A CONVEX ANALYSIS APPROACH TO THE METRIC MEAN DIMENSION: LIMITS OF SCALED PRESSURES AND VARIATIONAL PRINCIPLES

MARIA CARVALHO, GUSTAVO PESSIL, AND PAULO VARANDAS

Abstract. We introduce the concept of upper metric mean dimension of a one-parameter family of scaled pressure functions, which extends the corresponding notion for a single potential and satisfies a variational principle. This approach, supported by Convex Analysis, conveys a definition of measure-theoretic upper metric mean dimension, which is concave and upper semi-continuous, and therewith equilibrium states. In the context of dynamical systems, we establish a variational principle for the metric mean dimension with potential in terms of Katok entropy. As an application, we provide a simple formula for the upper metric mean dimension with potential for the shift on the space $\big([0,1]^D\big)^{\mathbb{N}}$, for every $D \in \mathbb{N}$, which links mean dimension theory with ergodic optimization.

Contents

Date: November 16, 2023.

2010 *Mathematics Subject Classification.* Primary: 37D35, 28D20, 37B40, 26A51;

Key words and phrases. Metric mean dimension; Pressure function; Katok entropy; Variational principle; Equilibrium state; Box dimension.

1. Introduction

Let (X, d) be a compact metric space and $T\colon X \to X$ be a continuous map. Denote by $C^0(X)$ the space of real valued continuous maps whose domain is X, endowed with the uniform norm; by $\mathfrak{B}$ the σ-algebra of Borel sets of X; 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 $\mathcal{E}_T(X)$ the set of its ergodic elements; and, given $\mu \in \mathcal{P}(X)$, let $\mathrm{supp}(\mu)$ stand for its support. Measure-theoretic and topological entropy are classical, comprehensive and well succeed invariants in the theory of dynamical systems. Yet, these are not complete invariants, so there are several recently developed entropy-like concepts to estimate the complexity of systems by innovative approaches. The *upper and lower metric mean dimensions* are labels for dynamical systems introduced by E. Lindenstrauss and B. Weiss in [15] to quantify the complexity of infinite entropy systems. We denote them by $\overline{\mathrm{mdim}}_M(X, d, T)$ and $\underline{\mathrm{mdim}}_M(X, d, T)$, respectively, to emphasize their dependence on the fixed metric d of 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 convey information about the dimension of the phase space and the action of the dynamical system: they primarily report on the speed at which the entropy at scale ε approaches $+\infty$ as this scale goes to zero. The choice of the metric d has impact precisely on the speed of such convergence, which is quantified by the upper and lower metric mean dimensions.

In [16], E. Lindenstrauss and M. Tsukamoto established, under mild conditions, a variational principle between the metric mean dimension and the L^p rate-distortion function $\mathcal{R}_{\mu,p}$ of each $\mu \in \mathcal{P}_T(X)$. More precisely, they showed that, for every $p \in \mathbb{N}$,

$$\overline{\mathrm{mdim}}_M(X, d, T) \,=\, \limsup_{\varepsilon \to 0^+} \,\sup_{\mu \in \mathcal{P}_T(X)} \frac{\mathcal{R}_{\mu,p}(\varepsilon)}{\log(1/\varepsilon)}. \tag{1}$$

Actually, Y. Gutman and A. Śpiewak showed in [11] that it suffices to take the previous supremum over $\mathcal{E}_T(X)$, and obtained a new variational principle linking the upper metric mean dimension to the metric entropy h_μ, namely

$$\overline{\mathrm{mdim}}_M(X, d, T) \,=\, \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \,\sup_{\mu \in \mathcal{E}_T(X)} \,\inf_{\mathrm{diam}(P) \le \varepsilon} h_\mu(P) \tag{2}$$

where the infimum is taken over the Borel partitions P of X with diameter at most ε and $h_\mu(P)$ stands for the entropy of P. Problem 3 in [11] asked whether the metric mean dimension could be expressed in terms of Brin-Katok local entropy h^{BK}. An affirmative answer to this problem was given by [11] and [20], where the authors proved that

$$\overline{\mathrm{mdim}}_M(X, d, T) \,=\, \limsup_{\varepsilon \to 0^+} \frac{1}{\log(1/\varepsilon)} \,\sup_{\mu \in \mathcal{E}_T(X)} \,h^{BK}_\mu(\varepsilon). \tag{3}$$

A recurrent question in the literature concerning the previous and similar definitions asks whether we may exchange the order of $\limsup_\varepsilon$ and $\sup_\mu$. It is known that this order can be exchanged under the marker property (cf. [17, 25, 29] and also [30, 31]), but there are relevant systems without this property, as shown in [26, 21]. Regarding this issue, we refer the reader to Example 10.3 and Remark 10.4.

Meanwhile, Tsukamoto introduced [25] the concept of *upper metric mean dimension with potential*, which we will denote by $\overline{\mathrm{mdim}}_M(X, d, T, \varphi)$ for the potential φ, and proved a double variational principle similar to the one obtained earlier in [17] for the (topological) mean dimension (see [10] for the definition). Tsukamoto's definition of upper metric mean dimension with potential is inspired by the topological pressure, though the potential may vary with the scale ε. One is led to ask whether there is a measure-theoretic upper metric mean dimension $\mathcal{H}\colon \mathcal{P}_T(X) \to \mathbb{R}$ satisfying a classical variational principle (like the one established in [28, Theorem 9.10]), that is,

$$\overline{\mathrm{mdim}}_M(X, d, T, \varphi) = \sup_{\mu \in \mathcal{P}_T(X)} \Big\{ \mathcal{H}(\mu) + \int \varphi \, d\mu \Big\}. \tag{4}$$

As happens with the topological pressure, there may exist several maps satisfying such a relation, and the selection of a particular one depends on the preferred mechanism to generate them, or on their applications. For the zero potential case, distinct notions of measure-theoretic metric mean dimension were shown to exist and be related through a variational principle to some version of entropy, such as Brin-Katok local entropy (cf. [11], [20]), Katok metric entropy (cf. [27], [20]), Shapira entropy (cf. [20]) and Rényi information dimension (cf. [11]), just to mention a few. Regarding variational principles for continuous potentials and the previous measure-theoretic notions, we refer the reader to [7]. We stress that in all these variational principles the maximization on the space of probability measures is done for a fixed scale, which only afterwards is made to decrease to zero. Therefore, their statements do not comply with formula (4).

The previous information suggests the existence of some unifying path of reasoning, and Convex Analysis methods arise naturally in this background. Actually, in [1], the authors established an abstract variational principle for the so called *pressure functions* acting on a Banach space of potentials on a compact metric space, for which equilibrium states always exist. The main result of [1] ensures the existence of a conjugate of the pressure function, which acts on the set of Borel finitely additive probability measures if the space of potentials is made up from bounded measurable maps, and on the space of Borel probability measures if the space of potentials is $C^0(X)$. The first aim of our work is to show that the upper metric mean dimension with potential is a pressure function, so [1, Theorem 1] provides a suitable definition of a measure-theoretic upper metric mean dimension, defined on $\mathcal{P}(X)$ whenever the potentials belong to $C^0(X)$. In subsequent sections we will deduce extra properties of this concept and relate it to the measure-theoretic metric mean dimension introduced in [29].

Motivated by the aforementioned equalities (1), (2) and (3) for the zero potential, we introduce a scaled version of the upper metric mean dimension with potential, determined by a one-parameter family of scaled pressure functions (definitions in Section 2). This scaled upper metric mean dimension turns out to be a pressure function as well, to which we may apply the arguments of [1]. This way we obtain a variational principle, equilibrium states and the natural query about the continuity of these objects with the scale, whose answer will be given in Section 3. Moreover, this approach conveys a notion of measure-theoretic upper metric mean dimension, whose properties will be established in Theorem A.

In the case of continuous maps acting on a compact metric space, we will discuss another definition of measure-theoretic metric mean dimension in terms of Katok metric entropy (cf. [28, Theorem 8.19]). This new notion turns out to be convex and upper semi-continuous, and satisfies a variational principle. See Theorem C for more information.

Inspired by the concepts of entropy point and local entropy for homeomorphisms introduced in [32], we define similarly a local metric mean dimension (precise definition in (15) of Section 3) and what is a metric mean dimension point. Afterwards, having fixed a homeomorphism T on a compact metric space (X, d), we relate through a variational principle the upper metric mean dimension of T with the average of the local metric mean dimension (cf. Theorem E).

A relevant case study to which Theorem A applies is the shift map on a space $Y^{\mathbb{N}}$, where Y is a dimensionally homogeneous compact metric space. We will show that, in this setting, the variational principle established by Theorem A becomes an instance of ergodic optimization (more details in Theorem D).

This paper is organized as follows. The purpose of Section 2 is to recall some definitions and to introduce the new concepts. In Section 3 we state the main results. In Section 4 we summon the axiomatic contributions of [1] which we will use further on. After proving our results in Sections 5-9, we present some examples and applications in Section 10 and end with a short list of open problems suggested by this work.

2. Main definitions

We start by briefly recalling the definition of upper metric mean dimension with potential and by introducing the notions of scaled pressure function and the upper metric mean dimension determined by a family of scaled pressure functions.

2.1. **Upper metric mean dimension with potential.** Let (X, d) be a compact metric space, $T\colon X \to X$ be a continuous map and $\varphi\colon X \to \mathbb{R}$ be a continuous potential. For each $n \in \mathbb{N}$, define the dynamical metric

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

which is equivalent to d, and the sum $S_n\varphi = \sum_{j=0}^{n-1} \varphi \circ T^j$.

Given a closed subset K of X and $\varepsilon > 0$, consider the following infimum

$$S(K, d, \varphi, \varepsilon) \,=\, \inf \left\{ \sum_{i=1}^{\ell} (1/\varepsilon)^{\sup_{U_i} \varphi} \colon\; \{U_i\}_{1 \leq i \leq \ell} \text{ finite open cover of } K,\; \operatorname{diam}(U_i, d) < \varepsilon \right\} \tag{5}$$

the average

$$A(K, d, T, \varphi, \varepsilon, n) \,=\, \frac{1}{n} \log S(K, d_n, S_n\varphi, \varepsilon)$$

and the limit

$$P(K, d, T, \varphi, \varepsilon) \,=\, \lim_{n \to +\infty} A(K, d, T, \varphi, \varepsilon, n)$$

which exists since the sequence $\big(S(K, d_n, S_n\varphi, \varepsilon)\big)_n$ is sub-additive in the variable n. The upper metric mean dimension with potential, as defined in [25], extends the concept of metric mean dimension introduced in [15] (corresponding to the particular case of $\varphi = 0$) as follows:

Definition 2.1. *The upper metric mean dimension with potential of (K, d, T, φ) is given by*

$$\overline{\operatorname{mdim}}_M(K, d, T, \varphi) \,=\, \limsup_{\varepsilon \to 0^+} \frac{P(K, d, T, \varphi, \varepsilon)}{\log(1/\varepsilon)}.$$

A subset $E \subset K$ is said to be ε-separated with respect to the metric d if $d(x, y) \geq \varepsilon$ for every $x, y \in E$; it is ε-spanning with respect to the metric d if for every $x \in K$ there exists some

$y \in E$ such that $d(x,y) \le \varepsilon$. The notion of upper metric mean dimension with potential can be equivalently defined if one replaces $S(K,d,\varphi,\varepsilon)$ by either

$$S_1(K,d,\varphi,\varepsilon) \,=\, \sup\Big\{ \sum_{x\,\in\, E} (1/\varepsilon)^{\varphi(x)} \colon\; E \subset K \text{ is } \varepsilon\text{-separated}\Big\}$$

or

$$S_2(K,d,\varphi,\varepsilon) \,=\, \inf\Big\{ \sum_{x\,\in\, E} (1/\varepsilon)^{\varphi(x)} \colon\; E \subset K \text{ is } \varepsilon\text{-spanning}\Big\}.$$

Remark 2.2. *Since the topological pressure of a continuous map is invariant by co-boundary, the same holds for the upper metric mean dimension with potential, that is,*

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \,=\, \overline{\mathrm{mdim}}_M(X,d,T,\varphi+\psi\circ T-\psi) \qquad \forall\, \varphi,\, \psi\, \in C^0(X).$$

2.2. **Box dimension.** Let (Y,d) be a compact metric space. The *upper box dimension* of Y is defined as

$$\overline{\dim}_B\, Y \,=\, \limsup_{\varepsilon\,\to\, 0^+} \frac{S(Y,d,0,\varepsilon)}{\log(1/\varepsilon)}.$$

We define similarly the lower box dimension of Y. If the upper and lower box dimensions of Y coincide, their value is the box dimension of Y.

Remark 2.3. *It is straightforward that if $T(K) \subset K$, then*

$$P(K,d,T,0,\varepsilon) \,\le\, \log S(K,d,0,\varepsilon) \qquad \forall\, \varepsilon \,>\, 0.$$

In particular,

$$\overline{\mathrm{mdim}}_M(K,d,T,0) \,\le\, \overline{\dim}_B(K,d).$$

2.3. **Pressure functions.** Let (X,d) be a locally compact metric space and $\mathbf{B}(X)$ be a Banach space over $\mathbb{R}$ equal to either

$$\begin{aligned} B_d(X) \,&=\, \{\phi\colon X \to \mathbb{R} : \phi \text{ is measurable and bounded}\} \\ \text{or } C_b(X) \,&=\, \{\phi \in B_d(X) : \phi \text{ is continuous}\} \\ \text{or else } C_c(X) \,&=\, \{\phi \in C_b(X) : \phi \text{ has compact support}\} \end{aligned}$$

endowed with the norm $\|\varphi\|_\infty = \sup_{x\,\in\, X}\, |\varphi(x)|$. In what follows, $\mathcal{P}_a(X)$ will stand for the set of normalized finitely additive set functions on the Borel σ-algebra of X, which we will simply call *finitely additive probability measures*, with the total variation norm, defined by

$$\|\mu-\nu\| \,=\, \sup\,\Big\{ \big| \int \psi\, d\nu - \int \psi\, d\mu \big| \colon\; \psi \in \mathbf{B}(X) \text{ and } \|\psi\|_\infty \le 1 \Big\}.$$

We recall the notion of pressure function used in [1].

Definition 2.4. *We say that a map $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ is a pressure function if it satisfies the following conditions:*

- *Monotonicity:* $\quad \varphi \le \psi \implies \Upsilon(\varphi) \le \Upsilon(\psi) \quad \forall\, \varphi,\psi \in \boldsymbol{B}(X)$
- *Translation invariance:* $\quad \Upsilon(\varphi+c) = \Upsilon(\varphi)+c \quad \forall\, \varphi \in \boldsymbol{B}(X),\, \forall\, c \in \mathbb{R}$
- *Convexity:* $\quad \Upsilon(t\varphi+(1-t)\psi) \le t\Upsilon(\varphi)+(1-t)\Upsilon(\psi) \quad \forall\, \varphi,\, \psi \in \boldsymbol{B}(X),\, \forall\, t \in (0,1).$

For instance, the topological pressure of a continuous self-map $T\colon X \to X$ of a compact metric space X, whose topological entropy is finite, is a pressure function (cf. [28, Theorem 9.7]). More generally, if the weighted topological entropy is finite, the weighted topological pressure function defined in [9] is a pressure function.

2.4. **Scaled pressure functions.** Let (X,d) be a locally compact metric space. In what follows we introduce a weakening of the notion of pressure function in Definition 2.4, which allows a weaker form of translation invariance.

Definition 2.5. *A map $\Gamma\colon \boldsymbol{B}(X) \to \mathbb{R}$ is said to be a scaled pressure function if there exists a constant $\alpha > 0$ (the scale) such that Γ satisfies the following conditions:*

- *Monotonicity:* $\varphi \le \psi \implies \Gamma(\varphi) \le \Gamma(\psi) \quad \forall\, \varphi, \psi \in \boldsymbol{B}(X)$
- *Scaled translation invariance:* $\Gamma(\varphi + c) = \Gamma(\varphi) + \alpha c \quad \forall\, \varphi \in \boldsymbol{B}(X), \forall\, c \in \mathbb{R}$
- *Convexity:* $\Gamma(t\varphi + (1-t)\psi) \le t\Gamma(\varphi) + (1-t)\Gamma(\psi) \quad \forall\, \varphi,\, \psi \in \boldsymbol{B}(X), \forall\, t \in (0,1)$.

It is clear that pressure functions are scaled pressure functions with constant $\alpha = 1$. Moreover we note that, if Γ is a scaled pressure function with constant α, then $\Upsilon = \Gamma/\alpha$ is a pressure function. For example, the functions $(P(X,d,T,\cdot,\varepsilon))_{0<\varepsilon<1}$, defined in Subsection 2.1 for a continuous self-map $T\colon X \to X$ of a compact metric space X, are scaled pressure functions with constant $\log 1/\varepsilon$. Indeed, the monotonicity and the convexity are immediate to show. Regarding the translation invariant condition, given $\varphi \in C^0(X)$ and $c \in \mathbb{R}$, one has for every $0 < \varepsilon < 1$

$$S(X,d,\varphi+c,\varepsilon) \,=\, \inf_{(U_i)_{1\le i\le n}} \left\{ \sum_{i=1}^{n} (1/\varepsilon)^{\sup_{U_i}(\varphi+c)} \right\} \,=\, (1/\varepsilon)^c \inf_{(U_i)_{1\le i\le n}} \left\{ \sum_{i=1}^{n} (1/\varepsilon)^{\sup_{U_i}\varphi} \right\}$$

and

$$A(X,d,T,\varphi+c,\varepsilon,n) \,=\, \frac{1}{n}\log S(X,d_n,S_n(\varphi+c),\varepsilon) \,=\, \frac{1}{n}\log\Big((1/\varepsilon)^{nc}\, S(X,d_n,S_n\varphi,\varepsilon)\Big).$$

Therefore,

$$\begin{aligned} P(X,d,T,\varphi+c,\varepsilon) &= \lim_{n\to+\infty} A(X,d,T,\varphi+c,\varepsilon,n) \\ &= \lim_{n\to+\infty} A(X,d,T,\varphi,\varepsilon,n) + \big(\log 1/\varepsilon\big)c \\ &= P(X,d,T,\varphi,\varepsilon) + \big(\log 1/\varepsilon\big)c. \end{aligned}$$

2.5. **Upper metric mean dimension determined by a family of scaled pressure functions.** In what follows, we will consider a compact metric space (X,d) and families of scaled pressure functions acting on the same Banach space, and associate to them a conjoint notion of upper metric mean dimension. Let $\boldsymbol{\Gamma} = (\Gamma_\varepsilon)_{0<\varepsilon<1}$ be a family of scaled pressure functions defined on the same Banach space $\mathbf{B(X)}$ and such that, for every $0 < \varepsilon < 1$, the map Γ_ε satisfies the scaled translation invariance condition with respect to the constant $\log 1/\varepsilon$.

Definition 2.6. *The upper metric mean dimension of $\boldsymbol{\Gamma}$ at $\varphi \in \boldsymbol{B}(X)$ is given by*

$$\overline{\mathrm{mdim}}_M(\boldsymbol{\Gamma}, d, \varphi) \,=\, \limsup_{\varepsilon\to 0^+} \frac{\Gamma_\varepsilon(\varphi)}{\log(1/\varepsilon)}. \tag{6}$$

The previous notion generalizes the concept of upper metric mean dimension with potential introduced in [25]. Indeed, if we consider a continuous self-map T acting on X and the family of scaled pressure functions $\mathbf{\Gamma} = (P(X,d,T,\cdot,\varepsilon))_{0<\varepsilon<1}$ defined in Subsection 2.1, then it is immediate to verify that

$$\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,\cdot) = \overline{\mathrm{mdim}}_M(X,d,T,\cdot).$$

To simplify the notation, if a scaled pressure function $\Gamma_\varepsilon\colon \mathbf{B(X)} \to \mathbb{R}$, for some $0<\varepsilon<1$, satisfies the scaled translation invariance condition with respect to the constant $\log 1/\varepsilon$, then we will call it an ε-pressure function.

3. Main results

We start with a consequence of [1], which provides a variational principle and equilibrium states for the upper metric mean dimension determined by a family $\mathbf{\Gamma} = (\Gamma_\varepsilon)_{0<\varepsilon<1}$ of ε-pressure functions.

