Manuscript submitted to AIMS Journal

doi:10.3934/dcdss.2026058

NONLINEAR METRIC MEAN DIMENSION

Alex J. Becker[1], Chunlin Liu*[2,3] and Fagner B. Rodrigues[4]

[1] Departamento de Matemática, Universidade Federal de Santa Maria, Av.Roraima 1000, CEP 97105-900, Santa Maria, RS - Brasil

[2] School of Mathematical Sciences, Dalian University of Technology, Dalian, 116024, P.R. China

[3] Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland

[4] Departamento de Matemática, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, CEP 91509-900, Porto Alegre, RS, Brazil

(Communicated by Handling Editor)

Abstract. Motivated by the nonlinear thermodynamic formalism introduced by Buzzi, Klonecker, and Leplaideur [4], we define the nonlinear metric mean dimension, designed to capture refined complexity beyond the reach of topological pressure. Our approach is explicitly convex-analytic: we formulate the nonlinear pressure and its scaled counterparts as convex, monotone, translation-equivariant functionals on spaces of energies, and we derive two variational principles via standard tools from convex analysis. Within this framework we also establish existence results for nonlinear equilibrium measures and present examples illustrating the scope and limitations of the theory.

1. **Introduction.**

The nonlinear thermodynamical formalism was formalized and explored by J. Buzzi, B. Kloeckner and R. Leplaideur in [4]. The authors considered a dynamics $f : X \to X$, a continuous map $\mathcal{E} : \mathcal{M}(X) \to \mathbb{R}$ (named energy) on the space of probability measures and defined the nonlinear pressure as

$$\Pi^{\mathcal{E}}(\mu) = h(\mu) + \mathcal{E}(\mu),$$

for any $\mu \in \mathcal{M}_f(X)$. When $\mathcal{E}(\mu) = \int \varphi \, d\mu$, for some $\varphi \in C^0(X)$, it is the classical linear pressure. Under suitable hypotheses, they showed the following variational principle

$$\Pi^{\mathcal{E}}_{\text{top}}(X, f) = \sup_{\mu \in \mathcal{M}_f(X)} \Pi^{\mathcal{E}}(\mu) \tag{1}$$

(see Section 2 for the definition of $\Pi^{\mathcal{E}}_{\text{top}}(X, f)$). The quantity $\Pi^{\mathcal{E}}_{\text{top}}(X, f)$ is called the *nonlinear topological pressure* and defined in terms of the topological properties of space X. In addition to the variational principle (1), the authors obtained some classical results of thermodynamical formalism theory as existence of equilibrium measures and properties of the equilibrium measures sets. Using this new approach, the authors expanded the results of [13, 14]. This techniques can deal with known cases (Curie–Weiss and Potts models) as well as with new examples (metastable phase transition). Recently, the work [2] introduced the nonlinear thermodynamical

---

*Corresponding author: Chunlin Liu.

formalism in higher-dimensional case and [10] established a connection between the nonlinear thermodynamical formalism theory and rotation theory.

However when examining the variational principle formula (1) for the nonlinear pressure, a crucial limitation is observed: it relies on the metric entropy $h(\mu)$. Given the compactness of the space $\mathcal{M}(X)$ and the existence of a vast class of systems whose topological entropy is infinite (for instance, they form a C^0-generic set in the space of homeomorphisms of a compact manifold [28] with dimension greater than one), the traditional approach becomes inadequate for measuring the complexity of these systems. This gap motivates the present study: it is necessary to consider a notion of nonlinear pressure that is effective for high-complexity systems. Therefore, introducing a nonlinear pressure that takes into account the *metric mean dimension* is a natural way to extend the nonlinear thermodynamical formalism to systems with infinite topological entropy, where the metric mean dimension serves as the discriminatory topological invariant.

The notion of metric mean dimension was introduced by Lindenstrauss and Weiss in [17], as metric-dependent analog of the *mean dimension*, another topological invariant of high-complexity maps which was introduced by Gromov [7]. The definition of metric mean dimension is a fusion of the definitions of topological entropy and Minkowski dimension and it provides a powerful method to obtain upper bounds on mean dimension. It has several applications, such as it was used in [25] for solving a problem, proposed by Gromov in [7], to estimate the mean dimension of a dynamical system in holomorphic curve theory . It also has an application to the study of expansive group actions [22]. Moreover, the metric mean dimension turned out to be useful in several contexts like in the study of compression [6, 8, 9, 18, 19, 20].

Inspired by these works, we extend the notion of nonlinear topological pressure to setting of dynamical systems with positive metric mean dimension. As in the case of the linear topological pressure, in [5] it was proved that the metric mean dimension with potential satisfies a variational principle. Examples of maps with infinite nonlinear topological pressure may be easily constructed and because of this, we believe that such a notion represents an interesting contribution to the study of continuous dynamics. We present several examples, and prove that, in certain setting, the nonlinear metric mean dimension can be computed via a variational principle.

Notice that unlike the metric mean dimension, finite nonlinear topological pressure does not necessarily imply that the nonlinear metric mean dimension vanishes (see Subsection 6.4). This fact was discuss in [4, Subsection 1.3], where the authors needed to use a property over the ergodic measures called abundance of ergodic measures (see also [2, 10]). In our work, for the sake of a cleaner presentation, we replace that hypothesis by assuming convexity of the energy functions. This assumption guarantees abundance of ergodic measures , though it does not, in general, imply entropy–density (see, e.g., [4, Section 1.3]).

This manuscript is organized as follows. In Section 2, we introduce many definitions. In Section 3, we present the main results. In Section 4, we present the proofs of the main results. In Section 5, we discuss a way to obtain equilibrium measures via variational principle. In Section 6, we give some examples.

2. **Preliminaries.**

2.1. **Basic notation.** Throughout this work, we denote *topological dynamical system* by (X, f), where X is a compact metric space and f is a continuous map on X. Given $n \in \mathbb{N}$, let $d_n(x,z) = \max_{0 \le i \le n-1} d(f^i(x), f^i(z))$ be the *Bowen metric*. It is well known that d_n is indeed a metric and generates the same topology as d. Furthermore, given $\varepsilon > 0$, $n \in \mathbb{N}$ and a point $x \in X$, we define the (n,ε)-*dynamical balls* of radius ε and length n around $x \in X$ as

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

Given $K \subset X$, we say that $E \subset K$ is an (n,ε)–*separated* subset of K if $d_n(x,z) > \varepsilon$ for every $x, z \in E$ with $x \neq z$. Moreover, we say that a set $F \subset K$ is an (n,ε)-*spanning* set of K if $K \subset \bigcup_{x \in F} B_n(x,\varepsilon)$. Given $n \in \mathbb{N}$ and $\varepsilon > 0$, let us denote by $s(f,n,\varepsilon)$ the maximal cardinality of all (n,ε)–separated subsets of X by f, and by $r(f,n,\varepsilon)$ the minimal cardinality of all (n,ε)–spanning subsets of X by f. Due to the compactness of X, they are finite for any $n \in \mathbb{N}$ and $\varepsilon > 0$.

In order to fix the notation,

- $\mathcal{M}(X)$ endowed with the weak$^*-$topology;
- $\mathcal{M}_f(X)$ is the set of $f-$invariant probability measures on X;
- $\mathcal{M}_f^{erg}(X)$ is the subset of all ergodic $f-$invariant measures on X.

A function $\mathcal{E} : \mathcal{M}(X) \to \mathbb{R}$ which is continuous in the weak$^*-$topology is an *energy*. A particular case of energy occurs when, given a continuous function $\mathcal{E} : X \to \mathbb{R}$ (called *potential*), it can be written as $\mathcal{E}(\mu) = F(\mu(\varphi))$, for some continuous function $F : I \to \mathbb{R}$ defined on interval $I \supset \varphi(X)$. In this, we say that $\mathcal{E}$ is an *energy with potential*. Moreover, an energy $\mathcal{E}$ is said *convex* if

$$\mathcal{E}\left(\int \mu \, d\mathbb{P}(\mu)\right) \le \int \mathcal{E}(\mu) d\mathbb{P}(\mu). \tag{2}$$

For instance, all energy $\mathcal{E}$ with potential is convex whenever F is. In particular, for $F(x) = x$, i.e., in the linear case.

2.2. **Metric mean dimension.** We recall the definition of metric mean dimension in [15, 17].

Definition 2.1. The *upper metric mean dimension* of f with respect to d is given by

$$\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big) = \limsup_{\varepsilon \to 0} \frac{h(f,\varepsilon)}{|\log \varepsilon|},$$

where

$$h(f,\varepsilon) = \limsup_{n \to \infty} \frac{1}{n} \log s(f,n,\varepsilon).$$

Similarly, the *lower metric mean dimension* of f with respect to d is given by

$$\underline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big) = \liminf_{\varepsilon \to 0} \frac{h(f,\varepsilon)}{|\log \varepsilon|}.$$

In the case when $\underline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big) = \overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big)$ this common value is called the *metric mean dimension* of f with respect to d and is denoted simply by $\mathrm{mdim}_{\mathrm{M}}\Big(X, f, d\Big)$.

Remark 2.2. Recall that the *topological entropy* of the map f is given by

$$h_{\mathrm{top}}(f) = \lim_{\varepsilon \to 0} h(f,\varepsilon).$$

Consequently, $\overline{\text{mdim}}_{\text{M}}\left(X,f,d\right)=\underline{\text{mdim}}_{\text{M}}\left(X,f,d\right)=0$ whenever the topological entropy of f is finite. In particular, the metric mean dimension is a suitable quantity to study systems with infinite topological entropy. For more on these quantities see [16, 17, 23, 26] and references therein.

2.3. **Measure theoretic metric mean dimension.** We introduce now some definitions for measure theoretic metric mean dimension inspired in [5, 21, 27].

2.3.1. *Katok-measure theoretic metric mean dimension.* The notion of measure theoretic metric mean dimension that we consider here is the one introduced in [5].

Given $\mu\in\mathcal{M}_f^{erg}(X)$, $\varepsilon>0$ and $\delta\in(0,1)$, let us denote by $N_\mu(\varepsilon,\delta,n)$ the minimal number (ε,n)-balls needed to cover a set of μ-measure bigger than $1-\delta$. Then, we define

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

where the superscript "K" in $h_\mu^K(f,\varepsilon,\delta)$ stands for "Katok", since this quantity comes from the description of the metric entropy given by Katok (see [11]).

The previous notion can be extended to non-ergodic probability measures in $\mathcal{M}_f(X)$ via integration: given $\mu\in\mathcal{M}_f(X)$, define

$$h_\mu^K(f,\varepsilon,\delta)=\int_{\mathcal{M}_f^{erg}(X)}h_m^K(f,\varepsilon,\delta)\,d\mathbb{P}_\mu(m),\tag{3}$$

where $\mu=\int_{\mathcal{M}_f^{erg}(X)}m\,d\mathbb{P}_\mu(m)$ is the ergodic decomposition of μ. Observe that, by the definition, the map $m\mapsto h_m^K(f,\varepsilon,\delta)$ is measurable and integrable. Consequently, the function

$$\mu\in\mathcal{M}_f(X)\mapsto h_\mu^K(f,\varepsilon,\delta)$$

is also affine.

Definition 2.3. Given $\delta\in(0,1)$, we define the map $H_\delta^K:\mathcal{M}_f(X)\to\mathbb{R}$ by

$$H_\delta^K(f,\mu)=\sup_{(\mu_\varepsilon)_\varepsilon\in\mathcal{M}(\mu)}\limsup_{\varepsilon\to0}\frac{h_{\mu_\varepsilon}^K(f,\varepsilon,\delta)}{|\log\varepsilon|},\tag{4}$$

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

By [5, Theorem C],

$$\overline{\text{mdim}}_{\text{M}}\left(X,f,d\right)=\max_{\mu\in\mathcal{M}_f^{erg}(X)}H_\delta^K(f,\mu).$$

Remark 2.4. It is important to notice that $h_\mu^K(f,\varepsilon,\delta)$ can be defined in terms of (ε,n)-spanning sets. More precisely,

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

where $b_\mu(\varepsilon,\delta,n)$ denotes the minimal cardinality of a (ε,n)-spanning set contained in a set of μ-measure bigger than $1-\delta$.

