D at a scale ε is the least positive integer N such that there exists $\nu \in \mathcal{P}(X)$ with $D(\mu, \nu) < \varepsilon$ and ν is supported on a set of cardinality N . The upper and lower quantization dimensions of μ are defined as*

$$\overline{\dim}_Q(\mu, D) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\log Q_{\mu, D}(\varepsilon)}{\log(1/\varepsilon)} \quad \text{and} \quad \underline{\dim}_Q(\mu, D) = \liminf_{\varepsilon \rightarrow 0^+} \frac{\log Q_{\mu, D}(\varepsilon)}{\log(1/\varepsilon)}.$$

It is known (cf. [1, Proposition 3.4]) that the quantization number of μ at scale ε is bounded from above by the ε -covering number of the space (see Subsection 3.3). Moreover, the quantization dimensions are related to the upper and lower box-counting dimensions of the space by the following variational principles (see [1, Proof of Theorem 3.9]):

$$\overline{\dim}_B(X, d) = \max_{\mu \in \mathcal{P}(X)} \overline{\dim}_Q(\mu, D) \quad \text{and} \quad \underline{\dim}_B(X, d) = \max_{\mu \in \mathcal{P}(X)} \underline{\dim}_Q(\mu, D). \quad (6)$$

2.2. Dynamical quantization. We now address the corresponding dynamical objects. Let $T: X \rightarrow X$ be a continuous map. Given $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ and $n \in \mathbb{N}$, we denote by D_n the Wasserstein or the Lévy-Prokhorov distance induced in $\mathcal{P}(X)$ by the Bowen metric d_n in X (defined in (16)).

Definition 2.2. *Given $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$, $\mu \in \mathcal{P}(X)$ and $\varepsilon > 0$, the upper and lower dynamical quantization numbers of $\mu \in \mathcal{P}(X)$ at scale ε are given by*

$$\overline{Q}_\mu(T, D, \varepsilon) = \limsup_{n \rightarrow +\infty} \frac{1}{n} \log Q_{\mu, D_n}(\varepsilon) \quad \text{and} \quad \underline{Q}_\mu(T, D, \varepsilon) = \liminf_{n \rightarrow +\infty} \frac{1}{n} \log Q_{\mu, D_n}(\varepsilon).$$

The following relation between the metric entropy and the dynamical quantization numbers of an ergodic measure was proved in [1].

Theorem 2.3. [1, Appendix A.3] *Let (X, d) be a compact metric space and $T: X \rightarrow X$ be a continuous map. Given $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ and $\mu \in \mathcal{E}_T(X)$, then*

$$h_\mu(T) = \lim_{\varepsilon \rightarrow 0^+} \overline{Q}_\mu(T, D, \varepsilon) = \lim_{\varepsilon \rightarrow 0^+} \underline{Q}_\mu(T, D, \varepsilon).$$

We observe that, applying [1, Proposition 3.4 & (1.1)] to the space (X, d_n) , we obtain

$$\overline{Q}_\mu(T, D, \varepsilon) \leq h_\varepsilon(X, d, T) \quad \forall \mu \in \mathcal{P}(X), \quad (7)$$

where

$$h_\varepsilon(X, d, T) = \limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon)$$

and $S(X, n, \varepsilon)$ is the maximal cardinality of the ε -separated subsets of X (see Subsection 3.2). Combining (7) with Theorem 2.3 and the classical variational principle (1), we get

$$\begin{aligned} h_{\text{top}}(T) &= \sup_{\mu \in \mathcal{E}_T(X)} \lim_{\varepsilon \rightarrow 0^+} \overline{Q}_\mu(T, D, \varepsilon) = \sup_{\mu \in \mathcal{E}_T(X)} \lim_{\varepsilon \rightarrow 0^+} \underline{Q}_\mu(T, D, \varepsilon) \\ &= \sup_{\mu \in \mathcal{P}_T(X)} \lim_{\varepsilon \rightarrow 0^+} \overline{Q}_\mu(T, D, \varepsilon) = \sup_{\mu \in \mathcal{P}_T(X)} \lim_{\varepsilon \rightarrow 0^+} \underline{Q}_\mu(T, D, \varepsilon). \end{aligned} \quad (8)$$

Definition 2.4. *The upper and lower mean quantization dimensions of a measure $\mu \in \mathcal{P}_T(X)$ with respect to the metric $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ are given, respectively, by*

$$\begin{aligned} \overline{\text{mdim}}_Q(X, d, T, D, \mu) &= \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)} \\ \underline{\text{mdim}}_Q(X, d, T, D, \mu) &= \liminf_{\varepsilon \rightarrow 0^+} \frac{\underline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)}. \end{aligned}$$

If they agree, the common value is denoted by $\text{mdim}_Q(X, d, T, D, \mu)$ and called mean quantization dimension.

We will show (cf. Lemmas 6.2 and 6.5) that, if μ is ergodic, then the mean quantization dimensions of μ do not change if one replaces $\overline{Q}_\mu(T, D, \varepsilon)$ by $\underline{Q}_\mu(T, D, \varepsilon)$ in their definitions. Also, by (7), we have

$$\overline{\text{mdim}}_Q(X, d, T, D, \mu) \leq \overline{\text{mdim}}_M(X, d, T) \quad \forall \mu \in \mathcal{P}_T(X). \quad (9)$$

Our first result states that the dynamical quantization numbers can be included in the set \mathcal{D} , that is, they satisfy the properties (3) and (4). Moreover, the mean quantization dimension satisfies a classical variational principle, where maximization is taken over the space of invariant probability measures, which can be seen both as a dimensional analog of (1) and as a dynamical analog of (6). Furthermore, measures with maximal mean quantization dimension always exist.

Theorem A. *Let (X, d) be a compact metric space, $T: X \rightarrow X$ be a homeomorphism and $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$. Then:*

(a) *For every $\mu \in \mathcal{E}_T(X)$ and $F \in \mathcal{D}$,*

$$\overline{\text{mdim}}_Q(X, d, T, D, \mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

(b) [Variational principle I]

$$\begin{aligned} \overline{\text{mdim}}_M(X, d, T) &= \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\overline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\underline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)} \\ &= \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{\overline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{\underline{Q}_\mu(T, D, \varepsilon)}{\log(1/\varepsilon)}. \end{aligned}$$

(c) [Variational principle II]

$$\begin{aligned}
\overline{\text{mdim}}_{\text{M}}(X, d, T) &= \max_{\mu \in \mathcal{P}_T(X)} \overline{\text{mdim}}_{\text{Q}}(X, d, T, D, \mu) \\
&= \max_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{Q}_{\mu}(T, D, \varepsilon)}{\log(1/\varepsilon)} \\
&= \max_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{Q_{\mu}(T, D, \varepsilon)}{\log(1/\varepsilon)}.
\end{aligned} \tag{10}$$

(d) *Unless T is uniquely ergodic, there exist uncountably many probabilities measures with maximal mean quantization dimension.*

The corresponding statements for the lower metric mean dimension also hold.

We observe that, in general, the maximum in (10) cannot be replaced by the supremum over the space of ergodic measures. Indeed, in Section 7 we discuss an example of a compact metric space (X_0, d_0) and a homeomorphism $T_0: X_0 \rightarrow X_0$ for which

$$\overline{\text{mdim}}_{\text{M}}(X_0, d_0, T_0) > \sup_{\mu \in \mathcal{P}_{T_0}(X_0)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_{\mu}^{BK}(\varepsilon)}{\log(1/\varepsilon)}.$$

By Theorem A(a), for any $F \in \mathcal{D}$ and $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ we have

$$\overline{\text{mdim}}_{\text{M}}(X_0, d_0, T_0) > \sup_{\mu \in \mathcal{E}_{T_0}(X_0)} \overline{\text{mdim}}_{\text{Q}}(X_0, d_0, T_0, D, \mu) = \sup_{\mu \in \mathcal{E}_{T_0}(X_0)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

2.3. Main property. The fundamental aspect that enables us to prove (10) is that the quantization numbers of an invariant measure are, up to normalization of scales by the weights appearing in its ergodic decomposition, at least the quantization numbers of its ergodic components. We express this idea using the following property.

Definition 2.5. *Let \mathcal{K} be a convex subset of $\mathcal{P}(X)$. A map $F: \mathcal{K} \times (0, +\infty) \rightarrow [0, +\infty]$ is said to satisfy the scaled monotonicity under measure domination (s.m.m.d. for short) in \mathcal{K} if there exists a constant $C > 0$ such that, for every $\mu, \nu \in \mathcal{K}$ with $\mu \geq t\nu$ for some $t \in (0, 1)$, one has*

$$F(\mu, t\varepsilon) \geq F(\nu, C\varepsilon) \quad \forall \varepsilon > 0.$$

We observe that, if $\mathcal{K} = \mathcal{P}_T(X)$ and $\mu \in \mathcal{K}$ is ergodic, then the condition $\mu \geq t\nu$ for some $\nu \in \mathcal{K} \setminus \{\mu\}$ is empty. Otherwise, ν would be absolutely continuous with respect to μ , hence ν is ergodic, and, as ν is T -invariant, $\nu = \mu$.