Theorem A. *Let (X,d) be a compact metric space and $\mathbf{\Gamma} = (\Gamma_\varepsilon)_{0<\varepsilon<1}$ be a family of ε-pressure functions defined on a Banach space $\boldsymbol{B}(X)$ and such that $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,\cdot) < +\infty$. Then*

$$\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,\varphi) = \max_{\mu\in\mathcal{P}_a(X)} \Big\{M(\mu) + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in \boldsymbol{B}(X) \tag{7}$$

where

$$M(\mu) = \inf_{\varphi\in\mathcal{C}_\Gamma} \int \varphi\, d\mu \quad \text{and} \quad \mathcal{C}_\Gamma = \{\varphi\in\boldsymbol{B}(X)\colon \overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi)\le 0\}. \tag{8}$$

The map M is concave, upper semi-continuous, bounded above by $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,0)$ and satisfies

$$M(\mu) = \inf_{\varphi\in\boldsymbol{B}(X)} \Big\{\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,\varphi) - \int \varphi\, d\mu\Big\} \quad \forall\, \mu\in\mathcal{P}_a(X).$$

In addition, if $\tau\colon \mathcal{P}_a(X)\to[0,+\infty]$ is another function taking the role of M in the equality (7), then $\tau\le M$. Moreover, $\mathcal{C}_\Gamma \supseteq \liminf_{\varepsilon\to 0^+} \mathcal{A}_{\Gamma_\varepsilon}$, where

$$\mathcal{A}_{\Gamma_\varepsilon} = \Big\{\varphi\in\boldsymbol{B}(X)\colon \Gamma_\varepsilon(-\varphi)\le 0\Big\}.$$

When $\boldsymbol{B}(X) = C^0(X)$, the maximum in (7) is attained in $\mathcal{P}(X)$.

It is straightforward to conclude that μ attains the infimum in (8) (that is, $M(\mu) = \int \varphi_0\, d\mu$ for some $\varphi_0\in\mathcal{C}_\Gamma$) if and only if $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi_0) = 0$ and μ is an equilibrium state of $-\varphi_0$. Indeed, if $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi_0) = 0$ and μ is an equilibrium state of $-\varphi_0$, then

$$0 = \overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi_0) = M(\mu) + \int -\varphi_0\, d\mu$$

so $M(\mu) = \int \varphi_0\, d\mu$. Conversely, if $M(\mu) = \int\varphi_0\, d\mu$ for some $\varphi_0\in\mathcal{C}_\Gamma$, then

$$0 \ge \overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi_0) \ge M(\mu) - \int \varphi_0 = 0$$

so $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi_0) = 0$ and μ is an equilibrium state of $-\varphi_0$.

After completing this work we became aware of the recent preprint [29] by Yang, Chen and Zhou, where the authors propose a notion of measure-theoretic upper metric mean dimension for continuous self-maps of compact metric spaces. Our strategy is different from, though akin to, the one we read in [29], whose authors define the measure-theoretic upper metric mean dimension for a continuous map $T\colon X \to X$ of a compact metric space (X, d) by

$$\mu \in \mathcal{P}(X) \quad \mapsto \quad F(\mu) \,=\, \inf_{\varphi \in \mathcal{C}} \int \varphi \, d\mu$$

where

$$\mathcal{C} \,=\, \{\varphi \in C^0(X)\colon \; \overline{\mathrm{mdim}}_M(X, d, T, -\varphi) \,=\, 0\}.$$

Moreover, the authors showed that the map F satisfies the variational principle

$$\overline{\mathrm{mdim}}_M(X, d, T, \varphi) \,=\, \sup_{\mu \in \mathcal{P}(X)} \Big\{F(\mu) + \int \varphi \, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X).$$

From the latter equality, we deduce by applying Theorem A to $(P(X, d, T, \cdot, \varepsilon))_{0<\varepsilon<1}$ that

$$F(\mu) \,\leq\, M(\mu) \qquad \forall\, \mu \in \mathcal{P}(X).$$

Yet, by definition, $\mathcal{C} \subseteq \mathcal{C}_\Gamma$; hence

$$M(\mu) \,\leq\, F(\mu) \qquad \forall\, \mu \in \mathcal{P}(X).$$

Thus, $M = F$, so we may add that F depends on the choice of the metric d (cf. Example 10.1) and only attains the maximum at invariant probability measures (cf. Lemma 5.4). Therefore, Theorem A recovers the main contribution of [29], besides extending it to scaled pressure functions. We note that Theorem A's statement is free from a dynamical context, which allows us to consider other settings, such as semigroup actions (cf. Subsection 11.1).

We proceed to discuss the computability of the map M. More precisely, take a family $\mathbf{\Gamma} = \big(\Gamma_\varepsilon)_{0<\varepsilon<1}$ of ε-pressure functions $\Gamma_\varepsilon\colon C^0(X) \to \mathbb{R}$ such that $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma}, d, \cdot) < +\infty$. Let $M\colon \mathcal{P}(X) \to \mathbb{R}$ be the map assigned by Theorem A to the family $\mathbf{\Gamma} = \big(\Gamma_\varepsilon)_{0<\varepsilon<1}$, which satisfies

$$\limsup_{\varepsilon \to 0^+} \frac{\Gamma_\varepsilon(\varphi)}{\log{(1/\varepsilon)}} \,=\, \max_{\mu \in \mathcal{P}(X)} \Big\{M(\mu) + \int \varphi \, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X). \tag{9}$$

For every $0 < \varepsilon < 1$, denote by $\mathfrak{h}_\varepsilon\colon \mathcal{P}(X) \to \mathbb{R}$ the map provided by the application of [1, Theorem 1] to the pressure function $\Gamma_\varepsilon / \log{(1/\varepsilon)}$, so that

$$\frac{\Gamma_\varepsilon(\varphi)}{\log{(1/\varepsilon)}} \,=\, \max_{\mu \in \mathcal{P}(X)} \Big\{\mathfrak{h}_\varepsilon(\mu) + \int \varphi \, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X). \tag{10}$$

Is M the $\limsup$, as ε go to 0^+, of the maps $\mathfrak{h}_\varepsilon$? To address this question, consider the function

$$\begin{array}{rccc} M^*\colon & \mathcal{P}(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon \in \mathcal{M}(\mu)} \, \limsup_{\varepsilon \to 0^+} \, \mathfrak{h}_\varepsilon(\mu_\varepsilon) \end{array}$$

where $\mathcal{M}(\mu)$ is the space of sequences of probability measures in $\mathcal{P}(X)$ which converge in the weak*-topology to μ.

Theorem B. *Let (X, d) be a compact metric space and $\mathbf{\Gamma} = \big(\Gamma_\varepsilon)_{0<\varepsilon<1}$ be a family of ε-pressure functions $\Gamma_\varepsilon\colon C^0(X) \to \mathbb{R}$ such that $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma}, d, \cdot) < +\infty$. Then*

$$\overline{\mathrm{mdim}}_M(\mathbf{\Gamma}, d, \varphi) \,=\, \max_{\mu \in \mathcal{P}(X)} \Big\{M^*(\mu) + \int \varphi \, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X).$$

Moreover, given $\varphi \in C^0(X)$, the previous maximum is attained at any accumulation point of the family of equilibrium states $(\mu_\varphi^{\varepsilon_n})_n$ for φ and the pressure functions $(\Gamma_{\varepsilon_n}/\log(1/\varepsilon_n))_n$, where the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ converges to zero and satisfies

$$\overline{\mathrm{mdim}}_M(\mathbf{\Gamma}, d, \varphi) = \lim_{n\to+\infty} \frac{\Gamma_{\varepsilon_n}(\varphi)}{\log(1/\varepsilon_n)}.$$

Theorem B is a consequence of a more general result that will be proved in Section 6. Theorems A and B may be applied to most definitions of pressure function available in the literature. We will illustrate this assertion in Subsection 11.1, within the context of topological pressures determined by dynamical systems or semigroup actions.

We will now address the dynamical setting. Let (X, d) be a compact metric space and $T\colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X, d, T) < +\infty$. Although Theorem A provides a map $M\colon \mathcal{P}(X) \to \mathbb{R}$ satisfying a variational principle, the variational nature of the definition of M prevents us from easily compute it in specific examples or draw dynamical information from it. In what follows, we introduce another suitable notion of measure-theoretic upper metric mean dimension, built on Katok's description of the metric entropy, which also satisfies a variational principle like (4).

Given $\mu \in \mathcal{P}_T(X)$ and $\delta \in]0,1[$, we define

$$\begin{array}{rccc} H_\delta^K\colon & \mathcal{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon\in\mathcal{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array} \tag{11}$$

where $\mathcal{M}(\mu)$ stands for the space of sequences of probability measures in $\mathcal{P}_T(X)$ which converge to μ in the weak*-topology, and $h^K_\nu(\varepsilon, \delta)$ is the δ-Katok entropy of ν at scale ε (see the definition in [14] and more details in Section 7).

Theorem C. *Let (X, d) be a compact metric space and $T\colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X, d, T) < +\infty$. Then, for every $\delta \in]0,1[$,*

$$\overline{\mathrm{mdim}}_M(X, d, T, \varphi) = \max_{\mu\in\mathcal{E}_T(X)} \left\{ H_\delta^K(\mu) + \int \varphi\, d\mu \right\} \qquad \forall\, \varphi \in C^0(X). \tag{12}$$

Our approach allows a generalization which will be established in Subsection 7.3.

We may improve the information conveyed by Theorems A and C if we consider full shifts on compact alphabets satisfying a box dimension homogeneity condition. Given a compact metric space (Y, d) and $\rho > 1$, we endow the space $Y^{\mathbb{N}}$ with the metric

$$d_\rho(x, y) = \sup_{n\in\mathbb{N}} \frac{d(x_n, y_n)}{\rho^{n-1}} \tag{13}$$

where $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$. For each $\varphi \in C^0(Y^{\mathbb{N}})$ the following result provides an exact formula for the upper metric mean dimension $\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi)$, yielding a reformulation of the variational principle (7) in this setting.

Theorem D. *Let (Y, d) be a compact metric space such that $\dim_B U = \dim_B Y$ for every nonempty open set $U \subset Y$. Then,*

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi) = \dim_B Y + \max_{\mu\in\mathcal{E}_\sigma(Y^{\mathbb{N}})} \int \varphi\, d\mu \qquad \forall\, \varphi \in C^0(Y^{\mathbb{N}}). \tag{14}$$

For instance, for every positive integer D and every $\rho > 1$,

$$\overline{\mathrm{mdim}}_M(([0,1]^D)^{\mathbb{N}}, d_\rho, \sigma, \varphi) \;=\; D + \max_{\mu \in \mathcal{E}_\sigma(([0,1]^D)^{\mathbb{N}})} \int \varphi \, d\mu \qquad \forall\, \varphi \in C^0(([0,1]^D)^{\mathbb{N}}).$$

For more information when $D = 1$, see Example 10.2. To the authors knowledge, this is the first explicit calculation of the upper metric mean dimension with potential of a dynamical system. We also note that the maximization in (14) is an issue of ergodic optimization. Regarding this subject, we refer the reader to Section 11.

In the remainder of this section, we consider the action of a homeomorphism T on a compact metric space (X, d). Define the *local metric mean dimension function* by

$$\begin{array}{rcl} \mathcal{D}\colon X & \to & \mathbb{R} \\ x & \mapsto & \inf \Big\{ \overline{\mathrm{mdim}}_M(\overline{U}, d, T) \colon U \text{ is an open neighborhood of } x \Big\}. \end{array} \tag{15}$$

As T is a homeomorphism, the map $\mathcal{D}$ is upper semi-continuous and T-invariant. In particular, it is constant almost everywhere of any ergodic probability measure. Moreover, it satisfies the following variational principle.

Theorem E. *Let (X, d) be a compact metric space and $T\colon X \to X$ be a homeomorphism such that $\overline{\mathrm{mdim}}_M(X, d, T) < +\infty$. Then*

$$\overline{\mathrm{mdim}}_M(X, d, T) \;=\; \max_{\mu \in \mathcal{P}_T(X)} \int \mathcal{D}(x)\, d\mu(x) \;=\; \max_{\mu \in \mathcal{E}_T(X)} \int \mathcal{D}(x)\, d\mu(x).$$

In addition, a measure $\mu \in \mathcal{P}_T(X)$ attains the previous maximum if and only if

$$\mathcal{D}|_{\mathrm{supp}(\mu)} \;\equiv\; \overline{\mathrm{mdim}}_M(X, d, T).$$

4. Convex Analysis and pressure functions

In this section we collect some information from [1] concerning pressure functions. We start by observing that a monotone and translation invariant map $\Upsilon\colon \mathbf{B}(X) \to \mathbb{R} \cup \{\pm\infty\}$ is finite valued or constantly ∞. Indeed, given $\varphi,\, \psi \in \mathbf{B}(X)$, then

$$\Upsilon(\varphi) - \|\varphi - \psi\|_\infty \;=\; \Upsilon(\varphi - \|\varphi - \psi\|_\infty) \;\le\; \Upsilon(\varphi) \;\le\; \Upsilon(\psi + \|\varphi - \psi\|_\infty) \;=\; \Upsilon(\psi) + \|\varphi - \psi\|_\infty.$$

Theorem 4.1 (Theorem 1 of [1]; see also [2])**.** *Let (X, d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. Then*

$$\Upsilon(\varphi) \;=\; \max_{\mu \in \mathcal{P}_a(X)} \Big\{ \mathfrak{h}(\mu) + \int \varphi \, d\mu \Big\} \qquad \forall\, \varphi \in \boldsymbol{B}(X) \tag{16}$$

where the map $\mathfrak{h} = \mathfrak{h}_{\Upsilon, \boldsymbol{B}(X)}\colon \mathcal{P}_a(X) \to \mathbb{R} \cup \{-\infty, +\infty\}$ satisfies

$$\mathfrak{h}(\mu) \;=\; \inf_{\varphi \in \mathcal{A}_\Upsilon} \int \varphi \, d\mu \qquad \text{and} \qquad \mathcal{A}_\Upsilon \;=\; \{\varphi \in \boldsymbol{B}(X) \,:\, \Upsilon(-\varphi) \le 0\}.$$

The map $\mathfrak{h}$ is concave, upper semi-continuous, bounded above by $\Upsilon(0)$ and, if $\tau\colon \mathcal{P}_a(X) \to [0, +\infty]$ takes the role of $\mathfrak{h}$ in (16), then $\tau \le \mathfrak{h}$. Moreover, one has

$$\mathfrak{h}(\mu) \;=\; \inf_{\varphi \in \boldsymbol{B}(X)} \Big\{ \Upsilon(\varphi) - \int \varphi \, d\mu \Big\} \qquad \forall\, \mu \,\in\, \mathcal{P}_a(X).$$

In case X is compact and $\boldsymbol{B}(X) = C^0(X)$, then the maximum is attained at $\mathcal{P}(X)$.

The variational principle (16) ensures that there always exist normalized finitely additive measures which attain the supremum; that is, the set

$$\mathcal{E}_\varphi(\Upsilon) \,=\, \big\{\mu \in \mathcal{P}_a(X)\colon\; \Upsilon(\varphi) \,=\, \mathfrak{h}(\mu) + \int \varphi \, d\mu\big\}$$

is non empty.

Definition 4.2. *Consider a pressure function $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ and a potential $\varphi \in \boldsymbol{B}(X)$. We say that $\mu \in \mathcal{P}_a(X)$ is a tangent functional to Υ at φ if*

$$\Upsilon(\varphi + \psi) - \Upsilon(\varphi) \,\geq\, \int \psi \, d\mu \qquad \forall \, \psi \in \boldsymbol{B}(X).$$

We denote by $\mathcal{T}_\varphi(\Upsilon)$ the set of tangent functionals to Υ at φ.

The next result states that, for every potential φ, the set $\mathcal{T}_\varphi(\Upsilon)$ coincides with the space of finitely additive equilibrium states of φ and established a sufficient condition for the uniqueness of such finitely additive equilibrium states.

Theorem 4.3 (Theorem 2 of [1])**.** *Let (X, d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. Then*

$$\mathcal{E}_\varphi(\Upsilon) \,=\, \mathcal{T}_\varphi(\Upsilon) \qquad \forall \, \varphi \in \boldsymbol{B}(X).$$

Moreover, if $\boldsymbol{B}(X) = C_b(X)$ or $\boldsymbol{B}(X) = C_c(X)$, then there exists a residual subset $\mathfrak{R} \subset \boldsymbol{B}(X)$ such that $\#\mathcal{E}_\varphi(\Upsilon) = 1$ for every $\varphi \in \mathfrak{R}$.

It is known that the uniqueness of equilibrium states for the topological pressure associated to a dynamical system is tied in with the differentiability of the pressure. In the wider setting of pressure functions one has the following generalization.

Definition 4.4. *A pressure function $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ is locally affine at $\varphi \in \boldsymbol{B}(X)$ if there exists a neighborhood $\mathcal{V}$ of $0 \in \boldsymbol{B}(X)$ and a unique $\mu_\varphi \in \mathcal{P}_a(X)$ such that*

$$\Upsilon(\varphi + \psi) - \Upsilon(\varphi) \,=\, \int \psi \, d\mu_\varphi \qquad \forall \, \psi \in \mathcal{V}.$$

Recall that $\Upsilon\colon \mathbf{B}(X) \to \mathbb{R}$ is *Fréchet differentiable at* $\varphi \in \boldsymbol{B}(X)$ if there exists a unique $\mu_\varphi \in \mathcal{P}_a(X)$ such that

$$\lim_{\psi \to 0} \frac{1}{\|\psi\|_\infty} \left|\Upsilon(\varphi + \psi) - \Upsilon(\varphi) - \int \psi \, d\mu_\varphi\right| \,=\, 0.$$

Theorem 4.5 (Theorem 3 of [1])**.** *Let (X, d) be a locally compact metric space and $\Upsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ be a pressure function. The following are equivalent:*

(a) Υ is locally affine at φ.

(b) There exists a unique tangent functional $\mu_\varphi \in \mathcal{T}_\varphi(\Upsilon)$ and

$$\lim_{\psi \to 0} \sup \left\{ \|\mu - \mu_\varphi\|\colon\; \mu \in \mathcal{T}_{\varphi+\psi}(\Upsilon) \right\} \,=\, 0$$

(c) Υ is Fréchet differentiable at φ.

In particular, the following are also equivalent:

(d) Υ *is affine.*

(e) $\bigcup_{\varphi \in \boldsymbol{B}(X)} \mathcal{T}_\varphi(\Upsilon)$ *is a singleton.*

(f) Υ *is everywhere Fréchet differentiable.*

As being affine is a rigid condition that often does not hold, one may consider the weaker notion of Gateaux differentiability. A pressure function $\Upsilon \colon \mathbf{B}(X) \to \mathbb{R}$ is *Gateaux differentiable at* $\varphi \in \boldsymbol{B}(X)$ if the directional pressure map $t \in \mathbb{R} \mapsto \Upsilon(\varphi + t\psi)$ is differentiable for every $\psi \in \mathbf{B}(X)$: that is, given $\psi \in \mathbf{B}(X)$, the limit

$$d\Upsilon(\varphi)(\psi) \, = \, \lim_{t \to 0} \frac{1}{t} \left[\Upsilon(\varphi + t\psi) - \Upsilon(\varphi)\right]$$

exists and is finite.

Theorem 4.6 (Corollary 4 of [1])**.** *A pressure function* $\Upsilon \colon \boldsymbol{B}(X) \to \mathbb{R}$ *is Gateaux differentiable at* φ *if and only if there exists a unique tangent functional in* $\mathcal{T}_\varphi(\Upsilon)$*.*

5. Proof of Theorem A

Part of the content of Theorem A is a direct consequence of [1, Theorem 1]. For instance, the map M is concave, since it is the infimum of affine maps, and bounded above by $\overline{\text{mdim}}_M(\mathbf{\Gamma}, d, 0)$ due to the translation invariance of Γ. We are left to verify that the upper metric mean dimension of a family $\mathbf{\Gamma} = (\Gamma_\varepsilon)_{0 < \varepsilon < 1}$ of ε-pressure functions (defined in Subsection 2.5) is a pressure function. Afterwards, we identify the admissible potentials in $\mathcal{C}_\Gamma$ (a subset defined by (8) in the statement of Theorem A) which determine the measure-theoretic upper metric mean dimension map M.