2.3.2. *Kolmogorov-Sinai measure theoretic metric mean dimension.* Let $\mu \in \mathcal{M}_f(X)$. We say that $\xi = \{C_1, \dots, C_k\}$ is a measurable partition of X if every C_i is a measurable set, $\mu\left(X \setminus \cup_{i=1}^k C_i\right) = 0$ and $\mu\left(C_i \cap C_j\right) = 0$ for every $i \neq j$. The *entropy* of ξ with respect to μ is given by

$$H_\mu(\xi) = -\sum_{i=1}^{k} \mu(C_i) \log(\mu(C_i)).$$

Given a measurable partition ξ, we consider $\xi^n = \bigvee_{j=0}^{n-1} f^{-j}\xi$. Then, the *metric entropy of* (f, μ) *with respect to* ξ is given by

$$h_\mu(f, \xi) = \lim_{n \to +\infty} \frac{1}{n} H_\mu(\xi^n).$$

Definition 2.5. *The Kolmogorov-Sinai measure theoretic metric mean dimension* is defined as

$$\overline{\text{Hmdim}}_{\text{M}}\,(X, f, d) = \limsup_{\varepsilon \to 0} \frac{1}{|\log \varepsilon|} \sup_{\mu \in \mathcal{M}_f(X)} \inf_{|\xi| < \varepsilon} h_\mu(f, \xi),$$

where $|\xi|$ denotes the diameter of the partition ξ and the infimum is taken over all finite measurable partitions of X satisfying $|\xi| < \varepsilon$. Moreover, we consider the quantity

$$H(f, \mu) = \sup_{(\mu_\varepsilon)_\varepsilon \in \mathcal{M}(\mu)} \limsup_{\varepsilon \to 0} \frac{\inf_{|\xi| < \varepsilon} h_{\mu_\varepsilon}(f, \xi)}{|\log \varepsilon|}.$$

2.3.3. *Brin-Katok measure theoretic metric mean dimension.* Firstly, we recall the Brin-Katok entropy in [12]. Let $\mu \in \mathcal{M}_f(X)$. Now, fix $\varepsilon > 0$ and for each $x \in X$, define

$$h_\mu^{BK}(f, x, \varepsilon) = \limsup_{n \to \infty} -\frac{1}{n} \log \mu(B_n(x, \varepsilon)).$$

If μ is ergodic we have that

$$h_\mu^{BK}(f, x) = \lim_{\varepsilon \to 0} h_\mu^{BK}(f, x, \varepsilon) = h_\mu(f), \text{ for } \mu \text{ a.e. } x \in X.$$

Definition 2.6. *The Brin-Katok measure theoretic metric mean dimension* is given by

$$H^{BK}(f, \mu) = \sup_{(\mu_\varepsilon)_\varepsilon \in \mathcal{M}(\mu)} \limsup_{\varepsilon \to 0} \frac{h_{\mu_\varepsilon}^{BK}(f, \varepsilon)}{|\log \varepsilon|}.$$

However, it is possible to define H^{BK} in more general sense. Using the ergodic decomposition theorem we can define, for any $\nu \in \mathcal{M}_f(X)$,

$$h_\nu^{BK}(f, \varepsilon) = \int_{\mathcal{M}_f^{erg}(X)} h_m^{BK}(f, \varepsilon)\; d\mathbb{P}_\mu(m),$$

where $\nu = \int_{\mathcal{M}_f^{erg}(X)} m \; d\mathbb{P}_\nu(m)$ is the ergodic decomposition of μ (since $h_m^{BK}(f, x, \varepsilon)$ is f-invariant, and then, m a.e. constant).

Remark 2.7. By [5, Theorem C] and [24, Lemma 3.1], we can prove that

$$\begin{aligned} \overline{\text{mdim}}_{\text{M}}\left(X, f, d\right) &= \sup\left\{H_\delta^K(\mu) : \mu \in \mathcal{M}_f(X)\right\} \\ &= \sup\left\{H(f, \mu) : \mu \in \mathcal{M}_f(X)\right\} \\ &= \sup\left\{H^{BK}(f, \mu) : \mu \in \mathcal{M}_f(X)\right\}. \end{aligned}$$

2.3.4. *Nonlinear pressure.* For every $x \in X$ and $n \in \mathbb{N}$, we consider the *empirical measure* defined by

$$\Delta_x^n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i(x)},$$

where δ_y is the Dirac-measure at the mass point $y \in X$. In particular, $\Delta_x^n(\varphi) = \frac{1}{n} S_n \varphi(x)$ is the averaged Birkhoff sum, for every potential φ. We recall that an *energy* is a continuous function $\mathcal{E} : \mathcal{M}(X) \to \mathbb{R}$, where $\mathcal{M}(X)$ is the space of all Borel probability measures on X with weak*-topology. Given any energy $\mathcal{E}$, we define the *nonlinear weight* of order n of a finite set $\mathcal{S} \subset X$ and the *nonlinear partition function* as

$$\omega_n(\mathcal{S}) = \sum_{x \in \mathcal{S}} e^{n\mathcal{E}(\Delta_x^n)} \qquad \text{and} \qquad \zeta(\varepsilon, n) = \sup_{\mathcal{S}} \omega_n(\mathcal{S}),$$

where the supremum is taken over all (ε, n)-separated sets $\mathcal{S}$. An (ε, n)-separated set $\mathcal{S}$ is said to be *adapted* if it realizes the maximum in $\zeta(\varepsilon, n)$.

Given a continuous function $f : X \to X$, and an energy $\mathcal{E}$, the *nonlinear topological pressure* is defined by

$$\Pi_{\text{top}}^{\mathcal{E}}(f) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log \zeta(\varepsilon, n).$$

For every $\mu \in \mathcal{M}_f(X)$, the *nonlinear pressure* is given by

$$\Pi^{\mathcal{E}}(f, \mu) = h(f, \mu) + \mathcal{E}(\mu),$$

where $h(f, \mu)$ is the measure theoretic entropy of f with respect to μ.

In [4, Theorem A] the authors proved that for systems with abundance of ergodic measures, the following variational principle

$$\sup_{\mu \in \mathcal{M}_f(X)} \Pi^{\mathcal{E}}(f, \mu) = \Pi_{\text{top}}^{\mathcal{E}}(f).$$

it holds.

3. **Main results.**

Building on the abstract framework of [3, 5], we develop a theory of *nonlinear* pressure. More precisely, for general pressure functionals and for the ε–pressure, we establish variational principles in which the nonlinearity is encoded by energy functionals $\mathcal{E}$. We also introduce and investigate nonlinear metric mean dimension and nonlinear box metric mean dimension, and we apply the abstract framework to derive several results.

3.1. **Nonlinear pressure functions and variational principle.** Initially, we consider the Banach spaces of energy functions given by

$$\mathbf{B}(X) := \{\mathcal{E} : \mathcal{M}(X) \to \mathbb{R} : \mathcal{E} \text{ is continuous in the weak}^*\text{–topology}\},$$

$$\mathbf{B_c}(X) := \{\mathcal{E} \in \mathbf{B}(X) : \mathcal{E} \text{ is convex}\},$$

where both sets are endowed with the norm $||\mathcal{E}|| = \sup_{\mu \in \mathcal{M}(X)} |\mathcal{E}(\mu)|$. Notice that, $\mathcal{M}(X)$ is compact in the weak*– topology and we can see $\mathbf{B}(X)$ as the space $C^0(Y)$ of the continuous functions defined in $Y = \mathcal{M}(X)$.

We denote by $\mathcal{M}_a(X)$ the set of *finitely additive probability measures* endowed with the total variation norm (see [3, Subsection 2.1]). This set is compact in the weak*–topology and $\mathcal{M}(X) \subset \mathcal{M}_a(X)$. In the next, we define the pressure function as [3] and [5].

Definition 3.1. A map $\Upsilon : \mathbf{B}(X) \to \mathbb{R}$ is called a *pressure function* if it satisfies

- Monotonicity: $\mathcal{E}_1 \leq \mathcal{E}_2 \implies \Upsilon(\mathcal{E}_1) \leq \Upsilon(\mathcal{E}_2), \ \forall\ \mathcal{E}_1, \mathcal{E}_2 \in \mathbf{B}(X)$;
- Translation invariance: $\Upsilon(\mathcal{E} + c) = \Upsilon(\mathcal{E}) + c, \ \forall\ \mathcal{E} \in B(X), \forall\ c \in \mathbb{R}$;
- Convexity: $\Upsilon(t\mathcal{E}_1 + (1-t)\mathcal{E}_2) \leq t\Upsilon(\mathcal{E}_1) + (1-t)\Upsilon(\mathcal{E}_2), \forall \mathcal{E}_1, \mathcal{E}_2 \in \mathbf{B}(X), \forall\ t \in (0,1)$.

The following example shows that the nonlinear topological pressure introduced in [4] is a nonlinear pressure function.

Example 3.2. Given $K \subset X$ a closed subset, an energy $\mathcal{E}$ and $\varepsilon > 0$ we consider

$$S(K, f, d, \mathcal{E}, \varepsilon, n) := \sup\left\{ \sum_{x \in \mathcal{C}} e^{n\mathcal{E}(\Delta_x^n)} : \mathcal{C} \text{ is a } (\varepsilon, n)\text{-separated set of } K \right\},$$

and

$$P(K, f, d, \mathcal{E}) = \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log S(K, f, d_n, \mathcal{E}, \varepsilon, n).$$

The family $P(X, f, d, \cdot)$ is a scaled pressure function. In fact, it is not difficult to see that the function $P(X, f, d, \cdot)$ is monotone and convex. The translation invariance is from

$$\begin{aligned} S(K, f, d, \mathcal{E} + c, \varepsilon, n) &= \sup\left\{ \sum_{x \in \mathcal{C}} e^{n(\mathcal{E}+c)(\Delta_x^n)} \right\} \\ &= \sup\left\{ \sum_{x \in \mathcal{C}} e^{n\mathcal{E}(\Delta_x^n)} \right\} \cdot e^{cn}, \end{aligned}$$

where the supremum are taken over all $(\varepsilon, n)-$separated set, and hence

$$\begin{aligned} P(K, f, d, \mathcal{E} + c) &= P(K, f, d, \mathcal{E}) + \limsup_{n \to \infty} \frac{1}{n}(cn) \\ &= P(K, f, d, \mathcal{E}) + c. \end{aligned}$$

First, we present a variational principle of pressure functions inspired in [3]. We emphasize that to obtain the equality in the variational principle we need to assume convexity for the energy functions (see (2)).

Theorem A. *Let (X, d) be a compact metric space and $\Upsilon : B(X) \to \mathbb{R}$ be a pressure function. Then*

$$\Upsilon(\mathcal{E}) \geq \sup_{\mu \in \mathcal{M}_a(X)} \Big\{ \mathcal{H}(\mu) + \mathcal{E}(\mu) \Big\}, \tag{5}$$

where $\mathcal{H}(\mu) = \inf_{\mathcal{E} \in \mathcal{A}_\Upsilon} \mathcal{E}(\mu)$ and $\mathcal{A}_\Upsilon := \{\mathcal{E} \in \boldsymbol{B}(X) : \Upsilon(-\mathcal{E}) \leq 0\}$. Moreover, restricting the pressure function Γ restricted to set $B_c(X)$, follows that

$$\Upsilon(\mathcal{E}) = \sup_{\mu \in \mathcal{M}_a(X)} \Big\{ \mathcal{H}^*(\mu) + \mathcal{E}(\mu) \Big\}, \tag{6}$$

with $\mathcal{H}^*(\mu) = \inf_{\mathcal{E} \in \mathcal{A}^*_\Upsilon} \mathcal{E}(\mu)$ and $\mathcal{A}^*_\Upsilon := \{\mathcal{E} \in \boldsymbol{B}_c(X) : \Upsilon(-\mathcal{E}) \leq 0\}$. Indeed, $\mathcal{H}^*$ is limited above by $\Upsilon(0)$, concave, upper semi-continuous and for every $\mu \in \mathcal{M}_a(X)$

satisfies

$$\mathcal{H}^*(\mu) = \inf_{\mathcal{E}\in \boldsymbol{B}_c(X)} \Big\{\Upsilon(\mathcal{E}) - \mathcal{E}(\mu)\Big\}. \tag{7}$$

Notice that, from (6) we obtain that $\mathcal{H}^*$ is limited above $\Upsilon(0)$. Indeed, it is concave, because it is an infimum of concave functions. Moreover, $\mathcal{H}^*$ is upper semi continuous due the continuity of energy functions $\mathcal{E}$.

3.2. **Scaled pressure functions and variational principle.** Now, we present the definition of scaled pressure functions as [5].