The s.m.m.d. property is essentially incompatible with convexity. For example, assume that map $F(\cdot, \varepsilon)$ is convex in \mathcal{K} for every $\varepsilon > 0$ and satisfies the s.m.m.d. property. Then, for any $\mu, \nu \in \mathcal{K}$, we have

$$F(\nu, C\varepsilon) \leq F((\mu + \nu)/2, \varepsilon/2) \leq [F(\mu, \varepsilon/2) + F(\nu, \varepsilon/2)]/2.$$

Hence,

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\nu, \varepsilon)}{\log(1/\varepsilon)} \leq \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

and we conclude that the map

$$\mu \mapsto \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

is constant in \mathcal{K} . Moreover, the s.m.m.d. property implies that, whenever $\mu = t\mu_1 + (1-t)\mu_2$ for some $t \in (0, 1)$ and $\mu_1, \mu_2 \in \mathcal{P}(X)$, then

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \geq \max \left\{ \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu_1, \varepsilon)}{\log(1/\varepsilon)}, \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu_2, \varepsilon)}{\log(1/\varepsilon)} \right\} \quad (11)$$

which is a rather strong form of concavity.

Theorem B. *Let (X, d) be a compact metric space and \mathcal{K} be a convex subset of $\mathcal{P}(X)$. Assume that $F: \mathcal{K} \times (0, +\infty) \rightarrow [0, +\infty]$ satisfies the s.m.m.d. property in \mathcal{K} . Then:*

(a)

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{K}} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}. \quad (12)$$

(b) *If, in addition, F is non-increasing in the variable ε , then*

$$\liminf_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{K}} \liminf_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}. \quad (13)$$

(c) *The space $\mathcal{K}_{max} \subset \mathcal{K}$ of maximizing measures in (12) (resp. (13)) satisfies the following property:*

$$\mu \in \mathcal{K}_{max}, \eta \in \mathcal{K}, \lambda \in [0, 1] \implies (1-\lambda)\mu + \lambda\eta \in \mathcal{K}_{max}.$$

We note that if the map $\mu \in \mathcal{K} \mapsto \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$ is concave, then the space \mathcal{K}_{max} is convex. Item (c) above states that under the s.m.m.d. property, if \mathcal{K} is not a singleton, then the set \mathcal{K}_{max} is huge.

The previous result provides a generalization of item (c) of Theorem A that also answers the next question.

Question 2.6. *Under what assumptions on $F: \mathcal{P}_T(X) \times (0, +\infty) \rightarrow [0, +\infty]$ can we ensure that*

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \sup_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} ?$$

Corollary 2.7. *Let (X, d) be a compact metric space and $T: X \rightarrow X$ be a homeomorphism. If $F \in \mathcal{D}$ satisfies the s.m.m.d. property in $\mathcal{P}_T(X)$, then*

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \max_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}. \quad (14)$$

Since $F \in \mathcal{D}$ is non-increasing in the variable ε , the corresponding lower version of (14) also holds. Moreover, unless T is uniquely ergodic, there exist uncountably many probabilities measures with maximal metric mean dimension.

2.4. Variational principles for maps in \mathcal{D} . Since all maps $F \in \mathcal{D}$, as well as the dynamical quantization numbers, have the same growth rate at ergodic measures, it is worthwhile to investigate how each individual F acts on invariant measures, as this is the main key to obtaining the variational principle (10). In spite of the common probabilistic and geometric hybrid nature of these concepts, some of them satisfy the s.m.m.d. property (as happens with Katok ε -entropy and Shapira ε -entropy; cf. Remark 7.1), while others are convex (like Brin-Katok ε -entropy; cf. Lemma 4.7). This is the content of the next result.

Theorem C. *The following assertions hold:*

- (a) *Let (X, d) be compact metric space and $T: X \rightarrow X$ be a homeomorphism. Then, for every $\mu \in \mathcal{P}_T(X)$,*

$$\overline{\text{mdim}}_Q(X, d, T, LP, \mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_\mu^K(\varepsilon, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_\mu^S(\varepsilon, \varepsilon)}{\log(1/\varepsilon)}$$

Consequently, if $F \in \{\overline{h}_\mu^K(\varepsilon), \overline{h}_\mu^S(\varepsilon)\}$, then

$$\overline{\text{mdim}}_M(X, d, T) = \max_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

The corresponding statements for the lower metric mean dimension also hold.

- (b) *There exists a compact metric space (X_0, d_0) and a homeomorphism $T_0: X_0 \rightarrow X_0$ such that*

$$\text{mdim}_M(X_0, d_0, T_0) > \sup_{\mu \in \mathcal{P}_{T_0}(X_0)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

for every

$$F \in \left\{ \begin{array}{l} \underline{h}_\mu^K(\varepsilon, \delta), \overline{h}_\mu^K(\varepsilon, \delta), \underline{h}_\mu^S(\varepsilon, \delta), \overline{h}_\mu^S(\varepsilon, \delta), \\ \overline{h}_\mu^{BK}(\varepsilon), \inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P) \end{array} \right\}.$$

2.5. Ergodic equilibrium states. The previous discussion suggests another question, previously addressed in [2].

Question 2.8. *Under what conditions on the space (X, d) , on the map $T: X \rightarrow X$, or else on the measure-theoretic ε -entropy map F can we guarantee that*

$$\overline{\text{mdim}}_M(X, d, T) = \sup_{\mu \in \mathcal{E}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} ?$$

The next two examples describe settings where this principle is valid.

Example 2.9. Let (X, d) be a compact connected smooth boundaryless manifold with dimension $\dim X \geq 2$, endowed with a Riemannian metric d . A Borel probability measure λ on X is said to be *good* if $\lambda(\{x\}) = 0$ for every $x \in X$ and $\lambda(U) > 0$ for every non-empty open subset U of X . Denote by $\text{Homeo}_\lambda(X)$ the space of homeomorphisms $T: X \rightarrow X$ that preserve λ . This is a complete metric space if one considers in $\text{Homeo}_\lambda(X)$ the topology induced by the metric

$$\rho(T_1, T_2) = \max \{d(T_1(x), T_2(x)), d(T_1^{-1}(x), T_2^{-1}(x))\}.$$

G. Lacerda and S. Rom ana proved in [6, Theorems A & B] that a C^0 -generic homeomorphism $T \in \text{Homeo}_\lambda(X)$ satisfies

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \dim X = \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_\lambda^K(\varepsilon)}{\log(1/\varepsilon)}. \quad (15)$$

Consider the residual subset \mathfrak{R} of $\text{Homeo}_\lambda(X)$ such that, if $T \in \mathfrak{R}$, then λ is ergodic (cf. [10]) and (15) holds. Combining this information with item (a) of Theorem A, we conclude that, for every $T \in \mathfrak{R}$, $F \in \mathcal{D}$ and $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$,

$$\overline{\text{mdim}}_{\text{M}}(X, d, T) = \max_{\mu \in \mathcal{E}_T(X)} \overline{\text{mdim}}_{\text{Q}}(X, d, T, D, \mu) = \max_{\mu \in \mathcal{E}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

with λ as a maximizing measure.

Example 2.10. Consider the set $A = [0, 1]$ with the Euclidean metric $|\cdot|$, the space $X = A^{\mathbb{Z}}$ endowed with the metric

$$d(x, y) = \sum_{n \in \mathbb{Z}} 2^{-|n|} |x_n - y_n|$$

and the standard shift $\sigma: X \rightarrow X$. Then $\overline{\text{mdim}}_{\text{M}}(X, d, \sigma) = 1$ (cf. [8]). Let $\lambda = \text{Leb}^{\mathbb{Z}}$ be the product measure on $[0, 1]^{\mathbb{Z}}$ of the Lebesgue measure on $[0, 1]$. Then λ is σ -invariant and ergodic; and, due to [8, E. Example] and Theorem A, for $D \in \{W_p: p \in [1, +\infty)\} \cup \{LP\}$ and every $F \in \mathcal{D}$,

$$\overline{\text{mdim}}_{\text{M}}(X, d, \sigma) = \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\lambda, \varepsilon)}{\log(1/\varepsilon)} = \overline{\text{mdim}}_{\text{Q}}(X, d, \sigma, D, \lambda).$$

Therefore,

$$\overline{\text{mdim}}_{\text{M}}(X, d, \sigma) = \max_{\mu \in \mathcal{E}_\sigma(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{E}_\sigma(X)} \overline{\text{mdim}}_{\text{Q}}(X, d, \sigma, D, \mu).$$

The previous example can be generalized. A compact metric space (A, ρ) is Ahlfors regular if there exists a Borel probability measure $\nu \in \mathcal{P}(A)$ satisfying

$$\exists \varepsilon_0, C_1, C_2 > 0: \quad C_1 \varepsilon^{\dim_{\text{B}}(A, \rho)} \leq \nu(B(a, \varepsilon)) \leq C_2 \varepsilon^{\dim_{\text{B}}(A, \rho)} \quad \forall a \in A, \forall 0 < \varepsilon < \varepsilon_0$$

where $B(a, \varepsilon)$ stands for the open ball in A centered at a with radius ε with respect to the metric ρ . For instance, a finite set with the normalized counting measure or the Euclidean cube with the Lebesgue measure are Ahlfors regular. Consider a Ahlfors regular space (A, ρ, ν) , the shift map on the space $X = A^{\mathbb{Z}}$ endowed with the metric d and the product measure $\lambda = \nu^{\mathbb{Z}} \in \mathcal{E}_\sigma(X)$. Given $F \in \mathcal{D}$, we already know that

$$\overline{\text{mdim}}_{\text{M}}(X, d, \sigma) \geq \sup_{\mu \in \mathcal{E}_\sigma(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

On the other hand, for any $x \in X$, $\varepsilon \in (0, \varepsilon_0)$ and $n \in \mathbb{N}$,

$$B_n(x, \varepsilon) \subset \{y \in X: y_j \in B(x_j, \varepsilon) \quad \forall 0 \leq j < n\}$$

so,

$$\lambda(B_n(x, \varepsilon)) \leq \prod_{0 \leq j < n} \nu(B(x_j, \varepsilon)) \leq C_2^n \cdot \varepsilon^{n \dim_{\text{B}}(A, \rho)}.$$