Consider the space

$$\mathcal{L}_{\mathbf{B(X)}} \, = \, \{\Gamma \colon \mathbf{B}(X) \to \mathbb{R} \mid \Gamma \text{ is a scaled pressure function}\} \cup \{\pm\infty\}.$$

This is a positive convex cone, that is,

(a) if $\Gamma \in \mathcal{L}_{\mathbf{B(X)}}$, then $\beta\Gamma \in \mathcal{L}_{\mathbf{B(X)}}$ for every $\beta > 0$;

(b) given $\Gamma_1, \Gamma_2 \in \mathcal{L}_{\mathbf{B(X)}}$ and $0 \le t \le 1$, then $t\Gamma_1 + (1-t)\Gamma_2 \in \mathcal{L}_{\mathbf{B(X)}}$.

The next lemma, whose proof is straightforward, describes another relevant property of this space.

Lemma 5.1. $\mathcal{L}_{\mathbf{B(X)}}$ *is closed under the* lim sup *operator, which preserves the scale. More precisely, let* $\left(G_t\right)_{0<t<1}$ *be a collection of elements of* $\mathcal{L}_{\mathbf{B(X)}}$ *with common scale* α*. Then the map*

$$\varphi \in \boldsymbol{B}(X) \quad \mapsto \quad \limsup_{t \to 0^+} G_t(\varphi)$$

is a scaled pressure function with scale α*.*

The next proposition establishes that $\overline{\text{mdim}}_M(\mathbf{\Gamma}, d, \cdot)$ is a pressure function.

Proposition 5.2. *Given a family* $\mathbf{\Gamma} = (\Gamma_\varepsilon)_{0<\varepsilon<1}$ *of* ε*-pressure functions, defined in the same Banach space* $\boldsymbol{B}(X)$*, the map*

$$\overline{\text{mdim}}_M(\mathbf{\Gamma}, d, \cdot) \colon \; \varphi \in \boldsymbol{B}(X) \quad \mapsto \quad \limsup_{\varepsilon \to 0^+} \frac{\Gamma_\varepsilon(\varphi)}{\log{(1/\varepsilon)}}$$

belongs to $\mathcal{L}_{\mathbf{B(X)}}$*.*

Proof. Fix $0<\varepsilon<1$. It is immediate to check that the renormalization

$$\Upsilon_\varepsilon\colon\ \varphi\in\mathbf{B}(X)\quad\mapsto\quad\frac{\Gamma_\varepsilon(\varphi)}{\log\,(1/\varepsilon)}$$

is a convex and monotone function. Moreover, for every $\varphi\in\mathbf{B}(\mathbf{X})$ and every $c\in\mathbb{R}$, one has

$$\Upsilon_\varepsilon(\varphi+c)\;=\;\frac{\Gamma_\varepsilon(\varphi+c)}{\log\,(1/\varepsilon)}\;=\;\frac{\Gamma_\varepsilon(\varphi)+\big(\log\,1/\varepsilon\big)c}{\log\,(1/\varepsilon)}\;=\;\Upsilon_\varepsilon(\varphi)+c.$$

Therefore, the map $\frac{\Gamma_\varepsilon}{\log\,(1/\varepsilon)}$ is in $\mathcal{L}_{\mathbf{B}(\mathbf{X})}$, with scale $\alpha=1$ for every $0<\varepsilon<1$. Consequently, by Lemma 5.1, the map $\overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,\cdot)$ is an element of $\mathcal{L}_{\mathbf{B}(\mathbf{X})}$, as claimed. □

We will now find suitable potentials.

Lemma 5.3. *Let* $\mathbf{\Gamma}\;=\;(\Gamma_\varepsilon)_{0\,<\,\varepsilon\,<\,1}$ *be a family of pressure functions and consider the sets*

$$\mathcal{C}_{\mathbf{\Gamma}}\;=\;\big\{\varphi\in\boldsymbol{B}(X)\colon\ \overline{\mathrm{mdim}}_M(\mathbf{\Gamma},d,-\varphi)\;\le\;0\big\}$$

and, for every $0<\varepsilon<1$,

$$\mathcal{A}_{\Gamma_\varepsilon}\;=\;\big\{\varphi\in\boldsymbol{B}(X)\colon\ \Gamma_\varepsilon(-\varphi)\;\le\;0\big\}.$$

Then

$$\mathcal{C}_{\mathbf{\Gamma}}\;\supseteq\;\liminf_{\varepsilon\,\to\,0^+}\,\mathcal{A}_{\Gamma_\varepsilon}.$$

Proof. Clearly, if $\Gamma_\varepsilon(-\varphi)\le 0$ for every small enough ε, then

$$\limsup_{\varepsilon\,\to\,0^+}\,\frac{\Gamma_\varepsilon(-\varphi)}{\log\,(1/\varepsilon)}\;\le\;0.$$

Consequently,

$$\bigcup_{n\,\in\,\mathbb{N}}\ \bigcap_{0\,<\,\varepsilon\,<\,1/n}\ \mathcal{A}_{\Gamma_\varepsilon}\;\subseteq\;\mathcal{C}_{\mathbf{\Gamma}}.$$

The proofs of Lemma 5.3 and Theorem A are complete. □

We conclude this section by showing that, in the case of a continuous map acting on a compact metric space with finite upper metric mean dimension, all equilibrium states are T-invariant probability measures.

Lemma 5.4. *Let* (X,d) *be a compact metric space,* $T\colon X\to X$ *be a continuous map such that* $\overline{\mathrm{mdim}}_M(X,d,T)<+\infty$ *and* $\varphi\in C^0(X)$ *be a continuous potential. Then any equilibrium state* μ_φ *of* φ *with respect to the variational principle* (7) *is invariant under* T.

Proof. Given $\varphi\in C^0(X)$, let $\mu_\varphi\in\mathcal{P}(X)$ be such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi)\;=\;M(\mu_\varphi)+\int\varphi\,d\mu_\varphi.$$

We will show that $\int(\psi\circ T)\,d\mu_\varphi=\int\psi\,d\mu_\varphi$ for every $\psi\in C^0(X)$. Fix $\psi\in C^0(X)$ and let μ_1 and μ_2 be equilibrium states associated to the potentials $\varphi+\psi\circ T-\psi$ and $\varphi-\psi\circ T+\psi$, respectively. Then

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi+\psi\circ T-\psi)\;=\;M(\mu_1)+\int\varphi\,d\mu_1+\int(\psi\circ T)\,d\mu_1-\int\psi\,d\mu_1$$

and

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi-\psi\circ T+\psi) \;=\; M(\mu_2)+\int \varphi\,d\mu_2-\int (\psi\circ T)\,d\mu_2+\int \psi\,d\mu_2.$$

Moreover, by Remark 2.2, we know that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \;=\; \overline{\mathrm{mdim}}_M(X,d,T,\varphi+\psi\circ T-\psi) \;=\; \overline{\mathrm{mdim}}_M(X,d,T,\varphi-\psi\circ T+\psi).$$

Therefore,

$$\begin{aligned} M(\mu_\varphi)+\int \varphi\,d\mu_\varphi &= M(\mu_1)+\int \varphi\,d\mu_1+\int (\psi\circ T)\,d\mu_1-\int \psi\,d\mu_1 \\ &\geq M(\mu_\varphi)+\int \varphi\,d\mu_\varphi+\int (\psi\circ T)\,d\mu_\varphi-\int \psi\,d\mu_\varphi \end{aligned}$$

which yields $\int (\psi\circ T)\,d\mu_\varphi \leq \int \psi\,d\mu_\varphi$. Analogously,

$$\begin{aligned} M(\mu_\varphi)+\int \varphi\,d\mu_\varphi &= M(\mu_2)+\int \varphi d\mu_2-\int (\psi\circ T)\,d\mu_2+\int \psi\,d\mu_2 \\ &\geq M(\mu_\varphi)+\int \varphi\,d\mu_\varphi-\int (\psi\circ T)\,d\mu_\varphi+\int \psi\,d\mu_\varphi \end{aligned}$$

so $\int (\psi\circ T)\,d\mu_\varphi \geq \int \psi\,d\mu_\varphi$. The proof is complete. Hence, the maximum in the variational principle provided by Theorem A can be computed on the set of T-invariant probability measures. □

6. Proof of Theorem B

Fix a compact metric space (X,d) and a family of pressure functions $\Upsilon_\varepsilon\colon C^0(X)\to\mathbb{R}$, for $0<\varepsilon<1$, such that $\boldsymbol{\Upsilon} = \limsup_{\varepsilon\to 0^+}\Upsilon_\varepsilon < +\infty$. Recall that Lemma 5.1 ensures that $\boldsymbol{\Upsilon}$ is a pressure function as well.

Lemma 6.1. *Consider $\varphi\in C^0(X)$ and a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to+\infty}\Upsilon_{\varepsilon_n}(\varphi) = \boldsymbol{\Upsilon}(\varphi)$. If, for every $n\in\mathbb{N}$, the measure $\mu_{\varepsilon_n}\in\mathcal{P}(X)$ is a tangent functional to Υ_{ε_n} at φ, then any accumulation point $\mu\in\mathcal{P}(X)$ of $(\mu_{\varepsilon_n})_{n\in\mathbb{N}}$ is a tangent functional to $\boldsymbol{\Upsilon}$ at φ.*

Proof. Given $\psi\in C^0(X)$,

$$\begin{aligned} \boldsymbol{\Upsilon}(\varphi+\psi)-\boldsymbol{\Upsilon}(\varphi) &\geq \limsup_{n\to+\infty}\Upsilon_{\varepsilon_n}(\varphi+\psi)-\lim_{n\to+\infty}\Upsilon_{\varepsilon_n}(\varphi) \\ &= \limsup_{n\to+\infty}\Big(\Upsilon_{\varepsilon_n}(\varphi+\psi)-\Upsilon_{\varepsilon_n}(\varphi)\Big) \\ &\geq \limsup_{n\to+\infty}\int \psi\,d\mu_{\varepsilon_n} \\ &\geq \int \psi\,d\mu. \end{aligned}$$

□

We notice that, from Theorem 4.3, it is known that the set of equilibrium states of the pressure function $\boldsymbol{\Upsilon}$ coincides with the set of its tangent functionals. For every $0<\varepsilon<1$, denote by

$\mathfrak{h}_\varepsilon \colon \mathcal{P}(X) \to \mathbb{R}$ the map given by the application of [1, Theorem 1] to the pressure function Υ_ε, which satisfies

$$\Upsilon_\varepsilon(\varphi) \,=\, \max_{\mu \in \mathcal{P}(X)} \Big\{\mathfrak{h}_\varepsilon(\mu) + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X).$$

Moreover, given $\varphi \in C^0(X)$ and $0 < \varepsilon < 1$, let μ_φ^ε be an equilibrium state of the pressure function Υ_ε for φ.

Theorem 6.2. *Let (X, d) be a compact metric space and $\big(\Upsilon_\varepsilon\big)_{0<\varepsilon<1}$ be a family of pressure functions $\Upsilon_\varepsilon \colon C^0(X) \to \mathbb{R}$ such that $\boldsymbol{\Upsilon} = \limsup_{\varepsilon\to 0^+} \Upsilon_\varepsilon < +\infty$. Then*

$$\boldsymbol{\Upsilon}(\varphi) \,=\, \max_{\mu \in \mathcal{P}(X)} \Big\{M^*(\mu) + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in C^0(X).$$

Moreover, given $\varphi \in C^0(X)$, the previous maximum is attained at any accumulation point of the family of equilibrium states $(\mu_\varphi^{\varepsilon_n})_n$ for φ and the pressure functions $(\Upsilon_{\varepsilon_n})_n$, where the sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ converges to zero and satisfies

$$\boldsymbol{\Upsilon}(\varphi) \,=\, \lim_{n\to+\infty} \Upsilon_{\varepsilon_n}(\varphi).$$

Theorem B is a particular instance of the previous statement. Indeed, given a compact metric space (X, d) and a family $\boldsymbol{\Gamma} = \big(\Gamma_\varepsilon\big)_{0<\varepsilon<1}$ of ε-pressure functions $\Gamma_\varepsilon \colon C^0(X) \to \mathbb{R}$ such that $\overline{\mathrm{mdim}}_M(\boldsymbol{\Gamma}, d, \cdot) < +\infty$, it is enough to apply Theorem 6.2 to the family of pressure functions defined, for every $0 < \varepsilon < 1$, by

$$\Upsilon_\varepsilon(\varphi) \,=\, \frac{\Gamma_\varepsilon(\varphi)}{\log(1/\varepsilon)} \qquad \forall\, \varphi \in C^0(X)$$

and the map

$$\boldsymbol{\Upsilon} \,=\, \overline{\mathrm{mdim}}_M\big(\big(\Upsilon_\varepsilon\big)_{0<\varepsilon<1}, d, \cdot\big).$$

Proof of Theorem 6.2. We will show that the map M is an upper bound for M^* and that they coincide at the equilibrium states μ provided by Lemma 6.1. Let $\mu \in \mathcal{P}(X)$ and assume that $\mu_\varepsilon \to \mu$ in the weak$*$-topology as ε goes to zero. Then, for every $\varphi \in C^0(X)$,

$$\begin{aligned}
\boldsymbol{\Upsilon}(\varphi) &= \limsup_{\varepsilon\to 0^+} \Upsilon_\varepsilon(\varphi) \\
&= \limsup_{\varepsilon\to 0^+} \max_{\nu\in\mathcal{P}(X)} \Big\{\mathfrak{h}_\varepsilon(\nu) + \int \varphi\, d\nu\Big\} \\
&\geq \limsup_{\varepsilon\to 0^+} \Big(\mathfrak{h}_\varepsilon(\mu_\varepsilon) + \int \varphi\, d\mu_\varepsilon\Big) \\
&= \big(\limsup_{\varepsilon\to 0^+} \mathfrak{h}_\varepsilon(\mu_\varepsilon)\big) + \int \varphi\, d\mu.
\end{aligned}$$

Thus

$$\boldsymbol{\Upsilon}(\varphi) - \int \varphi\, d\mu \,\geq\, M^*(\mu).$$

Taking the infimum over all $\varphi \in C^0(X)$ we get

$$M(\mu) \,=\, \inf_{\varphi\in C^0(X)} \Big\{\boldsymbol{\Upsilon}(\varphi) - \int \varphi\, d\mu\Big\} \,\geq\, M^*(\mu).$$

Therefore,

$$\max_{\mu \in \mathcal{P}(X)} \Big\{ M^*(\mu) + \int \varphi \, d\mu \Big\} \,\leq\, \mathbf{\Upsilon}(\varphi).$$

For the converse inequality, given $\varphi \in C^0(X)$, take a sequence $(\varepsilon_n)_{n \in \mathbb{N}}$ as in the Lemma 6.1 and such that the limit $\lim_{n \to +\infty} \mu_\varphi^{\varepsilon_n}$ in the weak*-topology exists. According to Lemma 6.1, such a limit is an equilibrium state for $\mathbf{\Upsilon}$ and φ. Denote it by μ_φ. Then

$$\begin{aligned}
\mathbf{\Upsilon}(\varphi) \,&=\, \lim_{n \to +\infty} \Upsilon_{\varepsilon_n}(\varphi) \\
&=\, \lim_{n \to \, +\infty} \Big(\mathfrak{h}_{\varepsilon_n}(\mu_\varphi^{\varepsilon_n}) + \int \varphi \, d\mu_\varphi^{\varepsilon_n} \Big) \\
&=\, \Big(\lim_{n \to +\infty} \mathfrak{h}_{\varepsilon_n}(\mu_\varphi^{\varepsilon_n}) \Big) + \int \varphi \, \mu_\varphi.
\end{aligned}$$

Consequently, as μ_φ is an equilibrium state of φ, the previous estimates imply that

$$M(\mu_\varphi) \,=\, \lim_{n \to +\infty} \mathfrak{h}_{\varepsilon_n}(\mu_\varphi^{\varepsilon_n}).$$

Thus

$$M(\mu_\varphi) \,=\, \lim_{n \to +\infty} \mathfrak{h}_{\varepsilon_n}(\mu_\varphi^{\varepsilon_n}) \,\leq\, M^*(\mu_\varphi)$$

and so

$$\mathbf{\Upsilon}(\varphi) \,=\, M(\mu_\varphi) + \int \varphi \, d\mu_\varphi \,\leq\, M^*(\mu_\varphi) + \int \varphi \, d\mu_\varphi \,\leq\, \max_{\mu \in \mathcal{P}(X)} \Big\{ M^*(\mu) + \int \varphi \, d\mu \Big\}.$$

This finishes the proof of Theorem 6.2. □

7. Proof of Theorem C

Let (X, d) be a compact metric space and $T \colon X \to X$ be a continuous map such that $\overline{\mathrm{mdim}}_M(X, d, T) < +\infty$. We start by recalling the definition of Katok metric entropy (cf. [14]).

Given $\mu \in \mathcal{E}_T(X)$, $\varepsilon > 0$, $\delta \in \,]0, 1[$ and $n \in \mathbb{N}$, consider the minimal number of open sets with diameter smaller than ε with respect to d_n needed to cover any subset $A \in \mathfrak{B}$ with measure μ greater than $1 - \delta$, that is,

$$N_\mu(\varepsilon, \delta, n) \,=\, \inf_{A \in \mathfrak{B}} \{ S(A, d_n, 0, \varepsilon) \colon \, \mu(A) \,>\, 1 - \delta \}.$$

Define

$$h_\mu^K(\varepsilon, \delta) \,=\, \limsup_{n \to +\infty} \frac{1}{n} \log N_\mu(\varepsilon, \delta, n).$$

We can extend the previous notion to non-ergodic probability measures in $\mathcal{P}_T(X)$ via integration: given $\mu \in \mathcal{P}_T(X)$, define

$$h_\mu^K(\varepsilon, \delta) \,=\, \int_{\mathcal{E}_T(X)} h_m^K(\varepsilon, \delta) \, d\mathbb{P}_\mu(m) \tag{17}$$

where $\mu = \int_{\mathcal{E}_T(X)} m \, d\mathbb{P}_\mu(m)$ is the ergodic decomposition of μ. We note that, by definition, the map h_m^K is measurable and m-integrable; this way, the function

$$\mu \in \mathcal{P}_T(X) \quad \mapsto \quad h_\mu^K(\varepsilon, \delta)$$

is also affine.

Definition 7.1. *Given $\mu \in \mathcal{P}_T(X)$ and $\delta \in \,]0,1[$, define*

$$\begin{array}{rccc} H^K_\delta\colon & \mathcal{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon \,\in\, \mathcal{M}(\mu)} \limsup_{\varepsilon \to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array} \tag{18}$$

where $\mathcal{M}(\mu)$ stands for the space of sequences of probability measures in $\mathcal{P}_T(X)$ which converge to μ in the weak-topology.*

Remark 7.2. *A priori, the map H^K_δ may depend on δ, though in several examples its value is independent of this parameter (cf. Section 10). Moreover, given $\mu \in \mathcal{P}_T(X)$,*

$$\lim_{\delta \to 0^+} H^K_\delta(\mu) \;=\; \sup_{\delta \,\in\,]0,1[} H^K_\delta(\mu) \tag{19}$$

since the map $\delta \in \,]0,1[\, \mapsto H^K_\delta(\mu)$ is non-increasing.

To introduce a potential in the previous concept, given a continuous map $\varphi\colon X \to \mathbb{R}$ and $\mu \in \mathcal{E}_T(X)$, take

$$N_\mu(\varphi,\varepsilon,\delta,n) \;=\; \inf\,\{S(A,d_n,S_n\varphi,\varepsilon)\colon\; \mu(A) \,>\, 1-\delta\}$$

and

$$P^K_\mu(\varphi,\varepsilon,\delta) \;=\; \limsup_{n \to +\infty} \frac{1}{n}\,\log N_\mu(\varphi,\varepsilon,\delta,n).$$

Similarly, we now extend the previous notion to $\mu \in \mathcal{P}_T(X)$ using its ergodic decomposition, that is,

$$P^K_\mu(\varphi,\varepsilon,\delta) \;=\; \int_{\mathcal{E}_T(X)} P^K_m(\varphi,\varepsilon,\delta)\, d\mathbb{P}_\mu(m)$$

if $\mu = \int_{\mathcal{E}_T(X)} m\, d\mathbb{P}_\mu(m) \in \mathcal{P}_T(X)$.

Remark 7.3. *Given $\delta \in \,]0,1[$ and $\mu \in \mathcal{E}_T(X)$, one has $N_\mu(\varphi,\varepsilon,\delta,n) \le S(X,d_n,S_n\varphi,\varepsilon)$ for every $n \in \mathbb{N}$ and $\varepsilon > 0$; so $P^K_\mu(\varphi,\varepsilon,\delta)$ is bounded above by $P(X,d,T,\varphi,\varepsilon)$.*

7.1. **Linking H^K_δ and $\overline{\mathrm{mdim}}_M(X,d,T,\varphi)$.** The following connection between h^K_μ and P^K_μ is a straightforward adaptation of Proposition 2.2 in [6].