Definition 3.3. We say that a map $\Gamma : \mathbf{B}(X) \to \mathbb{R}$ is a *scaled pressure function* if there exists a constant $\alpha > 0$ such that Γ satisfies

- Monotonicity: $\mathcal{E}_1 \leq \mathcal{E}_2 \implies \Gamma(\mathcal{E}_1) \leq \Gamma(\mathcal{E}_2),\ \forall\ \mathcal{E}_1, \mathcal{E}_2 \in \mathbf{B}(X)$;
- Scaled translation invariance: $\Gamma(\mathcal{E} + c) = \Gamma(\mathcal{E}) + \alpha c,\ \forall\ \mathcal{E} \in B(X), \forall\ c \in \mathbb{R}$;
- Convexity: $\Gamma(t\mathcal{E}_1 + (1-t)\mathcal{E}_2) \leq t\Gamma(\mathcal{E}_1) + (1-t)\Gamma(\mathcal{E}_2), \forall \mathcal{E}_1, \mathcal{E}_2 \in \mathbf{B}(X) \forall\ t \in (0, 1)$.

For instance, any pressure function is a scaled pressure function with constant $\alpha = 1$.

Remark 3.4. If Γ is a scaled pressure function with constant α, then $\Upsilon = \Gamma/\alpha$ is a pressure function. The monotonicity and convexity for Γ/α is straightforward from this properties of Γ. Moreover, for any $\mathcal{E} \in \mathbf{B}(X)$ and $c \in \mathbb{R}$

$$\Upsilon(\mathcal{E} + c) = \frac{\Gamma(\mathcal{E} + c)}{\alpha} = \frac{\Gamma(\mathcal{E}) + c\alpha}{\alpha} = \Upsilon(\mathcal{E}) + c.$$

Example 3.5. Given $K \subset X$ a closed subset, an energy $\mathcal{E}$ and $\varepsilon > 0$ we consider

$$S(K, f, d, \mathcal{E}, \varepsilon, n) = \sup\left\{\sum_{x\,\in\,\mathcal{C}} e^{n|\log\varepsilon|\mathcal{E}(\Delta_x^n)} : \mathcal{C} \text{ is a } (\varepsilon, n) - \text{separated set}\right\},$$

and

$$P(K, f, d, \mathcal{E}, \varepsilon) = \limsup_{n\to\infty} \frac{1}{n} \log S(K, f, d_n, \mathcal{E}, \varepsilon, n).$$

For each $0 < \varepsilon < 1$, the family $(P(X, f, d, \cdot, \varepsilon))_{0<\varepsilon<1}$ is a scaled pressure functions with constant $\log(1/\varepsilon)$. In fact, it is not difficult to see that the monotonicity and convexity of $(P(X, f, d, \cdot, \varepsilon))_{0<\varepsilon<1}$. Moreover,

$$\begin{aligned} S(K, f, d, \mathcal{E} + c, \varepsilon, n) &= \sup\left\{\sum_{x\,\in\,\mathcal{C}} e^{n|\log\varepsilon|(\mathcal{E}+c)(\Delta_x^n)}\right\} \\ &= \sup\left\{\sum_{x\,\in\,\mathcal{C}} e^{n|\log\varepsilon|\mathcal{E}(\Delta_x^n)}\right\} \cdot e^{cn|\log\varepsilon|}, \end{aligned}$$

where the supremum are taken over all $(\varepsilon, n)-$separated set. Therefore

$$\begin{aligned} P(K, f, d, \mathcal{E} + c, \varepsilon) &= P(K, f, d, \mathcal{E}, \varepsilon) + \limsup_{n\to\infty} \frac{1}{n} cn|\log\varepsilon| \\ &= P(K, f, d, \mathcal{E}, \varepsilon) + c\log(1/\varepsilon). \end{aligned}$$

The above example also motivates the study of scaled pressure functionals with scale $\log(1/\varepsilon)$, since the family $\big(P(X,f,d,\cdot,\varepsilon)\big)_{0<\varepsilon<1}$ naturally underlies the definition of the nonlinear upper metric mean dimension. Similar to [5], we will say that a family $\mathbf{\Gamma} := (\Gamma_\varepsilon)_{0<\varepsilon<1}$ is an $\varepsilon-$scaled pressure function if satisfies the scaled translation invariance with respect to the constant $\alpha = \log(1/\varepsilon)$. The next definition was given in [5].

Definition 3.6. *The nonlinear upper metric mean dimension of a family of $\varepsilon-$scaled pressure functions* $\mathbf{\Gamma} = (\Gamma_\varepsilon)_{0<\varepsilon<1}$ is defined as

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\mathbf{\Gamma}, d\Big) = \limsup_{\varepsilon\to 0}\frac{\Gamma_\varepsilon(\mathcal{E})}{|\log\varepsilon|}. \tag{8}$$

The following result gives a variational principle to the nonlinear upper metric mean dimensions for families of $\varepsilon-$pressure functions and, it is a direct consequence of Theorem A and Remark 3.4.

Theorem B. *Let (X,d) be a compact metric space and $\mathbf{\Gamma} := (\Gamma_\varepsilon)_{0<\varepsilon<1}$ be a $\varepsilon-$pressure functions defined on $B_c(X)$ and such that $\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\Gamma, d\Big) < \infty$. Then*

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\mathbf{\Gamma}, d\Big) = \sup_{\mu\in\mathcal{M}_a(X)}\Big\{\mathbb{H}(\mu)+\mathcal{E}(\mu)\Big\}, \tag{9}$$

where $\mathbb{H}(\mu) = \inf_{\mathcal{E}\in\mathbb{H}_\Gamma}\mathcal{E}(\mu)$ and $\mathbb{H}_{\mathbf{\Gamma}} := \{\mathcal{E}\in \boldsymbol{B}_c(X) : \Pi^{-\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\Gamma, d\Big)\le 0\}$. Moreover, the map $\mathbb{H}$ is concave, upper semi-continuous, bounded from above by $\Pi^{0}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\mathbf{\Gamma}, d\Big)$ (here 0 denotes the energy $\mathcal{E}\equiv 0$) and satisfies for every $\mu\in\mathcal{M}_a(X)$

$$\mathbb{H}(\mu) = \inf_{\mathcal{E}\in B_c(X)}\Big\{\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\mathbf{\Gamma}, d\Big) - \mathcal{E}(\mu)\Big\}. \tag{10}$$

3.3. **Nonlinear metric mean dimension.** Given $K\subset X$ a closed subset, an energy $\mathcal{E}$ and $\varepsilon>0$ we recall

$$S(K,f,d,\mathcal{E},\varepsilon,n) = \sup\left\{\sum_{x\in\mathcal{C}} e^{n|\log\varepsilon|\mathcal{E}(\Delta_x^n)} : \mathcal{C} \text{ is an } (n,\varepsilon)\text{-separated set of } K\right\},$$

and

$$P(K,f,d,\mathcal{E},\varepsilon) = \limsup_{n\to\infty}\frac{1}{n}\log S(K,f,d_n,\mathcal{E},\varepsilon,n).$$

Definition 3.7. *The nonlinear upper metric mean dimension* of a closed subset K on X is given as

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(K, f, d\Big) = \lim_{\varepsilon\to 0}\frac{P(K,f,d,\mathcal{E},\varepsilon)}{|\log\varepsilon|}. \tag{11}$$

We also can define the nonlinear metric mean dimension by spanning set. It is not hard to check as $\mathcal{E}$ is continuous, then these two definitions are equivalent.

If we take the energy with potential $\mathcal{E}(\mu) = F(\mu(\varphi))$, where $F(x) = x$, we obtain the definition of the upper metric mean dimension introduced in [5] Moreover, if $\mathcal{E}\equiv 0$, we have the usual definition of metric mean dimension (see [17]). Moreover, considering the family of ε-scaled pressure functions $\mathbf{\Gamma} = (P(X,f,d,\mathcal{E},\varepsilon))_{0<\varepsilon<1}$, we have that

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(\mathbf{\Gamma}, d\Big) = \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big).$$

We also define *the nonlinear measure theoretical metric mean dimension* of any measure $\mu \in \mathcal{M}_f(X)$ as

$$\Pi^{\mathcal{E}} H(f,\mu) = H(f,\mu) + \mathcal{E}(\mu), \tag{12}$$

where $H(f,\mu)$ is defined as in Section 2.

Remark 3.8. By the definition, positive nonlinear measure theoretical metric mean dimension imply infinity nonlinear pressure.

It is possible to consider another versions of the nonlinear measure theoretic metric mean dimension by another types of measure theoretic entropy (see definitions in Section 2) as

$$\Pi^{\mathcal{E}} H^K_\delta(f,\mu) = H^K_\delta(f,\mu) + \mathcal{E}(\mu), \text{ for } \delta \in (0,1),$$
$$\Pi^{\mathcal{E}} H^{BK}(f,\mu) = H^{BK}(f,\mu) + \mathcal{E}(\mu).$$

By [24] it is not difficult to see that

$$\sup_{\mu \in \mathcal{M}_f(X)} \Pi^{\mathcal{E}} H(f,\mu) = \sup_{\mu \in \mathcal{M}_f(X)} \Pi^{\mathcal{E}} H^K_\delta(f,\mu) = \sup_{\mu \in \mathcal{M}_f(X)} \Pi^{\mathcal{E}} H^{BK}(f,\mu).$$

3.4. **Nonlinear upper metric mean dimension.** With the help of Theorems A and B, we extend several results of [4] to the setting of *nonlinear upper metric mean dimension.* Our first statement provides an alternative variational representation. Define

$$\Pi^{\mathcal{E}} \mathrm{H}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) := \limsup_{\varepsilon\to 0} \frac{1}{|\log\varepsilon|} \sup_{\mu\in\mathcal{M}_f(X)} \big\{ \inf_{|\xi|<\varepsilon} h_\mu(f,\xi) + |\log\varepsilon|\,\mathcal{E}(\mu)\big\},$$

and, for $\delta \in (0,1)$,

$$\Pi^{\mathcal{E}}_\delta \overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) := \limsup_{\varepsilon\to 0} \frac{1}{|\log\varepsilon|} \sup_{\mu\in\mathcal{M}_f(X)} \big\{ h^K_\mu(f,\varepsilon,\delta) + |\log\varepsilon|\,\mathcal{E}(\mu)\big\}.$$

Theorem C. *Let $f : X \to X$ be a continuous map on a compact space X, and let $\mathcal{E} : \mathcal{M}_f(X) \to \mathbb{R}$ be an energy. Then*

$$\Pi^{\mathcal{E}} \mathrm{H}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \Pi^{\mathcal{E}}_\delta \overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d).$$

If, in addition, $\mathcal{E}$ is convex, then

$$\Pi^{\mathcal{E}} \overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \Pi^{\mathcal{E}} \mathrm{H}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \Pi^{\mathcal{E}}_\delta \overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d). \tag{13}$$

In the next, we present a variational principle by the nonlinear upper metric mean dimension setting.

Theorem D. *Let (X,d) be a compact metric space and $f : X \to X$ be a continuous map such that $\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_M(X,f,d) < \infty$. Then, for every $\delta \in (0,1)$ and energy $\mathcal{E} \in B_c(X)$*

$$\Pi^{\mathcal{E}} \overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X,f,d\Big) = \sup_{\mu\in\mathcal{M}_f(X)} \Pi^{\mathcal{E}} H^K_\delta(f,\mu).$$

Remark 3.9. If the energy is not convex, the conclusion may fail. Moreover — unlike the (linear) metric mean dimension — finite, nonlinear topological pressure does not necessarily imply that the nonlinear metric mean dimension vanishes, it can be negative (see Section 6.4).

As an immediate consequence of Theorem D we have the following:

Corollary A. *Let $f : X \to X$ be a continuous map of a compact space and let $\mathcal{E} : \mathcal{M}(X) \to \mathbb{R}$ be a convex energy. Then there exists at least one nonlinear equilibrium measure.*

3.5. **Nonlinear upper mean box dimension.** Let $X = [0,1]$, and set $Y := X^{\mathbb{Z}}$ endowed with the product (Tychonoff) metric

$$D\big((u_n)_{n\in\mathbb{Z}}, (v_n)_{n\in\mathbb{Z}}\big) := \sum_{n\in\mathbb{Z}} \frac{d(u_n, v_n)}{2^{|n|}}.$$

For a subshift $\mathbb{S} \subset Y$, Gutman and Śpiewak [8] introduced the *upper mean box dimension* by

$$\overline{\mathrm{mdim}}_B(\mathbb{S}) := \lim_{n\to\infty} \frac{\overline{\dim}_B\big(\pi_n(\mathbb{S})\big)}{n},$$

where π_n denotes the projection onto the first n coordinates and $\overline{\dim}_B\big(\pi_n(\mathbb{S})\big)$ is computed with respect to $\|\cdot\|_\infty$ on $[0,1]^n$, and also proved

$$\overline{\mathrm{mdim}}_M(\mathbb{S}, \sigma, D) \le \overline{\mathrm{mdim}}_B(\mathbb{S}). \tag{14}$$

In what follows we extend the previous result to the nonlinear metric mean dimension.