These estimates yield (see Definition 4.5)

$$\bar{h}_\lambda^{BK}(\varepsilon) \geq -\log C_2 + \dim_B(A, \rho) \cdot \log(1/\varepsilon).$$

Therefore,

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_\lambda^{BK}(\varepsilon)}{\log(1/\varepsilon)} \geq \dim_B(A, \rho) = \overline{\text{mdim}}_M(X, d, \sigma).$$

Thus,

$$\begin{aligned} \overline{\text{mdim}}_M(X, d, \sigma) &= \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\lambda, \varepsilon)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{E}_\sigma(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \\ &= \overline{\text{mdim}}_Q(X, d, \sigma, D, \lambda) = \max_{\mu \in \mathcal{E}_\sigma(X)} \overline{\text{mdim}}_Q(X, d, \sigma, D, \mu). \end{aligned}$$

Remark 2.11. Regarding the lower Brin-Katok entropy $\underline{h}_\lambda^{BK}(\varepsilon)$, R. Yang proved in [20, Theorem 3.9] that, if the Hausdorff and the upper box dimension of the alphabet (A, ρ) coincide, then

$$\overline{\text{mdim}}_M(A^{\mathbb{Z}}, d, \sigma) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_\sigma(A^{\mathbb{Z}})} \frac{\underline{h}_\mu^{BK}(\varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_\sigma(A^{\mathbb{Z}})} \frac{\underline{h}_\mu^{BK}(\varepsilon)}{\log(1/\varepsilon)}.$$

For instance, if (A, ρ) is Ahlfors regular then this condition is satisfied. Furthermore, the previous reasoning implies that

$$\overline{\text{mdim}}_M(A^{\mathbb{Z}}, d, \sigma) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\underline{h}_\lambda^{BK}(\varepsilon)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{E}_\sigma(A^{\mathbb{Z}})} \limsup_{\varepsilon \rightarrow 0^+} \frac{\underline{h}_\mu^{BK}(\varepsilon)}{\log(1/\varepsilon)}.$$

2.6. Organization of the paper. In Sections 3 and 4 we define the main concepts we use and gather some of their properties. After proving Theorems A and B in the ensuing sections, we pave the way to proving Theorem C by thoroughly discussing an example in Section 7.

3. BASIC DEFINITIONS

Let (X, d) be a compact metric space and $T: X \rightarrow X$ be a continuous map.

3.1. Bowen dynamical metrics. Given $n \in \mathbb{N}$, the dynamical metric $d_n: X \times X \rightarrow [0, +\infty)$, determined by the distance d in X and the map T , is defined by

$$d_n(x, z) = \max \left\{ d(x, z), d(T(x), T(z)), \dots, d(T^{n-1}(x), T^{n-1}(z)) \right\}. \quad (16)$$

It is easy to check that d_n is indeed a metric and that it generates the same topology as d .

Given $\varepsilon > 0$, $n \in \mathbb{N}$ and a point $x \in X$, the (n, ε) -ball centered at x with radius ε with respect to the metric d_n (also called Bowen's ball) is the set

$$B_n(x, \varepsilon) = \{y \in X: d_n(x, y) < \varepsilon\}. \quad (17)$$

3.2. Separated sets, spanning sets, open covers, closed covers. Given $\varepsilon > 0$ and $n \in \mathbb{N}$, the next three quantities count the number of orbit segments of length n that are distinguishable at scale ε .

One says that a set $E \subset X$ is (n, ε) -separated if two distinct points in E are at least ε apart in the metric d_n , that is, $d_n(x, z) > \varepsilon$ for every $x \neq z \in E$. We denote by $S(X, n, \varepsilon)$ the maximal cardinality of the (n, ε) -separated subsets of X . When $n = 1$, we simplify this notation by saying that the set E is ε -separated, and denoting the maximal cardinality of the ε -separated subsets of X by $S(X, d, \varepsilon)$. We note that, for every $\varepsilon > 0$ and $n \in \mathbb{N}$,

$$S(X, n, \varepsilon) = S(X, d_n, \varepsilon).$$

A set $Y \subset X$ is said to be (n, ε) -spanning if for any $x \in X$ there exists $z \in Y$ such that $d_n(x, z) < \varepsilon$. We denote by $R(X, n, \varepsilon)$ the minimal cardinality of the (n, ε) -spanning subsets of X . Again, when $n = 1$, we refer to ε -spanning sets, with minimal cardinality denoted $R(X, d, \varepsilon)$, and one has

$$R(X, n, \varepsilon) = R(X, d_n, \varepsilon) \quad \forall \varepsilon > 0 \quad \forall n \in \mathbb{N}.$$

Given $\varepsilon > 0$ and $n \in \mathbb{N}$, let $\mathcal{N}(X, n, \varepsilon)$ be the minimum cardinality of any cover of X by open balls with radius ε in the metric d_n . When $n = 1$, we refer to $\mathcal{N}(X, d, \varepsilon)$, so

$$\mathcal{N}(X, n, \varepsilon) = \mathcal{N}(X, d_n, \varepsilon) \quad \forall \varepsilon > 0 \quad \forall n \in \mathbb{N}.$$

In what follows, if we need to use closed balls instead of open ones, we change the notation to $\mathcal{C}(X, d, \varepsilon)$. We note that, given $\varepsilon > 0$,

$$\mathcal{N}(X, d, 2\varepsilon) \leq \mathcal{C}(X, d, \varepsilon) \leq \mathcal{N}(X, d, \varepsilon). \quad (18)$$

Due to the compactness of X , the values of $S(X, n, \varepsilon)$, $R(X, n, \varepsilon)$, $\mathcal{N}(X, n, \varepsilon)$ and $\mathcal{C}(X, n, \varepsilon)$ are finite for every $n \in \mathbb{N}$ and $\varepsilon > 0$. In addition, the previous three quantities are related by the following inequalities (cf. [19, §7]):

$$\mathcal{N}(X, n, 2\varepsilon) \leq R(X, n, \varepsilon) \leq S(X, n, \varepsilon) \leq R(X, n, \varepsilon/2) \leq \mathcal{N}(X, n, \varepsilon/2). \quad (19)$$

Moreover,

$$\mathcal{N}(X, n + m, \varepsilon) \leq \mathcal{N}(X, n, \varepsilon) \cdot \mathcal{N}(X, m, \varepsilon) \quad (20)$$

so, the sequence $(\log \mathcal{N}(X, n, \varepsilon))_{n \in \mathbb{N}}$ is sub-additive.

3.3. Box-counting dimension. Given $\varepsilon > 0$, recall that $\mathcal{N}(X, d, \varepsilon)$ stands for the smallest number of open balls in the metric d with radius ε that cover X . We observe that, for every $\varepsilon > 0$, one has

$$S(X, d, 2\varepsilon) \leq \mathcal{N}(X, d, \varepsilon) \leq S(X, d, \varepsilon/2). \quad (21)$$

The upper box-counting dimension of (X, d) is defined by

$$\overline{\dim}_{\text{B}}(X, d) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\log \mathcal{N}(X, d, \varepsilon)}{\log(1/\varepsilon)}.$$

The lower box-counting dimension $\underline{\dim}_{\text{B}}(X, d)$ is defined by taking \liminf instead of \limsup . Notice that, due to (21), the upper box-counting dimension may be computed using ε -separated sets, that is,

$$\overline{\dim}_{\text{B}}(X, d) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\log S(X, d, \varepsilon)}{\log(1/\varepsilon)}$$

(and similarly for the lower version).

3.4. Metric mean dimension. The *upper and lower metric mean dimensions* are labels for dynamical systems introduced by E. Lindenstrauss and B. Weiss in [7] to quantify the complexity of infinite entropy systems. We denote them by $\overline{\text{mdim}}_{\text{M}}(X, d, T)$ and $\underline{\text{mdim}}_{\text{M}}(X, d, T)$, respectively, to emphasize their dependence on the fixed metric d of the space X where the dynamics T acts. They are defined by

$$\begin{aligned}\overline{\text{mdim}}_{\text{M}}(X, d, T) &= \limsup_{\varepsilon \rightarrow 0^+} \frac{\limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon)}{\log(1/\varepsilon)} \\ \underline{\text{mdim}}_{\text{M}}(X, d, T) &= \liminf_{\varepsilon \rightarrow 0^+} \frac{\limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon)}{\log(1/\varepsilon)}.\end{aligned}$$

Due to (19), the metric mean dimension may be computed using $R(X, n, \varepsilon)$ or $\mathcal{N}(X, n, \varepsilon)$ instead of $S(X, n, \varepsilon)$. If $\overline{\text{mdim}}_{\text{M}}(X, d, T)$ and $\underline{\text{mdim}}_{\text{M}}(X, d, T)$ agree, then the common value is denoted by $\text{mdim}_{\text{M}}(X, d, T)$ and called metric mean dimension.

In what follows, we simplify the notation by denoting

$$h_{\varepsilon}(X, d, T) = \limsup_{n \rightarrow +\infty} \frac{1}{n} \log \mathcal{Z}(X, n, \varepsilon) \quad (22)$$

where \mathcal{Z} is chosen from the set $\{\mathcal{N}, R, S\}$ according to the context.