Lemma 7.4. *Let (X,d) be a compact metric space, $T\colon X \to X$ be a continuous map, $\varphi\colon X \to \mathbb{R}$ be a continuous potential and $\mu \in \mathcal{E}_T(X)$ be an ergodic probability measure. Then, given $\tau > 0$ and $1 > \delta_1 > \delta_2 > \delta_3 > 0$, one has for every $0 < \varepsilon < \tau$*

$$\frac{h^K_\mu(\varepsilon,\delta_1)}{\log(1/\varepsilon)} + \int \varphi\, d\mu - \tau \;\le\; \frac{P^K_\mu(\varphi,\varepsilon,\delta_2)}{\log(1/\varepsilon)} \;\le\; \frac{h^K_\mu(\varepsilon,\delta_3)}{\log(1/\varepsilon)} + \int \varphi\, d\mu + \tau. \tag{20}$$

This lemma has a main consequence: having fixed $\delta \in]0,1[$, if we define

$$\begin{array}{rccc} P^K_{\varphi,\delta}\colon & \mathcal{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon \,\in\, \mathcal{M}(\mu)} \limsup_{\varepsilon \to 0^+} \frac{P^K_{\mu_\varepsilon}(\varphi,\varepsilon,\delta)}{\log(1/\varepsilon)} \end{array} \tag{21}$$

then, taking $\limsup$ in (20), one gets

$$H^K_{\delta_1}(\mu) + \int \varphi\, d\mu \;\le\; P^K_{\varphi,\delta_2}(\mu) \;\le\; H^K_{\delta_3}(\mu) + \int \varphi\, d\mu \qquad \forall\, \mu \in \mathcal{P}_T(X).$$

Therefore, noticing that (see Remark 7.3)

$$P^K_{\varphi,\delta} \leq \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \qquad \forall\, \delta \in]0,1[\quad \forall\, \varphi \in C^0(X)$$

we conclude that:

Lemma 7.5. *For every* $\delta \in]0,1[$ *and* $\mu \in \mathcal{P}_T(X)$*, one has*

$$H^K_\delta(\mu) \leq \inf_{\varphi \in C^0(X)} \Big\{\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi\, d\mu\Big\} = M(\mu).$$

For future use, let us state two other properties of the map H^K_δ.

Lemma 7.6. *For every* $\delta \in]0,1[$*, the map* H^K_δ *is convex and upper semi-continuous.*

Proof. The convexity of H^K_δ is due to the fact that, for every $\delta \in]0,1[$ and $\varepsilon > 0$, the map

$$\mu \in \mathcal{P}_T(X) \quad \mapsto \quad h^K_\mu(\varepsilon,\delta)$$

is affine, hence the subsequent lim sup in the definition of H^K_δ yields a convex map in $\mathcal{P}_T(X)$. Besides, the supremum of convex maps is convex as well.

Regarding the upper semi-continuity of H^K_δ, consider $\mu \in \mathcal{P}_T(X)$ and a sequence $(\mu_n)_{n\in\mathbb{N}}$ in $\mathcal{P}_T(X)$ converging, in the weak*-topology, to μ. Given $n \in \mathbb{N}$, for every $\eta > 0$ there is a sequence $(\mu^{(n)}_{\varepsilon_k})_{k\in\mathbb{N}}$ in $\mathcal{M}(\mu_n)$ such that

$$H^K_\delta(\mu_n) - \eta < \limsup_{k\to+\infty} \frac{h^K_{\mu_{\varepsilon_k}}(\varepsilon_k,\delta)}{\log(1/\varepsilon_k)} \leq H^K_\delta(\mu_n).$$

Moreover, the sequence $(\mu^{(n)}_{\varepsilon_n})_{n\in\mathbb{N}}$ belongs to $\mathcal{M}(\mu)$. Thus,

$$H^K_\delta(\mu_n) < \limsup_{n\to+\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \eta \leq H^K_\delta(\mu) + \eta.$$

So, for every $\eta > 0$,

$$\limsup_{n\to+\infty} H^K_\delta(\mu_n) \leq H^K_\delta(\mu) + \eta$$

which implies that

$$\limsup_{n\to+\infty} H^K_\delta(\mu_n) \leq H^K_\delta(\mu).$$

□

We are ready to prove Theorem C.

7.2. **Proof.** Fix $\delta \in]0,1[$. The upper bound follows immediately from Lemma 7.5 and Theorem A:

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = \max_{\mu\in\mathcal{P}_T(X)} \Big\{M(\mu) + \int \varphi\, d\mu\Big\} \geq \max_{\mu\in\mathcal{E}_T(X)} \Big\{H^K_\delta(\mu) + \int \varphi\, d\mu\Big\}. \qquad (22)$$

We are left to show the other inequality. For that, we proceed by constructing $\mu_0 \in \mathcal{P}_T(X)$ such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \leq H^K_\delta(\mu_0) + \int \varphi\, d\mu_0.$$

It was proved in [7] that, for every $\delta \in]0,1[$ and every non-negative $\varphi \in C^0(X)$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \,=\, \limsup_{\varepsilon \to 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{P^K_\mu(\varphi,\varepsilon,\delta)}{\log(1/\varepsilon)}. \tag{23}$$

We note that the translation invariance property on both sides of the previous equality allows us to drop the condition $\varphi \geq 0$. Thus, given $\varphi \in C^0(X)$, applying Lemma 7.4 and (23) we obtain, for every $1 > \delta_1 > \delta_3 > 0$,

$$\limsup_{\varepsilon \to 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{h^K_\mu(\varepsilon,\delta_1)}{\log(1/\varepsilon)} + \int \varphi\, d\mu \leq \overline{\mathrm{mdim}}_M(X,d,T,\varphi) \leq \limsup_{\varepsilon \to 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{h^K_\mu(\varepsilon,\delta_3)}{\log(1/\varepsilon)} + \int \varphi\, d\mu.$$

These inequalities imply that, for every $\delta \in]0,1[$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \,=\, \limsup_{\varepsilon \to 0^+} \sup_{\mu \in \mathcal{E}_T(X)} \frac{h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)} + \int \varphi\, d\mu. \tag{24}$$

Therefore, we may find a sequence $(\varepsilon_n)_n$ of positive real numbers converging to 0 and such that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \,=\, \lim_{n \to +\infty} \sup_{\mu \in \mathcal{E}_T(X)} \frac{h^K_\mu(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu.$$

Thus, given $\eta > 0$, there is $p \in \mathbb{N}$ such that, for every $n \geq p$,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \eta \,<\, \sup_{\mu \in \mathcal{E}_T(X)} \frac{h^K_\mu(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu$$

and so, for every $n \geq p$ there exists $\mu_{\varepsilon_n} \in \mathcal{E}_T(X)$ satisfying

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \eta \,<\, \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\, d\mu_{\varepsilon_n}.$$

Taking a subsequence if necessary, we may assume that the sequence $\big(\mu_{\varepsilon_n}\big)_{n \geq p}$ converges in the weak*-topology to $\mu_0 \in \mathcal{P}_T(X)$. Then, by Definition 7.1, for every $\eta > 0$,

$$H^K_\delta(\mu_0) \,\geq\, \limsup_{n \to +\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \,\geq\, \overline{\mathrm{mdim}}_M(X,d,T,\varphi) - \int \varphi\, d\mu_0 - \eta. \tag{25}$$

Hence,

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \leq H^K_\delta(\mu_0) + \int \varphi\, d\mu_0.$$

To conclude, we note that even though the equilibrium state $\mu_0 \in \mathcal{P}_T(X)$ may be non-ergodic, by convexity and upper semi-continuity of the map

$$\mu \longmapsto \Big\{ H^K_\delta(\mu) + \int \varphi\, d\mu \Big\}$$

the maximum must also be attained at some ergodic measure. The proof of Theorem C is complete.

Remark 7.7. *From (22) and (25) we conclude that, up to a subsequence, the equilibrium state μ_0 satisfies*

$$H^K_\delta(\mu_0) \,=\, \lim_{n \to +\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)}.$$

Remark 7.8. *If one defines H_δ^K without considering measures varying with the scale, say*

$$\tilde{H}_\delta\colon\ \mu \quad\mapsto\quad \limsup_{\varepsilon\to 0^+} \frac{h_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)}$$

then the equality (12) *may not be valid: see Example 10.3 and Remark 10.4. This is related to the existence of probability measures capturing the dynamical complexity at all scales, which is a too strong requirement (see Section VIII in* [16]*). However, in some cases those probabilities do exist, as happens in Example 10.2.*

7.3. **Generalization.** The reasoning to prove Theorem C may be extended to maps

$$F\colon \mathcal{P}_T(X)\times\,]0,1[\,\to\mathbb{R}$$

satisfying

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \;=\; \limsup_{\varepsilon\to 0^+}\ \sup_{\mu\in\mathcal{P}_T(X)} \Big\{\frac{F(\mu,\varepsilon)}{\log(1/\varepsilon)} + \int \varphi\,d\mu\Big\}. \tag{26}$$

Actually, if we define

$$\begin{array}{rccc} F^*\colon & \mathcal{P}_T(X) & \to & \mathbb{R} \\ & \mu & \mapsto & \sup_{(\mu_\varepsilon)_\varepsilon\in\mathcal{M}(\mu)} \limsup_{\varepsilon\to 0^+} \frac{F(\mu_\varepsilon,\varepsilon)}{\log(1/\varepsilon)} \end{array}$$

where $\mathcal{M}(\mu)$ is the space of sequences of probability measures in $\mathcal{P}_T(X)$ which converge to μ in the weak*-topology, then the previous argument also shows that

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \;=\; \max_{\mu\in\mathcal{P}_T(X)} \Big\{F^*(\mu) + \int \varphi\,d\mu\Big\}.$$

For instance, F may be any of the examples referred to in (1), (2) and (3), or their corresponding versions with potential (see definitions in [7]).

Remark 7.9. *Example 10.3 provides a negative answer to the question of whether we can exchange the order of* $\limsup_\varepsilon$ *and* $\sup_\mu$ *in* (26)*. However, Theorem C shows that the regularization* F^* *of* F *is sufficient to surpass the obstructions to such an exchange of order.*

8. Proof of Theorem D

The estimates in this section will be done for unilateral sequences, though they are also valid for bilateral ones. Let (Y,d) be a compact metric space such that $\dim_B U = \dim_B Y$ for every nonempty open set $U\subset Y$. By Theorem A and Lemma 5.4 we have

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}},d_\rho,\sigma,\varphi) \;=\; \sup_{\mu\in\mathcal{P}_\sigma(Y^{\mathbb{N}})} \Big\{M(\mu) + \int\varphi\,d\mu\Big\} \qquad \forall\,\varphi\in C^0(X).$$

Since $\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}},d_\rho,\sigma) = \overline{\dim}_B Y$ (cf. [27, Theorem 5]), Theorem A also indicates that

$$\sup_{\mu\in\mathcal{P}_\sigma(Y^{\mathbb{N}})} M(\mu) \;=\; \overline{\dim}_B Y.$$

Thus,

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}},d_\rho,\sigma,\varphi) \;\le\; \overline{\dim}_B Y + \max_{\mu\in\mathcal{P}_\sigma(Y^{\mathbb{N}})} \int\varphi\,d\mu \qquad \forall\,\varphi\in C^0(X). \tag{27}$$

For the converse inequality, recall from Lemma 7.5 that, given $\delta>0$, we have

$$H_\delta^K(\mu) \;\le\; M(\mu) \;\le\; \overline{\mathrm{mdim}}_M(Y^{\mathbb{N}},d_\rho,\sigma) \;=\; \overline{\dim}_B Y \qquad \forall\mu\in\mathcal{P}_\sigma(Y^{\mathbb{N}}). \tag{28}$$

Theorem 8.1. *Let (Y, d) be a compact metric space such that $\dim_B U = \dim_B Y$ for every nonempty open set $U \subset Y$. Then, for every $\delta \in]0,1[$,*

$$H^K_\delta(\mu) = \dim_B Y \qquad \forall\, \mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}}).$$

Consequently,

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi) \geq \dim_B Y + \max_{\mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}})} \int \varphi\, d\mu \qquad \forall\, \varphi \in C^0(X) \tag{29}$$

and, bringing (27) and (29) together, the proof of Theorem D is complete, up to the proof of Theorem 8.1.

Proof of Theorem 8.1. To show that

$$H^K_\delta(\mu) \geq \dim_B Y \qquad \forall\, \mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}})$$

we start relating the map H^K_δ with the local box dimension of the alphabet set Y.

Lemma 8.2. *For every $\delta \in]0,1[$ and $\rho > 1$, given a fixed point $(y, y, \cdots) \in Y^{\mathbb{N}}$ one has*

$$\lim_{\gamma \to 0^+} \overline{\dim}_B B_\gamma(y) \leq H^K_\delta(\delta^{\mathbb{N}}_{\{y\}})$$

where $\delta^{\mathbb{N}}_{\{y\}} \in \mathcal{P}_\sigma(Y^{\mathbb{N}})$ is the product measure of the Dirac measure supported on y and $B_\gamma(y)$ stands for the open ball in Y centered at y with radius γ in the metric d_ρ.

Proof. Fix $\delta \in]0,1[$ and $\rho > 1$. Given $\varepsilon > 0$ and $\gamma > 0$, let $E = E(\varepsilon) \subset B_\gamma(y)$ be a maximal ε-separated subset of $B_\gamma(y)$. Take

$$\nu^\gamma_\varepsilon = \frac{1}{|E|} \sum_{e \in E} \delta_{\{e\}} \qquad \text{and} \qquad \mu^\gamma_\varepsilon = (\nu^\gamma_\varepsilon)^{\mathbb{N}}. \tag{30}$$

Then, clearly any two elements in the n-cylinder

$$z, w \in E \times \cdots \times E \times Y \times Y \times \cdots \subset Y^{\mathbb{N}}$$

which differ at some of the first n coordinates satisfy $d_n(z, w) \geq \varepsilon$. Let $L = L(n, \varepsilon, \delta)$ be the maximal positive integer such that $L|E|^{-n} < \delta$, where $|E|$ stands for the cardinal number of E. Then, any subset $A \subset Y^{\mathbb{N}}$ satisfying

$$\mu^\gamma_\varepsilon(A) > 1 - \delta$$

must contain an (n, ε)-separated subset $\mathcal{F}$ whose cardinal number satisfies

$$|\mathcal{F}| \geq |E|^n - L > (1 - \delta)|E|^n$$

since $Y^{\mathbb{N}} \setminus A$ can contain at most L sub-cylinders $\{y_1\} \times \cdots \times \{y_n\} \times Y \times Y \times \cdots$. Thus,

$$h^K_{\mu^\gamma_\varepsilon}(\varepsilon, \delta) \geq \log |E|.$$

Let $(\varepsilon^\gamma_n)_n$ be a sequence of positive real numbers converging to 0 such that

- $\overline{\dim}_B B_\gamma(y) = \lim_{n \to +\infty} \frac{\log |E(\varepsilon^\gamma_n)|}{\log(1/\varepsilon^\gamma_n)}$;
- $\lim_{n \to +\infty} \mu^\gamma_{\varepsilon^\gamma_n} = \mu^\gamma \in \mathcal{P}(Y^{\mathbb{N}})$ in the weak*-topology.

Then,

$$H^K_\delta(\mu^\gamma) \,\geq\, \limsup_{n\to+\infty} \frac{h^K_{\mu^\gamma_{\varepsilon^\gamma_n}}(\varepsilon^\gamma_n,\delta)}{\log(1/\varepsilon^\gamma_n)} \,\geq\, \limsup_{n\to+\infty} \frac{\log|E(\varepsilon^\gamma_n)|}{\log(1/\varepsilon^\gamma_n)} \,=\, \overline{\dim}_B\, B_\gamma(y).$$

As $\operatorname{supp}(\mu^\gamma) \subset B_\gamma(y)^{\mathbb{N}}$, we also have

$$\lim_{\gamma\to 0^+} \mu^\gamma \,=\, \delta^{\mathbb{N}}_{\{y\}}$$

so, by the upper semi-continuity of H^K_δ, we get

$$H^K_\delta(\delta^{\mathbb{N}}_{\{y\}}) \,\geq\, \limsup_{\gamma\to 0^+} H^K_\delta(\mu^\gamma) \,\geq\, \lim_{\gamma\to 0^+} \overline{\dim}_B\, B_\gamma(y).$$

□

The next proposition generalizes the previous information.

Proposition 8.3. *Given $\delta \in]0,1[$ and $\rho > 1$, for every periodic point by σ with minimal period p, say $\xi = (y_1,\cdots,y_p,y_1,\cdots,y_p,\cdots) \in Y^{\mathbb{N}}$, the Dirac periodic probability measure $\mu_\xi \in \mathcal{P}_\sigma(Y^{\mathbb{N}})$ supported on the orbit of ξ satisfies*

$$\frac{1}{p}\lim_{\gamma\to 0^+} \overline{\dim}_B\, B^\gamma_p(\xi) \,\leq\, H^K_\delta(\mu_\xi)$$

where $B^\gamma_p(y) = B_\gamma(y_1)\times\cdots\times B_\gamma(y_p) \subset Y^p$ and $B_\gamma(a)$ stands for the open ball in Y centered at a with radius γ.

Proof. Endow the space Y^p with the metric

$$d^{\max}\big((z_1,z_2,\cdots,z_p),(w_1,w_2,\cdots,w_p)\big) \,=\, \max\big\{d(z_1,w_1),\cdots,d(z_p,w_p)\big\}$$

and consider in $(Y^p)^{\mathbb{N}}$ the metric $d^{\max}_{\rho^p}$ given by (13) when applied to the metric space $(Y^p, d^{\max})$ and the value ρ^p:

$$d^{\max}_{\rho^p}(\alpha,\beta) \,=\, \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\leq i\leq p} d(\alpha_{i+np}-\beta_{i+np})}{\rho^{pn}}.$$

Then the map

$$\begin{array}{rccc} \Phi\colon & \big(Y^{\mathbb{N}}, d_\rho, \sigma^p\big) & \to & \big((Y^p)^{\mathbb{N}}, d^{\max}_{\rho^p}, \sigma\big) \\ & (z_1,z_2,\cdots) & \mapsto & \big((z_1,\cdots,z_p),(z_{p+1},\cdots,z_{2p}),\cdots\big) \end{array}$$

satisfies $\Phi\circ\sigma^p = \sigma\circ\Phi$ and is bi-Lipschitz. In fact, given $z = (z_1,z_2,\cdots)$ and $w = (w_1,w_2,\cdots)$ in $Y^{\mathbb{N}}$, one has

$$\begin{aligned} d^{\max}_{\rho^p}(\Phi(z),\Phi(w)) &= \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\leq i\leq p} d(z_{i+np}-w_{i+np})}{\rho^{pn}} \\ &\leq \rho^{p-1} \sup_{n\in\mathbb{N}\cup\{0\}}\ \max_{1\leq i\leq p} \frac{|z_{i+np}-w_{i+np}|}{\rho^{i+np-1}} \\ &= \rho^{p-1} d_\rho(z,w) \end{aligned}$$

and

$$\begin{aligned} d_{\rho^p}^{\max}(\Phi(z),\Phi(w)) &= \sup_{n\in\mathbb{N}\cup\{0\}} \frac{\max_{1\le i\le p} d(z_{i+np}-w_{i+np})}{\rho^{pn}} \\ &\ge \sup_{n\in\mathbb{N}\cup\{0\}} \max_{1\le i\le p} \frac{d(z_{i+np}-w_{i+np})}{\rho^{i+pn-1}} \\ &= d_\rho(z,w). \end{aligned}$$

Now, take a periodic point $\xi = (y_1,\cdots,y_p,y_1,\cdots,y_p,\cdots) \in Y^{\mathbb{N}}$ by σ, with minimal period p, and let μ_ξ be the Dirac periodic probability measure supported on the orbit of ξ. We notice that, for each $i \in \{1,\cdots,p\}$, the push-forward $\Phi_*(\delta_{\{\sigma^i(\xi)\}})$ is the Dirac measure $\delta_{\{\Phi(\sigma^i(\xi))\}}$ in $(Y^p)^{\mathbb{N}}$ supported on the fixed point

$$\Phi(\sigma^i(\xi)) \,=\, \Big((y_{1+i},\cdots,y_p,y_1,\cdots,y_i),(y_{1+i},\cdots,y_p,y_1,\cdots,y_i),\cdots\Big).$$