Fix a continuous *energy* $\mathcal{E} : \mathcal{M}(X) \to \mathbb{R}$, and for $m, n \in \mathbb{N}$ and $x_{-m}, \dots, x_n \in X$ define the empirical measure

$$\Delta_{-m,n}(x_{-m}, \dots, x_n) := \frac{1}{n+m+1} \sum_{i=-m}^{n} \delta_{x_i},$$

and

$$\Delta_n(x_1, \dots, x_n) := \frac{1}{n} \sum_{i=1}^{n} \delta_{x_i}.$$

Let $\pi_1 : Y \to X$ be the projection to the first coordinate, $\pi_1\big((x_n)_{n\in\mathbb{Z}}\big) := x_1$. It induces an energy $\hat{\mathcal{E}} : \mathcal{M}(Y) \to \mathbb{R}$ by $\hat{\mathcal{E}}(\mu) := \mathcal{E}\big((\pi_1)_*\mu\big)$. In particular, for any $x = (x_n)_{n\in\mathbb{Z}} \in Y$,

$$\hat{\mathcal{E}}\,(\Delta_n^x) = \mathcal{E}\big(\Delta(x_1, \dots, x_n)\big),$$

where $\Delta_x^n = \frac{1}{n}\sum_{i=0}^{n-1} \delta_{\sigma^i(x)}$.

Based on this, we now introduce the *nonlinear upper mean box dimension* as

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) = \limsup_{n\to\infty} \frac{1}{n} \limsup_{\varepsilon\to 0} \frac{1}{|\log\varepsilon|} \log \inf_E \left(\sum_{(x_1,\dots,x_n)\in E} e^{n|\log\varepsilon|\mathcal{E}(\Delta(x_1,\dots,x_n))} \right),$$

where E is taken over all ε-spanning subsets of $\pi_n(\mathbb{S})$. It is easy to see that if $\mathcal{E} \equiv 0$, then

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) = \overline{\mathrm{mdim}}_B(\mathbb{S}).$$

We say that the energy $\mathcal{E}$ on X is *subadditive* if for all $n, m \in \mathbb{N}$ and all *concatenable* words $u \in \pi_n(\mathbb{S})$, $v \in \pi_m(\mathbb{S})$ with $uv \in \pi_{n+m}(\mathbb{S})$,

$$(n+m)\,\mathcal{E}\big(\Delta_{m+n}(uv)\big) \;\le\; n\,\mathcal{E}\big(\Delta_n(u)\big) + m\,\mathcal{E}\big(\Delta_m(v)\big).$$

Remark 3.10. We notice that if the energy is convex or linear, it is subadditive.

Finally we have the following extension of the result of [8]:

Theorem E. *Let $\mathbb{S} \subset [0,1]^{\mathbb{Z}}$ be a subshift and $\mathcal{E}$ is a subadditive energy. Then*

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S},\sigma,D) \leq \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}).$$

4. Proofs of the main results.

Before we start the proofs we recall the definition of the Wasserstein metric on $\mathcal{M}(X)$. For any $\mu_1,\mu_2 \in \mathcal{M}(X)$, let the following *Wasserstein metric* be

$$\mathrm{W}(\mu_1,\mu_2) = \sup\Big\{\mu_1(f)-\mu_2(f)\colon f \text{ 1-Lipschitz function } X\mapsto\mathbb{R}\Big\}.$$

By "Kantorovich duality", the definition above is equivalent to

$$\mathrm{W}(\mu_1,\mu_2) = \inf\left\{\int d(x,y)\,\mathrm{d}\pi(x,y)\colon \pi\in\Gamma(\mu_1,\mu_2)\right\}$$

where d is the metric on X and $\Gamma(\mu_1,\mu_2)$ is the set of Borel probability measures on $X\times X$ with marginals μ_1 and μ_2 (called of *transport plans*). The weak$^*-$topology on $\mathcal{M}(X)$ is induced by Wasserstein metric due to the compactness of X, and holds that

$$\mathrm{W}(\mu_1,\mu_2) \leq \mathrm{diam}(X)\,\|\mu_1-\mu_2\|_{\mathrm{TV}}\,,$$

where $\|\cdot\|_{\mathrm{TV}}$ is the total variation norm.

The following result is from the Birkhoff's ergodic theorem (see e.g. [4, Lemma 2.1]).

Lemma 4.1. *Let $\mu\in\mathcal{M}_f(X)$ be ergodic. Then for μ-almost all $x\in X$, we have $\Delta_x^n\to\mu$ in the weak$^*-$topology.*

4.1. Proof of Theorem A.

For any fixed $\mathcal{E}\in\mathbf{B}(X)$, we will first show that

$$\Upsilon(\mathcal{E}) \geq \sup_{\mu\in\mathcal{M}_a(X)}\Big\{\mathcal{H}(\mu)+\mathcal{E}(\mu)\Big\}.$$

Consider $\tilde{\mathcal{E}} := \Upsilon(\mathcal{E})-\mathcal{E}$, for every pressure function Υ. By translation invariance, we have $\Upsilon(-\tilde{\mathcal{E}}) = \Upsilon(\mathcal{E}-\Upsilon(\mathcal{E})) = 0$ and so $\tilde{\mathcal{E}}\in\mathcal{A}_\Upsilon$. Therefore, for any $\mu\in\mathcal{M}_a(X)$

$$\mathcal{H}(\mu)\leq\tilde{\mathcal{E}}(\mu) = \Upsilon(\mathcal{E})-\mathcal{E}(\mu),$$

proving the inequality.

To prove the second part of the theorem, given any energy $\mathcal{E}\in\mathbf{B_c}(X)$, we will find a measure $\mu^*\in\mathcal{M}_a(X)$ such that

$$\Upsilon(\mathcal{E})\leq\mathcal{H}^*(\mu^*)+\mathcal{E}(\mu^*).$$

Notice that, it suffices to prove for $\Upsilon(\mathcal{E})=0$, because the general case follows taking $\mathcal{E}-\Upsilon(\mathcal{E})$ and using the translation invariance of Υ.

Assuming $\mathcal{E}$ is such that $\Upsilon(\mathcal{E})=0$, as $-(-\mathcal{E})=\mathcal{E}$ we have that the energy $-\mathcal{E}$ does not belong to the set

$$B_\Upsilon = \{\tilde{\mathcal{E}}\in\mathbf{B_c}(X) : \Upsilon(-\tilde{\mathcal{E}})<0\}.$$

The set B_Υ is convex and open and then, by geometric Hahn-Banach Theorem version there exists a continuous and not identically zero linear functional $L\colon\ \mathbf{B_c}(X)\to\ \mathbb{R}$ so that

$$L(\mathcal{E})+\inf_{\tilde{\mathcal{E}}\in B_\Upsilon}L(\tilde{\mathcal{E}})\geq 0.$$

It follows from [3, Lemma 3.1] that L is positive and $L(1) > 0$, where 1 is the constant function equal to one. Therefore, by Riez-Markov-Kakutani Theorem, there exists a finitely additive probability measure $\mathbb{P}_{\mathcal{E}}$ (note that $\mathbb{P}_{\mathcal{E}}$ is a measure on the set of measures) such that

$$\frac{L(\tilde{\mathcal{E}})}{L(1)} = \int \tilde{\mathcal{E}}(\mu)\, d\mathbb{P}_{\mathcal{E}}(\mu) \leq \tilde{\mathcal{E}}\left(\mu^*\right), \text{ for all } \tilde{\mathcal{E}} \in \mathbf{B_c}(X),$$

where $\mu^* = \int \mu \, d\mathbb{P}_{\mathcal{E}}(\mu) \in \mathcal{M}_a(X)$.

Notice that, the set $\mathcal{A}^*_{\Upsilon,\varepsilon} := \{\tilde{\mathcal{E}} + \varepsilon : \tilde{\mathcal{E}} \in \mathcal{A}^*_{\Upsilon}\}$ is contained in B_Υ, for all $\varepsilon > 0$. Then for all $\varepsilon > 0$

$$\mathcal{H}^*(\mu^*) = \inf_{\tilde{\mathcal{E}} \in \mathcal{A}^*_\Upsilon} \tilde{\mathcal{E}}(\mu^*) \leq \inf_{\tilde{\mathcal{E}} \in B_\Upsilon} \tilde{\mathcal{E}}(\mu^*) \leq \inf_{\tilde{\mathcal{E}} \in \mathcal{A}^*_{\Upsilon,\varepsilon}} \tilde{\mathcal{E}}(\mu^*) = \inf_{\tilde{\mathcal{E}} \in \mathcal{A}^*_\Upsilon} \tilde{\mathcal{E}}(\mu^*) + \varepsilon.$$

By the arbitrariness of $\varepsilon > 0$, $\mathcal{H}^*(\mu) = \inf_{\tilde{\mathcal{E}} \in B_\Upsilon} \tilde{\mathcal{E}}(\mu)$, and so

$$\mathcal{H}(\mu^*) + \mathcal{E}(\mu^*) \geq \inf_{\tilde{\mathcal{E}} \in B_\Upsilon} \tilde{\mathcal{E}}(\mu^*) + \frac{L(\mathcal{E})}{L(1)} \geq \frac{1}{L(1)} \left(L(\mathcal{E}) + \inf_{\tilde{\mathcal{E}} \in B_\Upsilon} L(\tilde{\mathcal{E}}) \right) \geq 0 = \Upsilon(\mathcal{E}).$$

To conclude, we are left to prove that

$$\mathcal{H}^*(\mu) = \inf_{\mathcal{E} \in \mathbf{B}_c(X)} \Big\{ \Upsilon(\mathcal{E}) - \mathcal{E}(\mu) \Big\}.$$

Fix any function α satisfying (6) playing the role of $\mathcal{H}^*$. Then for every $\mathcal{E} \in \mathbf{B}_c(X)$ and $\mu \in \mathcal{M}_a(X)$

$$\Upsilon(-\mathcal{E}) \geq \alpha(\mu) - \mathcal{E}(\mu)$$

which is equivalent to

$$\alpha(\mu) \leq \Upsilon(-\mathcal{E}) + \mathcal{E}(\mu).$$

Hence for all $\mu \in \mathcal{M}_a(X)$

$$\alpha(\mu) \leq \inf_{\mathcal{E} \in \mathbf{B}_c(X)} \Big\{ \Upsilon(-\mathcal{E}) + \mathcal{E}(\mu) \Big\}.$$

As $\mathcal{A}^*_\Upsilon \subset \mathbf{B}_c(X)$ and $\Upsilon(-\mathcal{E}) \leq 0$, for every $\mathcal{E} \in \mathcal{A}^*_\Upsilon$, we obtain that

$$\alpha(\mu) \leq \inf_{\mathcal{E} \in \mathbf{B}_c(X)} \Big\{ \Upsilon(-\mathcal{E}) + \mathcal{E}(\mu) \Big\} \leq \inf_{\mathcal{E} \in \mathcal{A}^*_\Upsilon} \Big\{ \Upsilon(-\mathcal{E}) + \mathcal{E}(\mu) \Big\} \leq \inf_{\mathcal{E} \in \mathcal{A}^*_\Upsilon} \mathcal{E}(\mu) = \mathcal{H}^*(\mu).$$

Taking in particular $\alpha = \mathcal{H}^*$ in expression above, we conclude that

$$\mathcal{H}^*(\mu) = \inf_{\mathcal{E} \in \mathbf{B}_c(X)} \Big\{ \Upsilon(-\mathcal{E}) + \mathcal{E}(\mu) \Big\} = \inf_{\mathcal{E} \in \mathbf{B}_c(X)} \Big\{ \Upsilon(\mathcal{E}) - \mathcal{E}(\mu) \Big\},$$

proving so (7) and ending the proof of Theorem. □

4.2. **Proof of Theorem C.** First, we prove the second equation. Notice that, for every $\delta \in (0,1)$ and $\varepsilon > 0$, in [24, Lemma 3.1], the authors proved that

$$h^{BK}_\mu(f, 2\varepsilon) \leq \inf_{|\xi| < \varepsilon} h_\mu(f, \xi) \leq h^K_\mu(f, \varepsilon/4, \delta) \leq h^{BK}_\mu(f, \varepsilon/64).$$

for all $\mu \in \mathcal{M}^{erg}_f(X)$. Then, using the ergodic decomposition for any $\nu \in \mathcal{M}_f(X)$, it follows that

$$h^{BK}_\nu(f, 2\varepsilon) \leq \inf_{|\xi| < \varepsilon} h_\nu(f, \xi) \leq h^K_\nu(f, \varepsilon/4, \delta) \leq h^{BK}_\nu(f, \varepsilon/64),$$