3.5. W_p and LP metrics on $\mathcal{P}(X)$. It is known that the space $\mathcal{P}(X)$ of the Borel probability measures on X is nonempty; and it is compact if endowed with the weak*-topology. Moreover, there are metrics on $\mathcal{P}(X)$ inducing this topology, the classic ones being the *Wasserstein distances* and the *Lévy-Prokhorov distance*. The former are defined by

$$W_p(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \left(\int_{X \times X} (d(x, y))^p d\pi(x, y) \right)^{1/p}$$

where $p \in [1, +\infty)$ and $\Pi(\mu, \nu)$ denotes the set of probability measures on the product space $X \times X$ with marginals μ and ν (see [16] and references therein for more details). The latter is defined by

$$LP(\mu, \nu) = \inf \left\{ \varepsilon > 0: \forall E \subset X \quad \forall \varepsilon\text{-neighborhood } V_{\varepsilon}(E) \text{ of } E \text{ one has} \right. \\ \left. \nu(E) \leq \mu(V_{\varepsilon}(E)) + \varepsilon \quad \text{and} \quad \mu(E) \leq \nu(V_{\varepsilon}(E)) + \varepsilon \right\}.$$

These distances are compared by (cf. [1, p. 11])

$$\begin{aligned}W_q &\leq W_p \leq \text{diam}(X, d)^{1-\frac{q}{p}} W_q^{\frac{q}{p}} & \forall 1 \leq q \leq p \\ LP^{1+\frac{1}{p}} &\leq W_p \leq (1 + \text{diam}(X, d)^p)^{\frac{1}{p}} LP^{\frac{1}{p}} & \forall p \in [1, +\infty).\end{aligned} \quad (23)$$

We refer the reader to [3, 16] where more information can be found on these metrics.

4. MEASURE-THEORETIC ε -ENTROPIES

Let (X, d) be a compact metric space, $T: X \rightarrow X$ a homeomorphism and $\mu \in \mathcal{P}_T(X)$. Recall from Subsection 3.1 the definitions of dynamical metric d_n and (n, ε) -ball $B_n(x, \varepsilon)$.

4.1. (ε, δ) -**Katok entropy**. For $0 < \delta < 1$, $n \in \mathbb{N}$ and $\varepsilon > 0$, we denote by $\mathcal{N}_\mu(n, \varepsilon, \delta)$ the minimum number of open (n, ε) -balls with radius ε in the dynamical metric d_n needed to cover any set of μ -measure strictly bigger than $1 - \delta$.

Definition 4.1. The upper and lower (ε, δ) -Katok entropy of μ are given, respectively, by

$$\bar{h}_\mu^K(\varepsilon, \delta) = \limsup_{n \rightarrow +\infty} \frac{\log \mathcal{N}_\mu(n, \varepsilon, \delta)}{n} \quad \text{and} \quad \underline{h}_\mu^K(\varepsilon, \delta) = \liminf_{n \rightarrow +\infty} \frac{\log \mathcal{N}_\mu(n, \varepsilon, \delta)}{n}. \quad (24)$$

The upper and lower Katok entropy of μ at scale ε are defined by

$$\begin{aligned} \bar{h}_\mu^K(\varepsilon) &= \lim_{\delta \rightarrow 0^+} \bar{h}_\mu^K(\varepsilon, \delta) = \sup_{\delta \rightarrow 0^+} \bar{h}_\mu^K(\varepsilon, \delta) \\ \underline{h}_\mu^K(\varepsilon) &= \lim_{\delta \rightarrow 0^+} \underline{h}_\mu^K(\varepsilon, \delta) = \sup_{\delta \rightarrow 0^+} \underline{h}_\mu^K(\varepsilon, \delta). \end{aligned} \quad (25)$$

Remark 4.2. In the literature, we can find an alternative definition of Katok entropy, whose values for ergodic measures are as above but the extension to invariant probabilities is done by integrating over their ergodic decompositions. In particular, the resulting entropy-like map is convex. Thus, the corresponding measure-theoretic metric mean dimension does not satisfy a classical variational principle (cf. Proposition 7.2).

4.2. (ε, δ) -**Shapira entropy**. Given a finite open cover \mathcal{U} of X , the diameter of \mathcal{U} , which we denote by $\text{diam}(\mathcal{U}, d)$, is given by $\max_{U \in \mathcal{U}} \text{diam}(U, d)$, where the diameter $\text{diam}(A, d)$ of a set $A \subset X$ is the supremum of the distances between pairs of points in A .

Take $\mu \in \mathcal{P}_T(X)$. For $0 < \delta < 1$ and a finite open cover \mathcal{U} of X , we denote by $\mathcal{N}_\mu(\mathcal{U}, \delta)$ the minimum number of elements of \mathcal{U} whose union has a μ -measure strictly greater than $1 - \delta$. Set

$$\bar{h}_\mu^S(\mathcal{U}, \delta) = \limsup_{n \rightarrow +\infty} \frac{\log \mathcal{N}_\mu(\mathcal{U}^n, \delta)}{n} \quad \text{and} \quad \underline{h}_\mu^S(\mathcal{U}, \delta) = \liminf_{n \rightarrow +\infty} \frac{\log \mathcal{N}_\mu(\mathcal{U}^n, \delta)}{n}$$

where $\mathcal{U}^n = \bigvee_{j=0}^{n-1} T^{-j}(\mathcal{U})$.

U. Shapira proved in [11] that, if μ is ergodic, then the above limits exist and are independent of δ . However, since we are interested in all invariant measures and their interplay with δ , we keep it in the notation.

Definition 4.3. The upper and lower (ε, δ) -Shapira entropies of μ are given by

$$\bar{h}_\mu^S(\varepsilon, \delta) = \inf_{\text{diam}(\mathcal{U}, d) \leq \varepsilon} \bar{h}_\mu^S(\mathcal{U}, \delta) \quad \underline{h}_\mu^S(\varepsilon, \delta) = \inf_{\text{diam}(\mathcal{U}, d) \leq \varepsilon} \underline{h}_\mu^S(\mathcal{U}, \delta). \quad (26)$$

The next lemma describes the connection between the Katok and Shapira entropies.

Lemma 4.4. For every $\mu \in \mathcal{P}_T(X)$, $\varepsilon > 0$ and $0 < \delta < 1$,

$$\bar{h}_\mu^K(\varepsilon, \delta) \leq \bar{h}_\mu^S(\varepsilon, \delta) \leq \bar{h}_\mu^K(\varepsilon/4, \delta) \leq \bar{h}_\mu^K(\varepsilon/4, \delta/4).$$

and

$$\underline{h}_\mu^K(\varepsilon, \delta) \leq \underline{h}_\mu^S(\varepsilon, \delta) \leq \underline{h}_\mu^K(\varepsilon/4, \delta) \leq \underline{h}_\mu^K(\varepsilon/4, \delta/4).$$

Proof. We prove the inequalities in the first line; a similar reasoning shows the others. Given $\mu \in \mathcal{P}_T(X)$, $\varepsilon > 0$ and $0 < \delta < 1$, fix $n \in \mathbb{N}$ and \mathcal{U} such that $\text{diam}(\mathcal{U}, d) \leq \varepsilon$. For each nonempty $U \in \mathcal{U}^n$, take $x_U \in U$. Then $U \subset B_n(x_U, \varepsilon)$ and

$$\mathcal{N}_\mu(n, \varepsilon, \delta) \leq \min \{ \mathcal{F} \subset \mathcal{U}^n : \mu(\bigcup_{U \in \mathcal{F}} B_n(x_U, \varepsilon)) > 1 - \delta \} \leq \mathcal{N}_\mu(\mathcal{U}^n, \delta).$$

Thus,

$$\bar{h}_\mu^K(\varepsilon, \delta) \leq \bar{h}_\mu^S(\mathcal{U}, \delta)$$

and the first inequality follows by taking the infimum over \mathcal{U} .

Regarding the remaining inequalities, let \mathcal{U} be a cover of X satisfying $\text{diam}(\mathcal{U}, d) \leq \varepsilon$ and $\text{Leb}(\mathcal{U}) \geq \varepsilon/4$ (which always exists, cf. [5, Lemma 3.4]). Since every Bowen's ball $B_n(x, \varepsilon/4)$ is contained in some element of \mathcal{U}^n , we have $\mathcal{N}_\mu(\mathcal{U}^n, \delta) \leq \mathcal{N}_\mu(n, \varepsilon/4, \delta)$. Thus,

$$\bar{h}_\mu^S(\varepsilon, \delta) \leq \bar{h}_\mu^S(\mathcal{U}, \delta) \leq \bar{h}_\mu^K(\varepsilon/4, \delta) \leq \bar{h}_\mu^K(\varepsilon/4, \delta/4).$$

□

4.3. Brin-Katok ε -entropy. Given $\mu \in \mathcal{P}_T(X)$, this notion of complexity is based on the average decay rate of the size, measured by μ , of a dynamical ball in X .

Definition 4.5. For $x \in X$ and $\varepsilon > 0$, the upper and lower Brin-Katok entropies of $\mu \in \mathcal{P}_T(X)$ at x with respect to the scale ε are given by

$$\bar{h}_\mu^{BK}(\varepsilon, x) = \limsup_{n \rightarrow +\infty} -\frac{\log \mu(B_n(x, \varepsilon))}{n} \quad \text{and} \quad \underline{h}_\mu^{BK}(\varepsilon, x) = \liminf_{n \rightarrow +\infty} -\frac{\log \mu(B_n(x, \varepsilon))}{n}$$

and the upper and lower Brin-Katok entropies of μ at scale ε are given, respectively, by

$$\bar{h}_\mu^{BK}(\varepsilon) = \int \bar{h}_\mu^{BK}(\varepsilon, x) d\mu(x) \quad \text{and} \quad \underline{h}_\mu^{BK}(\varepsilon) = \int \underline{h}_\mu^{BK}(\varepsilon, x) d\mu(x).$$

The following two properties will be useful later on.

