Ergod. Th. & Dynam. Sys., page 1 of 37 © The Author(s), 2024. Published by Cambridge University Press.
doi:10.1017/etds.2024.35

Weighted topological pressure revisited

NIMA ALIBABAEI

Department of Mathematics, Kyoto University, Kyoto 606-8501, Japan
(e-mail: alibabaei.nima.28c@st.kyoto-u.ac.jp)

(*Received* 3 *September* 2023 *and accepted in revised form* 8 *April* 2024)

Abstract. Feng and Huang [Variational principle for weighted topological pressure. *J. Math. Pures Appl. (9)* **106** (2016), 411–452] introduced weighted topological entropy and pressure for factor maps between dynamical systems and established its variational principle. Tsukamoto [New approach to weighted topological entropy and pressure. *Ergod. Th. & Dynam. Sys.* **43** (2023), 1004–1034] redefined those invariants quite differently for the simplest case and showed via the variational principle that the two definitions coincide. We generalize Tsukamoto's approach, redefine the weighted topological entropy and pressure for higher dimensions, and prove the variational principle. Our result allows for an elementary calculation of the Hausdorff dimension of affine-invariant sets such as self-affine sponges and certain sofic sets that reside in Euclidean space of arbitrary dimension.

Key words: dynamical systems, weighted topological entropy, weighted topological pressure, variational principle, affine-invariant sets, self-affine sponges, sofic sets, Hausdorff dimension
2020 Mathematics Subject Classification: 28A80, 28D20, 37D35 (Primary); 37B40, 37A35, 37C45 (Secondary)

1. *Introduction*

1.1. *Dynamical systems and entropy.* Topological pressure and its variational principle have been significant in several fields, including the dimension theory of dynamical systems. Recently, Feng and Huang devised an innovative invariant called weighted topological pressure for factor maps between dynamical systems and proved its variational principle [**FH16**]. Their work inspired Tsukamoto to suggest a new definition of this invariant [**Tsu22**]. He also established a variational principle, revealing the non-trivial coincidence of the two definitions. Tsukamoto focused on the simplest case with two dynamical systems.

In this paper, we extend Tsukamoto's definition to the case of an arbitrary number of dynamical systems and prove its variational principle. With our result, we can plainly calculate the Hausdorff dimension of self-affine sponges, a topic studied by Kenyon and Peres [**KP96**]. Furthermore, we will show in §6 that we can determine the Hausdorff dimension of certain sofic sets embedded in higher-dimensional Euclidean space.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

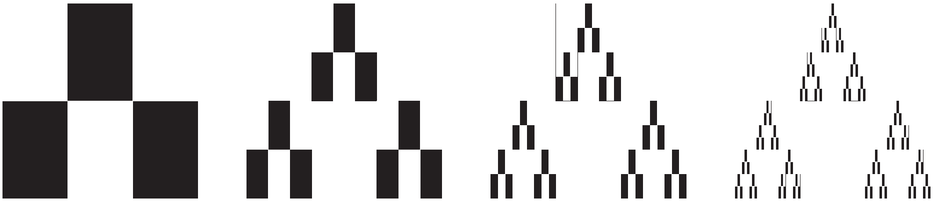


FIGURE 1. First four generations of a Bedford–McMullen carpet.

We review the basic notions of dynamical systems in this subsection. Refer to the book of Walters [Wal82] for the details.

A pair (X, T) is called a *dynamical system* if X is a compact metrizable space and $T : X \to X$ is a continuous map. A map $\pi : X \to Y$ between dynamical systems (X, T) and (Y, S) is said to be a *factor map* if π is a continuous surjection and $\pi \circ T = S \circ \pi$. We sometimes write as $\pi : (X, T) \to (Y, S)$ to clarify the dynamical systems in question.

For a dynamical system (X, T), denote its *topological entropy* by $h_{\mathrm{top}}(T)$. Let $P(f)$ be the *topological pressure* for a continuous function $f : X \to \mathbb{R}$ (see §2 for the definition of these quantities). Let $\mathscr{M}^T(X)$ be the set of T-invariant probability measures on X and $h_\mu(T)$ the *measure-theoretic entropy* for $\mu \in \mathscr{M}^T(X)$ (see §3.2). The variational principle then states that [Din70, Gm71, Gw69, Ru73, Wal75]

$$P(f) = \sup_{\mu \in \mathscr{M}^T(X)} \left(h_\mu(T) + \int_X f \, d\mu \right).$$

1.2. *Background.* We first look at *self-affine sponges* to understand the background of weighted topological entropy introduced by Feng and Huang. Let $m_1, m_2, \ldots, m_r$ be natural numbers with $m_1 \le m_2 \le \cdots \le m_r$. Consider an endomorphism T on $\mathbb{T}^r = \mathbb{R}^r/\mathbb{Z}^r$ represented by the diagonal matrix $A = \mathrm{diag}(m_1, m_2, \ldots, m_r)$. For $D \subset \prod_{i=1}^r \{0, 1, \ldots, m_i - 1\}$, define

$$K(T, D) = \left\{ \sum_{n=0}^{\infty} A^{-n} e_n \in \mathbb{T}^r \,\middle|\, e_n \in D \right\}.$$

This set is compact and T-invariant, that is, $TK(T, D) = K(T, D)$.

These sets for $r = 2$ are known as *Bedford–McMullen carpets* or *self-affine carpets*. Figure 1 exhibits a famous example, the case of $D = \{(0, 0), (1, 1), (0, 2)\} \subset \{0, 1\} \times \{0, 1, 2\}$. The analysis of these sets is complicated compared with ‘self-similar’ sets. Bedford [Bed84] and McMullen [McM84] independently studied these sets and showed that, in general, their Hausdorff dimension is strictly smaller than their Minkowski dimension (also known as box-counting dimension). Figure 1 has Hausdorff dimension $\log_2(1 + 2^{\log_3 2}) = 1.349 \cdots$ and Minkowski dimension $1 + \log_3 \frac{3}{2} = 1.369 \cdots$.

The sets $K(T, D)$ for $r \ge 3$ are called *self-affine sponges*. Kenyon and Peres [KP96] calculated their Hausdorff dimension for the general case (see Theorem 1.5 in this section). In addition, they showed the following variational principle for the Hausdorff dimension

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

of $K(T, D)$:

$$\dim_H K(T, D) = \sup_{\mu \in \mathscr{M}^T(\mathbb{T}^r)} \left\{ \frac{1}{\log m_r} h_\mu(T) + \sum_{i=2}^{r} \left(\frac{1}{\log m_{r-i+1}} - \frac{1}{\log m_{r-i+2}} \right) h_{\mu_i}(T_i) \right\}. \tag{1.1}$$

Here, the endomorphism T_i on $\mathbb{T}^{r-i+1}$ is defined from $A_i = \mathrm{diag}(m_1, m_2, \ldots, m_{r-i+1})$, and μ_i is defined as the push-forward measure of μ on $\mathbb{T}^{r-i+1}$ by the projection onto the first $r - i + 1$ coordinates. Feng and Huang's definition of weighted topological entropy of $K(T, D)$ equals $\dim_H K(T, D)$ with a proper setting.

1.3. *Original definition of the weighted topological pressure.* Motivated by the geometry of self-affine sponges described in the previous subsection, Feng and Huang introduced a generalized notion of pressure. Consider dynamical systems (X_i, T_i) $(i = 1, 2, \ldots, r)$ and factor maps $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r-1)$:

$$(X_1, T_1) \xrightarrow{\pi_1} (X_2, T_2) \xrightarrow{\pi_2} \cdots \xrightarrow{\pi_{r-1}} (X_r, T_r).$$

We refer to this as a *sequence of dynamical systems*. Let $\boldsymbol{w} = (w_1, w_2, \ldots, w_r)$ be a vector with $w_1 > 0$ and $w_i \geq 0$ for $i \geq 2$. Feng and Huang [**FH16**] ingeniously defined the $\boldsymbol{w}$-weighted topological pressure $P^{\boldsymbol{w}}_{\mathrm{FH}}(f)$ for a continuous function $f : X_1 \to \mathbb{R}$ and established the variational principle [**FH16**, Theorem 1.4]:

$$P^{\boldsymbol{w}}_{\mathrm{FH}}(f) = \sup_{\mu \in \mathscr{M}^{T_1}(X_1)} \left(\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i) + w_1 \int_{X_1} f \, d\mu \right). \tag{1.2}$$

Here, $\pi^{(i)}$ is defined by

$$\pi^{(0)} = \mathrm{id}_{X_1} : X_1 \to X_1,$$
$$\pi^{(i)} = \pi_i \circ \pi_{i-1} \circ \cdots \circ \pi_1 : X_1 \to X_{i+1},$$

and $\pi^{(i-1)}{}_*\mu$ is the push-forward measure of μ by $\pi^{(i-1)}$ on X_i. The $\boldsymbol{w}$-weighted topological entropy $h^{\boldsymbol{w}}_{\mathrm{top}}(T_1)$ is the value of $P^{\boldsymbol{w}}_{\mathrm{FH}}(f)$ when $f \equiv 0$. In this case, equation (1.2) becomes

$$h^{\boldsymbol{w}}_{\mathrm{top}}(T_1) = \sup_{\mu \in \mathscr{M}^{T_1}(X_1)} \left(\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i) \right). \tag{1.3}$$

We will explain here Feng and Huang's method of defining $h^{\boldsymbol{w}}_{\mathrm{top}}(T_1)$. For the definition of $P^{\boldsymbol{w}}_{\mathrm{FH}}(f)$, see their original paper [**FH16**].

Let n be a natural number and ε a positive number. Let $d^{(i)}$ be a metric on X_i. For $x \in X_1$, define the *nth $\boldsymbol{w}$-weighted Bowen ball of radius ε centered at x* by

$$B^{\boldsymbol{w}}_n(x, \varepsilon) = \left\{ y \in X_1 \,\middle|\, \begin{array}{l} d^{(i)}(T_i^j(\pi^{(i-1)}(x)), T_i^j(\pi^{(i-1)}(y))) < \varepsilon \text{ for every} \\ 0 \leq j \leq \lceil (w_1 + \cdots + w_i)n \rceil \text{ and } 1 \leq i \leq k. \end{array} \right\}.$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Consider $\Gamma = \{B^{\boldsymbol{w}}_{n_j}(x_j, \varepsilon)\}_j$, an at-most countable cover of X_1 by weighted Bowen balls. Let $n(\Gamma) = \min_j n_j$. For $s \geq 0$ and $N \in \mathbb{N}$, let

$$\Lambda^{\boldsymbol{w},s}_{N,\varepsilon} = \inf \left\{ \sum_j e^{-sn_j} \,\middle|\, \Gamma = \{B^{\boldsymbol{w}}_{n_j}(x_j, \varepsilon)\}_j \text{ covers } X_1 \text{ and } n(\Gamma) \geq N \right\}.$$

This quantity is non-decreasing as $N \to \infty$. The following limit hence exists:

$$\Lambda^{\boldsymbol{w},s}_{\varepsilon} = \lim_{N\to\infty} \Lambda^{\boldsymbol{w},s}_{N,\varepsilon}.$$

There is a value of s where $\Lambda^{\boldsymbol{w},s}_{\varepsilon}$ jumps from ∞ to 0, which we will denote by $h^{\boldsymbol{w}}_{\text{top}}(T_1, \varepsilon)$:

$$\Lambda^{\boldsymbol{w},s}_{\varepsilon} = \begin{cases} \infty & (s < h^{\boldsymbol{w}}_{\text{top}}(T_1, \varepsilon)), \\ 0 & (s > h^{\boldsymbol{w}}_{\text{top}}(T_1, \varepsilon)). \end{cases}$$

The value $h^{\boldsymbol{w}}_{\text{top}}(T_1, \varepsilon)$ is non-decreasing as $\varepsilon \to 0$. Therefore, we can define the $\boldsymbol{w}$-weighted topological entropy $h^{\boldsymbol{w}}_{\text{top}}(T_1)$ by

$$h^{\boldsymbol{w}}_{\text{top}}(T_1) = \lim_{\varepsilon\to 0} h^{\boldsymbol{w}}_{\text{top}}(T_1, \varepsilon).$$

An important point about this definition is that in some dynamical systems, such as self-affine sponges, the quantity $h^{\boldsymbol{w}}_{\text{top}}(T_1)$ is directly related to the Hausdorff dimension of X_1.