Therefore, we obtain that

$$\Pi^{\mathcal{E}} \overline{\mathrm{Hmdim}}_{\mathrm{M}}(X, f, d) = \Pi^{\mathcal{E}}_{\delta} \overline{\mathrm{mdim}}_{\mathrm{M}}(X, f, d). \tag{15}$$

Next we prove the first equation. To that end, given any energy $\mathcal{E}$, $\mu \in \mathcal{M}_f^{erg}(X)$ and $\delta > 0$, we define

$$\mathcal{N}_\mu(\mathcal{E}, \varepsilon, \delta, n) = \inf \Big\{ S(A, f, d, \mathcal{E}, \varepsilon) : \mu(A) > 1 - \delta \Big\}$$

and

$$P_\mu^K(\mathcal{E}, \varepsilon, \delta) = \limsup_{n \to \infty} \frac{1}{n} \log \mathcal{N}_\mu(\mathcal{E}, \varepsilon, \delta, n).$$

We can extend the notion above for $\mu \in \mathcal{M}_f(X)$ using its ergodic decomposition

$$P_\mu^K(\mathcal{E}, \varepsilon, \delta) = \int_{\mathcal{M}_f^{erg}(X)} P_m^K(\mathcal{E}, \varepsilon, \delta)\ d\mathbb{P}(m),$$

where $\mu = \int_{\mathcal{M}_f^{erg}(X)} m\ d\mathbb{P}_\mu(m)$. We define

$$\Pi^{\mathcal{E}} P_\delta^K(\mu) := \sup_{(\mu_\varepsilon)_\varepsilon \in \mathcal{M}(\mu)} \limsup_{\varepsilon \to 0} \frac{P_\mu^K(\mathcal{E}, \varepsilon, \delta)}{|\log \varepsilon|}.$$

Remark 4.2. Notice that $P_\mu^K(\mathcal{E}, \varepsilon, \delta) \le P(X, f, d, \mathcal{E}, \varepsilon)$ and consequently, for all $\delta \in (0, 1)$ and $\mu \in \mathcal{M}_f(X)$

$$\Pi^{\mathcal{E}} P_\delta^K(\mu) \le \Pi^{\mathcal{E}} \overline{\mathrm{mdim}}_{\mathrm{M}} \Big(X, f, d \Big).$$

Lemma 4.3. *Let $\gamma > 0$, $\mathcal{E} \in \mathbf{B}(X)$, $\mu \in \mathcal{M}_f^{\mathrm{erg}}(X)$, and $1 > \delta_1 > \delta_2 > \delta_3 > 0$. Then, for all sufficiently small $\varepsilon > 0$,*

$$\frac{h_\mu^K(f, \varepsilon, \delta_1)}{|\log \varepsilon|} + \mathcal{E}(\mu) - \gamma \le \frac{P_\mu^K(\mathcal{E}, \varepsilon, \delta_2)}{|\log \varepsilon|} \le \frac{h_\mu^K(f, \varepsilon, \delta_3)}{|\log \varepsilon|} + \mathcal{E}(\mu) + \gamma. \quad (16)$$

If $\mu \in \mathcal{M}_f(X)$ (not necessarily ergodic) and $\mathcal{E} \in \mathbf{B}_c(X)$ is convex, then the left-hand inequality in (16) remains valid.

Proof. Since $\mathcal{E}$ is continuous on the compact space $\mathcal{M}_f(X)$ (with the weak* topology), there exists $\tau = \tau(\gamma) > 0$ such that

$$\mathrm{W}(\nu, \mu) \le 2\tau \implies |\mathcal{E}(\nu) - \mathcal{E}(\mu)| < \gamma \qquad (\forall\, \nu \in \mathcal{M}_f(X)). \quad (17)$$

By Lemma 4.1 there exist $Y_1 \subset X$ and $N_1 \in \mathbb{N}$ such that

$$\mu(Y_1) \ge 1 - \delta_1 \quad \text{and} \quad \mathrm{W}(\Delta_x^n, \mu) \le \tau \quad \text{for all } x \in Y_1,\ n \ge N_1.$$

Fix $0 < \varepsilon \ll 1$ and $0 < \delta_3 < \delta_2 < 1$. Let $Y_2 \subset X$ satisfy $\mu(Y_2) > 1 - \delta_2$, and let $\mathcal{S}$ be a minimal (n, ε)-spanning set of Y_2 . Then $Y_1 \cap Y_2 \neq \emptyset$ and $\mu(Y_1 \cap Y_2) > 1 - \delta_3$. Let $\mathcal{S}_1 \subset \mathcal{S}$ be a minimal (n, ε)-spanning set of $Y_1 \cap Y_2$. For each $y \in Y_1 \cap Y_2$ there exists $x \in \mathcal{S}_1$ with $d(f^i(x), f^i(y)) \le \varepsilon$ for $0 \le i < n$, hence

$$\mathrm{W}(\Delta_x^n, \Delta_y^n) \le \frac{1}{n} \sum_{i=0}^{n-1} d\big(f^i(x), f^i(y)\big) \le \varepsilon.$$

Combining with $\mathrm{W}(\Delta_y^n, \mu) \le \tau$ and the triangle inequality,

$$\mathrm{W}(\Delta_x^n, \mu) \le \varepsilon + \tau \le 2\tau. \quad (18)$$

By (17), $\mathcal{E}(\Delta_x^n) \le \mathcal{E}(\mu) + \gamma$, thus

$$\sum_{x \in \mathcal{S}_1} e^{n |\log \varepsilon| \mathcal{E}(\Delta_x^n)} \le |\mathcal{S}_1|\ e^{n |\log \varepsilon| (\mathcal{E}(\mu) + \gamma)} \le |\mathcal{S}|\ e^{n |\log \varepsilon| (\mathcal{E}(\mu) + \gamma)}.$$

Taking $\frac{1}{n}\log$, then $\limsup_{n\to\infty}$, and using the definitions of $P^K_\mu(\mathcal{E},\varepsilon,\delta_2)$ and $h^K_\mu(f,\varepsilon,\delta_3)$ yields

$$\frac{P^K_\mu(\mathcal{E},\varepsilon,\delta_2)}{|\log\varepsilon|} \;\le\; \frac{h^K_\mu(f,\varepsilon,\delta_3)}{|\log\varepsilon|}+\mathcal{E}(\mu)+\gamma.$$

To prove the lower bound, we let $Y_3\subset X$ satisfy $\mu(Y_3)>1-\delta_2$ and let $\mathcal{S}'$ be an (n,ε)-spanning set of Y_3. Let $\mathcal{S}'_1\subset\mathcal{S}'$ be a minimal (n,ε)-spanning set of $Y_1\cap Y_3$. Arguing as above, for $x\in\mathcal{S}'_1$ there is $y\in Y_1\cap Y_3$ with $\mathrm{W}(\Delta^n_x,\mu)\le 2\tau$, hence $\mathcal{E}(\Delta^n_x)\ge\mathcal{E}(\mu)-\gamma$. Therefore,

$$\sum_{x\in\mathcal{S}'} e^{n|\log\varepsilon|\mathcal{E}(\Delta^n_x)} \;\ge\; \sum_{x\in\mathcal{S}'_1} e^{n|\log\varepsilon|(\mathcal{E}(\mu)-\gamma)} \;\ge\; b_\mu(n,\varepsilon,\delta_1)\; e^{n|\log\varepsilon|(\mathcal{E}(\mu)-\gamma)},$$

where $b_\mu(n,\varepsilon,\delta_1)$ denotes the minimal cardinality of an (n,ε)-spanning set of a subset of X with μ-measure at least $1-\delta_1$. Taking $\frac{1}{n}\log$ and then $\limsup_{n\to\infty}$ gives

$$\frac{P^K_\mu(\mathcal{E},\varepsilon,\delta_2)}{|\log\varepsilon|} \;\ge\; \frac{h^K_\mu(f,\varepsilon,\delta_1)}{|\log\varepsilon|}+\mathcal{E}(\mu)-\gamma.$$

For the non-ergodic case, we use the ergodic decomposition. More precisely, if $\mu=\int m\,d\mathbb{P}_\mu(m)$ is the ergodic decomposition and $\mathcal{E}$ is convex, then

$$P^K_\mu(\mathcal{E},\varepsilon,\delta_2)\ge\int P^K_m(\mathcal{E},\varepsilon,\delta_2)\,d\mathbb{P}_\mu(m)\ge\int\Big(h^K_m(f,\varepsilon,\delta_1)+|\log\varepsilon|(\mathcal{E}(m)-\gamma)\Big)\,d\mathbb{P}_\mu(m),$$

which yields the asserted lower bound after dividing by $|\log\varepsilon|$ and applying Jensen's inequality. □

We also need the following lemma.

Lemma 4.4. *Let $\mathcal{E}$ is convex. For any $\varepsilon>0$ and $\delta\in(0,1)$,*

$$\sup_{\mu\in\mathcal{M}^{erg}_f(X)}\left\{h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu)\right\}=\sup_{\mu\in\mathcal{M}_f(X)}\left\{h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu)\right\}.$$

Proof. We only need to prove that

$$\sup_{\mu\in\mathcal{M}^{\mathrm{erg}}_f(X)}\left\{h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu)\right\}\ge\sup_{\mu\in\mathcal{M}_f(X)}\left\{h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu)\right\},$$

as the other direction is clear. Indeed, for any $\mu\in\mathcal{M}_f(X)$, let

$$\mu=\int_{M^{\mathrm{erg}}_f(X)} m d\mathbb{P}(m)$$

be its ergodic decomposition. Then by convexity of $\mathcal{E}$

$$\begin{aligned}
h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu) &\le \int_{M^{\mathrm{erg}}_f(X)} h^K_m(f,\varepsilon,\delta)d\mathbb{P}(m)+|\log\varepsilon|\mathcal{E}\left(\int_{M^{\mathrm{erg}}_f(X)} md\mathbb{P}(m)\right)\\
&\le \int_{M^{\mathrm{erg}}_f(X)} h^K_m(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}\,(m)\,d\mathbb{P}(m)\\
&\le \sup_{\mu\in\mathcal{M}^{\mathrm{erg}}_f(X)}\left\{h^K_\mu(f,\varepsilon,\delta)+|\log\varepsilon|\mathcal{E}(\mu)\right\}.
\end{aligned}$$

The proof is completed by the arbitrariness of $\mu\in\mathcal{M}_f(X)$. □

Let us proceed to prove the Theorem C. We recall $\omega_n(\mathcal{C})=\sum_{x\in\mathcal{C}}e^{n|\log\varepsilon|\mathcal{E}(\Delta^n_x)}$. Indeed, we emphasize that the next theorem constitutes an important tool to prove our results and it was proved in [4, Theorem 2.5].