Lemma 4.6. Let $\mu \in \mathcal{P}_T(X)$ and $\varepsilon > 0$. Then

$$\bar{h}_\mu^{BK}(\varepsilon, x) \leq \limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon/2) \quad \text{at } \mu\text{-a.e. } x \in X.$$

In particular,

$$\bar{h}_\mu^{BK}(\varepsilon) \leq \limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon/2).$$

Proof. To simplify notation, set

$$h_{\varepsilon/2} = \limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon/2).$$

For each $\gamma > 0$, let

$$A_\gamma = \{x \in X : \bar{h}_\mu^{BK}(\varepsilon, x) > h_{\varepsilon/2} + \gamma\}.$$

Claim: $\mu(A_\gamma) = 0$ for every $\gamma > 0$.

Proof of the Claim: Fix $\gamma > 0$. For each $n \in \mathbb{N}$, let

$$\begin{aligned} A_\gamma^{(n)} &= \{x \in X : -\frac{1}{n} \log \mu(B_n(x, \varepsilon)) > h_{\varepsilon/2} + \gamma\} \\ &= \{x \in X : \mu(B_n(x, \varepsilon)) < e^{-n(h_{\varepsilon/2} + \gamma)}\}. \end{aligned}$$

Take $n_\gamma \in \mathbb{N}$ such that

$$S(X, n, \varepsilon/2) \leq e^{n(h_{\varepsilon/2} + \frac{\gamma}{2})} \quad \forall n \geq n_\gamma.$$

Hence,

$$A_\gamma = \bigcap_{N \geq n_\gamma} \bigcup_{n \geq N} A_\gamma^{(n)}.$$

Fix $N \geq n_\gamma$. For each $n \geq N$, let E_n be a maximal $(n, \varepsilon/2)$ -separated subset of X . Then

$$A_\gamma^{(n)} \subset \bigcup_{y \in E_n} B_n(y, \varepsilon/2) \quad \text{and} \quad \#E_n \leq e^{n(h_{\varepsilon/2} + \frac{\gamma}{2})}.$$

For each $x \in A_\gamma^{(n)}$, take $y \in E_n$ satisfying $x \in B_n(y, \varepsilon/2)$. Then $B_n(y, \varepsilon/2) \subset B_n(x, \varepsilon)$, which implies that $\mu(B_n(y, \varepsilon/2)) < e^{-n(h_{\varepsilon/2} + \gamma)}$. Define

$$E'_n = \left\{ y \in E_n : \mu(B_n(y, \varepsilon/2)) < e^{-n(h_{\varepsilon/2} + \gamma)} \right\}.$$

Then $A_\gamma^{(n)} \subset \bigcup_{y \in E'_n} B_n(y, \varepsilon/2)$ and we get

$$\mu(A_\gamma^{(n)}) \leq \sum_{y \in E'_n} \mu(B_n(y, \varepsilon/2)) \leq \#E'_n e^{-n(h_{\varepsilon/2} + \gamma)} \leq \#E_n e^{-n(h_{\varepsilon/2} + \gamma)} \leq e^{-\frac{n\gamma}{2}}.$$

Thus,

$$\mu(A_\gamma) \leq \mu\left(\bigcup_{n \geq N} A_\gamma^{(n)}\right) \leq \sum_{n \geq N} e^{-\frac{n\gamma}{2}} = \frac{e^{-\frac{N\gamma}{2}}}{1 - e^{-\frac{\gamma}{2}}} \quad \forall N \geq n_\gamma.$$

We finish the proof of the claim by taking $N \rightarrow +\infty$. \square

Now, Lemma 4.6 is a straightforward consequence of the claim since, given $\mu \in \mathcal{P}_T(X)$ and $\varepsilon > 0$,

$$X \setminus \left\{ x \in X : \bar{h}_\mu^{BK}(\varepsilon, x) \leq \limsup_{n \rightarrow +\infty} \frac{1}{n} \log S(X, n, \varepsilon/2) \right\} \subseteq \bigcup_{\gamma \in \mathbb{Q} \cap (0, +\infty)} A_\gamma. \quad \square$$

Lemma 4.7. *For every $\varepsilon > 0$, the maps*

$$\mu \in \mathcal{P}_T(X) \mapsto \bar{h}_\mu^{BK}(\varepsilon) \quad \text{and} \quad \mu \in \mathcal{P}_T(X) \mapsto \underline{h}_\mu^{BK}(\varepsilon)$$

are convex.

Proof. We prove the assertion for the upper Brin-Katok entropy; the reasoning concerning the other map is similar. Fix $\varepsilon > 0$. Let $(\mu_k)_{k \in \mathbb{N}}$ be a sequence of invariant probability measures and $(t_k)_{k \in \mathbb{N}}$ be a sequence of positive real numbers such that $\sum_{k \in \mathbb{N}} t_k = 1$. Consider $\mu = \sum_{k \in \mathbb{N}} t_k \mu_k \in \mathcal{P}_T(X)$. For each $k \in \mathbb{N}$, $\mu \geq t_k \mu_k$. Therefore, for all $x \in X$ and $n \in \mathbb{N}$,

$$\log \mu(B_n(x, \varepsilon)) \geq \log \mu_k(B_n(x, \varepsilon)) + \log t_k$$

and so

$$\bar{h}_\mu^{BK}(\varepsilon, x) \leq \bar{h}_{\mu_k}^{BK}(\varepsilon, x) \quad \forall x \in X.$$

We finish the proof by integrating both sides of the previous inequality with respect to μ_k , multiplying both sides by t_k and adding up over $k \in \mathbb{N}$. \square

4.4. Rényi entropy. Denote by P any finite Borel measurable partition of X , by $\text{diam}(P)$ its diameter, defined by $\max_{A \in P} \text{diam } A$ and by P^n the refined partition $\bigvee_{j=0}^{n-1} T^{-j}(P)$. Given $\mu \in \mathcal{P}_T(X)$, the measure-theoretic entropy of μ with respect to P is defined by

$$h_\mu(P) = \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(P^n)$$

where $H_\mu(Q) = -\sum_{Q \in Q} \mu(Q) \log \mu(Q)$ stands for the Shannon entropy.

Definition 4.8. Given $\varepsilon > 0$, the Rényi ε -entropy of μ at scale ε is defined by

$$\inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P)$$

where the infimum is taken over all finite Borel measurable partitions of X with diameter at most ε .

We refer the reader to [5, Section 3] for more information on this concept.

5. PROOF OF THEOREM B

Assume that $F: \mathcal{K} \times (0, +\infty) \rightarrow [0, +\infty]$ satisfies the s.m.m.d. property and let $C > 0$ be the constant provided by Definition 2.5.

Clearly, for every $\mu \in \mathcal{K}$, one has

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)} \geq \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

and so

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)} \geq \sup_{\mu \in \mathcal{K}} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}. \quad (27)$$

Regarding the converse inequality, set

$$\mathcal{L} = \limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

If $\mathcal{L} = 0$, then (27) already ensures that

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)} = \sup_{\mu \in \mathcal{K}} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

If $\mathcal{L} > 0$, let $(\varepsilon'_\ell)_{\ell \in \mathbb{N}}$ be a sequence of scales converging to zero such that

$$\mathcal{L} = \lim_{\ell \rightarrow +\infty} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon'_\ell)}{\log(1/\varepsilon'_\ell)}.$$

Taking a subsequence if necessary, we can assume that

$$\varepsilon'_\ell < 2^{-\ell} \quad \forall \ell \in \mathbb{N}.$$

Fix a sequence $(t_\ell)_{\ell \in \mathbb{N}}$ of positive real numbers with polynomial decay to zero as ℓ goes to $+\infty$ and such that $\sum_{\ell \in \mathbb{N}} t_\ell = 1$. For example, we may choose $t_\ell = 6/(\pi \ell)^2$ for every $\ell \in \mathbb{N}$. Then, since $\varepsilon'_\ell < 2^{-\ell}$ for every $\ell \in \mathbb{N}$, the sequence $\varepsilon_\ell = t_\ell \varepsilon'_\ell / C$ satisfies

$$\lim_{\ell \rightarrow +\infty} \frac{\log(1/\varepsilon_\ell)}{\log(1/\varepsilon'_\ell)} = 1$$

and so

$$\mathcal{L} = \lim_{\ell \rightarrow +\infty} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon'_\ell)}{\log(1/\varepsilon'_\ell)} = \lim_{\ell \rightarrow +\infty} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, C\varepsilon_\ell/t_\ell)}{\log(1/\varepsilon_\ell)}.$$

Case 1: $\mathcal{L} < +\infty$.