For every $\gamma > 0$ and $\varepsilon > 0$, consider the probability measure $\mu_\varepsilon^\gamma(i)$ in $(Y^p)^{\mathbb{N}}$ defined in (30), within the proof of Lemma 8.2, but now with respect to the fixed point $\Phi(\sigma^i(\xi))$ by σ. Then,

$$h^K_{\mu_\varepsilon^\gamma(i)}(\varepsilon,\delta) \,\ge\, \log|E_i(\varepsilon)|$$

where $E_i(\varepsilon)$ is a maximal ε-separated subset of

$$B_i^\gamma \,=\, B_\gamma(y_{i+1},\cdots,y_p,y_1,\cdots,y_i) \,=\, B_\gamma(y_{i+1})\times\cdots\times B_\gamma(y_p)\times B_\gamma(y_1)\times\cdots\times B_\gamma(y_i).$$

Therefore,

$$h^K_{\frac{1}{p}\sum_{i=0}^{p-1}\mu_\varepsilon^\gamma(i)}(\varepsilon,\delta) \,=\, \frac{1}{p}\sum_{i=0}^{p-1} h^K_{\mu_\varepsilon^\gamma(i)}(\varepsilon,\delta) \,\ge\, \frac{1}{p}\sum_{i=0}^{p-1}\log|E_i(\varepsilon)|.$$

We now observe that $B_i^\gamma(\xi)$ is given by permutations in the coordinates of $B_p^\gamma(\xi)$ for every $i = 1,...,p$. Hence they all have the same upper box dimension, which can be estimated using the same sequence of scales, say $(\varepsilon_n^\gamma)_n$. Assume (by taking a subsequence if necessary) that those scales are such that the sequence $\big(\mu^\gamma_{\varepsilon_n^\gamma}(i)\big)_n$ converges in the weak*-topology, say

$$\lim_{n\to+\infty} \mu^\gamma_{\varepsilon_n^\gamma}(i) \,=\, \mu^\gamma(i) \qquad \forall\, i\in\{1,\cdots,p\}.$$

Consequently,

$$H_\delta^K\Big(\frac{1}{p}\sum_{i=1}^p \mu^\gamma(i)\Big) \,\ge\, \limsup_{n\to+\infty} \frac{h^K_{\frac{1}{p}\sum_{i=0}^{p-1}\mu^\gamma_{\varepsilon_n^\gamma}(i)}(\varepsilon_n^\gamma,\delta)}{\log(1/\varepsilon_n^\gamma)} \,\ge\, \limsup_{n\to+\infty}\frac{1}{p}\sum_{i=1}^p \frac{\log|E_i(\varepsilon_n^\gamma)|}{\log(1/\varepsilon_n^\gamma)} \,=\, \overline{\dim}_B\, B_p^\gamma(\xi).$$

Moreover, as $\operatorname{supp}(\mu^\gamma(i)) \subset B_i^\gamma$ for every i,

$$\lim_{\gamma\to 0^+} \frac{1}{p}\sum_{i=1}^p \mu^\gamma(i) \,=\, \frac{1}{p}\sum_{i=1}^p \delta_{\{\Phi(\sigma^i(\xi))\}} \,=\, \frac{1}{p}\sum_{i=1}^p \Phi_*\delta_{\{\sigma^i(\xi)\}} \,=\, \Phi_*(\mu_\xi).$$

Finally, by the upper semi-continuity of H_δ^K,

$$H_\delta^K(\Phi_*(\mu_\xi)) \,\ge\, \lim_{\gamma\to 0^+} \overline{\dim}_B\, B_p^\gamma(\xi).$$

We are left to relate $H_\delta^K(\Phi_*(\mu_\xi))$ with $H_\delta^K(\mu_\xi)$. The following lemma explains why H_δ^K is preserved under bi-Lipschitz conjugations.

Lemma 8.4. *Let $T\colon X\to X$ and $T'\colon X'\to X'$ be continuous maps on compact metric spaces (X,d) and (X',d'), and $\Phi\colon X\to X'$ be a bi-Lipschitz conjugation between them. Then,*

$$H^K_\delta(X,d,T,\mu) \;=\; H^K_\delta(X',d',T',\Phi_*(\mu)) \qquad \forall\,\mu\in\mathcal{P}_T(X).$$

Proof. We start by recalling a few simple facts about the push-forward map $\mu\mapsto\Phi_*\mu=\mu\circ\Phi^{-1}$:
(a) A sequence $(\mu_\varepsilon)_\varepsilon$ in $\mathcal{P}_T(X)$ converges to $\mu\in\mathcal{P}_T(X)$ if and only if $\Phi_*\mu_\varepsilon$ in $\mathcal{P}_{T'}(X')$ converges to $\Phi_*\mu\in\mathcal{P}_{T'}(X')$.
(b) $m\in\mathcal{E}_T(X)$ if and only if $\Phi_*m\in\mathcal{E}_{T'}(X')$.
(c) $\nu=\int m\,d\mathbb{P}_\nu(m)$ is the ergodic decomposition of $\nu\in\mathcal{P}_T(X)$ if and only if $\Phi_*\nu=\int\Phi_*(m)\,d\mathbb{P}_\nu(m)$ is the ergodic decomposition of $\Phi_*\nu\in\mathcal{P}_{T'}(X')$.

Take now constants $C_1,C_2>0$ such that

$$C_1\,d(x,x')\;\le\;d'(\Phi(x),\Phi(x'))\;\le\;C_2\,d(x,y)\qquad\forall\,x,\,x'\in X.$$

Then, by considering separated subsets, we get for any $A\subset X$

$$S(A,d_n,0,\varepsilon)\;\le\;S(\Phi(A),d'_n,0,C_1\varepsilon)\;\le\;S(A,d_n,0,C_3\varepsilon)$$

where $C_3=C_1/C_2$. Thus, for every $m\in\mathcal{E}_T(X)$,

$$\begin{aligned}N_m(\varepsilon,\delta,n)&=\;\inf\{S(A,d_n,0,\varepsilon)\;:\;m(A)>1-\delta\}\\&\le\;\inf\{S(\Phi(A),d'_n,0,C_1\varepsilon)\;:\;\Phi_*m(\Phi(A))>1-\delta\}\\&=N_{\Phi_*m}(C_1\varepsilon,\delta,n)\\&\le\inf\{S(A,d_n,0,C_3\varepsilon)\;:\;m(A)>1-\delta\}\\&=N_m(C_3\varepsilon,\delta,n).\end{aligned}$$

Hence,

$$h^K_m(\varepsilon,\delta)\;\le\;h^K_{\Phi_*m}(C_1\varepsilon,\delta)\;\le\;h^K_m(C_3\varepsilon,\delta).$$

Consequently, for every $\nu\in\mathcal{P}_T(X)$ with ergodic decomposition given by $\nu=\int m\,d\mathbb{P}_\nu(m)$ we have

$$\begin{aligned}h^K_\nu(\varepsilon,\delta)&=\int h^K_m(\varepsilon,\delta)\,d\mathbb{P}_\nu(m)\\&\le\int h^K_{\Phi_*m}(C_1\varepsilon,\delta)\,d\mathbb{P}_\nu(m)\\&=h^K_{\Phi_*\nu}(C_1\varepsilon,\delta)\\&\le\int h^K_m(C_3\varepsilon,\delta)\,d\mathbb{P}_\nu(m)\\&=h^K_\nu(C_3\varepsilon,\delta).\end{aligned}$$

Thus,

$$h^K_\nu(\varepsilon,\delta)\;\le\;h^K_{\Phi_*\nu}(C_1\varepsilon,\delta)\;\le\;h^K_\nu(C_3\varepsilon,\delta).$$

We now proceed to evaluate H^K_δ. Given $\mu\in\mathcal{P}_T(X)$, take sequences $\mu_n\to\mu$ and $\varepsilon_n\to0^+$ such that

$$H^K_\delta(\mu)\;=\;\lim_{n\to+\infty}\frac{h^K_{\mu_n}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)}.$$

Then,

$$H_\delta^K(\mu) \,\leq\, \limsup_{n\to+\infty} \frac{h^K_{\Phi_*\mu_n}(C_1\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \,\leq\, H_\delta^K(\Phi_*\mu).$$

For the converse inequality, take sequences $\mu_n \to \mu$ and $\varepsilon_n \to 0^+$ such that

$$H_\delta^K(\Phi_*\mu) \,=\, \lim_{n\to+\infty} \frac{h^K_{\Phi_*\mu_n}(C_1\varepsilon_n,\delta)}{\log(1/C_1\varepsilon_n)}.$$

Then,

$$H_\delta^K(\Phi_*\mu) \,\leq\, \limsup_{n\to+\infty} \frac{h^K_{\mu_n}(C_3\varepsilon_n,\delta)}{\log(1/C_1\varepsilon_n)} \,\leq\, H_\delta^K(\mu).$$

□

It is equally straightforward to show the following power rule for the map H_δ^K.

Lemma 8.5. *Let (X,d,T) be as above and $p \in \mathbb{N}$. Then*

$$H_\delta^K(X,d,T^p,\mu) \,\leq\, p\, H_\delta^K(X,d,T,\mu) \qquad \forall\, \mu \in \mathcal{P}_T(X).$$

To complete the proof of Proposition 8.3, we summon Lemmas 8.4 and 8.5 to deduce that

$$\lim_{\gamma\to 0^+} \dim_B B_p^\gamma(x) \,\leq\, H_\delta^K\Big((Y^p)^{\mathbb{N}}, d_{\rho^p}^{\max}, \sigma, \Phi_*\mu_\xi\Big) \,=\, H_\delta^K\Big(Y^{\mathbb{N}}, d_\rho, \sigma^p, \mu_\xi\Big) \,\leq\, pH_\delta^K\Big(Y^{\mathbb{N}}, d_\rho, \sigma, \mu_\xi\Big).$$

□

Let us resume the proof of Theorem 8.1. Under the assumption that every nonempty open set in $U \subset Y$ satisfies $\dim_B U = \dim_B Y$, we know that

$$\dim_B\, B_p^\gamma(\xi) \,=\, \dim_B\, \Big(B_\gamma(y_1) \times \cdots \times B_\gamma(y_p)\Big) \,=\, \sum_{i=1}^p \dim_B\, B_\gamma(y_i).$$

Therefore, by Proposition 8.3, for every periodic probability measure μ_ξ supported on a periodic point $\xi = (y_1, \cdots, y_p, y_1, \cdots)$ with minimal period p, we have

$$\dim_B Y \,=\, \frac{1}{p}\sum_{i=1}^p \lim_{\gamma\to 0^+} \dim_B\, B_\gamma(y_i) \,=\, \frac{1}{p}\lim_{\gamma\to 0^+} \dim_B\, B_p^\gamma(\xi) \,\leq\, H_\delta^K\Big(Y^{\mathbb{N}}, d_\rho, \sigma, \mu_\xi\Big).$$

In addition, by the specification property of the shift map (cf. [23, Proposition 2]), the set of periodic probability measures is dense in $\mathcal{P}_\sigma(Y^{\mathbb{N}})$. Thus, from the upper semi-continuity of H_δ^K we obtain

$$H_\delta^K(\mu) \,\geq\, \dim_B Y \qquad \forall\, \mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}}).$$

We end the proof of Theorem D by bringing together the last inequality and (28). □

Remark 8.6. *We note that, under the assumptions of Theorem D, one has*

$$M(\mu) \,=\, \dim_B Y \qquad \forall\, \mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}}).$$

8.1. **Dimension homogeneity assumption.** We stress that some homogeneity hypothesis on the box dimension structure of the space Y is indeed necessary. For instance, if a space of positive upper box dimension contains an isolated point (as happens with $Y = \{0\} \cup \{1/n\colon\, n \in \mathbb{N}\}$ endowed with the Euclidean metric, whose box dimension is $1/2$; cf. [13, Lemma 3.1]), then Theorem 8.1 does not hold for the Dirac probability measure supported on that point. Consequently, Theorem D is not valid for a well chosen potential, as the following result illustrates.

Proposition 8.7. *Let (Y, d) be a compact metric space and σ the shift map on $(Y^{\mathbb{N}}, d_\rho)$.*

(a) Given $y \in Y$, one has

$$H^K_\delta(\delta^{\mathbb{N}}_{\{y\}}) \;=\; M(\delta^{\mathbb{N}}_{\{y\}}) \;=\; \lim_{\gamma\to 0^+} \overline{\dim}_B\, B_\gamma(y).$$

(b) Assume that (Y, d) has positive upper box dimension and contains a point y_0 such that

$$\lim_{\gamma\to 0^+} \overline{\dim}_B\, B_\gamma(y_0) \;<\; \overline{\dim}_B\, Y.$$

Then there exists $\varphi \in C^0(Y^{\mathbb{N}})$ such that

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi) \;<\; \overline{\dim}_B Y + \max_{\mu\in\mathcal{E}_\sigma(Y^{\mathbb{N}})} \int \varphi\, d\mu.$$

Proof. (a) Fix $y \in Y$. For each $\gamma > 0$ consider a map $\varphi_\gamma \in C^0(Y^{\mathbb{N}})$ satisfying

$$\varphi_\gamma(x_1, x_2, \cdots) \;=\; \begin{cases} 0 & \text{if } x_1 \in B_{\gamma/2}(y) \\ -\overline{\dim}_B Y - 1 & \text{if } x_1 \notin B_\gamma(y) \end{cases}$$

and always bounded above by 0. Given $\varepsilon > 0$, let $U_1, \cdots, U_N$ be a minimal open cover of $B_\gamma(y)$ with diameter at most ε. Complete it to an open cover of the whole space Y by adding a minimal ε-cover $U_{N+1}, \cdots, U_{N+P}$ of the complement of $\cup_{i=1}^n U_i$. Without loss of generality we may suppose that none of the sets $U_{N+1}, \cdots, U_{N+P}$ intersects $B_\gamma(y)$. Take $\ell = \ell(\varepsilon)$ such that $\mathrm{diam}(Y, d)/\rho^{\ell-1} < \varepsilon$ and consider the open cover of $Y^{\mathbb{N}}$ given by

$$\alpha \;=\; \big\{U_{j_1} \times \cdots \times U_{j_{\ell+n}} \times Y \times Y \times \cdots \;:\; 1 \le j_1, \cdots, j_{\ell+n} \le N + P\big\}.$$

Then the diameter of α with respect to the metric $d_{\rho,n}$ satisfies $\mathrm{diam}(\alpha, d_{\rho,n}) \le \varepsilon$ and

$$\begin{aligned}
\sum_{U\in\alpha} (1/\varepsilon)^{\sup_U S_n\varphi_\gamma} &= \sum_{U_{j_1}\times\cdots\times U_{j_{\ell+n}}} (1/\varepsilon)^{\sum_{v=1}^{\ell+n} \sup_{U_{j_v}} \varphi_\gamma} \\
&= (N+P)^\ell \sum_{U_{j_1}\times\cdots\times U_{j_n}} (1/\varepsilon)^{\sum_{v=1}^{n} \sup_{U_{j_v}} \varphi_\gamma} \\
&\le (N+P)^\ell \sum_{U_{j_1}\times\cdots\times U_{j_n}} (1/\varepsilon)^{-(\overline{\dim}_B Y+1)\,\#\{v\colon\, j_v\in\{N+1,\cdots,N+P\}\}} \\
&= (N+P)^\ell \Big(\sum_{k=0}^{n} \binom{n}{k} P^k N^{n-k} \,(\varepsilon^{\overline{\dim}_B Y+1})^k\Big) \\
&= (N+P)^\ell \Big(N + P\varepsilon^{\overline{\dim}_B Y+1}\Big)^n.
\end{aligned}$$

We observe that as P is smaller than the ε-covering number of Y then one has $P \leq \varepsilon^{-(\overline{\dim}_B Y + 1/2)}$ for small enough $\varepsilon > 0$. Thus, for those values of ε,

$$P(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi_\gamma, \varepsilon) \,\leq\, \log\Big(N + P\varepsilon^{\overline{\dim}_B Y + 1}\Big) \,\leq\, \log\Big(N \,+\, \varepsilon^{1/2}\Big)$$

and so

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi_\gamma) \,\leq\, \limsup_{\varepsilon \to 0^+} \frac{\log\Big(N \,+\, \varepsilon^{1/2}\Big)}{\log(1/\varepsilon)} \,=\, \overline{\dim}_B\, B_\gamma(y). \tag{31}$$

On the other hand,

$$\max_{\mu \in \mathcal{E}_\sigma(Y^{\mathbb{N}})} \int \varphi_\gamma \, d\mu \,=\, \int \varphi_\gamma \, d\delta^{\mathbb{N}}_{\{y\}} = 0 \qquad \forall \gamma > 0. \tag{32}$$

Therefore

$$\begin{aligned} M(\delta^{\mathbb{N}}_{\{y\}}) &= \inf_{\psi \in C^0(Y^{\mathbb{N}})} \Big\{\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d, \sigma, \psi) - \int \psi \, d\delta^{\mathbb{N}}_{\{y\}}\Big\} \\ &\leq \lim_{\gamma \to 0^+} \overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d, \sigma, \varphi_\gamma) \\ &\leq \lim_{\gamma \to 0^+} \overline{\dim}_B\, B_\gamma(y). \end{aligned}$$

To complete the proof of item (a) we summon Lemmas 8.2 and 7.5.

(b) Let $y_0 \in Y$ be such that $\lim_{\gamma \to 0^+} \overline{\dim}_B\, B_\gamma(y_0) \,<\, \overline{\dim}_B\, Y$ and $\gamma_0 > 0$ such that

$$\overline{\dim}_B\, B_{\gamma_0}(y_0) \,<\, \overline{\dim}_B\, Y.$$

Consider φ_{γ_0} associated to y_0 as previously defined. Then, taking into account (31) and (32) we conclude that

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi_{\gamma_0}) \,<\, \overline{\dim}_B\, Y \,+\, \max_{\mu \in \mathcal{E}_\sigma(Y^{\mathbb{N}})} \int \varphi_{\gamma_0} \, d\mu.$$

□

9. Proof of Theorem E

Recall from Section 3 that the map $\mathcal{D}\colon X \to \mathbb{R}$ is given by

$$\mathcal{D}(x) \,=\, \inf\Big\{\overline{\mathrm{mdim}}_M(\overline{U}, d, T) \,:\, U \text{ is an open neighborhood of } x\Big\}.$$

Even though the main results we show in this subsection assume invertibility of the dynamics, we define $\mathcal{D}$ for possibly non-injective maps. Unless stated otherwise, in what follows we will assume that T is just continuous. As one can see in Examples (10.3) and (10.5), which are far from being injective, the above concept is closely related to the map H^K_δ. We will show this is true in general if the map T is a homeomorphism.

We start by establishing some useful properties of the map $\mathcal{D}$.

Lemma 9.1. *The map $\mathcal{D}$ is upper semi-continuous and*

$$0 \,\leq\, \mathcal{D} \,\leq\, \overline{\mathrm{mdim}}_M(X, d, T).$$

In particular, $\mathcal{D}$ is measurable and μ-integrable with respect to every $\mu \in \mathcal{P}(X)$.

Proof. Given $x \in X$ and an open neighborhood U of x, let $(x_n)_{n\in\mathbb{N}}$ be a sequence of elements in U converging to x. Then, for sufficiently large n, one has $x_n \in U$ and

$$\mathcal{D}(x_n) \,\leq\, \overline{\mathrm{mdim}}_M(\overline{U}, d, T).$$

We conclude the proof by taking $\limsup_n$ and $\inf_U$. □

Lemma 9.2. *For any $\mu \in \mathcal{E}_T(X)$, the map $\mathcal{D}$ is constant μ-almost everywhere.*

Proof. We begin by showing that $\mathcal{D}$ is increasing along orbits. Given an open nonempty subset U of X, let $E \subset \overline{U}$ be $(n+1,\varepsilon)$-separated. Then, by the Pigeonhole Principle, there is an (n,ε)-separated subset $A \subset T(E) \subset \overline{T(U)}$ with cardinality $\#A \geq \#E/S_1(X,d,0,\varepsilon)$. Hence,

$$S_1(\overline{T(U)}, d_n, 0, \varepsilon) \,\geq\, S_1(\overline{U}, d_{n+1}, 0, \varepsilon)/S_1(X,d,0,\varepsilon)$$

and so $\mathcal{D}(x) \leq \mathcal{D}(T(x))$ for all $x \in X$.