Theorem 4.5. *Assume that f is continuous and consider $\varepsilon > 0$, an increasing sequence of positive integers $(n_k)_k$ and a sequence of (ε, n_k)-separated sets $\mathcal{C}_k$ such that $\frac{\log \omega_{n_k}(\mathcal{C}_k)}{n_k}$ converges.*

Then there exist partitions $\mathbf{D}_k = (\mathcal{D}_{k,i})_{1\le i\le N_k}$ of $\mathcal{C}_k$ and $I_k \subset [\![1, N_k]\!]$ non-empty sets with the following properties:

i. for every sequence $(i_k)_k \in \prod_k I_k$, every accumulation point μ_∞ of $(\mu_{\mathcal{D}_{k,i_k}})_k$ is T-invariant and

$$h_{\mu_\infty}(f, \alpha) + |\log \varepsilon|\mathcal{E}(\mu_\infty) \ge \lim_k \frac{\log \omega_{n_k}(\mathcal{C}_k)}{n_k},$$

for every finite partition α of X into subsets of diameter less than ε and with negligible boundaries with respect to μ_∞.

ii. As $k \to \infty$,

$$\sum_{i\in[\![1,N_k]\!]\setminus I_k} \omega_{n_k}(\mathcal{D}_{k,i}) = o\big(\omega_{n_k}(\mathcal{C}_k)\big)$$

Proof of the Theorem C. We now assume that $\mathcal{E}$ is convex to prove (13). From Lemma 4.3, for any $\gamma > 0$, $\mathcal{E} \in \mathbf{B}(X)$, $\mu \in \mathcal{M}_f^{erg}(X)$ and $1 > \delta_1 > \delta_2 > 0$, it follows that

$$h^K_\mu(f, \varepsilon, \delta_1) + |\log \varepsilon|\big(\mathcal{E}(\mu) - \gamma\big) \le P^K_\mu(\mathcal{E}, \epsilon, \delta_2).$$

Consequently, for every $\delta \in (0, 1)$, by Lemma 4.4 and Remark 4.2, one has

$$\limsup_{\varepsilon\to 0} \frac{1}{|\log \varepsilon|} \sup_{\mu\in\mathcal{M}_f(X)} \left\{h^K_\mu(f, \varepsilon, \delta) + |\log \varepsilon|\mathcal{E}(\mu)\right\} \le \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big).$$

Let $\delta > 0$, take a sequence $\mathcal{C}_k$ of (ε, n_k)-separated sets for some $\varepsilon \in (0, 1)$ such that

$$\lim_{k\to\infty} \frac{\log \omega_{n_k}(\mathcal{C}_k)}{n_k} \ge (\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big) - \delta)|\log \varepsilon|.$$

By Theorem 4.5, for any sequence $i_k \in I_k$ and any accumulation point μ_∞ of $(\mu_{\mathcal{D}_{k,i_k}})$, for any finite measurable partition with $|\alpha| < \varepsilon$,

$$h_{\mu_\infty}(f, \alpha) + |\log \varepsilon|\mathcal{E}(\mu_\infty) \ge \lim_{k\to\infty} \frac{\log \omega_{n_k}(\mathcal{C}_k)}{n_k},$$

where $|\alpha|$ is the maximal diameter of elements in α. So,

$$\frac{1}{|\log \varepsilon|}\left(\inf_{|\xi|<\varepsilon} h_{\mu_\infty}(f, \xi) + |\log \varepsilon|\mathcal{E}(\mu_\infty)\right) \ge \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big) - \delta,$$

which together with the arbitrariness of $\delta > 0$,

$$\limsup_{\varepsilon\to 0} \frac{1}{|\log \varepsilon|} \sup_{\mu\in\mathcal{M}_f(X)} \left\{\inf_{|\xi|<\varepsilon} h_\mu(f, \xi) + |\log \varepsilon|\mathcal{E}(\mu)\right\} \ge \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X, f, d\Big).$$

Combining this with (15), we finish the proof. □

4.3. **Proof of Theorem D.**

We start showing the inequality

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X, f, d) \ge \sup_{\mu\in\mathcal{M}_f(X)} \Pi^{\mathcal{E}} H^K_\delta(f, \mu).$$

By the first inequality of the Lemma 4.3, we obtain that

$$\Pi^{\mathcal{E}} P^K_{\delta_2}(\mu) \ge H^K_{\delta_1}(f, \mu) + \mathcal{E}(\mu), \tag{19}$$

for any $\mu \in \mathcal{M}_f(X)$, $\mathcal{E} \in \mathbf{B}_c(X)$ and $1 > \delta_1 > \delta_2 > 0$, which together with Remark 4.2, implies that

$$H_\delta^K(f,\mu) \leq \inf_{\mathcal{E}\in\mathbf{B}_c(X)} \Big\{\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\Big(X,f,d\Big) - \mathcal{E}(\mu)\Big\} = \mathbb{H}(\mu),$$

for every $\delta \in (0,1)$ and $\mu \in \mathcal{M}_f(X)$. Hence by Theorem B, we obtain

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \sup_{\mu\in\mathcal{M}_a(X)} \Big\{\mathbb{H}(\mu) + \mathcal{E}(\mu)\Big\} \geq \sup_{\mu\in\mathcal{M}_f(X)} \Pi^{\mathcal{E}} H_\delta^K(f,\mu),$$

proving the first inequality.

To prove the reverse inequality, we construct a measure $\mu_0 \in \mathcal{M}_f(X)$ such that

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) \leq \Pi^{\mathcal{E}} H_\delta^K(f,\mu_0).$$

It follows from Theorem C that

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \limsup_{\varepsilon\to 0} \frac{1}{|\log\varepsilon|} \sup_{\mu\in\mathcal{M}_f^{erg}(X)} \Big\{h_\mu^K(f,\varepsilon,\delta) + |\log\varepsilon|\mathcal{E}(\mu)\Big\}.$$

We can find a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ of positive real numbers where $\varepsilon_n \to 0$ as $n \to \infty$ satisfying

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) = \lim_{n\to\infty} \frac{1}{|\log\varepsilon_n|} \sup_{\mu\in\mathcal{M}_f(X)} \Big\{h_\mu^K(f,\varepsilon_n,\delta) + |\log\varepsilon_n|\mathcal{E}(\mu)\Big\}.$$

Then, given $\gamma > 0$, there exists a positive integer N such that for every $n \geq N$,

$$\Big(\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) - \gamma\Big) < \frac{1}{|\log\varepsilon_n|} \sup_{\mu\in\mathcal{M}_f(X)} \Big\{h_\mu^K(f,\varepsilon_n,\delta) + |\log\varepsilon_n|\mathcal{E}(\mu)\Big\}.$$

Moreover, for every $n \geq N$, there exist measures $\mu_{\varepsilon_n} \in \mathcal{M}_f(X)$ where

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) - \gamma < \frac{h_{\mu_{\varepsilon_n}}^K(f,\varepsilon_n,\delta)}{|\log(\varepsilon_n)|} + \mathcal{E}(\mu_{\varepsilon_n}).$$

Taking a subsequence if necessary, we have that the sequence (μ_{ε_n}) converges to a measure $\mu_0 \in \mathcal{M}_f(X)$ in the weak$^*-$topology. From Definition 2.3 and the continuity of energy $\mathcal{E}$ in the weak$^*-$topology, we obtain

$$H_\delta^K(f,\mu_0) \geq \lim_{n\to\infty} \frac{h_{\mu_{\varepsilon_n}}^K(f,\varepsilon_n,\delta)}{|\log(\varepsilon_n)|} > \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) - \mathcal{E}(\mu_0) - 2\gamma$$

for every $\gamma > 0$. Therefore

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X,f,d) \leq H_\delta^K(f,\mu_0) + \mathcal{E}(\mu_0) = \Pi^{\mathcal{E}} H_\delta^K(f,\mu_0),$$

ending the proof of the theorem. □

4.4. **Proof of Theorem E.** We start by showing that, in the case of the shift space $Y = X^{\mathbb{Z}}$ and $X = [0,1]$, the nonlinear metric mean dimension admits a more canonical expression.

For $n \in \mathbb{N}$ and $\varepsilon > 0$, denote

$$Z_n^{\mathrm{sep}}(\varepsilon) := \sup_E \left(\sum_{(x_1,\ldots,x_n)\in E} e^{n|\log\varepsilon|\mathcal{E}(\Delta(x_1,\ldots,x_n))}\right).$$

where the supremum ranges over all ε-separated sets $E \subset \pi_n(\mathbb{S})$ with respect to $\|\cdot\|_\infty$ on X^n.

Proposition 4.6. *Let $\mathbb{S} \subset Y$ be a subshift and $\mathcal{E}$ is subadditive. Then*

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S}, \sigma, D) = \limsup_{\varepsilon \to 0} \frac{\limsup_{n\to\infty} \frac{1}{n} \log Z_n^{\mathrm{sep}}(\varepsilon)}{|\log \varepsilon|}.$$

Remark 4.7. By the standard argument,

$$\limsup_{\varepsilon \to 0} \frac{\limsup_{n\to\infty} \frac{1}{n} \log Z_n^{\mathrm{span}}(\varepsilon)}{|\log \varepsilon|} = \limsup_{\varepsilon \to 0} \frac{\limsup_{n\to\infty} \frac{1}{n} \log Z_n^{\mathrm{sep}}(\varepsilon)}{|\log \varepsilon|}, \tag{20}$$

where

$$Z_n^{\mathrm{span}}(\varepsilon) := \inf_E \left(\sum_{(x_1,\ldots,x_n)\in E} \exp\big(n|\log\varepsilon|\,\mathcal{E}(\Delta(x_1,\ldots,x_n))\big) \right),$$

and the infimum runs over all ε-spanning sets $E \subset \pi_n(\mathbb{S})$ with respect to the norm $\|\cdot\|_\infty$ on X^n.

Consequently, Proposition 4.6 admits the spanning-set formulation

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S}, \sigma, D) = \limsup_{\varepsilon \to 0} \frac{\limsup\limits_{n\to\infty} \dfrac{1}{n} Z_n^{\mathrm{span}}(\varepsilon)}{|\log \varepsilon|}.$$

To prove Proposition 4.6, we need the following submultiplicative property of $Z_n^{\mathrm{sep}}(\varepsilon)$.

Proposition 4.8. *If the energy $\mathcal{E}$ is subadditive, then for every $\varepsilon > 0$ and $n, m \in \mathbb{N}$,*

$$Z_{n+m}^{\mathrm{sep}}(\varepsilon) \le Z_n^{\mathrm{sep}}(\varepsilon)\, Z_m^{\mathrm{sep}}(\varepsilon).$$

Consequently, $\log Z_n^{\mathrm{sep}}(\varepsilon)$ *is subadditive in n, and thus*

$$\lim_{n\to\infty} \frac{1}{n} \log Z_n^{\mathrm{sep}}(\varepsilon) = \inf_{n\in\mathbb{N}} \frac{1}{n} \log Z_n^{\mathrm{sep}}(\varepsilon).$$

Proof. Let $F \subset \pi_{n+m}(\mathbb{S})$ be an arbitrary ε-separated set (in $\|\cdot\|_\infty$ on $[0,1]^{n+m}$). Write each $w \in F$ uniquely as $w = uv$ with $u \in \pi_n(\mathbb{S})$ and $v \in \pi_m(\mathbb{S})$, and set

$$P := \{(u,v) : \ uv \in F\} \subset \pi_n(\mathbb{S}) \times \pi_m(\mathbb{S}).$$

Endow the product with the sup metric $d_\infty((u,v),(u',v')) := \max\{\|u-u'\|_\infty, \|v-v'\|_\infty\}$. Then P is ε-separated in d_∞ (since $\|uv - u'v'\|_\infty \ge \varepsilon$ implies $\max\{\|u-u'\|_\infty, \|v-v'\|_\infty\} \ge \varepsilon$).

By subadditivity,

$$e^{(n+m)|\log\varepsilon|\,\mathcal{E}(\Delta_{n+m}(uv))} \le e^{n|\log\varepsilon|\,\mathcal{E}(\Delta_n(u))} \cdot e^{m|\log\varepsilon|\,\mathcal{E}(\Delta_m(v))} =: f(u)\,g(v).$$

Hence

$$\sum_{uv\in F} e^{(n+m)|\log\varepsilon|\,\mathcal{E}(\Delta_{n+m}(uv))} \le \sum_{(u,v)\in P} f(u)\,g(v).$$

We now apply a greedy selection on the first coordinates. From $\mathrm{proj}_1(P) \subset \pi_n(\mathbb{S})$ build a set E as follows: iteratively pick a point e maximizing $f(\cdot)$ among the remaining candidates, add e to E, and remove all candidates in the $\varepsilon/2$-ball around e. Then

1. E is $\varepsilon/2$-separated;
2. for every $u \in \mathrm{proj}_1(P)$ there exists $e \in E$ with $\|u-e\|_\infty < \varepsilon/2$ and $f(u) \le f(e)$.

Divide P into disjoint part

$$P = \bigsqcup_{e\in E} P_e, \qquad P_e := \{(u,v)\in P:\ \|u-e\|_\infty \le \varepsilon/2\}.$$

We claim that for each fixed $e\in E$, the set

$$V_e := \{v\in \pi_m(\mathbb{S}):\ \text{there exists } u \text{ with } (u,v)\in P_e\}$$

is ε-separated in $\|\cdot\|_\infty$. Indeed, if $v_1, v_2 \in V_e$ with $\|v_1-v_2\|_\infty < \varepsilon$, we may take $(u_1,v_1),(u_2,v_2)\in P_e$; then $\|u_i-e\|_\infty \le \varepsilon/2$, so $\|u_1-u_2\|_\infty < \varepsilon$, hence $d_\infty\big((u_1,v_1),(u_2,v_2)\big) < \varepsilon$, contradicting the ε-separation of P.