For every $N \in \mathbb{N}$, there exists $L(N) \in \mathbb{N}$ such that

$$\sup_{\mu \in \mathcal{K}} F(\mu, C\varepsilon_\ell/t_\ell) > \left(1 - \frac{1}{N}\right) \cdot \mathcal{L} \cdot \log(1/\varepsilon_\ell) \quad \forall \ell \geq L(N). \quad (28)$$

Without loss of generality, we may assume that

$$L(N+1) > L(N) \quad \forall N \in \mathbb{N}.$$

Consider the subsequence $(\varepsilon_{L(N)})_{N \in \mathbb{N}}$ of $(\varepsilon_\ell)_{\ell \in \mathbb{N}}$ and a sequence of measures $(\mu_N)_{N \in \mathbb{N}}$ in \mathcal{K} such that

$$F(\mu_N, C\varepsilon_{L(N)}/t_{L(N)}) > \left(1 - \frac{1}{N}\right) \cdot \mathcal{L} \cdot \log(1/\varepsilon_{L(N)}) \quad \forall N \in \mathbb{N}. \quad (29)$$

Let $\mu_0 \in \mathcal{K}$ be defined by

$$\mu_0 = \beta \cdot \sum_{N \in \mathbb{N}} t_{L(N)} \mu_N$$

where β is a normalization constant that is greater than or equal to 1 since each natural number appears at most once in the sequence $(L(N))_{N \in \mathbb{N}}$. Then

$$\mu_0 \geq t_{L(N)} \mu_N \quad \forall N \in \mathbb{N}$$

so, by the s.m.m.d. property in \mathcal{K} and the inequality (29), we get

$$F(\mu_0, \varepsilon_{L(N)}) \geq F(\mu_N, C\varepsilon_{L(N)}/t_{L(N)}) > \left(1 - \frac{1}{N}\right) \cdot \mathcal{L} \cdot \log(1/\varepsilon_{L(N)}) \quad \forall N \in \mathbb{N}.$$

By dividing the previous estimates by $\log(1/\varepsilon_{L(N)})$ and taking $N \rightarrow +\infty$, we obtain

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu_0, \varepsilon)}{\log(1/\varepsilon)} \geq \limsup_{N \rightarrow +\infty} \frac{F(\mu_0, \varepsilon_{L(N)})}{\log(1/\varepsilon_{L(N)})} \geq \mathcal{L}. \quad (30)$$

Bringing together (27) and (30), we conclude that

$$\mathcal{L} = \max_{\mu \in \mathcal{K}} \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}.$$

Case 2: $\mathcal{L} = +\infty$.

The proof in this case proceeds in a similar way after replacing $(1 - \frac{1}{N}) \cdot \mathcal{L}$ on the right-hand side of (28) by N .

Assuming that F is non-increasing in ε , the limits

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \quad \text{and} \quad \limsup_{\varepsilon \rightarrow 0^+} \frac{\sup_{\mu \in \mathcal{K}} F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

are attained at any sequence $\varepsilon_k \rightarrow 0^+$ satisfying

$$\lim_{k \rightarrow +\infty} \frac{\log \varepsilon_{k+1}}{\log \varepsilon_k} = 1.$$

The equality (13) is deduced by an analogous reasoning.

The strong convex geometry of \mathcal{K}_{max} described in (c) is an immediate consequence of (11).

6. PROOF OF THEOREM A

This section is divided into two parts, corresponding to the choice of the Wasserstein or Lévy-Prokhorov distance to metrize the weak*-topology.

6.1. W_p -mean quantization dimension. Definition 2.1 admits the following reformulation when D is a p -Wasserstein metric for some $p \in [1, +\infty)$.

Lemma 6.1. [1, Proposition 3.2] *The quantization number $Q_{\mu, W_p}(\varepsilon)$ for the metric W_p is the minimal cardinality of a set $F \subset X$ satisfying*

$$\int d(x, F)^p d\mu(x) \leq \varepsilon^p$$

where $d(x, F) = \min_{y \in F} d(x, y)$.

The next lemma relates the Katok entropies to the dynamical quantization numbers with respect to the Wasserstein distances.

Lemma 6.2. *Let (X, d) be a compact metric space and $T: X \rightarrow X$ be a continuous map. For every $\mu \in \mathcal{P}_T(X)$ and $p \in [1, +\infty)$,*

$$\begin{aligned} \bar{h}_\mu^K(2\varepsilon, 3/4) &\leq \bar{Q}_\mu(T, W_1, \varepsilon) \leq \bar{Q}_\mu(T, W_p, \varepsilon) \leq \bar{h}_\mu^K(2^{-1/p}\varepsilon) \\ \underline{h}_\mu^K(2\varepsilon, 3/4) &\leq \underline{Q}_\mu(T, W_1, \varepsilon) \leq \underline{Q}_\mu(T, W_p, \varepsilon) \leq \underline{h}_\mu^K(2^{-1/p}\varepsilon). \end{aligned}$$

Proof. We shall prove the statement for the upper limits, starting with the first inequality. Given $n \in \mathbb{N}$, let $F \subset X$ be such that

$$\#F = Q_{\mu, W_{1, n}(\varepsilon)} \quad \text{and} \quad \int d_n(x, F) d\mu(x) \leq \varepsilon.$$

By Markov's inequality ,

$$\mu(d_n(\cdot, F) \geq 2\varepsilon) \leq \frac{1}{2\varepsilon} \int d_n(x, F) d\mu(x) \leq \frac{1}{2}.$$

Thus,

$$\mu\left(\bigcup_{y \in F} B_n(y, 2\varepsilon)\right) \geq 1/2 > 1 - 3/4$$

and so

$$Q_{\mu, W_{1, n}(\varepsilon)} \geq \mathcal{N}_\mu(n, 2\varepsilon, 3/4).$$

We conclude the argument by applying the logarithm to both sides of this inequality, dividing by n and taking $n \rightarrow +\infty$.

The second inequality is immediate since the map $p \mapsto W_p$ is non-decreasing (cf. [16, p. 210]) or (23)). Regarding the third inequality, fix

$$0 < \delta < \frac{1}{2} \left(\frac{\varepsilon}{\text{diam}(X, d)} \right)^p$$

and

$$s > \bar{h}_\mu^K(2^{-1/p}\varepsilon, \delta).$$

Then, for sufficiently large $n \in \mathbb{N}$, we have

$$\mathcal{N}_\mu(n, 2^{-1/p}\varepsilon, \delta) < e^{ns}.$$

Take $F \subset X$ such that

$$\mu\left(\bigcup_{y \in F} B_n(y, 2^{-1/p}\varepsilon)\right) > 1 - \delta \quad \text{and} \quad \#F < e^{ns}.$$

Thus,

$$\begin{aligned} \int d_n(\cdot, F)^p d\mu &\leq \int \bigcup_{y \in F} B_n(y, 2^{-1/p}\varepsilon) d_n(\cdot, F)^p d\mu + \int_{X \setminus \bigcup_{y \in F} B_n(y, 2^{-1/p}\varepsilon)} d_n(\cdot, F)^p d\mu \\ &\leq \frac{\varepsilon^p}{2} \mu\left(\bigcup_{y \in F} B_n(y, 2^{-1/p}\varepsilon)\right) + \delta \cdot \text{diam}(X, d_n)^p \\ &\leq \frac{\varepsilon^p}{2} + \delta \cdot \text{diam}(X, d)^p \leq \varepsilon^p. \end{aligned}$$

Thus,

$$Q_{\mu, D_n}(\varepsilon) \leq \#F < e^{ns}$$

and so

$$\overline{Q}_\mu(T, W_p, \varepsilon) \leq s.$$

To finish the proof, we just have to make $s \rightarrow \overline{h}_\mu^K(2^{-1/p}\varepsilon, \delta)$ and take the limit as $\delta \rightarrow 0^+$. \square

By (3), (4) and the fact that

$$\left\{ \overline{h}_\mu^K(2\varepsilon, 3/4), \overline{h}_\mu^K(\varepsilon), \underline{h}_\mu^K(2\varepsilon, 3/4), \underline{h}_\mu^K(\varepsilon) \right\} \subset \mathcal{D}$$

it is easy to see that Lemma 6.2 yields:

(i) For every $\mu \in \mathcal{E}_T(X)$, $F \in \mathcal{D}$ and $p \in [1, +\infty)$,

$$\overline{\text{mdim}}_Q(X, d, T, W_p, \mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

and, for every $p \in [1, +\infty)$,

$$\overline{\text{mdim}}_M(X, d, T) = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\overline{Q}_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{Q_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)}.$$

Combining the previous equalities with (7), we obtain:

(ii) For every $p \in [1, +\infty)$,

$$\begin{aligned} \overline{\text{mdim}}_M(X, d, T) &= \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{\overline{Q}_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{Q_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)} \\ &= \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{\overline{Q}_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \sup_{\mu \in \mathcal{P}_T(X)} \frac{\overline{Q}_\mu(T, W_p, \varepsilon)}{\log(1/\varepsilon)}. \end{aligned}$$

To finish the proof of Theorem A, we need further information.

Lemma 6.3. [1, Lemma 3.18] *For every $\mu, \nu \in \mathcal{P}(X)$ such that $\mu \geq t\nu$ for some $t > 0$, one has*

$$Q_{\mu, W_1}(t\varepsilon) \geq Q_{\nu, W_1}(\varepsilon) \quad \forall \varepsilon > 0.$$

Applying the previous lemma to (X, d_n) , we obtain

$$\begin{aligned} \overline{Q}_\mu(T, W_1, t\varepsilon) &\geq \overline{Q}_\nu(T, W_1, \varepsilon) \\ \underline{Q}_\mu(T, W_1, t\varepsilon) &\geq \underline{Q}_\nu(T, W_1, \varepsilon) \end{aligned}$$

whenever $\mu \geq t\nu$ for some $t > 0$. This means that the W_1 -dynamical quantization numbers satisfy the s.m.m.d. property.