We observe now that if $x \in X$ is recurrent, that is, $\lim_{k\to+\infty} T^{n_k}(x) = x$ for some subsequence of the orbit of x by T, then by the upper semi-continuity of $\mathcal{D}$ one has

$$\limsup_{k\to+\infty} \mathcal{D}(T^{n_k}(x)) \leq \mathcal{D}(x).$$

As $\mathcal{D}$ is increasing along orbits, the previous inequality implies that $\mathcal{D}$ is constant along orbits of recurrent points. Thus, given $\mu \in \mathcal{E}_T(X)$, by the Poincaré Recurrence Theorem we deduce that $\mathcal{D}$ is constant μ-almost everywhere. □

9.1. **Linking $\mathcal{D}$ and $\overline{\mathrm{mdim}}_M(X,d,T,\varphi)$.** In the remaining of this section, we assume that the map $T\colon X \to X$ is a homeomorphism. We will prove a general result, taking into account the role of the potentials, of which Theorem E is an immediate consequence.

Theorem 9.3. *Let (X,d) be a compact metric space and $T\colon X \to X$ be a homeomorphism such that $\overline{\mathrm{mdim}}_M(X,d,T) < +\infty$. Then, for every $\varphi \in C^0(X)$,*

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) \,\leq\, \max_{\mu\,\in\,\mathcal{P}_T(X)} \int \big(\mathcal{D}+\varphi\big)(x)\,d\mu(x) \,=\, \max_{\mu\,\in\,\mathcal{E}_T(X)} \int \big(\mathcal{D}+\varphi\big)(x)\,d\mu(x).$$

Proof. Fix $\delta \in]0,1[$ and let μ_0 be the equilibrium state constructed in the proof of Theorem C, namely the probability measure which is the weak*-limit of the sequence $\big(\mu_{\varepsilon_n}\big)_{n\,\in\,\mathbb{N}}$ in $\mathcal{E}_T(X)$ as $(\varepsilon_n)_n$ goes to 0 and satisfies

$$\overline{\mathrm{mdim}}_M(X,d,T,\varphi) = H^K_\delta(\mu_0) + \int \varphi\,d\mu_0 = \lim_{n\to\infty} \frac{h^K_{\mu_{\varepsilon_n}}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + \int \varphi\,d\mu_0.$$

Note that, given an open set $U \subset X$ such that $\mu(U) > 0$, for large enough n one has $\mu_{\varepsilon_n}(U) > 0$.

Lemma 9.4. [32, Theorem 3.7] *If $T\colon X \to X$ is a homeomorphism on a compact metric space (X,d), for every ergodic probability measure $\mu \in \mathcal{E}_T(X)$, every open set $U \subset X$ such that $\mu(U) > 0$, every $\varepsilon > 0$ and every $\delta \in]0,1[$ we have*

$$P(\overline{U}, d, T, 0, \varepsilon) \,\geq\, h^K_\mu(\varepsilon,\delta).$$

By Lemma 9.4 and Remark 7.7 we have

$$H^K_\delta(\mu_0) \,=\, \lim_{n\,\to\,+\infty} \frac{h^K_{\mu_{\varepsilon_n}(\varepsilon_n,\delta)}}{\log(1/\varepsilon_n)} \,\leq\, \overline{\mathrm{mdim}}_M(\overline{U}, d, T).$$

Thus, for any $x \in \text{supp}(\mu_0)$,

$$H^K_\delta(\mu_0) \,\leq\, \inf_{x\,\in\, U}\, \overline{\text{mdim}}_M(\overline{U}, d, T) \,=\, \mathcal{D}(x).$$

Therefore,

$$H^K_\delta(\mu_0) \leq \inf_{x\,\in\,\text{supp}(\mu_0)} \mathcal{D}(x) \leq \int \mathcal{D}\, d\mu_0.$$

Hence,

$$\overline{\text{mdim}}_M(X, d, T, \varphi) \,\leq\, \int (\mathcal{D} + \varphi)\, d\mu_0 \,\leq\, \max_{\mu\,\in\,\mathcal{P}_T(X)} \int (\mathcal{D} + \varphi)(x)\, d\mu(x).$$

As $\mu \mapsto \int \mathcal{D} d\mu$ is affine, the previous maximum is also attained at some ergodic probability measure. This ends the proof of Theorem 9.3.

□

Let us resume the proof of Theorem E. The equality in its statement follows from Theorem 9.3, when $\varphi \equiv 0$, and the fact that $\mathcal{D} \leq \overline{\text{mdim}}_M(X, d, T)$ (cf. Lemma 9.1).

Now consider a measure $\mu \in \mathcal{P}_T(X)$ satisfying

$$\overline{\text{mdim}}_M(X, d, T) \,=\, \int \mathcal{D}\, d\mu.$$

Then, as $\mathcal{D} \leq \overline{\text{mdim}}_M(X, d, T)$, there exists a full measure set $Z \subset X$ such that $\mathcal{D}|_Z \equiv \overline{\text{mdim}}_M(X, d, T)$. Therefore, since $\text{supp}(\mu) \subset \overline{Z}$ and $\mathcal{D}$ is upper semi-continuous, we deduce that $\mathcal{D}|_{\text{supp}(\mu)} \equiv \overline{\text{mdim}}_M(X, d, T)$. The proof of Theorem E is complete. □

Remark 9.5. *A strict inequality may happen in Theorem 9.3. For instance, if one considers the shift map on $Y^{\mathbb{Z}}$, we always have*

$$\mathcal{D} \equiv \overline{\dim}_B Y$$

and so, if (Y, d) has positive upper box dimension and contains an isolated point, then using the potential φ provided by Proposition 8.7 we get

$$\overline{mdim}_M(Y^{\mathbb{Z}}, d_\rho, \sigma, \varphi) < \max_{\mu\,\in\,\mathcal{E}_\sigma(Y^{\mathbb{Z}})} \int (\mathcal{D} + \varphi)(x)\, d\mu(x).$$

Another consequence of Theorem 9.3 is the following relation between $\mu \mapsto \int \mathcal{D}\, d\mu$ and $\mu \mapsto M(\mu)$ (see (8)).

Corollary 9.6. *Let (X, d) be a compact metric space and $T\colon X \to X$ be a homeomorphism such that $\overline{\text{mdim}}_M(X, d, T) < +\infty$. Then,*

$$M(\mu) \,\leq\, \int \mathcal{D}\, d\mu \qquad \forall\, \mu \,\in\, \mathcal{P}_T(X).$$

In addition, the equality in Theorem 9.3 holds for every $\varphi \in C^0(X)$ if and only if

$$M(\mu) \,=\, \int \mathcal{D}\, d\mu \qquad \forall\, \mu \,\in\, \mathcal{P}_T(X).$$

Proof. We recall that M is given by $M(\mu) = \inf_{\varphi \in \mathcal{C}_\Gamma} \int \varphi \, d\mu$, where

$$\mathcal{C}_\Gamma = \{\varphi \in C^0(X) : \overline{\mathrm{mdim}}_M(X, d, T, -\varphi) \leq 0\}.$$

Now observe that, by Theorem 9.3, for every continuous potential satisfying $\psi \geq \mathcal{D}$ we have

$$\overline{\mathrm{mdim}}_M(X, d, T, -\psi) \leq \max_{\mu \in \mathcal{P}_T(X)} \int \big(\mathcal{D} - \psi\big) \, d\mu \leq 0$$

that is, $\psi \in \mathcal{C}_\Gamma$. By upper semi-continuity of $\mathcal{D}$, there exists a decreasing sequence of continuous potentials φ_n converging to $\mathcal{D}$ as $n \to \infty$. Then, by the Monotone Convergence Theorem,

$$M(\mu) = \inf_{\varphi \in \mathcal{C}_\Gamma} \int \varphi \, d\mu \leq \lim_{n \to \infty} \int \varphi_n \, d\mu = \int \mathcal{D} \, d\mu \qquad \forall \, \mu \in \mathcal{P}_T(X).$$

To conclude, assume that the inequality in Theorem 9.3 is indeed an equality for any potential $\varphi \in C^0(X)$. Then, $\mu \mapsto \int \mathcal{D} \, d\mu$ satisfies the variational principle (7). By the maximality of $\mu \mapsto M(\mu)$ among all those maps satisfying this variational principle (provided by Theorem A), we get

$$M(\mu) \geq \int \mathcal{D} \, d\mu \qquad \forall \, \mu \in \mathcal{P}_T(X).$$

□

9.2. **Metric mean dimension points.** An element $x \in X$ is said to be as a *metric mean dimension point* if $\mathcal{D}(x) > 0$; it is a *full metric mean dimension point* if $\mathcal{D}(x) = \overline{\mathrm{mdim}}_M(X, d, T)$. Denote the set of such points by $D_p(X, d, T)$ and $D_p^f(X, d, T)$, respectively. It was shown in [19] that $D_p^f(X, d, T) \neq \emptyset$. The following is an immediate consequence of Theorem E:

Corollary 9.7. *Under the assumptions of Theorem E, there exists an ergodic probability measure* $\mu \in \mathcal{E}_T(X)$ *such that* $\mathrm{supp}(\mu) \subseteq D_p^f(X, d, T)$.

10. Examples

The first example confirms that the measure-theoretic upper metric mean dimension defined in the statement of Theorem A depends on the metric.

Example 10.1. Consider the symbolic space $Y = \{0, 1\}^{\mathbb{N}}$ with the product topology. This topology is generated by both distances

$$d_1(x, y) = \sum_{n \in \mathbb{N}} \frac{|x_n - y_n|}{2^n} \quad \text{and} \quad d_2(x, y) = \sum_{n \in \mathbb{N}} \frac{|x_n - y_n|}{3^n}.$$

Endowing the product space $Y^{\mathbb{N}}$ with the metrics D_j, for $j = 1, 2$, given by

$$D_j(a, b) = \sup_{n \in \mathbb{N}} \frac{d_j(a_n, b_n)}{2^n}$$

which induce the same topology in $Y^{\mathbb{N}}$, and considering in $Y^{\mathbb{N}}$ the shift map σ, we get (cf. [15])

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, D_1, \sigma, 0) = \dim_B(Y, d_1) = 1$$

while

$$\overline{\mathrm{mdim}}_M(Y^{\mathbb{N}}, D_2, \sigma, 0) = \dim_B(Y, d_2) = \log 2 / \log 3.$$

Let M_1 and M_2 denote the measure-theoretic upper metric mean dimension maps assigned by Theorem A to $(Y^{\mathbb{N}}, D_1, \sigma)$ and $(Y^{\mathbb{N}}, D_2, \sigma)$, respectively, when applied to the families of ε-pressure functions $\big(P(Y^{\mathbb{N}}, D_1, \sigma, \cdot, \varepsilon)\big)_{0<\varepsilon<1}$ and $\big(P(Y^{\mathbb{N}}, D_2, \sigma, \cdot, \varepsilon)\big)_{0<\varepsilon<1}$ introduced in Subsection 2.1. Thus

$$\max_{\mu \in \mathcal{P}(Y^{\mathbb{N}})} M_1(\mu) \;=\; \frac{\log 2}{\log 3} \;<\; 1 \;=\; \max_{\mu \in \mathcal{P}(Y^{\mathbb{N}})} M_2(\mu).$$

The next two examples comprise continuous maps with the same positive upper metric mean dimension, though their dynamical traits and corresponding sources of complexity are different.

Example 10.2. Denote by $|\cdot|$ the Euclidean distance in $[0,1]$ and, given any $\rho > 1$, consider the space $[0,1]^{\mathbb{N}}$ with the metric d_ρ (defined in (13)). Let $\sigma\colon [0,1]^{\mathbb{N}} \to [0,1]^{\mathbb{N}}$ be the shift map. From Theorem 8.1 we already know that, for every $\delta \in]0,1[$,

$$H^K_\delta(\mu) = 1 \qquad \forall\, \mu \in \mathcal{P}_\sigma([0,1]^{\mathbb{N}}). \tag{33}$$

Nevertheless, we can improve the argument in the proof of Lemma 8.2 when computing H^K_δ at a Dirac measure supported on a fixed point, thereby showing that each approximating measure μ_ε detects separation at *all* scales.

Given $x \in [0,1]$ and $\varepsilon > 0$, denote by $I(\varepsilon)$ the amount of odd numbers between 1 and $\lceil 1/\varepsilon \rceil$ (thus $I(\varepsilon) \geq 1/2\varepsilon$) and by $B = B_\varepsilon(x) \subset [0,1]$ the open interval of radius ε centered at x. Partition B into $\lceil 1/\varepsilon \rceil$ intervals with the same length, say $B(1), \cdots, B(\lceil 1/\varepsilon \rceil)$. Iterate by partitioning $B(i_1, \cdots, i_n)$ into $B(i_1, \cdots, i_n, 1), \cdots, B(i_1, \cdots, i_n, \lceil 1/\varepsilon \rceil)$ so that the length of $B(i_1, \cdots, i_n)$ is at least ε^{n+1}. We observe that, if

$$\begin{aligned} y &\in B(i^1_1, \cdots, i^1_n) \times B(i^2_1, \cdots, i^2_n) \times \cdots \times B(i^k_1, \cdots, i^k_n) \\ z &\in B(j^1_1, \cdots, j^1_n) \times B(j^2_1, \cdots, j^2_n) \times \cdots \times B(j^k_1, \cdots, j^k_n) \end{aligned}$$

where the $i's$ and $j's$ are all odd, then $d_k(y,z) \geq \varepsilon^{n+1}$.

Let ν_ε be a probability measure in $[0,1]$ determined by

- $\nu_\varepsilon(B) = 1$;
- $\nu_\varepsilon(B(i)) = I(\varepsilon)^{-1}$, for every odd i;
- $\nu_\varepsilon(B(i_1, \cdots, i_n)) = I(\varepsilon)^{-n}$, for every odd $i's$.

Define $\mu_\varepsilon = \nu_\varepsilon^{\mathbb{N}}$, which is clearly ergodic, converges to $\delta^{\mathbb{N}}_{\{x\}}$ as $\varepsilon \to 0^+$ and satisfies

$$\mu_\varepsilon\big(B(i^1_1, \cdots, i^1_n) \times B(i^2_1, \cdots, i^2_n) \times \cdots \times B(i^k_1, , i^k_n)\big) \;=\; I(\varepsilon)^{-kn} \qquad \forall \text{ odd } i's.$$

Let $L = L(\varepsilon, k, n, \delta) \in \mathbb{N}$ be the maximal integer such that $L I(\varepsilon)^{-kn} < \delta$. Then, any set A satisfying $\mu_\varepsilon(A) > 1 - \delta$ must intersect at least $I(\varepsilon)^{kn} - L$ cylinders, with odd indices, of the partition level $B(i^1_1, \cdots, i^1_n) \times B(i^2_1, \cdots, i^2_n) \times \cdots \times B(i^k_1, \cdots, i^k_n)$, where

$$I(\varepsilon)^{kn} - L \;>\; (1-\delta) I(\varepsilon)^{kn}$$

since $[0,1]^{\mathbb{N}} \setminus A$ can contain at most L of such sets. Hence, A must have a (d_k, ε^{n+1})-separated subset with cardinality $(1-\delta) I(\varepsilon)^{kn}$. Thus,

$$h^K_{\mu_\varepsilon}(\varepsilon^{n+1}, \delta) \;\geq\; \limsup_{k \to +\infty} \frac{1}{k} \log\big((1-\delta) I(\varepsilon)^{kn}\big) \;=\; \log I(\varepsilon)^n \;\geq\; \log\,(1/2\varepsilon)^n.$$

Fix $\varepsilon > 0$. Given $\varepsilon' > 0$, let $n = n(\varepsilon, \varepsilon')$ be such that $\varepsilon^{n+1} \le \varepsilon' \le \varepsilon^n$. Then

$$\begin{aligned} H_\delta^K(\mu_\varepsilon) &\ge \limsup_{\varepsilon'\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon',\delta)}{\log(1/\varepsilon')} \\ &\ge \liminf_{\varepsilon'\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon',\delta)}{\log(1/\varepsilon')} \\ &\ge \liminf_{n\to+\infty} \frac{h^K_{\mu_\varepsilon}(\varepsilon^n,\delta)}{\log(1/\varepsilon^{n+1})} \\ &\ge \liminf_{n\to+\infty} \frac{n-1}{n+1}\Big(\frac{\log(1/\varepsilon)-\log 2}{\log(1/\varepsilon)}\Big) \\ &= 1 - \frac{\log 2}{\log(1/\varepsilon)}. \end{aligned} \tag{34}$$

Finally, the upper semi-continuity of H_δ^K yields

$$H_\delta^K(\delta^{\mathbb{N}}_{\{x\}}) \ge \limsup_{\varepsilon\to 0^+} H_\delta^K(\mu_\varepsilon) = 1.$$

We emphasize that the interesting feature of this construction, in contrast to the proof of Lemma 8.2, is the fact that the lower bound (34) of $H_\delta^K(\mu_\varepsilon)$ is done by using the measure μ_ε for all scales $\varepsilon' > 0$. More precisely, the inequality

$$\liminf_{\varepsilon'\to 0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon',\delta)}{\log(1/\varepsilon')} \ge 1 - \frac{\log 2}{\log(1/\varepsilon)}$$

indicates that the complexity at every scale $\varepsilon' > 0$ (in terms of metric mean dimension) captured by the measure μ_ε is at least $1-\frac{\log 2}{\log(1/\varepsilon)}$. This means precisely that μ_ε detects separation of points at all scales.

Example 10.3. Consider the interval $[0,1]$ with the metric $|\cdot|$, the map $f\colon [0,1]\to[0,1]$ given by $f(x) = |1-|3x-1||$ and the sequence $(a_n)_{n\in\mathbb{N}\cup\{0\}}$ of numbers in $[0,1[$ whose general term is $a_0 = 0$ and $a_n = \sum_{k=1}^n \frac{6}{\pi k^2}$. For each $n\in\mathbb{N}$, take the interval $J_n = [a_{n-1}, a_n]$ and let $T_n\colon J_n \to [0,1]$ be the unique increasing affine map from J_n onto $[0,1]$. Define

$$\begin{array}{cccc} T\colon & [0,1] & \to & [0,1] \\ & x\in J_n & \mapsto & T_n^{-1}\circ f^n \circ T_n \\ & x=1 & \mapsto & 1 \end{array}$$

which is illustrated in Figure 1.

It is known (cf. [27] or [5]) that $\overline{\text{mdim}}_M([0,1],|\cdot|,T) = 1$, although, for every $n\in\mathbb{N}$, one has $\overline{\text{mdim}}_M(J_n,|\cdot|,T_{|J_n}) = 0$. Let us redo this computation by applying (12) in order to show that the Dirac mass $\delta_{\{1\}}$ is the unique probability measure which maximizes H_δ^K.

Given $n\in\mathbb{N}$, the set J_n can be partitioned into 3^n intervals $J_n(1),\cdots,J_n(3^n)$ with the same length in such a way that

$$T(J_n(i)) = J_n \qquad \forall\, i\in\{1,\cdots,3^n\}.$$

Similarly, $J_n(i)$ can be partitioned into 3^n intervals $J_n(i,1),\cdots,J_n(i,3^n)$ with the same length so that

$$T^2(J_n(i,j)) = J_n \qquad \forall\, i,j\in\{1,\cdots,3^n\}.$$

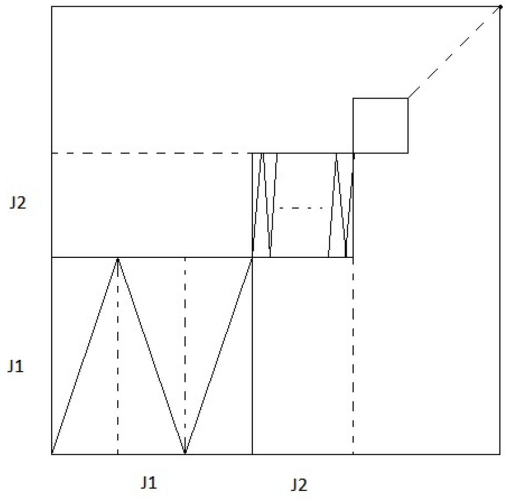

FIGURE 1. Graph of the map T.