Example 1.1. Consider the self-affine sponges introduced in §1.2. Define $p_i : \mathbb{T}^{r-i+1} \to \mathbb{T}^{r-i}$ by

$$p_i(x_1, x_2, \ldots, x_{r-i}, x_{r-i+1}) = (x_1, x_2, \ldots, x_{r-i}).$$

Let $X_1 = K(T, D)$, $X_i = p_{i-1} \circ p_i \circ \cdots \circ p_1(X_1)$, and $T_i : X_i \to X_i$ be the endomorphism defined by $A_i = \text{diag}(m_1, m_2, \ldots, m_{r-i+1})$. Define the factor maps $\pi_i : X_i \to X_{i+1}$ as the restrictions of p_i. Let

$$\boldsymbol{w} = \left(\frac{\log m_1}{\log m_r}, \quad \frac{\log m_1}{\log m_{r-1}} - \frac{\log m_1}{\log m_r}, \ldots, \quad \frac{\log m_1}{\log m_2} - \frac{\log m_1}{\log m_3}, \quad 1 - \frac{\log m_1}{\log m_2} \right). \tag{1.4}$$

Then each nth $\boldsymbol{w}$-weighted Bowen ball is approximately a square of side length εm_1^{-n}. Therefore,

$$\dim_H K(T, D) = \frac{h^{\boldsymbol{w}}_{\text{top}}(T_1)}{\log m_1}. \tag{1.5}$$

1.4. *Tsukamoto's approach and its extension.* Following the work of Feng and Huang [**FH16**] described in §1.3, Tsukamoto [**Tsu22**] published an intriguing approach to these invariants. There, he gave a new definition of the weighted topological pressure for a factor map between two dynamical systems:

$$(X_1, T_1) \xrightarrow{\pi} (X_2, T_2).$$

He then proved the variational principle using his definition, showing the surprising coincidence of the two definitions. His definition of weighted topological entropy allowed for relatively easy calculations for sets like self-affine carpets.

We will extend Tsukamoto's idea, redefine the weighted topological pressure for a sequence of dynamical systems of arbitrary length, and establish the variational principle. Here we will explain our definition in the case $f \equiv 0$. See §2 for the general setting. We will not explain Tsukamoto's definition itself since it is obtained by letting $r = 2$ in the following argument.

Consider a sequence of dynamical systems:

$$(X_1, T_1) \xrightarrow{\pi_1} (X_2, T_2) \xrightarrow{\pi_2} \cdots \xrightarrow{\pi_{r-1}} (X_r, T_r).$$

Take a metric $d^{(i)}$ on X_i. Let $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$ with $0 \le a_i \le 1$ for each i. Let N be a natural number and ε a positive number. We define a new metric $d_N^{(i)}$ on X_i by

$$d_N^{(i)}(x_1, x_2) = \max_{0\le n<N} d^{(i)}(T_i{}^n x_1, T_i{}^n x_2).$$

We inductively define a quantity $\#_i^{\boldsymbol{a}}(\Omega, N, \varepsilon)$ for $\Omega \subset X_i$. For $\Omega \subset X_1$, set

$$\#_1^{\boldsymbol{a}}(\Omega, N, \varepsilon) = \min\left\{ n \in \mathbb{N} \,\middle|\, \begin{array}{l}\text{There exists an open cover } \{U_j\}_{j=1}^n \text{ of } \Omega \\ \text{with } \operatorname{diam}(U_j, d_N^{(1)}) < \varepsilon \text{ for all } 1 \le j \le n\end{array}\right\}.$$

(The quantity $\#_1^{\boldsymbol{a}}(\Omega, N, \varepsilon)$ is independent of the parameter $\boldsymbol{a}$. However, we use this notation for the convenience of what follows.) Let $\Omega \subset X_{i+1}$. Suppose $\#_i^{\boldsymbol{a}}$ is already defined. We set

$$\begin{aligned}&\#_{i+1}^{\boldsymbol{a}}(\Omega, N, \varepsilon)\\ &\quad = \min\left\{ \sum_{j=1}^{n} (\#_i^{\boldsymbol{a}}(\pi_i^{-1}(U_j), N, \varepsilon))^{a_i} \,\middle|\, \begin{array}{l} n \in \mathbb{N}, \{U_j\}_{j=1}^n \text{ is an open cover of } \Omega \\ \text{with } \operatorname{diam}(U_j, d_N^{(i+1)}) < \varepsilon \text{ for all } 1 \le j \le n\end{array}\right\}.\end{aligned}$$

We define the *topological entropy of* $\boldsymbol{a}$*-exponent* $h^{\boldsymbol{a}}(\boldsymbol{T})$, where $\boldsymbol{T} = (T_i)_i$, by

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \lim_{\varepsilon\to 0}\left(\lim_{N\to\infty} \frac{\log \#_r^{\boldsymbol{a}}(X_r, N, \varepsilon)}{N}\right).$$

This limit exists since $\log \#_r^{\boldsymbol{a}}(X_r, N, \varepsilon)$ is sub-additive in N and non-decreasing as ε tends to 0.

From $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$, we define a probability vector (that is, all entries are non-negative, and their sum is 1) $\boldsymbol{w_a} = (w_1, \ldots, w_r)$ by

$$\begin{cases} w_1 = a_1 a_2 a_3 \cdots a_{r-1},\\ w_2 = (1-a_1) a_2 a_3 \cdots a_{r-1},\\ w_3 = (1-a_2) a_3 \cdots a_{r-1},\\ \vdots \\ w_{r-1} = (1-a_{r-2}) a_{r-1},\\ w_r = 1 - a_{r-1}.\end{cases}$$

The following theorem is a direct consequence of our main result in Theorem 2.1.

THEOREM 1.2. *For* $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$ *with* $0 \le a_i \le 1$ *for each i,*

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \sup_{\mu\in\mathscr{M}^{T_1}(X_1)} \left(\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i)\right). \tag{1.6}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

The strategy of the proof is adopted from Tsukamoto's paper. However, there are some additional difficulties. Let $h^{\boldsymbol{a}}_{\mathrm{var}}(\boldsymbol{T})$ be the right-hand side of equation (1.6). We use the 'zero-dimensional trick' for proving $h^{\boldsymbol{a}}(\boldsymbol{T}) \leq h^{\boldsymbol{a}}_{\mathrm{var}}(\boldsymbol{T})$, meaning we reduce the proof to the case where all dynamical systems are zero-dimensional. Merely taking a zero-dimensional extension for each X_i does not work. Therefore, we realize this by taking step by step extensions of the whole sequence of dynamical systems (see §3.3). Then we show $h^{\boldsymbol{a}}(\boldsymbol{T}) \leq h^{\boldsymbol{a}}_{\mathrm{var}}(\boldsymbol{T})$ by using an appropriate measure, the definition of which is quite sophisticated (see σ_N in the proof of Theorem 4.1). In proving $h^{\boldsymbol{a}}(\boldsymbol{T}) \geq h^{\boldsymbol{a}}_{\mathrm{var}}(\boldsymbol{T})$, the zero-dimensional trick can not be used. The proof, therefore, requires a detailed estimation of these quantities for arbitrary covers, which is more complicated than the original argument in [**Tsu22**].

Theorem 1.2 and Feng and Huang's version of variational principle in equation (1.3) yield the following corollary.

COROLLARY 1.3. *For* $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$ *with* $0 < a_i \leq 1$ *for each i,*

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = h^{\boldsymbol{w_a}}_{\mathrm{top}}(T_1).$$

This corollary is rather profound, connecting the two seemingly different quantities. We can calculate the Hausdorff dimension of self-affine sponges using this result as in the following example. Additionally, we will show in §6 that we can now determine the Hausdorff dimension of certain sofic sets in higher-dimensional Euclidean space.

Example 1.4. Let us take another look at self-affine sponges. Kenyon and Peres [**KP96**, Theorem 1.2] calculated their Hausdorff dimension as follows. Recall the notation in §1.2 and that $m_1 \leq m_2 \leq \cdots \leq m_r$.

THEOREM 1.5. *Define a sequence of real numbers* $(Z_j)_j$ *as follows. Let* Z_r *be the indicator of D, namely,* $Z_r(i_1, \ldots, i_r) = 1$ *if* $(i_1, \ldots, i_r) \in D$ *and* 0 *otherwise. Define* Z_{r-1} *by*

$$Z_{r-1}(i_1, \ldots, i_{r-1}) = \sum_{i_r=0}^{m_r-1} Z_r(i_1, \ldots, i_{r-1}, i_r).$$

More generally, if Z_{j+1} *is already defined, let*

$$Z_j(i_1, \ldots, i_j) = \sum_{i_{j+1}=0}^{m_{j+1}-1} Z_{j+1}(i_1, \ldots, i_j, i_{j+1})^{\log m_{j+1}/\log m_{j+2}}.$$

Then

$$\dim_H K(T, D) = \frac{\log Z_0}{\log m_1}.$$

We can prove this result in a fairly elementary way by Corollary 1.3 without requiring measure theory on the surface. Set $a_i = \log_{m_{r-i+1}} m_{r-i}$ for each $1 \leq i \leq r-1$, then $\boldsymbol{w_a}$ equals $\boldsymbol{w}$ in equation (1.4). Combining equation (1.5) and Corollary 1.3, we have

$$\dim_H K(T, D) = \frac{h^{\boldsymbol{w_a}}_{\mathrm{top}}(T_1)}{\log m_1} = \frac{h^{\boldsymbol{a}}(\boldsymbol{T})}{\log m_1}.$$

Hence, we need to show the following claim.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

CLAIM 1.6. *We have*

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \log Z_0.$$

Proof. Observe first that taking the infimum over closed covers instead of open ones in the definition of $h^{\boldsymbol{a}}(\boldsymbol{T})$ does not change its value. Define a metric $d^{(i)}$ on each X_i by

$$d^{(i)}(x, y) = \min_{n \in \mathbb{Z}^{r-i+1}} |x - y - n| .$$

Let

$$D_j = \{(e_1, \ldots, e_j) | \text{ there are } e_{j+1}, \ldots, e_r \text{ with } (e_1, \ldots, e_r) \in D\}.$$

Define $p_i : D_{r-i+1} \to D_{r-i}$ by $p_i(e_1, \ldots, e_{r-i+1}) = (e_1, \ldots, e_{r-i})$. Fix $0 < \varepsilon < 1/m_r$ and take a natural number n with $m_1^{-n} < \varepsilon$. Fix a natural number N and let $\psi_i : D_{r-i+1}^{N+n} \to D_{r-i}^{N+n}$ be the product map of p_i, that is, $\psi_i(v_1, \ldots, v_{N+n}) = (p_i(v_1), \ldots, p_i(v_{N+n}))$.