Since the map $p \in [1, +\infty) \mapsto W_p$ is non-decreasing (see (23)), combining Corollary 2.7 with (9) we conclude that, for every $p \in [1, +\infty)$,

$$\begin{aligned} \overline{\text{mdim}}_M(X, d, T) &= \max_{\mu \in \mathcal{P}_T(X)} \overline{\text{mdim}}_Q(X, d, T, W_1, \mu) \\ &\leq \max_{\mu \in \mathcal{P}_T(X)} \overline{\text{mdim}}_Q(X, d, T, W_p, \mu) \\ &\leq \overline{\text{mdim}}_M(X, d, T). \end{aligned} \tag{31}$$

6.2. LP-mean quantization dimension. Definition 2.1 admits a reformulation when D is the Lévy-Prokhorov metric.

Lemma 6.4. [1, Proposition 3.3] *The quantization number $Q_{\mu, LP}(\varepsilon)$ for the LP metric is the least number of closed balls with radius ε that cover any subset of X with μ measure at least $1 - \varepsilon$.*

Observe that if we consider the space (X, d_n) , the above formulation resembles the cover number $\mathcal{N}_\mu(n, \varepsilon, \delta)$ from the definition of the (ε, δ) -Katok entropy when $\delta = \varepsilon$, the only difference being that the balls in the covers are closed. Therefore:

Lemma 6.5. *For every $\mu \in \mathcal{P}_T(X)$ and $\varepsilon > 0$,*

$$\begin{aligned} \overline{h}_\mu^K(2\varepsilon, \varepsilon) &\leq \overline{Q}_\mu(T, LP, \varepsilon) \leq \overline{h}_\mu^K(\varepsilon, \varepsilon) \\ \underline{h}_\mu^K(2\varepsilon, \varepsilon) &\leq \underline{Q}_\mu(T, LP, \varepsilon) \leq \underline{h}_\mu^K(\varepsilon, \varepsilon). \end{aligned}$$

In particular, for every $0 < \varepsilon < 1/2$, we have

$$\begin{aligned} \overline{h}_\mu^K(2\varepsilon, 1/2) &\leq \overline{Q}_\mu(T, LP, \varepsilon) \leq \overline{h}_\mu^K(\varepsilon) \\ \underline{h}_\mu^K(2\varepsilon, 1/2) &\leq \underline{Q}_\mu(T, LP, \varepsilon) \leq \underline{h}_\mu^K(\varepsilon) \end{aligned}$$

and by a reasoning similar to that in the previous subsection, we prove items (a) and (b) of Theorem A when $D = LP$.

To conclude, recall that, since the metric LP is at least a uniform constant factor of W_1 (see (23)), we have

$$\begin{aligned} \overline{\text{mdim}}_M(X, d, T) &= \max_{\mu \in \mathcal{P}_T(X)} \overline{\text{mdim}}_Q(X, d, T, W_1, \mu) \\ &\leq \max_{\mu \in \mathcal{P}_T(X)} \overline{\text{mdim}}_Q(X, d, T, LP, \mu) \\ &\leq \overline{\text{mdim}}_M(X, d, T). \end{aligned} \tag{32}$$

The property (d) of Theorem A is a straightforward consequence of the item (c) of Theorem B. The proof of Theorem A regarding the upper metric mean dimension is finished. A similar argument yields the corresponding statement for the lower metric mean dimension.

7. PROOF OF THEOREM C

Since the (ε, δ) -Katok entropy is non-increasing with respect to δ , Lemma 6.5 implies that, for every $\mu \in \mathcal{P}_T(X)$ and $\varepsilon > 0$,

$$\bar{h}_\mu^K(2\varepsilon, 2\varepsilon) \leq \bar{h}_\mu^K(2\varepsilon, \varepsilon) \leq \bar{Q}_\mu(T, d, LP, \varepsilon) \leq \bar{h}_\mu^K(\varepsilon, \varepsilon).$$

Combining this information with Lemma 4.4, we obtain, for every $\mu \in \mathcal{P}_T(X)$,

$$\overline{\text{mdim}}_Q(X, d, T, LP, \mu) = \limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_\mu^K(\varepsilon, \varepsilon)}{\log(1/\varepsilon)} = \limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_\mu^S(\varepsilon, \varepsilon)}{\log(1/\varepsilon)}$$

Remark 7.1. By definition, for a fixed $\delta > 0$, the (ε, δ) -Katok entropy satisfies

$$\bar{h}_\mu^K(\varepsilon, \delta) \geq \bar{h}_\nu^K(\varepsilon, \delta/t) \geq \bar{h}_\nu^K(\varepsilon/t, \delta/t)$$

whenever $\mu \geq t\nu$ for some $t \in (0, 1)$. In particular, $\bar{h}_\mu^K(\varepsilon, \varepsilon)$ has the s.m.m.d. property, and the map $\bar{h}_\mu^K(\varepsilon) = \lim_{\delta \rightarrow 0^+} \bar{h}_\mu^K(\varepsilon, \delta)$ satisfies the following stronger version of that property:

$$\bar{h}_\mu^K(\varepsilon) \geq \bar{h}_\nu^K(\varepsilon) \geq \bar{h}_\nu^K(\varepsilon/t).$$

The corresponding statements for the Shapira entropies also hold.

7.1. Example. Consider the interval $[0, 1]$ with the Euclidean distance, the space $[0, 1]^{\mathbb{Z}}$ endowed with the following distance, which is compatible with the product topology,

$$d((x_n)_{n \in \mathbb{Z}}, (y_n)_{n \in \mathbb{Z}}) = \sum_{n \in \mathbb{Z}} \frac{|x_n - y_n|}{2^{|n|}}$$

and the shift map $\sigma((x_n)_{n \in \mathbb{Z}}) = (x_{n+1})_{n \in \mathbb{Z}}$. We refer to a closed shift-invariant subset $Y \subset [0, 1]^{\mathbb{Z}}$ as a *subshift*.

Fix a sequence $a = (a_k)_{k \geq 0}$ of real numbers and a sequence $b = (b_k)_{k \geq 1}$ of positive integers that meet the following conditions:

- (C1) $0 = a_0 < a_1 < \dots < a_k \rightarrow 1$.
- (C2) $(a_k - a_{k-1})_k$ decreases to zero.
- (C3) $(b_k)_k$ is strictly increasing.

To simplify the notation, we write

$$\varepsilon_k = \frac{|a_k - a_{k-1}|}{b_k} \tag{33}$$

and, for each $k \in \mathbb{N}$,

$$J_k = \left\{ a_{k-1} + j \varepsilon_k : j \in \{0, \dots, b_k - 1\} \right\}.$$

Then,

$$J_i \cap J_j = \emptyset \quad \forall i \neq j \quad \text{and} \quad |x - y| \geq \varepsilon_k \quad \forall x, y \in J_k.$$

Denote $X_k = (J_k)^{\mathbb{Z}}$ and the subshift

$$X = \bigcup_{k=1}^{+\infty} X_k \cup \{1^{\mathbb{Z}}\} \subset [0, 1]^{\mathbb{Z}}.$$

By construction, every ergodic measure of the system (X, d, σ) is either the Dirac mass $\delta_{1^{\mathbb{Z}}}$ or gives full mass to X_k , for some $k \in \mathbb{N}$. Thus, by the ergodic decomposition theorem, we may write any invariant probability measure $\mu \in \mathcal{P}_\sigma(X)$ as

$$\mu = \sum_{k \in \mathbb{N}} t_k \mu_k + t_0 \delta_{\{1^{\mathbb{Z}}\}}, \quad (34)$$

where $t_0 = \mu(\{1^{\mathbb{Z}}\})$ and

$$\mu_k \in \mathcal{P}_\sigma(X), \quad \mu_k(X_k) = 1, \quad t_k = \mu(X_k) \quad \forall k \in \mathbb{N}.$$

We recall that, given $k \in \mathbb{N}$, the topological entropy of the shift map restricted to X_k is $\log b_k$. So, the topological entropy of (X, σ) is infinite and there exists an invariant probability measure with maximal entropy, but it is not ergodic.

In what follows, we assume that the parameters are chosen so that

$$\text{mdim}_M(X, d, \sigma) = 1 \quad \text{and} \quad \lim_{k \rightarrow +\infty} \frac{\log \varepsilon_k}{\log \varepsilon_{k+1}} = 1.$$

This is the case, for example, of $a_k = \sum_{n=1}^k 6/\pi n^2$ and $b_k = 3^k$.

7.2. Brin-Katok and Rényi entropies do not satisfy the variational principle (14). We start by relating the behavior of convex entropy-like functions and the fixed point $1^{\mathbb{Z}}$.

Proposition 7.2. *Consider the system (X, d, σ) and a map*

$$F: (\mu, \varepsilon) \in \mathcal{P}_\sigma(X) \times (0, +\infty) \mapsto F(\mu, \varepsilon).$$

Assume that:

(i) *The map*

$$\mu \in \mathcal{P}_\sigma(X) \mapsto \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)}$$

is convex.