Using the fact $f(u)\le f(e)$ for $(u,v)\in P_e$ and the separation of V_e,

$$\sum_{(u,v)\in P_e} f(u)\,g(v) \ \le\ f(e)\sum_{v\in V_e} g(v) \ \le\ f(e)\ Z_m^{\mathrm{sep}}(\varepsilon).$$

Summing over $e\in E$ and recalling that E is ε-separated,

$$\sum_{(u,v)\in P} f(u)\,g(v) \ \le\ Z_m^{\mathrm{sep}}(\varepsilon)\sum_{e\in E} f(e) \ \le\ Z_m^{\mathrm{sep}}(\varepsilon)\ Z_n^{\mathrm{sep}}(\varepsilon).$$

Taking the supremum over all ε-separated $F\subset \pi_{n+m}(\mathbb{S})$ yields

$$Z_{n+m}^{\mathrm{sep}}(\varepsilon) \le Z_n^{\mathrm{sep}}(\varepsilon) Z_m^{\mathrm{sep}}(\varepsilon).$$

The proof is completed. □

The proof follows the argument of Gutman and Śpiewak [8], based on their Lemma 12.1.1, which we recall in the form below. For $m\le n$ integers, write $\pi_m^n : Y \to X^{n-m+1}$ for the window projection $\pi_m^n\big((x_i)_{i\in\mathbb{Z}}\big) = (x_m,\dots,x_n)$. If $m=1$, we denote $\pi_1^n = \pi_n$ for simplicity.

Lemma 4.9. *Let $\mathbb{S}\subset Y$ be a subshift. Fix $\varepsilon>0$ and $m\in\mathbb{N}$ with $2^{-m+2}<\varepsilon$. Then for any $(8\varepsilon,n)$-separated set A in $(\mathbb{S},D)$, $\pi_{-(m-1)}^{n+m}(A)$ is a ε-separated set of $\pi_{-(m-1)}^{n+m}(\mathbb{S})$ with $|A| = |\pi_{-(m-1)}^{n+m}(A)|$.*

Now we are able to prove Proposition 4.6.

Proof of Proposition 4.6. By the definitions, one has for any finite $A\in\mathbb{S}$,

$$\sum_{x\in A} e^{n|\log\varepsilon|\hat{\mathcal{E}}(\Delta_x^n)} \ge \sum_{(x_1,\dots,x_n)\in\pi_n(A)} e^{n|\log\varepsilon|\mathcal{E}(\Delta(x_1,\dots,x_n))}. \tag{21}$$

If in addition, assume that π_n is injective, then

$$\sum_{x\in A} e^{n|\log\varepsilon|\hat{\mathcal{E}}(\Delta_x^n)} = \sum_{(x_1,\dots,x_n)\in\pi_n(A)} e^{n|\log\varepsilon|\mathcal{E}(\Delta(x_1,\dots,x_n))}. \tag{22}$$

On the one hand, notice that for $x,y\in Y$, the inequality $D_n(x,y)<\varepsilon$ implies $\|\pi_n(x)-\pi_n(y)\|_\infty < 2\varepsilon$, and hence for each (n,ε)-spanning set A of $\mathbb{S}$, $\pi_n(A)$ is a 2ε-spanning set. Thus, (21) implies that

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S},\sigma,D) \ge \limsup_{\varepsilon\to 0}\frac{\limsup_{n\to\infty}\frac{1}{n}\log Z_n^{\mathrm{span}}(\varepsilon)}{|\log\varepsilon|}.$$

By (20), one has

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S},\sigma,D) \ge \limsup_{\varepsilon\to 0}\frac{\limsup_{n\to\infty}\frac{1}{n}\log Z_n^{\mathrm{sep}}(\varepsilon)}{|\log\varepsilon|}.$$

On the other hand, let $E \subset \mathbb{S}$ be an $(8\varepsilon, n)$–separated set. By Lemma 4.9, the image $\pi^{n+m}_{-(m-1)}(E)$ is an ε–separated subset of $\pi^{n+m}_{-(m-1)}(\mathbb{S})$ and $|E| = |\pi^{n+m}_{-(m-1)}(E)|$. Therefore, using (22), we obtain

$$\sum_{x\in E} \exp\big(n|\log\varepsilon|\,\hat{\mathcal{E}}(\Delta^n_x)\big) = \sum_{(x_{-(m-1)},\ldots,x_{n+m})\in\pi^{n+m}_{-(m-1)}(E)} \exp\big(n|\log\varepsilon|\,\mathcal{E}(\Delta(x_1,\ldots,x_n))\big).$$

Combining this with the submultiplicativity in Proposition 4.8 yields

$$\sum_{x\in E} \exp\big(n|\log\varepsilon|\,\hat{\mathcal{E}}(\Delta^n_x)\big) \;\le\; \mathrm{Sep}\left(\varepsilon, \pi^{0}_{-(m-1)}(\mathbb{S})\ \right)\cdot \mathrm{Sep}\left(\varepsilon, \pi^{n+m}_{n+1}(\mathbb{S})\right)\cdot Z^{\mathrm{sep}}_n(\varepsilon),$$

where $\mathrm{Sep}(\varepsilon, A)$ denotes the maximal cardinality of an ε–separated subset of A with respect to the norm $\|\cdot\|_\infty$.

Taking the supremum over all $(8\varepsilon, n)$–separated E, then letting $n \to \infty$ followed by $\varepsilon \to 0$, we conclude that

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S},\sigma,D) \;\le\; \limsup_{\varepsilon\to 0}\frac{1}{|\log\varepsilon|}\limsup_{n\to\infty}\cdot\frac{1}{n}\log Z^{\mathrm{sep}}_n(\varepsilon).$$

This completes the proof. ☐

We now prove that (14) also holds for subadditive energies.

Proof of Theorem E. Fix $\eta > 0$. Take $N \in \mathbb{N}$ large enough such that

$$\frac{1}{N}\limsup_{\varepsilon\to 0}\frac{1}{|\log\varepsilon|}\log\inf_E\left(\sum_{(x_1,\ldots,x_N)\in E} e^{N|\log\varepsilon|\mathcal{E}(\Delta(x_1,\ldots,x_N))}\right) \le \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) + \eta.$$

Choose $\varepsilon_0 > 0$ small enough such that

$$\frac{1}{N}\log Z^{\mathrm{sep}}_N(\varepsilon) \le |\log\varepsilon|(\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) + \eta) \qquad \text{for } 0<\varepsilon<\varepsilon_0.$$

For any $\varepsilon < \varepsilon_0$, by Proposition 4.8, one has

$$\lim_{n\to\infty}\frac{1}{n}\log Z^{\mathrm{sep}}_n(\varepsilon) \le \frac{1}{N}\log Z^{\mathrm{sep}}_N(\varepsilon) \le |\log\varepsilon|(\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) + \eta)$$

By Proposition 4.6, it holds that

$$\Pi^{\hat{\mathcal{E}}}\overline{\mathrm{mdim}}_M(\mathbb{S},\sigma,D) \le \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_B(\mathbb{S}) + \eta.$$

The proof is completed from the arbitrariness of $\eta > 0$. ☐

5. **Nonlinear equilibrium states.**

Remembering that $\mu(\varphi) = \int \varphi\, d\mu$, we consider the *rotation set* associated the potential φ as

$$\mathrm{rot}(\varphi) := \Big\{\mu(\varphi) : \mu \in \mathcal{M}_f(X)\Big\}.$$

It is straightforward that $\mathrm{rot}(\varphi)$ is a compact and convex subset of $\mathbb{R}$. Now, we assume that $\mathcal{E}(\mu) = F(\mu(\varphi))$, where F is convex and defined in an open set $U \supset \mathrm{rot}(\varphi)$ of $\mathbb{R}$.

We denote the *nonlinear equilibrium measure set* as follows

$$\mathscr{EM} := \big\{\mu \in \mathcal{M}_f(X) : H^K_\delta(f,\mu) + F(\mu(\varphi)) \text{ is maximal}\big\},$$

and we define

$$\mathcal{R}(\varphi) := \Big\{\mu(\varphi) : \mu \in \mathscr{EM}\Big\} \subset \mathrm{rot}(\varphi).$$

Due to Theorem A, it follows that $\mathscr{EM}$ and $\mathcal{R}(\varphi)$ are both nonempty. Using the upper semicontinuity of map $\mu \mapsto H_{\delta}^{K}(\mu)$, holds that $\mathcal{R}(\varphi)$ is a compact set. We also consider the level set

$$\mathcal{M}(z) := \{\mu \in \mathcal{M}_f(X) : \mu(\varphi) = z\}.$$

Definition 5.1. Given a continuous dynamical system $f : X \to X$ with potential φ, the *finite-dimensional metric mean dimension function* $\mathbf{H}_{\delta}^{K} : \mathbb{R} \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$\mathbf{H}_{\delta}^{K}(z) := \sup_{\mu \in \mathcal{M}(z)} \mathrm{H}_{\delta}^{K}(f, \mu).$$

for any $\delta \in [0, 1]$.

Remark 5.2. Given any $z \in \mathrm{rot}(\varphi)$, we have that there exists $\mu \in \mathcal{M}_f(X)$ such that $\mu(\varphi) = z$. Thus $\mathcal{M}(z) \neq \emptyset$. Moreover, $\mathcal{M}(z)$ is compact and there exists $\mu \in \mathcal{M}(z)$ maximizing H_{δ}^{K} by upper semicontinuity. Therefore, the function $\mathbf{H}_{\delta}^{K}$ is well defined.

Remark 5.3. We notice that for a fixed $\alpha \in \mathbb{R}$, if $f : X \to X$ is continuous and satisfies the specification property then, by [1],

$$\mathbf{H}_{\delta}^{K}(\alpha) = \overline{\mathrm{mdim}}_{\mathrm{M}}\Big(K_{\alpha}, f, d\Big),$$

where $K_{\alpha} = \Big\{x \in X : \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \varphi(f^{i}(x)) = \alpha\Big\}$, for any continuous potential $\varphi : X \to \mathbb{R}$. This shows that the function $\mathbf{H}_{\delta}^{K}(\alpha)$ is related with the multifractal formalism induced by the metric mean dimension.

Consider the function

$$P(z) := \mathbf{H}_{\delta}^{K}(z) + F(z).$$

From Remark 5.2, one has

$$P(z) \leq \sup_{\mu \in \mathcal{M}_f(X)} \Bigg\{ H_{\delta}^{K}(f, \mu) + F(\mu(\varphi)) \Bigg\}. \tag{23}$$

By the following proposition, it is possible to obtain equilibrium measures using the function $P(z) = \mathbf{H}_{\delta}^{K}(z) + F(z)$. Thus, combining this proposition with the Theorem A, we obtain a tool to compute the nonlinear metric mean dimension $\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}(X, f, d)$ as we illustrate in next section.

Proposition 5.4. *$z \in \mathcal{R}(\varphi)$ if only if $z \in rot(\varphi)$ maximizes the function $P(z)$.*

In this case,

$$P(z) = \sup_{\mu \in \mathcal{M}_f(X)} \Bigg\{ H_{\delta}^{K}(f, \mu) + F(\mu(\varphi)) \Bigg\}.$$

Proof. Assume that $z \in \mathrm{rot}(\varphi)$. Then, there exists a measure $\mu \in \mathscr{EM}$ such that $\mu(\varphi) = z$. Then, as μ is an equilibrium measure, we obtain that

$$P(z) = \mathbf{H}_{\delta}^{K}(z) + F(z) \geq H_{\delta}^{K}(f, \mu) + F(\mu(\varphi)) = \sup_{\mu \in \mathcal{M}_f(X)} \Bigg\{ H_{\delta}^{K}(f, \mu) + F(\mu(\varphi)) \Bigg\}.$$

Thus, combining this with (23), one has z maximize the function $P(z)$.