For $x \in D_{r-i+1}^{N+n}$, define (recall that $A_i = \operatorname{diag}(m_1, m_2, \ldots, m_{r-i+1})$)

$$U_x^{(i)} = \left\{ \sum_{k=0}^{\infty} A_i^{-k} e_k \in X_i \,\middle|\, e_k \in D_{r-i+1} \text{ for each } k \text{ and } (e_1, \ldots, e_{N+n}) = x \right\}.$$

Then $\{U_x^{(i)}\}_{x \in D_{r-i+1}^{N+n}}$ is a closed cover of X_i with $\operatorname{diam}(U_x^{(i)}, d_N^{(i)}) < \varepsilon$. For $x, y \in D_{r-i+1}^{N+n}$, we write $x \backsim y$ if and only if $U_x^{(i)} \cap U_y^{(i)} \neq \varnothing$. We have for any i and $x \in D_{r-i}^{N+n}$,

$$\pi_i^{-1}(U_x^{(i+1)}) \subset \bigcup_{\substack{x' \in D_{r-i}^{N+n} \\ x' \backsim x}} \bigcup_{y \in \psi_i^{-1}(x')} U_y^{(i)}.$$

Notice that for each $x \in D_{r-i}^{N+n}$, the number of $x' \in D_{r-i}^{N+n}$ with $x' \backsim x$ is not more than 3^r. Therefore, for every $v = (v_1^{(1)}, \ldots, v_{N+n}^{(1)}) \in D_{r-1}^{N+n}$, there are $(v_1^{(k)}, \ldots, v_{N+n}^{(k)}) \in D_{r-1}^{N+n}$, $k = 2, 3, \ldots, L$, and $L \leq 3^r$, with

$$\#_1^{\boldsymbol{a}}(\pi_1^{-1}(U_v^{(2)}), N, \varepsilon) \leq \sum_{k=1}^{L} Z_{r-1}(v_1^{(k)}) \cdots Z_{r-1}(v_{N+n}^{(k)}).$$

We inductively continue while considering that the multiplicity is at most 3^r and obtain

$$\begin{aligned}
&\#_r^{\boldsymbol{a}}(X_r, N, \varepsilon) \\
&\leq 3^{r(r-1)} \sum_{x_1 \in D_1^{N+n}} \Bigg(\sum_{x_2 \in \psi_{r-1}{}^{-1}(x_1)} \Bigg(\cdots \Bigg(\sum_{x_{r-2} \in \psi_3{}^{-1}(x_{r-3})} \\
&\quad \Bigg(\sum_{\substack{(v_1, \ldots, v_{N+n}) \in \psi_2{}^{-1}(x_{r-2}) \\ v_j \in D_{r-1} \text{ for each } j}} (Z_{r-1}(v_1) \cdots Z_{r-1}(v_{N+n}))^{a_1} \Bigg)^{a_2} \Bigg)^{a_3} \cdots \Bigg)^{a_{r-2}} \Bigg)^{a_{r-1}} \\
&= 3^{r(r-1)} \left\{ \sum_{x_1 \in D_1} \Bigg(\sum_{x_2 \in p_{r-1}^{-1}(x_1)} \Bigg(\cdots \Bigg(\sum_{x_{r-1} \in p_2^{-1}(x_{r-2})} Z_{r-1}(x_1, \ldots, x_{r-1})^{a_1} \Bigg)^{a_2} \cdots \Bigg)^{a_{r-2}} \Bigg)^{a_{r-1}} \right\}^{N+n} \\
&= 3^{r(r-1)} Z_0{}^{N+n}.
\end{aligned}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Therefore,

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \lim_{\varepsilon\to 0}\bigg(\lim_{N\to\infty}\frac{\log \#_r^{\boldsymbol{a}}(X_r, N, \varepsilon)}{N}\bigg) \le \log Z_0.$$

Next, we prove $h^{\boldsymbol{a}}(\boldsymbol{T}) \ge \log Z_0$. We fix $0 < \varepsilon < 1/m_r$ and use ε-separated sets. Take and fix $\boldsymbol{s} = (t_1, \ldots, t_r) \in D$, and set $\boldsymbol{s}_i = (t_1, \ldots, t_{r-i+1})$. Fix a natural number N and let $\psi_i : D^N_{r-i+1} \to D^N_{r-i}$ be the product map of p_i as in the previous definition. Define

$$Q_i = \bigg\{\sum_{k=1}^{N} {A_i}^{-k} e_k + \sum_{k=N+1}^{\infty} {A_i}^{-k}\boldsymbol{s}_i \in X_i \,\bigg|\, e_1, \ldots, e_N \in D_{r-i+1}\bigg\}.$$

Then Q_i is an ε-separated set with respect to the metric $d_N^{(i)}$ on X_i. Consider an arbitrary open cover $\mathscr{F}^{(i)}$ of X_i for each i with the following properties (this $(\mathscr{F}^{(i)})_i$ is defined as *a chain of open* (N, ε)*-covers* of $(X_i)_i$ in Definition 3.1).

(1) For every i and $V \in \mathscr{F}^{(i)}$, we have $\operatorname{diam}(V, d_N^{(i)}) < \varepsilon$.

(2) For each $1 \le i \le r-1$ and $U \in \mathscr{F}^{(i+1)}$, there is $\mathscr{F}^{(i)}(U) \subset \mathscr{F}^{(i)}$ such that

$$\pi_i^{-1}(U) \subset \bigcup \mathscr{F}^{(i)}(U)$$

and

$$\mathscr{F}^{(i)} = \bigcup_{U\in\mathscr{F}^{(i+1)}} \mathscr{F}^{(i)}(U).$$

We have $\#(V \cap Q_i) \le 1$ for each $V \in \mathscr{F}^{(i)}$ by (1). Let $(e_1^{(2)}, e_2^{(2)}, \ldots, e_N^{(2)}) \in D^N_{r-1}$ and suppose $U \in \mathscr{F}^{(2)}$ satisfies

$$\sum_{k=1}^{N} {A_2}^{-k} e_k^{(2)} + \sum_{k=N+1}^{\infty} {A_2}^{-k}\boldsymbol{s}_2 \in U \cap Q_2.$$

Then $\pi_1^{-1}(U)$ contains at least $Z_{r-1}(e_1^{(2)}) \cdots Z_{r-1}(e_N^{(2)})$ points of Q_1. Hence,

$$\#_1^{\boldsymbol{a}}(\pi_1^{-1}(U), N, \varepsilon) \ge Z_{r-1}(e_1^{(2)}) \cdots Z_{r-1}(e_N^{(2)}).$$

We continue this reasoning inductively and get

$$\begin{aligned}
&\#_r^{\boldsymbol{a}}(X_r, N, \varepsilon)\\
&\ge \sum_{e^{(r)}\in D_1^N}\Bigg(\sum_{e^{(r-1)}\in{\psi_{r-1}}^{-1}(e^{(r)})}\Bigg(\cdots\Bigg(\sum_{e^{(3)}\in{\psi_3}^{-1}(e^{(4)})}\\
&\quad\Bigg(\sum_{\substack{(e_1^{(2)},\ldots,e_N^{(2)})\in{\psi_2}^{-1}(e^{(3)})\\ e_j^{(2)}\in D_{r-1}\text{ for each } j}} (Z_{r-1}(e_1^{(2)})\cdots Z_{r-1}(e_N^{(2)}))^{a_1}\Bigg)^{a_2}\Bigg)^{a_3}\cdots\Bigg)^{a_{r-2}}\Bigg)^{a_{r-1}}\\
&= \Bigg\{\sum_{x_1\in D_1}\Bigg(\sum_{x_2\in p_{r-1}^{-1}(x_1)}\Bigg(\cdots\Bigg(\sum_{x_{r-1}\in p_2^{-1}(x_{r-2})} Z_{r-1}(x_1,\ldots,x_{r-1})^{a_1}\Bigg)^{a_2}\cdots\Bigg)^{a_{r-2}}\Bigg)^{a_{r-1}}\Bigg\}^N\\
&= {Z_0}^N.
\end{aligned}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

This implies

$$h^{\boldsymbol{a}}(\boldsymbol{T}) \geq \log Z_0.$$

We conclude that

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \log Z_0. \qquad \square$$

We would like to mention the work of Barral and Feng [**BF12**, **Fe11**], and of Yayama [**Ya11**]. These papers studied the related invariants when $(X_i, T_i)(i = 1, \ldots, r)$ are subshifts over finite alphabets. In this subshift case, our definition of $h^{\boldsymbol{a}}(\boldsymbol{T})$ (and its pressure version in §2) is essentially the same as that given in [**BF12**, Theorem 3.1]. Hence, we can say that our definition generalizes the approach in [**BF12**, Theorem 3.1] from subshifts to general dynamical systems.

2. *Weighted topological pressure*

Here, we introduce the generalized, new definition of weighted topological pressure. Let (X_i, T_i) $(i = 1, 2, \ldots, r)$ be dynamical systems and $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r-1)$ factor maps. For a continuous function $f : X_1 \to \mathbb{R}$ and a natural number N, set

$$S_N f(x) = f(x) + f(T_1 x) + f(T_1^2 x) + \cdots + f(T_1^{N-1} x).$$

Let $d^{(i)}$ be a metric on X_i. Recall that we defined a new metric $d_N^{(i)}$ on X_i by

$$d_N^{(i)}(x_1, x_2) = \max_{0 \leq n < N} d^{(i)}(T_i{}^n x_1, T_i{}^n x_2).$$

We may write these as $S_N^{T_1} f$ or $d_N^{T_i}$ to clarify the maps T_1 and T_i in the definitions above.

Let $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$ with $0 \leq a_i \leq 1$ for each i and ε a positive number. We inductively define a quantity $P_i^{\boldsymbol{a}}(\Omega, f, N, \varepsilon)$ for $\Omega \subset X_i$. For $\Omega \subset X_1$, set

$$P_1^{\boldsymbol{a}}(\Omega, f, N, \varepsilon) = \inf \left\{ \sum_{j=1}^{n} \exp(\sup_{U_j} S_N f) \;\middle|\; \begin{array}{l} n \in \mathbb{N}, \{U_j\}_{j=1}^n \text{ is an open cover of } \Omega \\ \text{with } \operatorname{diam}(U_j, d_N^{(1)}) < \varepsilon \text{ for all } 1 \leq j \leq n \end{array} \right\}.$$

Let $\Omega \subset X_{i+1}$. If $P_i^{\boldsymbol{a}}$ is already defined, let

$$P_{i+1}^{\boldsymbol{a}}(\Omega, f, N, \varepsilon) = \inf \left\{ \sum_{j=1}^{n} (P_i^{\boldsymbol{a}}(\pi_i^{-1}(U_j), f, N, \varepsilon))^{a_i} \;\middle|\; \begin{array}{l} n \in \mathbb{N}, \{U_j\}_{j=1}^n \text{ is an open cover of } \Omega \\ \text{with } \operatorname{diam}(U_j, d_N^{T_{i+1}}) < \varepsilon \text{ for all } 1 \leq j \leq n \end{array} \right\}.$$

We define the *topological pressure of $\boldsymbol{a}$-exponent* $P^{\boldsymbol{a}}(f)$ by

$$P^{\boldsymbol{a}}(f) = \lim_{\varepsilon \to 0} \left(\lim_{N \to \infty} \frac{\log P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon)}{N} \right).$$

This limit exists since $\log P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon)$ is sub-additive in N and non-decreasing as ε tends to 0. When $r = 1$, this coincides with the standard definition of the topological pressure $P(f)$ on (X_1, T_1). The topological entropy $h_{\text{top}}(T_1)$ is the value of $P(f)$ when $f \equiv 0$. When we want to clarify the maps T_i and π_i used in the definition of $P^{\boldsymbol{a}}(f)$, we will denote it by $P^{\boldsymbol{a}}(f, \boldsymbol{T})$ or $P^{\boldsymbol{a}}(f, \boldsymbol{T}, \boldsymbol{\pi})$ with $\boldsymbol{T} = (T_i)_{i=1}^r$ and $\boldsymbol{\pi} = (\pi_i)_{i=1}^r$.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Recall that we defined a probability vector $\boldsymbol{w_a} = (w_1, \ldots, w_r)$ from $\boldsymbol{a} = (a_1, a_2, \ldots, a_{r-1})$ by

$$\begin{cases} w_1 = a_1 a_2 a_3 \cdots a_{r-1}, \\ w_2 = (1 - a_1) a_2 a_3 \cdots a_{r-1}, \\ w_3 = (1 - a_2) a_3 \cdots a_{r-1}, \\ \vdots \\ w_{r-1} = (1 - a_{r-2}) a_{r-1}, \\ w_r = 1 - a_{r-1}. \end{cases} \tag{2.1}$$

Let

$$\pi^{(0)} = \mathrm{id}_{X_1} : X_1 \to X_1,$$
$$\pi^{(i)} = \pi_i \circ \pi_{i-1} \circ \cdots \circ \pi_1 : X_1 \to X_{i+1}.$$

We can now state the main result of this paper.