(ii) *For every $\varepsilon > 0$ and $\mu \in \mathcal{P}_\sigma(X)$,*

$$F(\mu, \varepsilon) \leq h_\varepsilon(\text{supp}(\mu), d|_{\text{supp}(\mu)}, \sigma|_{\text{supp}(\mu)}).$$

Then

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} \leq \mu(\{1^{\mathbb{Z}}\}) \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\delta_{\{1^{\mathbb{Z}}\}}, \varepsilon)}{\log(1/\varepsilon)} \quad \forall \mu \in \mathcal{P}_\sigma(X). \quad (35)$$

Proof. Take $\mu \in \mathcal{P}_\sigma(X)$ and consider its decomposition (34), say

$$\mu = \sum_{k \in \mathbb{N}} t_k \mu_k + t_0 \delta_{\{1^{\mathbb{Z}}\}}.$$

Fix $K \in \mathbb{N}$. Then

$$\begin{aligned}
 \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu, \varepsilon)}{\log(1/\varepsilon)} &\leq \sum_{k=1}^K t_k \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu_k, \varepsilon)}{\log(1/\varepsilon)} + \sum_{k>K} t_k \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\mu_k, \varepsilon)}{\log(1/\varepsilon)} \\
 &\quad + t_0 \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\delta_{\{1^{\mathbb{Z}}\}}, \varepsilon)}{\log(1/\varepsilon)} \\
 &\leq \sum_{k=1}^K t_k \limsup_{\varepsilon \rightarrow 0^+} \frac{\log b_K}{\log(1/\varepsilon)} + \sum_{k>K} t_k \limsup_{\varepsilon \rightarrow 0^+} \frac{h_\varepsilon(X, d, \sigma)}{\log(1/\varepsilon)} \\
 &\quad + t_0 \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\delta_{\{1^{\mathbb{Z}}\}}, \varepsilon)}{\log(1/\varepsilon)} \\
 &= \mu\left(\bigcup_{k>K} J_k\right) \text{mdim}_M(X, d, \sigma) + t_0 \limsup_{\varepsilon \rightarrow 0^+} \frac{F(\delta_{\{1^{\mathbb{Z}}\}}, \varepsilon)}{\log(1/\varepsilon)}.
 \end{aligned}$$

Making $K \rightarrow +\infty$, we get (35) since $\lim_{K \rightarrow +\infty} \mu\left(\bigcup_{k>K} J_k\right) = 0$. \square

Taking into account that

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_{\delta_{\{1^{\mathbb{Z}}\}}}^{BK}(\varepsilon)}{\log(1/\varepsilon)} = 0$$

and using the properties of \bar{h}_μ^{BK} described in Lemmas 4.6 and 4.7, Proposition 7.2 yields:

Corollary 7.3. *For the system (X, d, σ) we have*

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_\mu^{BK}(\varepsilon)}{\log(1/\varepsilon)} = 0 \quad \forall \mu \in \mathcal{P}_\sigma(X).$$

Therefore, as $\overline{\text{mdim}}_M(X, d, \sigma) = 1$,

$$\overline{\text{mdim}}_M(X, d, \sigma) > \sup_{\mu \in \mathcal{P}_\sigma(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\bar{h}_\mu^{BK}(\varepsilon)}{\log(1/\varepsilon)}.$$

We have a similar statement for the Rényi entropy.

Corollary 7.4. *For every $\mu \in \mathcal{P}_\sigma(X)$,*

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P)}{\log(1/\varepsilon)} = 0.$$

Therefore, as $\overline{\text{mdim}}_M(X, d, \sigma) = 1$,

$$\overline{\text{mdim}}_M(X, d, \sigma) > \sup_{\mu \in \mathcal{P}_\sigma(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P)}{\log(1/\varepsilon)}.$$

Proof. We are left to show that the Rényi entropy satisfies the assumptions of Proposition 7.2. We start by proving that the map

$$\mu \in \mathcal{P}_\sigma(X) \longmapsto \limsup_{\varepsilon \rightarrow 0^+} \frac{\inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P)}{\log(1/\varepsilon)} \tag{36}$$

is convex. The next property was first established for ergodic measures in [5] and then extended to the invariant probability measures in [20].

Claim: [5, 20] For every $\mu \in \mathcal{P}_\sigma([0, 1]^\mathbb{Z})$,

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\inf_{\text{diam}(P) \leq \varepsilon} h_\mu(P)}{\log(1/\varepsilon)} = \limsup_{m \rightarrow +\infty} \frac{h_\mu(P_m)}{\log m} \quad (37)$$

where $m \in \mathbb{N}$,

$$P_m = \pi_0^{-1} \left\{ \left[\frac{i}{m}, \frac{i+1}{m} \right) : i = 0, \dots, m-1 \right\}$$

and $\pi_0: [0, 1]^\mathbb{Z} \rightarrow [0, 1]$ is the projection on the 0-coordinate.

In particular, the partitions considered on the limit on the right-hand side of (37) do not depend on the measure. Since, for each $m \in \mathbb{N}$, the map $\mu \mapsto h_\mu(P_m)$ is affine (cf. [19, Proof of Theorem 8.1]), and the lim sup of affine functions is convex, we conclude that the map (36) is convex.

The second assumption of Proposition 7.2 was established in [5, Section 3]. \square

7.3. (ε, δ) –Katok/Shapira entropy with fixed δ . Fix $\delta_0 \in (0, 1)$. Given $\mu \in \mathcal{P}_\sigma(X)$, there exists $K = K(\mu, \delta_0)$ such that $\mu(\bigcup_{k > K} J_k) < \delta_0$. Since σ restricted to $\bigcup_{k > K} J_k \cup \{1^\mathbb{Z}\}$ has a finite topological entropy, one has

$$\limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_\mu^K(\varepsilon, \delta_0)}{\log(1/\varepsilon)} = 0.$$

Therefore, since this equality is valid for every $\mu \in \mathcal{P}_\sigma(X)$, we conclude that

$$\overline{\text{mdim}}_M(X, d, T) > \sup_{\mu \in \mathcal{P}_T(X)} \limsup_{\varepsilon \rightarrow 0^+} \frac{\overline{h}_\mu^K(\varepsilon, \delta_0)}{\log(1/\varepsilon)}.$$

By Lemma 4.4, the above inequality also holds for the (ε, δ_0) -Shapira entropy.

Acknowledgments. Gustavo Pessil is grateful to Prof. Yonatan Gutman for his mentoring at IMPAN. The authors thank Adam Śpiewak, Chunlin Liu and Rui Yang for insightful comments on this work. Maria Carvalho was partially supported by CMUP, member of LASI, which is financed by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project UID/00144/2025, <https://doi.org/10.54499/UID/00144/2025>.

REFERENCES

- [1] P. Berger and J. Bochi. *On emergence and complexity of ergodic decompositions*. Adv. Math. 390 (2021), 107904. [2.1](#), [2.2](#), [2.3](#), [2.2](#), [3.5](#), [6.1](#), [6.3](#), [6.4](#)
- [2] E. Chen, R. Yang and X. Zhou. *Measure-theoretic metric mean dimension*. Studia Math. 280(1) (2025), 1–25. [1](#), [2.5](#)
- [3] A. L. Gibbs and F. E. Su. *On choosing and bounding probability metrics*. Int. Stat. Rev. 70(3) (2002), 419–435. [3.5](#)
- [4] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Lecture Notes in Mathematics 1730, Springer-Verlag, Berlin, 2000. [2.1](#)
- [5] Y. Gutman and A. Śpiewak. *Around the variational principle for metric mean dimension*. Studia Math. 261(3) (2021), 345–360. [1](#), [4.2](#), [4.4](#), [7.2](#), [7.2](#)
- [6] G. Lacerda and S. Romána. *Mean topological dimension*. IEEE Trans. Inform. Theory 70(11) (2024), 7664–7672. [2.9](#)

- [7] E. Lindenstrauss and B. Weiss. *Mean topological dimension*. Israel J. Math. 115 (2000), 1–24. [1](#), [3.4](#)
- [8] E. Lindenstrauss and M. Tsukamoto. *From rate distortion theory to metric mean dimension: variational principle*. IEEE Trans. Inform. Theory 64 (2018), 3590–3609. [1](#), [1.1](#), [2.10](#)
- [9] E. Lindenstrauss and M. Tsukamoto. *Double variational principle for mean dimension*. Geom. Funct. Anal. 29 (2019), 1048–1109. [1](#)
- [10] J. C. Oxtoby and S. M. Ulam. *Measure-preserving homeomorphisms and metrical transitivity*. Ann. Math. 42(4) (1941), 874–920. [2.9](#)
- [11] U. Shapira. *Measure theoretical entropy of covers*. Israel. J. Math. 158(1) (2007), 225–247. [4.2](#)
- [12] R. Shi. *On variational principles for metric mean dimension*. IEEE Trans. Inform. Theory 68 (2022), 4282–4288. [1](#)
- [13] R. Shi. *Finite mean dimension and marker property*. Trans. Amer. Math. Soc. 376 (2023), 6123–6139. [1](#)
- [14] M. Tsukamoto. *Rate distortion dimension and ergodic decomposition for \mathbb{R}^d -actions*. Preprint arXiv:2503.06851, 2025. [1.2](#), [1](#)
- [15] M. Tsukamoto, M. Tsutaya and M. Yoshinaga. *G-index, topological dynamics and the marker property*. Israel J. Math. 251 (2022) 737–764. [1](#)
- [16] C. Villani. *Topics in Optimal Transportation*. Graduate Studies in Mathematics 58, American Mathematical Society, Providence, RI, 2003. [3.5](#), [3.5](#), [6.1](#)
- [17] A. Velozo and R. Velozo. *Rate distortion theory, metric mean dimension and measure theoretic entropy*. Preprint, arXiv:1707.05762, 2017. [1.1](#)
- [18] T. Wang. *Variational relations for metric mean dimension and rate distortion dimension*. Discrete Contin. Dyn. Syst. 41(10) (2021), 4593–4608. [1](#)
- [19] P. Walters. *An Introduction to Ergodic Theory*. Graduate Texts in Mathematics 79, Springer-verlag, New York, Berlin, Heidelberg, 1982. [3.2](#), [7.2](#)
- [20] R. Yang. *Mean dimension and rate-distortion function revisited*. Preprint, arXiv:2510.08051, 2025. [2.11](#), [7.2](#)

CMUP & DEPARTAMENTO DE MATEMÁTICA, FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DO PORTO, RUA DO CAMPO ALEGRE 687, PORTO, PORTUGAL.

Email address: mpcarval@fc.up.pt

INSTITUTE OF MATHEMATICS OF THE POLISH ACADEMY OF SCIENCES, ŚNIADECKICH 8, WARSAW, 00-656, POLAND.

Email address: Gustavo.pessil@outlook.com