On the other hand, let $z \in \mathrm{rot}(\varphi)$ that maximize $P(z)$. By Remark 5.2, there exists some $\mu \in \mathcal{M}(z)$ such that $\mathbf{H}_\delta^K(z) = H_\delta^K(f,\mu)$. Then

$$\mathbf{H}_\delta^K(z) + F(z) - H_\delta^K(f,\mu) + F(\mu(\varphi)) = \sup_{\mu \in \mathcal{M}_f(X)} \left\{ H_\delta^K(f,\mu) + F(\mu(\varphi)) \right\}.$$

Therefore, $\mu \in \mathscr{EM}$ and consequently, we conclude that $\mu \in \mathcal{R}(\varphi)$. $\square$

6. Examples.

6.1. **Shifts on compact homogeneous alphabets.** Let (X,d) be a compact metric space and consider $Y = X^{\mathbb{N}}$ and the distance on Y given by

$$D_\rho((u_n)_{n\in\mathbb{N}}, (v_n)_{n\in\mathbb{N}}) := \sup_{n\in\mathbb{N}} \frac{d(u_n,v_n)}{\rho^{n-1}},$$

for some $\rho > 1$ fixed. In this case we have that if X is so that for any open subset $U \subset X$, $\dim_{\mathrm{B}}(U) = \dim_{\mathrm{B}}(X)$, then, by [5, Theorem 8.1], the metric mean dimension of any shift-invariant probability measure is given by

$$H_\delta^K(f,\mu) = \dim_{\mathrm{B}}(X), \text{ for every } \delta \in (0,1) \text{ and } \mu \in \mathcal{M}_\sigma(Y).$$

Consequently by Theorem D, the variational principle in this setting is reduced to the analysis of the energy map $\mathcal{E}$. Moreover, we have that

$$\sup_{\mu \in \mathcal{M}_\sigma(X)} \mathcal{E}(\mu) = \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}\left(X^{\mathbb{N}}, \sigma, D_\rho\right) - \dim_{\mathrm{B}}(X).$$

6.2. **Shifts on non-homogeneous compact alphabets.** Let $X = \{0\} \cup \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$. If we consider the space $Y = X^{\mathbb{N}}$ we have that X does not satisfy the condition presented in the previous example. Moreover, by [5] it is possible to conclude that

$$H_\delta^K(f, \delta_{\{0^{\mathbb{N}}\}}) = \dim_{\mathrm{B}}(X) = \frac{1}{2},$$

but for every $x \neq 0$

$$H_\delta^K(f, \delta_{\{x^{\mathbb{N}}\}}) = 0.$$

6.3. **Phase transitions.** Let $g : [0,1] \to [0,1]$ be given by $g(x) = |1 - |3x-1||$ and the sequence $(a_n)_{n\in\mathbb{N}_0}$ on interval $[0,1)$ where $a_0 = 0$ and $a_n = \sum_{k=1}^{n} \frac{6}{\pi^2 k^2}$. For each $n \in \mathbb{N}$, we take the interval $J_n = [a_{n-1}, a_n]$ and let $T_n : J_n \to [0,1]$ be the unique increasing affine map from J_n into $[0,1]$. We define $f : [0,1] \to [0,1]$ as $f(x) = T_n^{-1} \circ g \circ T_n(x)$ if $x \in J_n$, and $T(1) = 1$. It was proved in [5, Example 10.3] that

$$H_\delta^K(f,\mu) = \mu(\{1\}),$$

for all $\mu \in \mathcal{M}_f([0,1])$ and H_δ^K is maximized by Dirac measure $\delta_{\{1\}}$. Now, we take the energy $\mathcal{E}(\mu) = F(\mu(\varphi))$, with $F(z) := -\frac{1}{2}\beta z^2$, for every $\beta \geq 0$, and the potential $\varphi : [0,1] \to \mathbb{R}$ given by $\varphi(x) = x$. We want to maximize the function $P(z) = \mathbf{H}_\delta^K(z) + F(z)$.

For any $z \in [0,1]$. We consider invariant measures $\mu \in \mathcal{M}(z)$, i.e., such that $\int \varphi d\mu = z$ and maximize $H_\delta^K(f,\mu)$ with respect this measures. Fix any $z \in (0,1)$. Notice that, the measure $\mu = z\delta_1 + (1-z)\delta_0$ belongs to $\mathcal{M}(z)$. Then

$$\mathbf{H}_\delta^K(z) \geq H_\delta^K(f,\mu) = \mu(\{1\}) = z.$$

On the other hand, any ergodic measure μ assigns full measure to some interval $[a_{n-1}, a_n)$, i.e. $\mu([a_{n-1}, a_n)) = 1$. Consequently, $H_\delta^K(\mu) = \mu(\{1\}) = 0$, for all

$\mu \in \mathcal{M}_f^{erg}(X)$. Indeed, we can write any $\nu \in \mathcal{M}_f([0,1]) \setminus \{\delta_1\}$ non ergodic as $v = z\delta_1 + (1-z)\mu$, with $\mu(\{1\}) = 0$. Consequently, any invariant measure $v = z\delta_1 + (1-z)\mu$, with $\mu(\{1\}) = 0$, maximizes H_δ^K on $\mathcal{M}(z)$ and so, $\mathbf{H}_\delta^K(z) = z$. It is not difficult to see that $\mathbf{H}_\delta^K(0) = 0$ and $\mathbf{H}_\delta^K(1) = 1$.

Now, given $\beta \geq 0$, we are left with maximizing the function

$$\phi_\beta(z) := \mathbf{H}_\delta^K(z) + F(z) = z - \frac{1}{2}\beta z^2.$$

for $z \in [0,1]$. Consider the cases:
(i) if $0 \leq \beta \leq 1$, the maximum is attained in $z = 1$. Thus, the unique equilibrium measure is $\mu = \delta_1$ and $\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X, f, d) = 2$;
(ii) for each $\beta > 1$, a simply computation gives that $z = \frac{1}{\beta}$ is the unique critical point of ϕ_β and it is a maximum. Therefore, the unique equilibrium measure is

$$\nu = \frac{1}{\beta}\delta_1 + \left(1 - \frac{1}{\beta}\right)\mu,$$

with $\mu(\{1\}) = 0$ and we can conclude

$$\Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_{\mathrm{M}}(X, f, d) = \frac{1}{2\beta}.$$

6.4. **Finite nonlinear topological pressure does not imply zero nonlinear metric mean dimension.** In the classical (linear) setting, if a topological dynamical system (X, f) has finite topological entropy, then its metric mean dimension is zero. The following example shows that this implication may fail in the nonlinear framework.

Let $X = \{p, q\}$ and let f fix both points. Define $\phi \in C(X)$ by $\phi(p) = 1$ and $\phi(q) = -1$, and set the energy

$$\mathcal{E}(\mu) := F\big(\mu(\phi)\big) \quad \text{with} \quad F(z) = -z^2.$$

Then $\mathcal{M}_f(X) = \{t\delta_p + (1-t)\delta_q : t \in [0,1]\}$ and $\mathcal{M}_f^{\mathrm{erg}}(X) = \{\delta_p, \delta_q\}$. We have:

- $h_\nu(f) = 0$ for every $\nu \in \mathcal{M}_f(X)$, hence $H_\delta^K(f, \nu) = 0$ for every ν;
- $\mathcal{E}(\delta_p) = \mathcal{E}(\delta_q) = -1$ and, more generally, $\mathcal{E}\big(t\delta_p + (1-t)\delta_q\big) = -(2t-1)^2 \leq 0$ with $\sup_\nu \mathcal{E}(\nu) = 0$.

Since the only orbits are the fixed points p and q, all empirical measures along orbits equal either δ_p or δ_q, so

$$\Pi_{\mathrm{top}}^{\mathcal{E}}(X, f) = \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_M(X, f, d) = -1.$$

Consequently, a finite nonlinear topological pressure does **NOT** imply that the nonlinear metric mean dimension is zero (indeed, it can be negative in this example).

Meanwhile, one has

$$\sup_{\nu \in \mathcal{M}_f(X)} \big(H_\delta^K(f, \nu) + \mathcal{E}(\nu)\big) = 0 > \Pi^{\mathcal{E}}\overline{\mathrm{mdim}}_M(X, f, d).$$

This shows that the convexity of energies is necessary.

Let us end this paper with a question.

Question 1. Is there any condition on the phase space or on the potential (or on both of them) under which finite nonlinear pressure implies zero nonlinear metric mean dimension?

Acknowledgments. The authors gratefully acknowledges valuable suggestions from Gustavo Pessil and Changlin Wang. We thank the referee for this helpful comment, which prompted us to clarify this point.

Funding. A. J. Becker was supported by Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES (Brazil) - PRAPD - DRI under Grant Number 88887.005391/2024-00. C. Liu was supported by the Postdoctoral Fellowship Program and China Postdoctoral Science Foundation under Grant Number BX20250067, and the China Postdoctoral Science Foundation under Grant Number 2025M773074.

REFERENCES

[1] L. Backes and F. B. Rodrigues, *A variational principle for the metric mean dimension of level sets*, IEEE Transactions on Information Theory, **69** (2023), 5485–5496.

[2] Barreira, L., Holanda, C. *Higher-Dimensional Nonlinear Thermodynamic Formalism.* J Stat Phys 187, 18 (2022).

[3] Biś, A., Carvalho, M., Mendes, M. et al. *A Convex Analysis Approach to Entropy Functions, Variational Principles and Equilibrium States.* Commun. Math. Phys. 394, 215–256 (2022).

[4] J. Buzzi, B. Kloeckner and R. Leplaideur, *Nonlinear thermodynamical formalism.* Annales Henri Lebesgue, Volume 6 (2023), pp. 1429-1477.

[5] M. Carvalho, G. Pessil, P. Varandas, *A convex analysis approach to the metric mean dimension: Limits of scaled pressures and variational principles*, Advances in Mathematics **436** (2024), 109407.

[6] D. Cheng, Z. Liand and B. Selmi. *Upper metric mean dimensions with potential on subsets.* Nonlinearity **34** (2021), 852–867.

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

[8] Y. Gutman, A. Śpiewak, *Metric mean dimension and analog compression*, IEEE Transactions on Information Theory, 66 (2020), 6977–6998.

[9] Y. Gutman, A. Śpiewak, *Around the variational principle for metric mean dimension*, Studia Mathematica 261 (2021), 345-360.

[10] Tamara Kucherenko. *Nonlinear thermodynamic formalism through the lens of rotation theory.* Discrete and Continuous Dynamical Systems, 2024, 44(12): 3760-3773.

[11] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Etudes Sci. Publ. Math. **51** (1980), 137–173.

[12] M. Brin and A. B. Katok, On local entropy, in *Geometric dynamics (Rio de Janeiro, 1981)*, 30–38, Lecture Notes in Math., 1007, Springer, Berlin, ; MR0730261

[13] R. Leplaideur and F. Watbled. *Generalized Curie-Weiss model and quadratic pressure in Ergodic Theory.* Bull. Soc. Math. Fr. 147 (2), 2019, p. 197–219. 1, 1.3, 1.3, 1.4.2, 5.1.

[14] R. Leplaideur and F. Watbled. *Curie–Weiss type models for general spin spaces and quadratic pressure in ergodic theory.* J. Stat. Phys. 181 (1), 2020, 263–292. 1.4.4, 5.3.

[15] E. Lindenstrauss, *Mean dimension, small entropy factors and an embedding theorem*, Inst. Hautes Études Sci. Publ. Math., (1999), pp. 227–262.

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

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

[18] C. Liu and F. B. Rodrigues, *Metric mean dimension via preimage structures*, J. Stat. Phys., **191** (2024), no. 2, 31.

[19] Y. Liu, B. Selmi and Z. Li, *Some properties of a new mean measure*, J. Math. Anal. Appl., **555** (2026), 130051.

[20] Y. Liu, B. Selmi and Z. Li, *On the mean fractal dimensions of the Cartesian product sets*, Chaos, Solitons & Fractals, **180** (2024), 114503.

[21] R. Shi, *On variational principles for metric mean dimension*, IEEE Transactions on Information Theory. **68** (2022), 4282–4288.

[22] T. Meyerovitch and M. Tsukamoto, Expansive multiparameter actions and mean dimension, Trans. Amer. Math. Soc. **371** (2019), no. 10, 7275–7299; MR3939578

[23] M. Tsukamoto. *Double variational principle for mean dimension with potential*, Advances in Mathematics **361** (2020), 106935, 53 pp.

[24] R. Yang, E. Chen and X. Zhou. *Measure-theoretic metric mean dimension*, Studia Math. **280** (2025), no. 1, 1–25.

[25] Tsukamoto, Masaki. *Mean dimension of the dynamical system of Brody curves*. Invent. Math. **211** (2018), no. 3, 935–968.

[26] E. Lindenstrauss and M. Tsukamoto, Double variational principle for mean dimension, Geom. Funct. Anal. **29** (2019), no. 4, 1048–1109; MR3990194

[27] W. Wu, *On relative metric mean dimension with potential and variational principles*, J. Dynam. Differential Equations 35, **3** (2023), 2313–2335.

[28] K. Yano, *A remark on the topological entropy of homeomorphisms*, Invent. Math. **59** (1980), 215–220.

Received xxxx 20xx; revised xxxx 20xx; early access xxxx 20xx.