THEOREM 2.1. *Let (X_i, T_i) $(i = 1, 2, \ldots, r)$ be dynamical systems and $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r - 1)$ factor maps. For any continuous function $f : X_1 \to \mathbb{R}$,*

$$P^{\boldsymbol{a}}(f) = \sup_{\mu \in \mathscr{M}^{T_1}(X_1)} \bigg(\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i) + w_1 \int_{X_1} f \, d\mu \bigg). \tag{2.2}$$

We define $P^{\boldsymbol{a}}_{\mathrm{var}}(f)$ to be the right-hand side of this equation, where 'var' is the abbreviation of 'variational'. Then we need to prove

$$P^{\boldsymbol{a}}(f) = P^{\boldsymbol{a}}_{\mathrm{var}}(f).$$

3. *Preparation*

In this section, we prepare several tools which will be used in the proof of Theorem 2.1.

3.1. *Basic properties and tools.*

Let (X_i, T_i) $(i = 1, 2, \ldots, r)$ be dynamical systems, $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r - 1)$ factor maps, $\boldsymbol{a} = (a_1, \ldots, a_{r-1}) \in [0, 1]^{r-1}$, and $f : X_1 \to \mathbb{R}$ a continuous function.

We will use the following notions in §§3.3 and 5.

Definition 3.1. Consider a cover $\mathscr{F}^{(i)}$ of X_i for each i. For a natural number N and a positive number ε, the family $(\mathscr{F}^{(i)})_i$ is said to be *a chain of* $(\boldsymbol{N}, \boldsymbol{\varepsilon})$*-covers* of $(X_i)_i$ if the following conditions are true.

(1) For every i and $V \in \mathscr{F}^{(i)}$, we have $\mathrm{diam}(V, d_N^{(i)}) < \varepsilon$.
(2) For each $1 \le i \le r - 1$ and $U \in \mathscr{F}^{(i+1)}$, there is $\mathscr{F}^{(i)}(U) \subset \mathscr{F}^{(i)}$ such that

$$\pi_i^{-1}(U) \subset \bigcup \mathscr{F}^{(i)}(U)$$

and

$$\mathscr{F}^{(i)} = \bigcup_{U \in \mathscr{F}^{(i+1)}} \mathscr{F}^{(i)}(U).$$

Moreover, if all the elements of each $\mathscr{F}^{(i)}$ are open/closed/compact, we call $(\mathscr{F}^{(i)})_i$ *a chain of open/closed/compact* $(\boldsymbol{N}, \boldsymbol{\varepsilon})$*-covers* of $(X_i)_i$.

Remark 3.2. Note that we can rewrite $P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon)$ using chains of open covers as follows. For a chain of (N, ε)-covers $(\mathscr{F}^{(i)})_i$ of $(X_i)_i$, let

$$\mathscr{P}^{\boldsymbol{a}}\bigg(f, N, \varepsilon, (\mathscr{F}^{(i)})_i\bigg) = \sum_{U^{(r)} \in \mathscr{F}^{(r)}} \bigg(\sum_{U^{(r-1)} \in \mathscr{F}^{(r-1)}(U^{(r)})} \bigg(\cdots \bigg(\sum_{U^{(1)} \in \mathscr{F}^{(1)}(U^{(2)})} e^{\sup_{U^{(1)}} S_N f} \bigg)^{a_1} \cdots \bigg)^{a_{r-2}} \bigg)^{a_{r-1}}.$$

Then

$$P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon) = \inf \{\mathscr{P}^{\boldsymbol{a}}(f, N, \varepsilon, (\mathscr{F}^{(i)})_i) | (\mathscr{F}^{(i)})_i \text{ is a chain of open } (N, \varepsilon)\text{-covers of } (X_i)_i\}.$$

Just like the classic notion of pressure, we have the following property.

LEMMA 3.3. *For any natural number m,*

$$P^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m) = m P^{\boldsymbol{a}}(f, \boldsymbol{T}),$$

where $\boldsymbol{T}^m = ({T_i}^m)_{i=1}^r$.

Proof. Fix $\varepsilon > 0$. It is obvious from the definition of $P_1^{\boldsymbol{a}}$ that for any $\Omega_1 \subset X_1$ and a natural number N,

$$P_1^{\boldsymbol{a}}(\Omega_1, S_m^{T_1} f, \boldsymbol{T}^m, N, \varepsilon) \le P_1^{\boldsymbol{a}}(\Omega_1, f, \boldsymbol{T}, mN, \varepsilon).$$

Let $\Omega_{i+1} \subset X_{i+1}$. By induction on i, we have

$$P_i^{\boldsymbol{a}}(\Omega_{i+1}, S_m^{T_1} f, \boldsymbol{T}^m, N, \varepsilon) \le P_i^{\boldsymbol{a}}(\Omega_{i+1}, f, \boldsymbol{T}, mN, \varepsilon).$$

Thus,

$$P_r^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m, N, \varepsilon) \le P_r^{\boldsymbol{a}}(f, \boldsymbol{T}, mN, \varepsilon). \tag{3.1}$$

There exists $0 < \delta < \varepsilon$ such that for any $1 \le i \le r$,

$$d^{(i)}(x, y) < \delta \implies d_m^{T_i}(x, y) < \varepsilon \quad (\text{for } x, y \in X_i).$$

Then

$$d_N^{T_i^m}(x, y) < \delta \implies d_{mN}^{T_i}(x, y) < \varepsilon \quad (\text{for } x, y \in X_i \text{ and } 1 \le i \le r). \tag{3.2}$$

Let $i = 1$ in equation (3.2), then we have for any $\Omega_1 \subset X_1$,

$$P_1^{\boldsymbol{a}}(\Omega_1, f, \boldsymbol{T}, mN, \varepsilon) \le P_1^{\boldsymbol{a}}(\Omega_1, S_m^{T_1} f, \boldsymbol{T}^m, N, \delta).$$

Take $\Omega_{i+1} \subset X_{i+1}$. Again by induction on i and by equation (3.2), we have

$$P_i^{\boldsymbol{a}}(\Omega_{i+1}, f, \boldsymbol{T}, mN, \varepsilon) \le P_i^{\boldsymbol{a}}(\Omega_{i+1}, S_m^{T_1} f, \boldsymbol{T}^m, N, \delta).$$

Hence,

$$P_r^{\boldsymbol{a}}(f, \boldsymbol{T}, mN, \varepsilon) \le P_r^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m, N, \delta).$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Combining with equation (3.1), we have

$$P_r^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m, N, \varepsilon) \le P_r^{\boldsymbol{a}}(f, \boldsymbol{T}, mN, \varepsilon) \le P_r^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m, N, \delta).$$

Therefore,

$$P^{\boldsymbol{a}}(S_m^{T_1} f, \boldsymbol{T}^m) = m P^{\boldsymbol{a}}(f, \boldsymbol{T}).$$ □

We will later use the following standard lemma of calculus.

LEMMA 3.4

(1) *For* $0 \le a \le 1$ *and non-negative numbers* x, y,

$$(x+y)^a \le x^a + y^a.$$

(2) *Suppose that non-negative real numbers* $p_1, p_2, \ldots, p_n$ *satisfy* $\sum_{i=1}^n p_i = 1$. *Then for any real numbers* $x_1, x_2, \ldots, x_n$, *we have*

$$\sum_{i=1}^n (-p_i \log p_i + x_i p_i) \le \log \sum_{i=1}^n e^{x_i}.$$

In particular, letting $x_1 = x_2 = \cdots = x_n = 0$ *gives*

$$\sum_{i=1}^n (-p_i \log p_i) \le \log n.$$

Here, $0 \cdot \log 0$ *is defined as* 0.

The proof for item (1) is elementary. See [**Wal82**, §9.3, Lemma 9.9] for item (2).

3.2. *Measure theoretic entropy.* In this subsection, we will introduce the classical measure-theoretic entropy (also known as Kolmogorov–Sinai entropy) and state some of the basic lemmas we need to prove Theorem 2.1. The main reference is the book of Walters [**Wal82**].

Let (X, T) be a dynamical system and $\mu \in \mathscr{M}^T(X)$. A set $\mathscr{A} = \{A_1, \ldots, A_n\}$ is called a finite partition of X with measurable elements if $X = A_1 \cup \cdots \cup A_n$, each A_i is a measurable set, and $A_i \cap A_j = \varnothing$ for $i \ne j$. In this paper, a partition is always finite and consists of measurable elements.

Let $\mathscr{A}$ and $\mathscr{A}'$ be partitions of X. We define a new partition $\mathscr{A} \vee \mathscr{A}'$ by

$$\mathscr{A} \vee \mathscr{A}' = \{A \cap A' | A \in \mathscr{A} \text{ and } A' \in \mathscr{A}'\}.$$

For a natural number N, we define a refined partition $\mathscr{A}_N$ of $\mathscr{A}$ by

$$\mathscr{A}_N = \mathscr{A} \vee T^{-1}\mathscr{A} \vee T^{-2}\mathscr{A} \vee \cdots \vee T^{-(N-1)}\mathscr{A},$$

where $T^{-i}\mathscr{A} = \{T^{-i}(A) | A \in \mathscr{A}\}$ is a partition for $i \in \mathbb{N}$.

For a partition $\mathscr{A}$ of X, let

$$H_\mu(\mathscr{A}) = -\sum_{A \in \mathscr{A}} \mu(A) \log(\mu(A)).$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

We set

$$h_\mu(T,\mathscr{A}) = \lim_{N\to\infty}\frac{H_\mu(\mathscr{A}_N)}{N}.$$

This limit exists since $H_\mu(\mathscr{A}_N)$ is sub-additive in N. The *measure theoretic entropy* $h_\mu(T)$ is defined by

$$h_\mu(T) = \sup\{h_\mu(T,\mathscr{A})|\mathscr{A} \text{ is a partition of } X\}.$$

Let $\mathscr{A}$ and $\mathscr{A}'$ be partitions. Their *conditional entropy* is defined by

$$H_\mu(\mathscr{A}|\mathscr{A}') = -\sum_{\substack{A'\in\mathscr{A}'\\ \mu(A')\neq 0}}\mu(A')\sum_{A\in\mathscr{A}}\frac{\mu(A\cap A')}{\mu(A')}\log\bigg(\frac{\mu(A\cap A')}{\mu(A')}\bigg).$$

LEMMA 3.5

(1) *$H_\mu(\mathscr{A})$ is sub-additive in $\mathscr{A}$: that is, for partitions $\mathscr{A}$ and $\mathscr{A}'$,*

$$H_\mu(\mathscr{A}\vee\mathscr{A}') \le H_\mu(\mathscr{A}) + H_\mu(\mathscr{A}').$$

(2) *$H_\mu(\mathscr{A})$ is concave in μ: that is, for $\mu,\nu\in\mathscr{M}^T(X)$ and $0\le t\le 1$,*

$$H_{(1-t)\mu+t\nu}(\mathscr{A}) \ge (1-t)H_\mu(\mathscr{A}) + tH_\nu(\mathscr{A}).$$

(3) *For partitions $\mathscr{A}$ and $\mathscr{A}'$,*

$$h_\mu(T,\mathscr{A}) \le h_\mu(T,\mathscr{A}') + H_\mu(\mathscr{A}'|\mathscr{A}).$$

For the proof, confer with [**Wal82**, Theorem 4.3(viii), §4.5] for item (1), [**Wal82**, Remark, §8.1] for item (2), and [**Wal82**, Theorem 4.12, §4.5] for item (3).