Inductively, for every $k \in \mathbb{N}$ and $(i_1, \cdots, i_k)$ such that $i_j \in \{1, \cdots, 3^n\}$, we can split $J_n(i_1, \cdots, i_k)$ into 3^n intervals

$$J_n(i_1, \cdots, i_k, 1),\ J_n(i_1, \cdots, i_k, 2),\ \cdots,\ J_n(i_1, \cdots, i_k, 3^n)$$

with the same length and satisfying

$$T^{k+1} J_n(i_1, \cdots, i_k, i) \;=\; J_n \qquad \forall\, i \in \{1, \cdots, 3^n\}.$$

We may rename these subsets so that the order of the intervals $J_n(i_1, \cdots, i_k, i)$ partitioning $J_n(i_1, \cdots, i_k)$ is increasing in i if i_1 is odd, and decreasing in i if i_1 is even. This choice fits the fact that f^n is increasing in $J_n(i_1)$ if i_1 is odd, and decreasing otherwise. This way, for every $1 \leq j \leq k$, each $x \in J_n(i_1, \cdots, i_k)$ belongs to $J_n(i_j, i_{j+1}, \cdots, i_k)$ after $j-1$ iterates.

We note that each $J_n(i_1, \cdots, i_k)$ has length $|J_n|/3^{kn}$ for every $k \in \mathbb{N}$. Take

$$\varepsilon_n \;=\; |J_n|/3^n \;=\; 6/(\pi^2 n^2 3^n). \tag{35}$$

If $x \in J_n(i_1, \cdots, i_k)$ and $y \in J_n(j_1, \cdots, j_k)$, where $(i_1, \cdots, i_k) \neq (j_1, \cdots, j_k)$ and the i's and the j's are all odd, in at most k iterates their images lie in $J(i_l)$ and $J(j_l)$, respectively, whose distance is at least ε_n; hence $d_k(x, y) \geq \varepsilon_n$.

We are ready to define a suitable sequence of invariant probability measures $(\mu_n)_n$ which converges to $\delta_{\{1\}}$. We start by recalling that, for every $n \in \mathbb{N}$,

$$\#\{1 \leq i \leq 3^n \,:\, i \text{ is odd}\} \;=\; \frac{3^n + 1}{2}.$$

Let μ_n be defined by

- $\mu_n(J_n) = 1$;
- $\mu_n(J_n(i)) = \frac{2}{3^n+1}$ for every odd $1 \leq i \leq 3^n$;
- $\mu_n(J_n(i_1, \cdots, i_k)) = (\frac{2}{3^n+1})^k$ for every odd $1 \leq i_1, \cdots, i_k \leq 3^n$ and every $k \in \mathbb{N}$.

Clearly, $(\mu_n)_n$ converges to $\delta_{\{1\}}$ and each μ_n is T-invariant. In fact,

$$T^{-1}J_n(i_1, \cdots, i_k) \;=\; \bigcup_{i=1}^{3^n} J_n(i, i_1, \cdots, i_k)$$

and

$$\mu(T^{-1}J_n(i_1, \cdots, i_k)) \;=\; \sum_{1\leq i\leq 3^n;\, i \text{ odd}} \Big(\frac{2}{3^n+1}\Big)^{k+1} \;=\; \Big(\frac{2}{3^n+1}\Big)^{k} \;=\; \mu(J_n(i_1, \cdots, i_k)).$$

Moreover, being a Bernoulli probability measure, μ_n is ergodic.

Let $L = L(\delta, n, k) \in \mathbb{N}$ be the maximal positive integer satisfying

$$L\Big(\frac{2}{3^n+1}\Big)^{k} \;<\; \delta.$$

Then, any subset $A \subset [0,1]$ with $\mu_n(A) > 1-\delta$ intersects at least $(\frac{3^n+1}{2})^k - L$ sets $J_n(i_1, \cdots, i_k)$ with odd indices and

$$\Big(\frac{3^n+1}{2}\Big)^{k} - L \;>\; \Big(\frac{3^n+1}{2}\Big)^{k}(1-\delta)$$

since $[0,1] \setminus A$ contains at most L of such intervals. Therefore, A must contain an (ε_n, d_k)-separated subset of cardinality bigger than $\Big(\frac{3^n+1}{2}\Big)^{k}(1-\delta)$ and

$$h^K_{\mu_n}(\varepsilon_n, \delta) \;\geq\; \limsup_{k\to+\infty} \frac{1}{k}\,\log\Big(\Big(\frac{3^n+1}{2}\Big)^{k}(1-\delta)\Big) \;=\; \log\Big(\frac{3^n+1}{2}\Big).$$

Hence, for every $\delta \in]0,1[$,

$$H^K_\delta(\delta_{\{1\}}) \;\geq\; \limsup_{n\to+\infty} \frac{\log(\frac{3^n+1}{2})}{\log(1/\varepsilon_n)} \;=\; \limsup_{n\to+\infty} \frac{\log(\frac{3^n+1}{2})}{\log(\frac{\pi^2 n^2 3^n}{6})} \;=\; 1. \tag{36}$$

Moreover, by Remark 2.3,

$$1 \;= \overline{dim}_B([0,1], |\cdot|) \;\geq\; \overline{\mathrm{mdim}}_M([0,1], |\cdot|, T) \;\geq\; H^K_\delta(\delta_{\{1\}}).$$

Thus,

$$H^K_\delta(\delta_{\{1\}}) \;=\; 1.$$

Let us now use the previous information to evaluate H^K_δ for all T-invariant measures.

Claim: *For every $\delta \in]0,1[$,*

$$H^K_\delta(\mu) \;=\; \mu(\{1\}) \qquad \forall\, \mu \in \mathcal{P}_T([0,1]).$$

In particular, the map H^K_δ is affine and the unique probability measure which maximizes H^K_δ is the Dirac mass $\delta_{\{1\}}$.

Proof of the Claim. We will start by showing that, given $\mu \in \mathcal{P}_T(X)$,

$$\mu(\{1\}) \;=\; 0 \quad\;\Rightarrow\quad\; H^K_\delta(\mu) \;=\; 0.$$

Take one such a μ and $\mu_\varepsilon \in \mathcal{M}(\mu)$. Denote by $\mu_\varepsilon = \int_{\mathcal{E}_T(X)} m\, d\mathbb{P}_\varepsilon(m)$ its ergodic decomposition. Given $\delta \in]0,1[$, select $k = k(\delta)$ such that $\mu([0, a_k[) > 1-\delta$. Then $\mu_\varepsilon([0, a_k[) > 1-\delta$ for small enough $\varepsilon > 0$. As every ergodic probability measure other than $\delta_{\{1\}}$ gives full mass to some $J'_n = [a_n, a_{n+1})$, we can write

$$\mu_\varepsilon \,=\, \sum_{n=1}^{\infty} \int_{m(J'_n)\,=\,1} m\, d\mathbb{P}_\varepsilon(m) + \mu_\varepsilon(\{1\})\, \delta_{\{1\}}.$$

Moreover, for every positive integer $n \leq k$, if $m(J'_n) = 1$ then $m([0, a_k[) = 1$. Yet, as T restricted to $[0, a_k[$ has finite topological entropy, we know that

$$\limsup_{\varepsilon\,\to\,0^+} \frac{1}{\log(1/\varepsilon)} \sum_{n=1}^{k} \int_{m(J'_n)\,=\,1} h^K_m(\varepsilon, \delta)\, d\mathbb{P}_\varepsilon(m) \,=\, 0.$$

Consequently, to estimate $H^K_\delta(\mu)$, we are reduced to the ergodic components m of μ_ε such that $m([a_k, 1]) \,=\, 1$. So,

$$\begin{aligned}
\limsup_{\varepsilon\,\to\,0^+} \frac{h^K_{\mu_\varepsilon}(\varepsilon, \delta)}{\log(1/\varepsilon)} &= \limsup_{\varepsilon\,\to\,0^+} \frac{1}{\log(1/\varepsilon)} \int_{m([a_k,1])\,=\,1} h^K_m(\varepsilon, \delta)\, d\mathbb{P}_\varepsilon(m) \\
&\leq \limsup_{\varepsilon\,\to\,0^+} \frac{P(X, d, T, 0, \varepsilon)}{\log(1/\varepsilon)}\, \mu_\varepsilon([a_k, 1]) \\
&< \overline{\text{mdim}}_M([0, 1], |\cdot|, T)\, \delta \\
&= \delta.
\end{aligned}$$

Therefore,

$$0 \,\leq\, H^K_\delta(\mu) \,\leq\, \delta \qquad \forall\, \delta \in\,]0, 1[$$

and so

$$0 \,\leq\, H^K_\delta(\mu) \,\leq\, \sup_{0\,<\,\delta'\,<\,1} H^K_{\delta'}(\mu) \,=\, \lim_{\delta'\,\to\,0^+} H^K_{\delta'}(\mu) \,=\, 0.$$

(cf. Remark 7.2). Consequently,

$$H^K_\delta(\mu) \,=\, 0 \,=\, \mu(\{1\}). \tag{37}$$

From the previous equality and the convexity of H^K_δ, we further deduce that, for $t \in [0, 1]$ and $\mu \in \mathcal{P}_T([0, 1])$ such that $\mu(\{1\}) = 0$, if $\nu = t\delta_{\{1\}} + (1 - t)\mu$ then

$$H^K_\delta(\nu) \,=\, H^K_\delta(t\delta_{\{1\}} + (1 - t)\mu) \,\leq\, tH^K_\delta(\delta_{\{1\}}) + (1 - t)H^K_\delta(\mu) \,=\, t \,=\, \nu(\{1\}).$$

We are left to prove the reverse inequality, that is, $H^K_\delta(\nu) \,\geq\, t$. Take the sequence $(\varepsilon_n)_n$ as in (35) and the sequence of probability measures $(\mu_n)_n$ in $\mathcal{M}(\delta_{\{1\}})$ used in the previous page to compute $H^K_\delta(\delta_{\{1\}})$, which satisfy

$$\limsup_{n\,\to\,+\infty} \frac{h^K_{\mu_n}(\varepsilon_n, \delta)}{\log(1/\varepsilon_n)} \,=\, H^K_\delta(\delta_{\{1\}}) \,=\, 1.$$

Let $(\nu_n)_n = \big(t\mu_n + (1-t)\mu\big)_n$ in $\mathcal{M}(\nu)$. As $\eta \mapsto h^K_\eta(\varepsilon,\delta)$ is affine (cf. (17)), then

$$\begin{aligned} H^K_\delta(\nu) &\geq \limsup_{n\to+\infty} \frac{h^K_{\nu_n}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \qquad \text{by definition of } H^K_\delta(\nu) \\ &= \limsup_{n\to+\infty} \Big(t\,\frac{h^K_{\mu_n}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} + (1-t)\frac{h^K_{\mu}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)}\Big) \\ &= \limsup_{n\to+\infty} t\,\frac{h^K_{\mu_n}(\varepsilon_n,\delta)}{\log(1/\varepsilon_n)} \qquad \text{by (37)} \\ &= t. \end{aligned}$$

The proof of the claim is complete. □

Remark 10.4. *Example 10.3 evinces an obstruction if one tries to prove (12) using the map*

$$\widetilde{H}_\delta\colon\ \mu \in \mathcal{E}_T(X) \quad \mapsto \quad \limsup_{\varepsilon\to 0^+} \frac{h^K_\mu(\varepsilon,\delta)}{\log(1/\varepsilon)}$$

instead of H^K_δ. *Take* $\mu \in \mathcal{E}_T(X)$ *and, for* $\delta \in]0,1[$, *consider* $k = k(\delta)$ *large enough so that the set* $[0,a_k]\cup\{1\} = \bigcup_{1\leq j\leq k} J_j \cup \{1\}$ *satisfies the condition*

$$\mu([0,a_k]\cup\{1\}) \,>\, 1-\delta.$$

Then,

$$N_\mu(\varepsilon,\delta,n) \,\leq\, S([0,a_k]\cup\{1\}, d_n, 0, \varepsilon) \,\leq\, k \max_{1\leq j\leq k} S(J_j, d_n, 0,\varepsilon) + 1$$

and so

$$h^K_\mu(\varepsilon,\delta) \,\leq\, \max_{1\leq j\leq k} P(J_j, d, T|_{J_j}, 0, \varepsilon).$$

Therefore, as the restriction of T *to any* J_j *has finite topological entropy, for every* $\delta \in]0,1[$ *one has*

$$\widetilde{H}_\delta(\mu) \,=\, 0 \qquad \forall\, \mu \in \mathcal{P}_T(X).$$

The next example generalizes the last one.

Example 10.5. Given $\alpha \in]0,1[$, let $T_\alpha\colon [0,1] \to [0,1]$ be as in the previous example but whose invariant intervals $J^\alpha_n = [a^\alpha_{n-1}, a^\alpha_n]$ are determined by the sequence with general term

$$n \in \mathbb{N} \quad \mapsto \quad a^\alpha_n \,=\, \sum_{i=1}^n C(\alpha)(3^{i(1-1/\alpha)})$$

where $C(\alpha) = (\sum_{i=1}^\infty 3^{i(1-1/\alpha)})^{-1}$.

Claim: *For every* $\delta \in]0,1[$,

$$H^K_\delta(\mu) \,=\, \alpha\,\mu(\{1\}) \qquad \forall\, \mu \in \mathcal{P}_T([0,1]).$$

In particular, the map H^K_δ *is affine and the unique probability measure which maximizes* H^K_δ *is the Dirac mass* $\delta_{\{1\}}$.

Proof of the Claim. The argument to prove the claim is completely analogous to the one in Example 10.3 up to checking the upper bound

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_\alpha) \,\leq\, \alpha$$

which in the previous example was a trivial consequence of Remark 2.3.

For each $n \in \mathbb{N}$, let

$$\varepsilon_n^\alpha \,=\, |J_n^\alpha|/3^n \,=\, C(\alpha)\,3^{-n/\alpha}.$$

Consider a positive integer $L \geq 1$ large enough so that $1 - a_{n+L}^\alpha \,\leq\, \varepsilon_n^\alpha$, and take the following partition of the $[0,1]$:

$$[0,1] \,=\, [0,a_n^\alpha] \,\cup\, \bigcup_{\ell=1}^{L-1} J_{n+\ell}^\alpha \,\cup\, [a_{n+L}^\alpha, 1].$$

Thus, by Remark 2.3, for every $\ell = 1, ..., L-1$ we have

$$P(J_{n+\ell}^\alpha,\, |\cdot|,\, T_\alpha,\, 0,\, \varepsilon_n^\alpha) \,\leq\, \log S(J_{n+\ell}^\alpha,\, |\cdot|,\, 0,\, \varepsilon_n^\alpha) = \log\lceil |J_{n+\ell}^\alpha|/\varepsilon_n^\alpha\rceil \,=\, \log\lceil 3^{n-(\frac{1}{\alpha}-1)\ell}\rceil \,\leq\, \log 3^n.$$

Moreover,

$$P([0,a_n^\alpha],\, |\cdot|,\, T_\alpha,\, 0,\, \varepsilon_n^\alpha) \,\leq\, h_{top}(T_\alpha|_{[0,\,a_n^\alpha]}) \,=\, \log 3^n.$$

Hence,

$$\begin{aligned} P([0,1],\, |\cdot|,\, T_\alpha,\, 0\,,\varepsilon_n^\alpha) \;&=\; \max\big\{P\big([0,a_n^\alpha],\, |\cdot|,\, T_\alpha,\, 0,\, \varepsilon_n^\alpha\big),\; P\big(J_{n+\ell}^\alpha,\, |\cdot|,\, T_\alpha,\, 0,\, \varepsilon_n^\alpha\big) \mid \ell = 1,\cdots,L\big\} \\ &\leq\; \log 3^n. \end{aligned}$$

Now, for each $\varepsilon > 0$, let $n \in \mathbb{N}$ be such that $\varepsilon_{n+1}^\alpha \,\leq\, \varepsilon \,\leq\, \varepsilon_n^\alpha$. Then,

$$\frac{P([0,1],|\cdot|,T_\alpha,0,\varepsilon)}{\log(1/\varepsilon)} \,\leq\, \frac{P([0,1],|\cdot|,T_\alpha,0,\varepsilon_{n+1}^\alpha)}{\log(1/\varepsilon_n^\alpha)} \,\leq\, \frac{\log 3^{n+1}}{\log 3^{n/\alpha} - \log C(\alpha)}$$

and so

$$\overline{\mathrm{mdim}}_M([0,1],|\cdot|,T_\alpha) \,\leq\, \limsup_{n\to+\infty} \frac{\log 3^{n+1}}{\log 3^{n/\alpha} - \log C(\alpha)} \,=\, \alpha.$$

□

More generally, given $\mathcal{N} \in \mathbb{N}$ and $\alpha_1, \cdots, \alpha_\tau \in \,]0,1]$, consider the function

$$\begin{aligned} \mathcal{T}\colon \qquad [0,\mathcal{N}] \quad &\to \qquad [0,\mathcal{N}] \\ x \in [\ell-1,\ell] \;&\mapsto\; T_{\alpha_\ell}(x-\ell+1)+\ell-1 \end{aligned}$$

for $\ell \in \{1,\cdots,\mathcal{N}\}$, where T_1 stands for the map of Example 10.3. Then,

$$\overline{\mathrm{mdim}}_M([0,\mathcal{N}],|\cdot|,\mathcal{T}) \,=\, \max\{\alpha_1,\cdots,\alpha_{\mathcal{N}}\}$$

and, for every $\delta \in]0,1[$,

$$H_\delta^K(\mu) \,=\, \sum_{\ell=1}^{\mathcal{N}} \alpha_\ell\, \mu(\{\ell\}) \,=\, \int \mathcal{D}\, d\mu \qquad \forall\, \mu \in \mathcal{P}_{\mathcal{T}}([0,\mathcal{N}]).$$

In particular, H_δ^K is affine and the unique probability measures which maximize H_δ^K are the convex combinations of the Dirac measures $\{\delta_{\{\ell\}}\}_{\ell\in\mathcal{I}}$, where

$$\mathcal{I} \,=\, \big\{\ell \in \{1,\cdots,\mathcal{N}\} \,:\, \alpha_\ell = \max\{\alpha_1,\cdots,\alpha_{\mathcal{N}}\}\big\}.$$

Summoning Theorem C, we obtain the following simple formula for the upper metric mean dimension with potential of $\mathcal{T}$:

$$\overline{\mathrm{mdim}}_M([0,\mathcal{N}], |\cdot|, \mathcal{T}, \varphi) \,=\, \max_{\mu \,\in\, \mathcal{E}_{\mathcal{T}}([0,\mathcal{N}])} \int \big(\mathcal{D} + \varphi\big)\, d\mu \qquad \forall\, \varphi \in C^0([0,\mathcal{N}])$$

where

$$\mathcal{D}(x) \,=\, \begin{cases} \alpha_\ell & \text{if } x = \ell \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, the space $\mathcal{C}_\Gamma$ described in Theorem A is, in this family of examples, given by

$$\mathcal{C}_\Gamma \,=\, \Big\{\varphi \in C^0(X)\colon\; \int \mathcal{D}\, d\mu \,\leq\, \int \varphi\, d\mu, \quad \forall\, \mu \in \mathcal{P}_{\mathcal{T}}([0,\mathcal{N}])\Big\}.$$

Therefore,

$$M(\mu) \,=\, H^K_\delta(\mu) \,=\, \int \mathcal{D}\, d\mu \,=\, \sum_{\ell=1}^{\mathcal{N}} \alpha_\ell\, \mu(\{\ell\}) \qquad \forall\, \mu \in \mathcal{P}_{\mathcal{T}}([0,\mathcal{N}]).$$

11. Final comments

In this section we comment on some related topics and address a few questions that are suggested by our main results.

11.1. **Semigroup actions.** We start by exploring the generality of our results by applying them to the non-dynamical context of finitely generated semigroup actions for which a notion of topological pressure is already defined (see [3] for an account of these notions). Consider $m \in \mathbb{N}$ and a semigroup $\mathcal{G}$ generated by a family of $m+1$ continuous self-maps $G = \{id, g_1, \cdots, g_m\}$ of a compact metric space (X, d), with the composition operation. The semigroup action of $\mathcal{G}$ on X is the continuous map $\mathcal{S}\colon \mathcal{G} \times X \to X$ defined by $(g, x) \mapsto g(x)$ for every $g \in G$ and $x \in X$.