3.3. *Zero-dimensional principal extension.* Here we will see how we can reduce the proof of $P^{\boldsymbol{a}}(f) \le P^{\boldsymbol{a}}_{\mathrm{var}}(f)$ to the case where all dynamical systems are zero- dimensional.

First, we review the definitions and properties of (zero-dimensional) principal extension. The introduction here closely follows Tsukamoto's paper [**Tsu22**] and the book of Downarowicz [**Dow11**]. Suppose $\pi : (Y,S)\to(X,T)$ is a factor map between dynamical systems. Let d be a metric on Y. We define the *conditional topological entropy* of π by

$$h_{\mathrm{top}}(Y,S|X,T) = \lim_{\varepsilon\to 0}\bigg(\lim_{N\to\infty}\frac{\sup_{x\in X}\log\#(\pi^{-1}(x),N,\varepsilon)}{N}\bigg).$$

Here,

$$\#(\pi^{-1}(x),N,\varepsilon) = \min\left\{n\in\mathbb{N}\;\middle|\;\begin{array}{l}\text{There exists an open cover } \{U_j\}_{j=1}^n \text{ of } \pi^{-1}(x)\\ \text{with } \mathrm{diam}(U_j,d_N)<\varepsilon \text{ for all } 1\le j\le n\end{array}\right\}.$$

A factor map $\pi : (Y,S)\to(X,T)$ between dynamical systems is said to be a *principal factor map* if

$$h_{\mathrm{top}}(Y,S|X,T) = 0.$$

Also, (Y,S) is called a *principal extension* of (X,T).

The following theorem is from [**Dow11**, Corollary 6.8.9].

THEOREM 3.6. *Suppose $\pi : (Y, S) \to (X, T)$ is a principal factor map. Then π preserves measure-theoretic entropy, namely,*

$$h_\mu(S) = h_{\pi_*\mu}(T)$$

for any S-invariant probability measure μ on Y.

More precisely, it is proved in [**Dow11**, Corollary 6.8.9] that π is a principal factor map if and only if it preserves measure-theoretic entropy, provided that $h_{\mathrm{top}}(X, T) < \infty$.

Suppose $\pi : (X_1, T_1) \to (X_2, T_2)$ and $\phi : (Y, S) \to (X_2, T_2)$ are factor maps between dynamical systems. We define a fiber product $(X_1 \times_{X_2} Y, T_1 \times S)$ of (X_1, T_1) and (Y, S) over (X_2, T_2) by

$$X_1 \times_{X_2} Y = \{(x, y) \in X_1 \times Y | \pi(x) = \phi(y)\},$$

$$T_1 \times S : X_1 \times_{X_2} Y \ni (x, y) \longmapsto (T_1(x), S(y)) \in X_1 \times_{X_2} Y.$$

We have the following commutative diagram:

$$\begin{array}{ccc} X_1 \times_{X_2} Y & \xrightarrow{\ \psi\ } & X_1 \\ {\scriptstyle \pi'} \downarrow & & \downarrow {\scriptstyle \pi} \\ Y & \xrightarrow[\ \phi\ ]{} & X_2 \end{array} \tag{3.3}$$

Here, π' and ψ are restrictions of the projections onto Y and X_1, respectively:

$$\pi' : X_1 \times_{X_2} Y \ni (x, y) \longmapsto y \in Y,$$

$$\psi : X_1 \times_{X_2} Y \ni (x, y) \longmapsto x \in X_1.$$

Since π and ϕ are surjective, both π' and ψ are factor maps. The following lemma is proved in [**Tsu22**, Lemma 5.3].

LEMMA 3.7. *If ϕ is a principal extension in the diagram in equation (3.3), then ψ is also a principal extension.*

A dynamical system (Y, S) is said to be *zero-dimensional* if there is a clopen basis of the topology of Y, where clopen means any element in the basis is both closed and open. A basic example of a zero-dimensional dynamical system is the Cantor set $\{0, 1\}^{\mathbb{N}}$ with the shift map.

A principal extension (Y, S) of (X, T) is called a *zero-dimensional principal extension* if (Y, S) is zero-dimensional. The following important theorem can be found in [**Dow11**, Theorem 7.6.1].

THEOREM 3.8. *For any dynamical system, there is a zero-dimensional principal extension.*

Let (Y_i, R_i) $(i = 1, 2, \ldots, m)$ be dynamical systems, $\pi_i : Y_i \to Y_{i+1}$ $(i = 1, 2, \ldots, m - 1)$ factor maps, and $\boldsymbol{a} = (a_1, \ldots, a_{m-1}) \in [0, 1]^{m-1}$. Fix $2 \le k \le m - 1$ and take a zero-dimensional principal extension $\phi_k : (Z_k, S_k) \to (Y_k, R_k)$. For each $1 \le i \le k - 1$,

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

let $(Y_i \times_{Y_k} Z_k, R_i \times S_k)$ be the fiber product and $\phi_i : Y_i \times_{Y_k} Z_k \to Y_i$ be the restriction of the projection as in the earlier definition. We have

$$\begin{array}{ccc} Y_i \times_{Y_k} Z_k & \xrightarrow{\ \phi_i\ } & Y_i \\ \downarrow & & \downarrow {\scriptstyle \pi_{k-1}\circ\pi_{k-2}\circ\cdots\circ\pi_i} \\ Z_k & \xrightarrow[\ \phi_k\ ]{} & Y_k \end{array}$$

By Lemma 3.7, ϕ_i is a principal factor map. We define $\Pi_i : Y_i \times_{Y_k} Z_k \to Y_{i+1} \times_{Y_k} Z_k$ by $\Pi_i(x, y) = (\pi_i(x), y)$ for each $1 \le i \le k-2$, and $\Pi_{k-1} : Y_{k-1} \times_{Y_k} Z_k \to Z_k$ as the projection. Then we have the following commutative diagram:

$$\begin{array}{ccc} Y_1 \times_{Y_k} Z_k & \xrightarrow{\ \phi_1\ } & Y_1 \\ {\scriptstyle \Pi_1}\downarrow & & \downarrow{\scriptstyle \pi_1} \\ Y_2 \times_{Y_k} Z_k & \xrightarrow{\ \phi_2\ } & Y_2 \\ {\scriptstyle \Pi_2}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ \vdots & & \vdots \\ {\scriptstyle \Pi_{k-2}}\downarrow & & \downarrow{\scriptstyle \pi_{k-2}} \\ Y_{k-1} \times_{Y_k} Z_k & \xrightarrow{\ \phi_{k-1}\ } & Y_{k-1} \\ {\scriptstyle \Pi_{k-1}}\downarrow & & \downarrow{\scriptstyle \pi_{k-1}} \\ Z_k & \xrightarrow{\ \phi_k\ } & Y_k \\ & \searrow^{\pi_k\circ\phi_k} & \downarrow{\scriptstyle \pi_k} \\ & & Y_{k+1} \\ & & \downarrow{\scriptstyle \pi_{k+1}} \\ & & \vdots \\ & & \downarrow{\scriptstyle \pi_{m-1}} \\ & & Y_m \end{array} \tag{3.4}$$

Let

$$(Z_i, S_i) = (Y_i \times_{Y_k} Z_k, R_i \times S_k) \quad \text{for } 1 \le i \le k-1, \quad (Z_i, S_i) = (Y_i, R_i) \quad \text{for } k+1 \le i \le m,$$
$$\Pi_k = \pi_k \circ \phi_k : Z_k \to Y_{k+1}, \Pi_i = \pi_i : Z_i \to Z_{i+1} \quad \text{for } k+1 \le i \le m-1,$$
$$\phi_i = \mathrm{id}_{Z_i} : Z_i \to Z_i \quad \text{for } k+1 \le i \le m.$$

LEMMA 3.9. *In the settings above,*

$$P^{\boldsymbol{a}}_{\mathrm{var}}(f, \boldsymbol{R}, \boldsymbol{\pi}) \ge P^{\boldsymbol{a}}_{\mathrm{var}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi})$$

and

$$P^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}) \le P^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}).$$

Here, $\boldsymbol{R} = (R_i)_i$, $\boldsymbol{\pi} = (\pi_i)_i$, $\boldsymbol{S} = (S_i)_i$ *and* $\boldsymbol{\Pi} = (\Pi_i)_i$.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Proof. We remark that the following proof does not require Z_k to be zero-dimensional. Let

$$\pi^{(0)} = \mathrm{id}_{Y_1} : Y_1 \to Y_1,$$
$$\pi^{(i)} = \pi_i \circ \pi_{i-1} \circ \cdots \circ \pi_1 : Y_1 \to Y_{i+1},$$

and

$$\Pi^{(0)} = \mathrm{id}_{Z_1} : Z_1 \to Z_1,$$
$$\Pi^{(i)} = \Pi_i \circ \Pi_{i-1} \circ \cdots \circ \Pi_1 : Z_1 \to Z_{i+1}.$$

Let $\nu \in \mathscr{M}^{S_1}(Y_1)$ and $1 \le i \le m$. Since all the horizontal maps in equation (3.4) are principal factor maps, we have

$$h_{\Pi^{(i-1)}{}_*\nu}(S_i) = h_{(\phi_i)_*\Pi^{(i-1)}{}_*\nu}(R_i) = h_{\pi^{(i-1)}{}_*(\phi_1)_*\nu}(R_i).$$

It follows that

$$\begin{aligned} P^{\boldsymbol{a}}_{\mathrm{var}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}) &= \sup_{\nu \in \mathscr{M}^{S_1}(Z_1)} \bigg(\sum_{i=1}^m w_i h_{\Pi^{(i-1)}{}_*\nu}(S_i) + w_1 \int_{Z_1} f \circ \phi_1 \, d\nu \bigg) \\ &= \sup_{\nu \in \mathscr{M}^{S_1}(Z_1)} \bigg(\sum_{i=1}^m w_i h_{\pi^{(i-1)}{}_*(\phi_1)_*\nu}(R_i) + w_1 \int_{Y_1} f \, d((\phi_1)_*\nu) \bigg) \\ &\le \sup_{\mu \in \mathscr{M}^{T_1}(Y_1)} \bigg(\sum_{i=1}^m w_i h_{\pi^{(i-1)}{}_*\mu}(R_i) + w_1 \int_{Y_1} f \, d\mu \bigg) \\ &= P^{\boldsymbol{a}}_{\mathrm{var}}(f, \boldsymbol{R}, \boldsymbol{\pi}). \end{aligned}$$

(The reversed inequality is generally true by the surjectivity of factor maps, yielding equality. However, we do not use this fact.)

Let d^i be a metric on Y_i for each i and $\widetilde{d}^k$ a metric on Z_k. We define a metric $\widetilde{d}^i$ on (Z_i, S_i) for $1 \le i \le k-1$ by

$$\begin{aligned} &\widetilde{d}^i((x_1, y_1), (x_2, y_2)) \\ &\quad = \max\{d^i(x_1, x_2), \widetilde{d}^k(y_1, y_2)\} \quad ((x_1, y_1), (x_2, y_2) \in Z_i = Y_i \times_{Y_k} Z_k). \end{aligned}$$

Set $\widetilde{d}^i = d^i$ for $k+1 \le i \le m$. Take an arbitrary positive number ε. There exists $0 < \delta < \varepsilon$ such that for every $1 \le i \le m$,

$$\widetilde{d}^i(x, y) < \delta \implies d^i(\phi_i(x), \phi_i(y)) < \varepsilon \quad (x, y \in Z_i). \tag{3.5}$$

Let N be a natural number. We claim that

$$P^{\boldsymbol{a}}_r(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon) \le P^{\boldsymbol{a}}_r(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta).$$

Take $M > 0$ with

$$P^{\boldsymbol{a}}_r(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta) < M.$$

Then there exists a chain of open (N, δ)-covers $(\mathscr{F}^{(i)})_i$ of $(Z_i)_i$ (see Definition 3.1 and Remark 3.2) with

$$\mathscr{P}^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta, (\mathscr{F}^{(i)})_i) < M.$$

We can find a compact set $C_U \subset U$ for each $U \in \mathscr{F}^{(m)}$ such that $\bigcup_{U \in \mathscr{F}^{(m)}} C_U = Z_m$. Let $\mathscr{K}^{(m)} := \{C_U | U \in \mathscr{F}^{(m)}\}$. Since $\Pi_{m-1}^{-1}(C_U) \subset \Pi_{m-1}^{-1}(U)$ is compact for each $U \in \mathscr{F}^{(m)}$, we can find a compact set $E_V \subset V$ for each $V \in \mathscr{F}^{(m-1)}(U)$ such that $\Pi_{m-1}^{-1}(C_U) \subset \bigcup_{V \in \mathscr{F}^{(k)}(U)} E_V$. Let $\mathscr{K}^{(m-1)}(C_U) := \{E_V | V \in \mathscr{F}^{(m-1)}(U)\}$ and $\mathscr{K}^{(m-1)} := \bigcup_{C \in \mathscr{K}^{(m)}} \mathscr{K}^{(m-1)}(C)$. We continue likewise and obtain a chain of compact (N, δ)-covers $(\mathscr{K}^{(i)})_i$ of $(Z_i)_i$ with

$$\mathscr{P}^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta, (\mathscr{K}^{(i)})_i) \leq \mathscr{P}^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta, (\mathscr{F}^{(i)})_i) < M.$$

Let $\phi_i(\mathscr{K}^{(i)}) = \{\phi_i(C) | C \in \mathscr{K}^{(i)}\}$ for each i. Note that for any $\Omega \subset Z_i$,

$$\pi_{i-1}^{-1}(\phi_i(\Omega)) = \phi_{i-1}(\Pi_{i-1}^{-1}(\Omega)).$$

This and equation (3.5) assure that $(\phi_i(\mathscr{K}^{(i)}))_i$ is a chain of compact (N, ε)-covers of $(Y_i)_i$. We have

$$\mathscr{P}^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon, (\phi_i(\mathscr{K}^{(i)}))_i) = \mathscr{P}^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta, (\mathscr{K}^{(i)})_i) < M.$$

Since f is continuous and each $\phi_i(\mathscr{K}^{(i)})$ is a closed cover, we can slightly enlarge each set in $\phi_i(\mathscr{K}^{(i)})$ and create a chain of open (N, ε)-covers $(\mathscr{O}^{(i)})_i$ of $(Y_i)_i$ satisfying

$$\mathscr{P}^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon, (\mathscr{O}^{(i)})_i) < M.$$

Therefore,

$$P_r^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon) \leq \mathscr{P}^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon, (\mathscr{O}^{(i)})_i) < M.$$

Since $M > P_r^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta)$ was chosen arbitrarily, we have

$$P_r^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}, N, \varepsilon) \leq P_r^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}, N, \delta).$$

This implies

$$P^{\boldsymbol{a}}(f, \boldsymbol{R}, \boldsymbol{\pi}) \leq P^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}, \boldsymbol{\Pi}). \qquad \square$$

The following proposition reduces the proof of $P^{\boldsymbol{a}}(f) \leq P^{\boldsymbol{a}}_{\mathrm{var}}(f)$ in the next section to the case where all dynamical systems are zero-dimensional.

PROPOSITION 3.10. *For all dynamical systems (X_i, T_i) $(i = 1, 2, \ldots, r)$ and factor maps $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r-1)$, there are zero-dimensional dynamical systems (Z_i, S_i) $(i = 1, 2, \ldots, r)$ and factor maps $\Pi_i : Z_i \to Z_{i+1}$ $(i = 1, 2, \ldots, r-1)$ with the following property; for every continuous function $f : X_1 \to \mathbb{R}$, there exists a continuous function $g : Z_1 \to \mathbb{R}$ with*

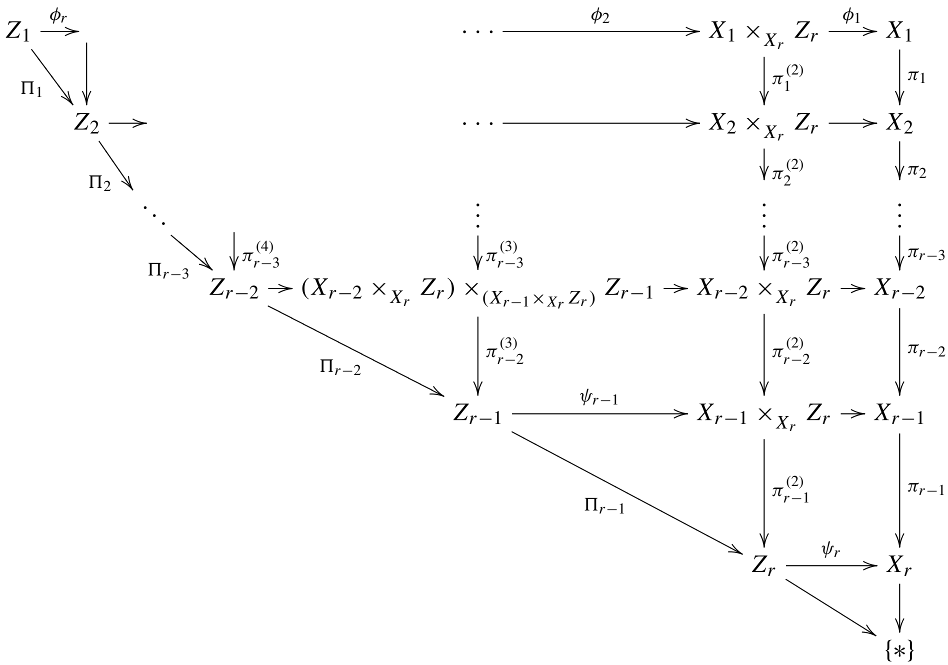
$$P^{\boldsymbol{a}}_{\mathrm{var}}(f, \boldsymbol{T}, \boldsymbol{\pi}) \geq P^{\boldsymbol{a}}_{\mathrm{var}}(g, \boldsymbol{S}, \boldsymbol{\Pi})$$

and

$$P^{\boldsymbol{a}}(f, \boldsymbol{T}, \boldsymbol{\pi}) \leq P^{\boldsymbol{a}}(g, \boldsymbol{S}, \boldsymbol{\Pi}).$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Proof. We will first construct zero-dimensional dynamical systems (Z_i, S_i) $(i = 1, 2, \ldots, r)$ and factor maps $\Pi_i : Z_i \to Z_{i+1}$ $(i = 1, 2, \ldots, r-1)$ alongside the following commutative diagram of dynamical systems and factor maps:



(3.6)

where all the horizontal maps are principal factor maps.

By Theorem 3.8, there is a zero-dimensional principal extension $\psi_r : (Z_r, S_r) \to (X_r, T_r)$. The set $\{*\}$ is the trivial dynamical system, and the maps $X_r \to \{*\}$ and $Z_r \to \{*\}$ send every element to $*$. For each $1 \le i \le r-1$, the map $X_i \times_{X_r} Z_r \to X_i$ in the following diagram is a principal factor map by Lemma 3.7:

$$\begin{array}{ccc} X_i \times_{X_r} Z_r & \longrightarrow & X_i \\ \downarrow & & \downarrow{\scriptstyle \pi_{r-1} \circ \pi_{r-2} \circ \cdots \circ \pi_i} \\ Z_r & \xrightarrow[\psi_r]{} & X_r \end{array}$$

For $1 \le i \le r-2$, define $\pi_i^{(2)} : X_i \times_{X_r} Z_r \to X_{i+1} \times_{X_r} Z_r$ by

$$\pi_i^{(2)}(x, z) = (\pi_i(x), y).$$

Then every horizontal map in the right two rows of diagram (3.6) is a principal factor map. Next, take a zero-dimensional principal extension $\psi_{r-1} : (Z_{r-1}, S_{r-1}) \to (X_{r-1} \times_{X_r} Z_r, T_{r-1} \times S_r)$ and let $\Pi_{r-1} = \pi_{r-1}^{(2)} \circ \psi_{r-1}$. The rest of diagram (3.6) is constructed similarly, and by Lemma 3.7, each horizontal map is a principal factor map.

Let $f : X_1 \to \mathbb{R}$ be a continuous map. Applying Lemma 3.9 to the right two rows of diagram (3.6), we get

$$P^{\boldsymbol{a}}_{\mathrm{var}}(f, \boldsymbol{T}, \boldsymbol{\pi}) \geq P^{\boldsymbol{a}}_{\mathrm{var}}(f \circ \phi_1, \boldsymbol{S}^{(2)}, \boldsymbol{\Pi}^{(2)})$$

and

$$P^{\boldsymbol{a}}(f, \boldsymbol{T}, \boldsymbol{\pi}) \leq P^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}^{(2)}, \boldsymbol{\Pi}^{(2)})$$

for $\boldsymbol{\Pi}^{(2)} = (\pi_i^{(2)})_i$ and $\boldsymbol{S}^{(2)} = (T_i \times S_r)_i$. Again by Lemma 3.9,

$$P^{\boldsymbol{a}}_{\mathrm{var}}(f \circ \phi_1, \boldsymbol{S}^{(2)}, \boldsymbol{\Pi}^{(2)}) \geq P^{\boldsymbol{a}}_{\mathrm{var}}(f \circ \phi_1 \circ \phi_2, \boldsymbol{S}^{(3)}, \boldsymbol{\Pi}^{(3)})$$

and

$$P^{\boldsymbol{a}}(f \circ \phi_1, \boldsymbol{S}^{(2)}, \boldsymbol{\Pi}^{(2)}) \leq P^{\boldsymbol{a}}(f \circ \phi_1 \circ \phi_2, \boldsymbol{S}^{(3)}, \boldsymbol{\Pi}^{(3)})$$

where $\boldsymbol{\Pi}^{(3)} = ((\pi_i^{(3)})_{i=1}^{r-2}, \Pi_{r-1})$, and $\boldsymbol{S}^{(3)}$ is the collection of maps associated with Z_r and the third row from the right of diagram (3.6). We continue inductively and obtain the desired inequalities, where g is taken as $f \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_r$. □

4. *Proof of* $P^{\boldsymbol{a}}(f) \leq P^{\boldsymbol{a}}_{\mathrm{var}}(f)$

Let $\boldsymbol{a} = (a_1, \ldots, a_{r-1}) \in [0, 1]^{r-1}$. Recall that we defined $(w_1, \ldots, w_r)$ by

$$\begin{cases} w_1 = a_1 a_2 a_3 \cdots a_{r-1}, \\ w_2 = (1 - a_1) a_2 a_3 \cdots a_{r-1}, \\ w_3 = (1 - a_2) a_3 \cdots a_{r-1}, \\ \qquad \vdots \\ w_{r-1} = (1 - a_{r-2}) a_{r-1}, \\ w_r = 1 - a_{r-1} \end{cases}$$

and $P^{\boldsymbol{a}}_{\mathrm{var}}(f)$ by

$$P^{\boldsymbol{a}}_{\mathrm{var}}(f) = \sup_{\mu \in \mathscr{M}^{T_1}(X_1)} \bigg(\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i) + w_1 \int_{X_1} f \, d\mu \bigg),$$

where

$$\pi^{(0)} = \mathrm{id}_{X_1} : X_1 \to X_1,$$
$$\pi^{(i)} = \pi_i \circ \pi_{i-1} \circ \cdots \circ \pi_1 : X_1 \to X_{i+1}.$$

The following theorem suffices by Proposition 3.10 in proving $P^{\boldsymbol{a}}(f) \leq P^{\boldsymbol{a}}_{\mathrm{var}}(f)$ for arbitrary dynamical systems.