Given $n \in \mathbb{N}$ and $\underline{g} = (g_{i_1}, g_{i_2}, \cdots, g_{i_n}) \in G^n$, consider the metric

$$d_{\underline{g},n}(x,y) \,=\, \max_{0 \,\leq\, j \,\leq\, n} \; d(g_{i_j} g_{i_{j-1}} \cdots g_{i_1}(x),\, g_{i_j} g_{i_{j-1}} \cdots g_{i_1}(y))$$

which is equivalent to d. For every $\varphi \in C^0(X)$, $n \in \mathbb{N}$ and $\varepsilon > 0$, take the average

$$S(X, d, G, \varphi, \varepsilon, n) \,=\, \frac{1}{m^n} \sum_{|\underline{g}| \,=\, n} S(X, d_{\underline{g},n},\, S_{\underline{g},n}\varphi, \varepsilon)$$

where $S(X, d_{\underline{g},n},\, S_{\underline{g},n}\varphi, \varepsilon)$ is as defined in (5) and

$$S_{\underline{g},n}\varphi(x) \,=\, \sum_{j=0}^{n} \varphi(g_{i_j} g_{i_{j-1}} \cdots g_{i_1}(x)).$$

Then

$$P(X, d, G, \cdot, \varepsilon)\colon \quad \varphi \quad \mapsto \quad \limsup_{n \,\to\, +\infty} \frac{1}{n} \log S(X, d, G, \varphi, \varepsilon, n)$$

is an ε-pressure function. This motivates the following notion of upper metric mean dimension with potential for the semigroup action $\mathcal{S}$.

Definition 11.1. *Given* $\varphi \in C^0(X)$*, the* upper metric mean dimension with potential*, of the action* $\mathcal{S}$ *of* $\mathcal{G}$ *on* X *and* φ*, is given by*

$$\overline{\text{mdim}}_M(X, d, G, \varphi) \;=\; \limsup_{\varepsilon \to 0^+} \frac{P(X, d, G, \varphi, \varepsilon)}{\log(1/\varepsilon)}.$$

Some comments are in order. Firstly, as the notion of ε-scaled pressure function depends on the generating set G, so does the previous notion of upper metric mean dimension with potential. Secondly, according to Lemma 5.1, if $\limsup_{\varepsilon \to 0^+} \frac{P(X,d,G,\varphi,\varepsilon)}{\log(1/\varepsilon)} < +\infty$ at some (thus every) $\varphi \in C^0(X)$, then $\overline{\text{mdim}}_M(X, d, G, \cdot)$ is a pressure function on $C^0(X)$. Therefore, we obtain the following consequence of Theorem A.

Corollary 11.2. *Assume that* $\overline{\text{mdim}}_M(X, d, G, 0) < +\infty$*. Then*

$$\overline{\text{mdim}}_M(X, d, G, \varphi) \;=\; \max_{\mu \in \mathcal{P}(X)} \left\{ \mathcal{M}(\mu) + \int \varphi \, d\mu \right\} \qquad \forall \, \varphi \in C^0(X) \tag{38}$$

where

$$\mathcal{M}(\mu) \;=\; \inf_{\varphi \in \mathcal{A}} \int \varphi \, d\mu \quad \text{ and } \quad \mathcal{A} \;=\; \left\{ \varphi \in C^0(X) \colon \overline{\text{mdim}}_M(X, d, G, \varphi) \;\leq\; 0 \right\}.$$

Moreover,

$$\mathcal{M}(\mu) \;=\; \inf_{\varphi \in C^0(X)} \left\{ \overline{\text{mdim}}_M(X, d, G, \varphi) - \int \varphi \, d\mu \right\}.$$

The map $\mathcal{M}$ *is concave, upper semi-continuous and, if* $\gamma \colon \mathcal{P}(X) \to [0, +\infty]$ *is another function with the role of* $\mathcal{M}$ *in* (38)*, then* $\gamma \leq \mathcal{M}$*.*

Since there are finitely generated semigroup actions by continuous maps on compact metric spaces which admit no probability measure preserved by all the generators, one cannot expect to replace the space $\mathcal{P}(X)$ in (38) by another requesting such a notion of invariance. Whenever the semigroup action admits a probability measure preserved by all the generators, as happens when the semigroup is amenable, then Corollary 11.2 and Theorem E suggest the following question.

Question 1: Let $\mathcal{G}$ be a finitely generated amenable group of homeomorphisms acting on a compact metric space X such that $\overline{\text{mdim}}_M(X, d, G, 0) < +\infty$. Denote by $\mathcal{P}_{\mathcal{G}}(X)$ the space of probability measures on X preserved by all homeomorphisms $g \in \mathcal{G}$. Does there exist a function $\mathcal{D} \colon X \to [0, +\infty]$ such that

$$\overline{\text{mdim}}_M(X, d, G, 0) \;=\; \max_{\mu \in \mathcal{P}_{\mathcal{G}}(X)} \int \mathcal{D} \, d\mu \;?$$

If so:

(a) Does $\mathcal{M}(\mu) = \mathcal{D}(x)$ for μ-almost every point $x \in X$ and every probability measure $\mu \in \mathcal{P}_{\mathcal{G}}(X)$?

(b) Is $\mathcal{D}$ constant on the support of any ergodic probability measure $\mu \in \mathcal{P}_{\mathcal{G}}(X)$?

For instance, we may consider a finitely generated group $\mathcal{G}$ acting on the shift $X = \{0, 1\}^{\mathcal{G}}$ by

$$\big(g, \, (x_h)_{h \in \mathcal{G}}\big) \quad \mapsto \quad (x_{gh})_{h \in \mathcal{G}}$$

as these actions have plenty of invariant probability measures. A positive solution to the previous question would clarify the ergodic theory of amenable finitely generated semigroup actions with infinite topological entropy.

11.2. **Ergodic optimization.** In the literature we find several notions of topological pressure which are pressure functions. So, instead of defining an upper metric mean dimension as in Definition 2.1 or 2.6, we might have followed another strategy.

Definition 11.3. *Let* (X,d) *be a compact metric space and* $\boldsymbol{\Upsilon} = \big(\Upsilon_\varepsilon)_{0<\varepsilon<1}$ *be a family of pressure functions* $\Upsilon_\varepsilon\colon \boldsymbol{B}(X) \to \mathbb{R}$ *such that, for every* $\varphi \in C^0(X)$*,*

$$\limsup_{\varepsilon\to 0^+} \frac{\Upsilon_\varepsilon\big((\log 1/\varepsilon)\,\varphi\big)}{\log(1/\varepsilon)} < +\infty.$$

Let $\boldsymbol{\Gamma}$ *stand for the family of* ε*-pressure functions* $\Gamma_\varepsilon(\varphi) = \Upsilon_\varepsilon(\log(1/\varepsilon)\varphi)$*. The upper metric mean dimension of* $\boldsymbol{\Upsilon}$ *at* $\varphi \in \boldsymbol{B}(X)$ *is the limit*

$$\overline{\mathrm{mdim}}_M(\boldsymbol{\Upsilon}, d, \varphi) = \overline{\mathrm{mdim}}_M(\boldsymbol{\Gamma}, d, \varphi) = \limsup_{\varepsilon\to 0^+} \frac{\Upsilon_\varepsilon\big((\log 1/\varepsilon)\,\varphi\big)}{\log(1/\varepsilon)}. \tag{39}$$

Observe that, by Lemma 5.1, the map $\overline{\mathrm{mdim}}_M(\boldsymbol{\Upsilon}, d, \cdot)$ is a pressure function as well. Given $0<\varepsilon<1$, denote by E_ε the map assigned by Theorem 4.1 to the pressure function

$$\varphi \in \mathbf{B}(X) \quad \mapsto \quad \frac{\Upsilon_\varepsilon\big((\log 1/\varepsilon)\,\varphi\big)}{\log(1/\varepsilon)}$$

which satisfies

$$\frac{\Upsilon_\varepsilon\big((\log 1/\varepsilon)\varphi\big)}{\log(1/\varepsilon)} = \max_{\mu\in\mathcal{P}_a(X)} \Big\{E_\varepsilon(\mu) + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in \mathbf{B}(X). \tag{40}$$

Similarly, for each $0<\varepsilon<1$, let $\mathfrak{h}_\varepsilon$ be the map assigned by Theorem 4.1 to the pressure function Υ_ε, so that

$$\Upsilon_\varepsilon(\varphi) = \max_{\mu\in\mathcal{P}_a(X)} \Big\{\mathfrak{h}_\varepsilon(\mu) + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in \mathbf{B}(X). \tag{41}$$

Proposition 11.4. *For every* $0<\varepsilon<1$ *and* $\mu \in \mathcal{P}_a(X)$*, one has*

$$E_\varepsilon(\mu) = \frac{\mathfrak{h}_\varepsilon(\mu)}{\log(1/\varepsilon)}.$$

Proof. Applying the variational principle (41) to the potential $\big(\log 1/\varepsilon\big)\,\varphi \in \mathbf{B}(X)$, we get

$$\Upsilon_\varepsilon\big(\big(\log 1/\varepsilon\big)\,\varphi\big) = \max_{\mu\in\mathcal{P}_a(X)} \Big\{\mathfrak{h}_\varepsilon(\mu) + \int \big(\log 1/\varepsilon\big)\,\varphi\, d\mu\Big\} \qquad \forall\, \varphi \in \mathbf{B}(X)$$

so, dividing by $\log 1/\varepsilon$, we obtain

$$\frac{\Upsilon_\varepsilon\big((\log 1/\varepsilon)\,\varphi\big)}{\log(1/\varepsilon)} = \max_{\mu\in\mathcal{P}_a(X)} \Big\{\frac{\mathfrak{h}_\varepsilon(\mu)}{\log(1/\varepsilon)} + \int \varphi\, d\mu\Big\} \qquad \forall\, \varphi \in \mathbf{B}(X).$$

Therefore, by the maximality of E_ε among the maps that comply with the variational principle (40), we get

$$\frac{\mathfrak{h}_\varepsilon(\mu)}{\log(1/\varepsilon)} \le E_\varepsilon.$$

Analogously, by the variational principle relation (40) when applied to the potential $\frac{\varphi}{\log(1/\varepsilon)} \in \mathbf{B}(X)$, one has

$$\frac{\Upsilon_\varepsilon(\varphi)}{\log(1/\varepsilon)} = \max_{\mu \in \mathcal{P}_a(X)} \Big\{ E_\varepsilon(\mu) + \int \frac{\varphi}{\log(1/\varepsilon)}\, d\mu \Big\} \qquad \forall\, \varphi \in \mathbf{B}(X)$$

so

$$\Upsilon_\varepsilon(\varphi) = \max_{\mu \in \mathcal{P}_a(X)} \Big\{ \big(\log 1/\varepsilon\big)\, E_\varepsilon(\mu) + \int \varphi\, d\mu \Big\} \qquad \forall\, \varphi \in \mathbf{B}(X)$$

which implies, due to the maximality of $\mathfrak{h}_\varepsilon$ regarding the variational principle (41), that

$$\big(\log 1/\varepsilon\big)\, E_\varepsilon \le \mathfrak{h}_\varepsilon.$$

□

If $\boldsymbol{\Upsilon} = (\Upsilon)_\varepsilon$ is determined by a single pressure function Υ, the previous concept of upper metric mean dimension leads us to the realm of ergodic optimization. For instance, given a continuous map $T\colon X \to X$ on a compact metric space (X, d) with finite topological entropy, then the topological pressure P_{top} satisfies (cf. [12])

$$\limsup_{\varepsilon \to 0^+} \frac{P_{\text{top}}((\log(1/\varepsilon))\, \varphi)}{\log(1/\varepsilon)} = \sup_{\mu \in \mathcal{P}_T(X)} \int \varphi\, d\mu \qquad \forall\, \varphi \in C^0(X).$$

This is the case of $X = Y^{\mathbb{N}}$ and the shift map $\sigma\colon Y^{\mathbb{N}} \to Y^{\mathbb{N}}$, for which the space of ergodic σ-invariant probability measures is pathwise-connected as a consequence of the specification property [24]. Therefore, combining [22, Theorem B], [12, Theorem 2.4] and Theorem D we obtain the following additional information.

Corollary 11.5. *Let (Y, d) be a compact metric space such that $\dim_B U = \dim_B Y$ for every nonempty open set $U \subset Y$. Consider the shift map $\sigma\colon Y^{\mathbb{N}} \to Y^{\mathbb{N}}$. Then*

(a) There exists a Baire generic subset $\mathfrak{R} \subset C^0(Y^{\mathbb{N}})$ (resp. $\mathfrak{R}_\alpha \subset C^\alpha(Y^{\mathbb{N}})$ for each $\alpha > 0$) such that, given $\varphi \in \mathfrak{R}$ (resp. $\varphi \in \mathfrak{R}_\alpha$), there exists a unique $\mu_\varphi \in \mathcal{E}_\sigma(Y^{\mathbb{N}})$ satisfying

$$\overline{\text{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi) = \dim_B Y + \int \varphi\, d\mu_\varphi. \tag{42}$$

(b) There exists a dense subset $\mathfrak{D} \subset C^0(Y^{\mathbb{N}})$ such that, for every $\varphi \in \mathfrak{D}$, the subset of ergodic probability measures in the set

$$\Big\{ \mu \in \mathcal{P}_\sigma(Y^{\mathbb{N}}) \colon \overline{\text{mdim}}_M(Y^{\mathbb{N}}, d_\rho, \sigma, \varphi) = \dim_B Y + \int \varphi\, d\mu \Big\}$$

is uncountable.

In case the compact set Y is not finite, the shift map $\sigma\colon Y^{\mathbb{N}} \to Y^{\mathbb{N}}$ has infinite topological entropy and may no longer be expanding. It is known that hyperbolicity is a key assumption in many results concerning ergodic optimization and, to the best of our knowledge, ergodic optimization for dynamical systems with infinite topological entropy is seldom addressed. In view of [4, 8, 18], it is natural to ask the following questions, whose answers would convey relevant information on the upper metric mean dimension with potential for these shifts.

Question 2: Let (Y, d) be an infinite compact metric space and $\sigma\colon Y^{\mathbb{N}} \to Y^{\mathbb{N}}$ be corresponding shift map.

(a) Does there exist a Baire generic subset $\mathfrak{R} \subset C^0(Y^{\mathbb{N}})$ such that, for every $\varphi \in \mathfrak{R}$, the unique invariant probability measure μ_φ satisfying (42) has zero entropy and full support?

(b) Given $\alpha > 0$, does there exist an open and dense subset $\mathfrak{O} \subset C^\alpha(Y^{\mathbb{N}})$ such that, for each $\varphi \in \mathfrak{O}$, there is a unique invariant probability measure μ_φ supported on a periodic orbit by σ and satisfying (42) ?

Acknowledgments. We thank the anonymous reviewer for the careful reading of our manuscript and the many insightful comments and suggestions. MC and PV were partially supported by CMUP, which is financed by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the project with reference UIDB/00144/2020, and also acknowledges financial support from the project PTDC/MAT-PUR/4048/2021. PV benefited from the grant CEECIND/03721/2017 of the Stimulus of Scientific Employment, Individual Support 2017 Call, awarded by FCT, and from the project PTDC/MAT-PUR/29126/2017. GP has been awarded a PhD grant by FCT, with reference UI/BD/152212/2021.

References

[1] A. Biś, M. Carvalho, M. Mendes and P. Varandas. A convex analysis approach to entropy functions, variational principles and equilibrium states. Commun. Math. Phys 394 (2022), 215–256. 1, 2.3, 3, 3, 4, 4.1, 4.3, 4.5, 4.6, 5, 6

[2] A. Biś, M. Carvalho, M. Mendes, P. Varandas and X. Zhong. Correction: A convex analysis approach to entropy functions, variational principles and equilibrium states. Commun. Math. Phys 401:3 (2023), 3335–3342. 4.1

[3] A. Biś, M. Carvalho, M. Mendes and P. Varandas. Entropy functions for semigroup actions. Preprint, 2022. 11.1

[4] J. Brémont. Entropy and maximizing measures of generic continuous functions. C. R. Math. Acad. Sci. Sér. I 346 (2008), 199–201. 11.2

[5] M. Carvalho, F. Rodrigues and P. Varandas. A variational formula for the metric mean dimension of free semigroup actions. Ergodic Theory Dynam. Systems 42 (2021), 65–85. 10.2

[6] H. Chen, D. Cheng and Z. Li. Upper metric mean dimensions with potential. Results Math. 77 (2022), 54. 7.1

[7] D. Cheng and Z. Li. Scaled pressure of dynamical systems. J. Differential Equations 342 (2023), 441–471. 1, 7.2, 7.3

[8] G. Contreras. Ground states are generically a periodic orbit. Invent. Math. 205:2 (2016), 383–412. 11.2

[9] D.-J. Feng and W. Huang. Variational principle for weighted topological pressure. J. Math. Pures Appl. 106 (2016), 411–452. 2.3

[10] M. Gromov. Topological invariants of dynamical systems and spaces of holomorphic maps: I. Math. Phys. Anal. Geom. 2 (1999), 323–415. 1

[11] Y. Gutman and A. Śpiewak. Around the variational principle for metric mean dimension. Studia Math. 261:3 (2021), 345–360. 1, 1, 1

[12] O. Jenkinson. Survey: Ergodic optimization in dynamical systems. Ergodic Theory Dynam. Systems 39:10 (2019), 2593–2618. 11.2

[13] T. Kawabata and A. Dembo. The rate-distortion dimension of sets and measures. IEEE Trans. Inf. Theory 40:5 (1994), 1564–1572. 8.1

[14] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. Publ. Math. Inst. Hautes Études Sci. 51 (1980), 137–173. 3, 7

[15] E. Lindenstrauss and B. Weiss. Mean topological dimension. Israel J. Math. 115 (2000), 1–24. 1, 2.1, 10.1

[16] E. Lindenstrauss and M. Tsukamoto. From rate distortion theory to metric mean dimension: variational principle. IEEE Trans. Inform. Theory 64 (2018), 3590–3609. 1, 7.8

[17] E. Lindenstrauss and M. Tsukamoto. Double variational principle for mean dimension. Geom. Funct. Anal. 29 (2019), 1048–1109. 1

[18] I. Morris. Ergodic optimization for generic continuous functions. Discrete Contin. Dyn. Sys. 27 (2010), 383–388. 11.2

[19] E. Scopel. Contributions to the theory of metric mean dimension and mean Hausdorff dimension (in Portuguese). PhD thesis. Universidade Federal do Rio Grande do Sul - UFRGS (2021). 9.2

[20] R. Shi. On variational principles for metric mean dimension. IEEE Trans. Inform. Theory 68:7 (2022), 4282–4288. 1, 1

[21] R. Shi. Finite mean dimension and marker property. Trans. Amer. Math. Soc., electronically published on June 22, 2023. 1

[22] M. Shinoda. Uncountably many maximizing measures for a dense subset of continuous functions. Nonlinearity 31 (2018), 2192–2200. 11.2

[23] K. Sigmund. On dynamical systems with the specification property. Trans. Amer. Math. Soc. 190 (1974), 285–299. 8

[24] K. Sigmund. On the connectedness of ergodic systems. Manuscr. Math. 22 (1977), 27–32. 11.2

[25] M. Tsukamoto. Double variational principle for mean dimension with potential. Adv. Math. 361 (2020), 106935. 1, 2.1, 2.5

[26] M. Tsukamoto, M. Tsutaya and M. Yoshinaga. G-index, topological dynamics and the marker property. Israel J. Math. 251 (2022), 737–764. 1

[27] A. Velozo and R. Velozo. Rate distortion theory, metric mean dimension and measure theoretic entropy. Preprint, 2017, arXiv:1707.05762 1, 8, 10.2

[28] P. Walters. An Introduction to Ergodic Theory. Graduate Texts in Mathematics 79, Springer-verlag, New York, Berlin, Heidelberg, 1982. 1, 1, 2.3

[29] R. Yang, E. Chen and X. Zhou. On variational principle for upper metric mean dimension with potential. Preprint 2022, arXiv: 2207.01901. 1, 1, 3

[30] R. Yang, E. Chen and X. Zhou. Some notes on variational principle for metric mean dimension. IEEE Transactions on Information Theory 69:5 (2023), 2796–2800. 1

[31] R. Yang, E. Chen and X. Zhou. Bowen's equations for upper metric mean dimension with potential. Nonlinearity 35:9 (2022), 4905–4938. 1

[32] X. Ye and G. Zhang. Entropy points and applications. Trans. Amer. Math. Soc. 359:12 (2007), 6167–6186. 1, 9.4

CMUP & Departamento de Matemática, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, Porto, Portugal.
E-mail address: `mpcarval@fc.up.pt`

CMUP & Departamento de Matemática, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, Porto, Portugal.
E-mail address: `Gustavo.pessil@outlook.com`

CMUP & Departamento de Matemática, Universidade Federal da Bahia, Salvador, Brazil.
E-mail address: `paulo.varandas@ufba.br`