THEOREM 4.1. *Suppose (X_i, T_i) $(i = 1, 2, \ldots, r)$ are zero-dimensional dynamical systems and $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r-1)$ are factor maps. Then we have*

$$P^{\boldsymbol{a}}(f) \leq P^{\boldsymbol{a}}_{\mathrm{var}}(f)$$

for any continuous function $f : X_1 \to \mathbb{R}$.

Proof. Let $d^{(i)}$ be a metric on X_i for each $i = 1, 2, \ldots, r$. Take a positive number ε and a natural number N. First, we will backward inductively define a finite clopen partition

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

$\mathscr{A}^{(i)}$ of X_i for each i. Since X_r is zero-dimensional, we can take a sufficiently fine finite clopen partition $\mathscr{A}^{(r)}$ of X_r. That is, each $A \in \mathscr{A}^{(r)}$ is both open and closed, and $\operatorname{diam}(A, d_N^{(r)}) < \varepsilon$. Suppose $\mathscr{A}^{(i+1)}$ is defined. For each $A \in \mathscr{A}^{(i+1)}$, take a clopen partition $\mathscr{B}(A)$ of $\pi_i^{-1}(A) \subset X_i$ such that any $B \in \mathscr{B}(A)$ satisfies $\operatorname{diam}(B, d_N^{(i)}) < \varepsilon$. We let $\mathscr{A}^{(i)} = \bigcup_{A \in \mathscr{A}^{(i+1)}} \mathscr{B}(A)$. Then $\mathscr{A}^{(i)}$ is a finite clopen partition of X_i. We define

$$\mathscr{A}_N^{(i)} = \mathscr{A}^{(i)} \vee T_i^{-1}\mathscr{A}^{(i)} \vee T_i^{-2}\mathscr{A}^{(i)} \vee \cdots \vee T_i^{-(N-1)}\mathscr{A}^{(i)}.$$

We employ the following notation. For $i < j$ and $A \in \mathscr{A}_N^{(j)}$, let $\mathscr{A}_N^{(i)}(A)$ be the set of 'children' of A:

$$\mathscr{A}_N^{(i)}(A) = \{B \in \mathscr{A}_N^{(i)} | \pi_{j-1} \circ \pi_{j-2} \circ \cdots \circ \pi_i(B) \subset A\}.$$

Also, for $B \in \mathscr{A}_N^{(i)}$ and $i < j$, we denote by $\widetilde{\pi}_j B$ the unique 'parent' of B in $\mathscr{A}_N^{(j)}$:

$$\widetilde{\pi}_j B = A \in \mathscr{A}_N^{(j)} \quad \text{such that } \pi_{j-1} \circ \pi_{j-2} \circ \cdots \circ \pi_i(B) \subset A.$$

We will evaluate $P^{\boldsymbol{a}}(f, N, \varepsilon)$ from above using $\{\mathscr{A}^{(i)}\}$. Let $A \in \mathscr{A}_N^{(2)}$, and start by setting

$$Z_N^{(1)}(A) = \sum_{B \in \mathscr{A}_N^{(1)}(A)} e^{\sup_B S_N f}.$$

Let $A \in \mathscr{A}_N^{(i+1)}$. If $Z_N^{(i-1)}$ is already defined, set

$$Z_N^{(i)}(A) = \sum_{B \in \mathscr{A}_N^{(i)}(A)} (Z_N^{(i-1)}(B))^{a_{i-1}}.$$

We then define Z_N by

$$Z_N = \sum_{A \in \mathscr{A}_N^{(r)}} (Z_N^{(r-1)}(A))^{a_{r-1}}.$$

It is straightforward from the construction that

$$P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon) \leq Z_N.$$

Therefore, we only need to prove that there is a T_1-invariant probability measure μ on X_1 such that

$$\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i, \mathscr{A}^{(i)}) + w_1 \int_{X_1} f \, d\mu \geq \lim_{N \to \infty} \frac{\log Z_N}{N}.$$

Since each $A \in \mathscr{A}_N^{(1)}$ is closed, we can choose a point $x_A \in A$ so that

$$S_N f(x_A) = \sup_A S_N f.$$

We define a probability measure σ_N on X_1 by

$$\sigma_N = \frac{1}{Z_N} \sum_{A \in \mathscr{A}_N^{(1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} Z_N^{(r-2)}(\widetilde{\pi}_{r-1} A)^{a_{r-2}-1}$$
$$\times \cdots \times Z_N^{(2)}(\widetilde{\pi}_3 A)^{a_2-1} Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)} \delta_{x_A},$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

where δ_{x_A} is the Dirac measure at x_A. This is indeed a probability measure on X_1 since

$$\sigma_N(X_1) = \frac{1}{Z_N} \sum_{A \in \mathscr{A}_N^{(1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} Z_N^{(r-2)}(\widetilde{\pi}_{r-1} A)^{a_{r-2}-1}$$

$$\times \cdots \times Z_N^{(2)}(\widetilde{\pi}_3 A)^{a_2-1} Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)}$$

$$= \frac{1}{Z_N} \sum_{A_r \in \mathscr{A}_N^{(r)}} Z_N^{(r-1)}(A_r)^{a_{r-1}-1} \sum_{A_{r-1} \in \mathscr{A}_N^{(r-1)}(A_r)} Z_N^{(r-2)}(A_{r-1})^{a_{r-2}-1}$$

$$\cdots \sum_{A_3 \in \mathscr{A}_N^{(3)}(A_4)} Z_N^{(2)}(A_3)^{a_2-1} \sum_{A_2 \in \mathscr{A}_N^{(2)}(A_3)} Z_N^{(1)}(A_2)^{a_1-1} \underbrace{\sum_{A_1 \in \mathscr{A}_N^{(1)}(A_2)} e^{S_N f(x_{A_1})}}_{=Z_N^{(1)}(A_2)}$$

$$= \frac{1}{Z_N} \sum_{A_r \in \mathscr{A}_N^{(r)}} Z_N^{(r-1)}(A_r)^{a_{r-1}-1} \sum_{A_{r-1} \in \mathscr{A}_N^{(r-1)}(A_r)} Z_N^{(r-2)}(A_{r-1})^{a_{r-2}-1}$$

$$\cdots \sum_{A_3 \in \mathscr{A}_N^{(3)}(A_4)} Z_N^{(2)}(A_3)^{a_2-1} \underbrace{\sum_{A_2 \in \mathscr{A}_N^{(2)}(A_3)} Z_N^{(1)}(A_2)^{a_1}}_{=Z_N^{(2)}(A_3)}$$

$$= \cdots = \frac{1}{Z_N} \sum_{A_r \in \mathscr{A}_N^{(r)}} Z_N^{(j-1)}(A_r)^{a_{r-1}} = 1.$$

Although σ_N is not generally T_1-invariant, the following well-known trick allows us to create a T_1-invariant measure μ. We begin by setting

$$\mu_N = \frac{1}{N} \sum_{k=0}^{N-1} {T_1^k}_* \sigma_N.$$

Since X_1 is compact, we can take a sub-sequence of $(\mu_N)_N$ so that it weakly converges to a probability measure μ on X_1. Then μ is T_1-invariant by the definition of μ_N. We will show that this μ satisfies

$$\sum_{i=1}^{r} w_i h_{\pi^{(i-1)}{}_*\mu}(T_i, \mathscr{A}^{(i)}) + w_1 \int_{X_1} f \, d\mu \geq \lim_{N \to \infty} \frac{\log Z_N}{N}.$$

We first prove

$$\sum_{i=1}^{r} w_i H_{\pi^{(i-1)}{}_*\sigma_N}(\mathscr{A}_N^{(i)}) + w_1 \int_{X_1} S_N f \, d\mu = \log Z_N.$$

To simplify the notation, let

$$\sigma_N^{(i)} = \pi^{(i-1)}{}_* \sigma_N$$

$$= \frac{1}{Z_N} \sum_{B \in \mathscr{A}_N^{(1)}} Z_N^{(r-1)}(\widetilde{\pi}_r B)^{a_{r-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 B)^{a_1-1} e^{S_N f(x_B)} \delta_{\pi^{(i)}(x_B)}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

and

$$W_N^{(j)} = \sum_{A \in \mathscr{A}_N^{(j+1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(j+1)}(\widetilde{\pi}_{j+2} A)^{a_{j+1}-1} Z_N^{(j)}(A)^{a_j} \log (Z_N^{(j)}(A)).$$

CLAIM 4.2. *We have the following equations:*

$$H_{\sigma_N}(\mathscr{A}_N^{(1)}) = \log Z_N - \int_{X_1} S_N f \, d\sigma_N - \sum_{j=1}^{r-1} \frac{a_j - 1}{Z_n} W_N^{(j)},$$

$$H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) = \log Z_N - \frac{a_{i-1}}{Z_n} W_N^{(i-1)} - \sum_{j=i}^{r-1} \frac{a_j - 1}{Z_n} W_N^{(j)} \quad (\textit{for } 2 \leq i \leq r).$$

Here, $\sum_{j=r}^{r-1} ((a_j - 1)/Z_n) W_N^{(j)}$ *is defined to be* 0.

Proof. Let $A \in \mathscr{A}_N^{(1)}$. We have

$$\sigma_N(A) = \frac{1}{Z_N} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)}.$$

Then

$$\begin{aligned} H_{\sigma_N}(\mathscr{A}_N^{(1)}) &= - \sum_{A \in \mathscr{A}_N^{(1)}} \sigma_N(A) \log (\sigma_N(A)) \\ &= \log Z_N - \underbrace{\sum_{A \in \mathscr{A}_N^{(1)}} \sigma_N(A) S_N f(x_A)}_{\text{(I)}} \\ &\quad - \sum_{j=1}^{r-1} \frac{a_j - 1}{Z_N} \underbrace{\sum_{A \in \mathscr{A}_N^{(1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)} \log (Z_N^{(j)}(\widetilde{\pi}_{j+1} A))}_{\text{(II)}} . \end{aligned}$$

For term (I), we have

$$\begin{aligned} \int_{X_1} S_N f \, d\sigma_N &= \frac{1}{Z_N} \sum_{A \in \mathscr{A}_N^{(1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \\ &\qquad \cdots Z_N^{(2)}(\widetilde{\pi}_3 A)^{a_2-1} Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)} S_N f(x_A) \\ &= \text{(I)}. \end{aligned}$$

We will show that (II) $= W_N^{(j)}$. Let $A' \in \mathscr{A}_N^{(j+1)}$. Then any $A \in \mathscr{A}_N^{(1)}(A')$ satisfies $\widetilde{\pi}_{j+1} A = A'$. Hence,

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

$$
\begin{aligned}
\text{(II)} &= \sum_{A' \in \mathscr{A}_N^{(j+1)}} \sum_{A \in \mathscr{A}_N^{(1)}(A')} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)} \log\,(Z_N^{(j)}(\widetilde{\pi}_{j+1} A)) \\
&= \sum_{A' \in \mathscr{A}_N^{(j+1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A')^{a_{r-1}-1} \cdots Z_N^{(j+1)}(\widetilde{\pi}_{j+2} A')^{a_{j+1}-1} Z_N^{(j)}(A')^{a_j-1} \log\,(Z_N^{(j)}(A')) \\
&\quad \times \underbrace{\sum_{A \in \mathscr{A}_N^{(1)}(A')} Z_N^{(j-1)}(\widetilde{\pi}_j A)^{a_{j-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 A)^{a_1-1} e^{S_N f(x_A)}}_{\text{(II)}'}.
\end{aligned}
$$

The term (II)$'$ can be calculated similarly to how we showed $\sigma_N(X_1) = 1$. Namely,

$$
\begin{aligned}
\text{(II)}' &= \sum_{A_j \in \mathscr{A}_N^{(j)}(A')} Z_N^{(j-1)}(A_j)^{a_{j-1}-1} \sum_{A_{j-1} \in \mathscr{A}_N^{(j-1)}(A_j)} Z_N^{(j-2)}(A_{j-1})^{a_{j-2}-1} \\
&\quad \cdots \sum_{A_3 \in \mathscr{A}_N^{(3)}(A_4)} Z_N^{(2)}(A_3)^{a_2-1} \sum_{A_2 \in \mathscr{A}_N^{(2)}(A_3)} Z_N^{(1)}(A_2)^{a_1-1} \underbrace{\sum_{A_1 \in \mathscr{A}_N^{(1)}(A_2)} e^{S_N f(x_{A_1})}}_{= Z_N^{(1)}(A_2)} \\
&= \cdots = \sum_{A_j \in \mathscr{A}_N^{(j)}(A')} Z_N^{(j-1)}(A_j)^{a_{j-1}} = Z_N^{(j)}(A').
\end{aligned}
$$

Thus, we get

$$
\begin{aligned}
\text{(II)} &= \sum_{A \in \mathscr{A}_N^{(j+1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(j+1)}(\widetilde{\pi}_{j+2} A)^{a_{j+1}-1} \cdot Z_N^{(j)}(A)^{a_j} \log\,(Z_N^{(j)}(A)) \\
&= W_N^{(j)}.
\end{aligned}
$$

This completes the proof of the first assertion.

Next, let $2 \le i \le r$. For any $A \in \mathscr{A}_N^{(i)}$,

$$
\begin{aligned}
\sigma_N^{(i)}(A) &= \frac{1}{Z_n} \sum_{\substack{B \in \mathscr{A}_N^{(1)}, \\ \pi^{(i)}(x_B) \in A}} Z_N^{(r-1)}(\widetilde{\pi}_r B)^{a_{r-1}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 B)^{a_1-1} e^{S_N f(x_B)} \\
&= \frac{1}{Z_n} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(i-1)}(\widetilde{\pi}_i A)^{a_{i-1}-1} \\
&\quad \times \sum_{B \in \mathscr{A}_N^{(1)}(A)} Z_N^{(i-2)}(\widetilde{\pi}_{i-1} B)^{a_{i-2}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 B)^{a_1-1} e^{S_N f(x_B)}.
\end{aligned}
$$

As in the evaluation of term (II)$'$, we have

$$
\sum_{B \in \mathscr{A}_N^{(1)}(A)} Z_N^{(i-2)}(\widetilde{\pi}_{i-1} B)^{a_{i-2}-1} \cdots Z_N^{(1)}(\widetilde{\pi}_2 B)^{a_1-1} e^{S_N f(x_B)} = Z_N^{(i-1)}(A)^{a_{i-1}}.
$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Hence,

$$\sigma_N^{(i)}(A) = \frac{1}{Z_n} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1} \cdots Z_N^{(i)}(\widetilde{\pi}_{i+1} A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}}.$$

Therefore,

$$\begin{aligned}
H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) &= -\sum_{A\in\mathscr{A}_N^{(i)}} \sigma_N^{(i)}(A)\log\sigma_N^{(i)}(A) \\
&= \log Z_N - \frac{1}{Z_n}\sum_{A\in\mathscr{A}_N^{(i)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1}\cdots Z_N^{(i)}(\widetilde{\pi}_{i+1}A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}} \\
&\quad\times \log\,(Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1}\cdots Z_N^{(i)}(\widetilde{\pi}_{i+1}A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}}) \\
&= \log Z_N - \frac{a_{i-1}}{Z_n}\sum_{A\in\mathscr{A}_N^{(i)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1}\cdots Z_N^{(i)}(\widetilde{\pi}_{i+1}A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}}\log\,(Z_N^{(i-1)}(A)) \\
&\quad -\sum_{j=i}^{r-1}\frac{a_j-1}{Z_n}\underbrace{\sum_{A\in\mathscr{A}_N^{(i)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1}\cdots Z_N^{(i)}(\widetilde{\pi}_{i+1}A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}}\log(Z_N^{(j)}(\widetilde{\pi}_{j+1}A))}_{\text{(III)}}.
\end{aligned}$$

Note that we can calculate term (III) as

$$\begin{aligned}
&\sum_{A\in\mathscr{A}_N^{(i)}} Z_N^{(r-1)}(\widetilde{\pi}_r A)^{a_{r-1}-1}\cdots Z_N^{(i)}(\widetilde{\pi}_{i+1}A)^{a_i-1} Z_N^{(i-1)}(A)^{a_{i-1}}\log\,(Z_N^{(j)}(\widetilde{\pi}_{j+1}A)) \\
&= \sum_{A_{j+1}\in\mathscr{A}_N^{(j+1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A_{j+1})^{a_{r-1}-1}\cdots Z_N^{(j+1)}(\widetilde{\pi}_{j+2}A_{j+1})^{a_j-1} Z_N^{(j)}(A_{j+1})^{a_{j-1}-1}\log(Z_N^{(j)}(A_{j+1})) \\
&\quad\times\sum_{A_j\in\mathscr{A}_N^{(j)}(A_{j+1})} Z_N^{(j-1)}(A_j)^{a_{j-2}-1}\cdots\sum_{A_{i+1}\in\mathscr{A}_N^{(i+1)}(A_{i+2})} Z_N^{(i)}(A_{i+1})^{a_{i+1}-1}\underbrace{\sum_{A_i\in\mathscr{A}_N^{(i)}(A_{i+1})} Z_N^{(i-1)}(A_i)^{a_{i-1}}}_{=Z_N^{(i)}(A_{i+1})} \\
&=\cdots=\sum_{A_{j+1}\in\mathscr{A}_N^{(j+1)}} Z_N^{(r-1)}(\widetilde{\pi}_r A_{j+1})^{a_{r-1}-1} \\
&\quad\times\cdots\times Z_N^{(j+1)}(\widetilde{\pi}_{j+2}A_{j+1})^{a_j-1} Z_N^{(j)}(A_{j+1})^{a_{j-1}}\log\,(Z_N^{(j)}(A_{j+1})).
\end{aligned}$$

We conclude that

$$H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) = \log Z_N - \frac{a_{i-1}}{Z_n}W_N^{(i-1)} - \sum_{j=i}^{r-1}\frac{a_j-1}{Z_n}W_N^{(j)}.$$

This completes the proof of the claim. □

By this claim,

$$\sum_{i=1}^{r} w_i H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) + w_1\int_{X_1} S_N f\,d\mu = \log Z_N - \sum_{i=2}^{r}\frac{w_i a_{i-1}}{Z_n}W_N^{(i-1)} - \sum_{i=1}^{r-1}\sum_{j=i}^{r-1}\frac{w_i(a_j-1)}{Z_n}W_N^{(j)}.$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

However, we have

$$\sum_{i=2}^{r} w_i a_{i-1} W_N^{(i-1)} + \sum_{i=1}^{r-1} \sum_{j=i}^{r-1} w_i (a_j - 1) W_N^{(j)} = 0.$$

Indeed, the coefficient of $W_N^{(k)}$ $(1 \leq k \leq r-1)$ is

$$\begin{aligned} w_{k+1} a_k + (a_k - 1) \sum_{i=1}^{k} w_i &= w_{k+1} a_k + (a_k - 1) a_k a_{k+1} \cdots a_{r-1} \\ &= a_k \{ w_{k+1} - (1 - a_k) a_{k+1} a_{k+2} \cdots a_{r-1} \} = 0. \end{aligned}$$

Thus, we have

$$\sum_{i=1}^{r} w_i H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) + w_1 \int_{X_1} S_N f \, d\mu = \log Z_N. \tag{4.1}$$

Let $\mu^{(i)} = \pi^{(i-1)}{}_* \mu$ and $\mu_N^{(i)} = \pi^{(i-1)}{}_* \mu_N$.

LEMMA 4.3. *Let N and M be natural numbers. For any $1 \leq i \leq r$,*

$$\frac{1}{M} H_{\mu_N^{(i)}}(\mathscr{A}_M^{(i)}) \geq \frac{1}{N} H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) - \frac{2M \log |\mathscr{A}^{(i)}|}{N}.$$

Here, $|\mathscr{A}^{(i)}|$ is the number of elements in $\mathscr{A}^{(i)}$.

Suppose this is true, and let N and M be natural numbers. Together with equation (4.1), we obtain the following evaluation:

$$\begin{aligned} \sum_{i=1}^{r} \frac{w_i}{M} H_{\mu_N^{(i)}}(\mathscr{A}_M^{(i)}) + w_1 \int_{X_1} f \, d\mu_N &\geq \sum_{i=1}^{r} \frac{w_i}{N} H_{\sigma_N^{(i)}}(\mathscr{A}_N^{(i)}) \\ &\quad - \sum_{i=1}^{r} \frac{2M \log |\mathscr{A}^{(i)}|}{N} + \frac{w_1}{N} \int_{X_1} S_N f \, d\sigma_N \\ &= \frac{\log Z_N}{N} - \sum_{i=1}^{r} \frac{2M \log |\mathscr{A}^{(i)}|}{N}. \end{aligned}$$

Let $N = N_k \to \infty$ along the sub-sequence (N_k) for which $\mu_{N_k} \rightharpoonup \mu$. This yields

$$\sum_{i=1}^{r} \frac{w_i}{M} H_{\mu^{(i)}}(\mathscr{A}_M^{(i)}) + w_1 \int_{X_1} f \, d\mu \geq \varlimsup_{N \to \infty} \frac{\log Z_N}{N}.$$

We let $M \to \infty$ and get

$$\sum_{i=1}^{r} w_i h_{\mu^{(i)}}(T_i, \mathscr{A}^{(i)}) + w_1 \int_{X_1} f \, d\mu \geq \varlimsup_{N \to \infty} \frac{\log Z_N}{N}.$$

Hence,

$$P_{\mathrm{var}}^{\boldsymbol{a}}(f) \geq P^{\boldsymbol{a}}(f).$$

We are left to prove Lemma 4.3.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Proof of Lemma 4.3. This statement appears in the proof of variational principle in [**Wal82**, Theorem 8.6], and Tsukamoto also proves it in [**Tsu22**, Claim 6.3]. The following proof is taken from the latter. We will explain for $i = 1$; the same argument works for all i.

Let $\mathscr{A} = \mathscr{A}^{(1)}$. Recall that $\mu_N = (1/N) \sum_{k=0}^{N-1} {T_1}^k{}_*\sigma_N$. Since the entropy function is concave (Lemma 3.5), we have

$$H_{\mu_N}(\mathscr{A}_M) \geq \frac{1}{N} \sum_{k=0}^{N-1} H_{{T_1}^k{}_*\sigma_N}(\mathscr{A}_M) = \frac{1}{N} \sum_{k=0}^{N-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}_M).$$

Let $N = qM + r$ with $0 \leq r < M$, then

$$\begin{aligned}\sum_{k=0}^{N-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}_M) &= \sum_{s=0}^{q} \sum_{t=0}^{M-1} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) - \sum_{k=N}^{qM+M-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}_M) \\ &\geq \sum_{t=0}^{M-1} \sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) - M \log |\mathscr{A}_M| \\ &\geq \sum_{t=0}^{M-1} \sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) - M^2 \log |\mathscr{A}|. \qquad (4.2)\end{aligned}$$

We will evaluate $\sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M)$ from below for each $0 \leq t \leq M-1$. First, observe that

$$T_1^{-sM-t}\mathscr{A}_M = \bigvee_{j=0}^{M-1} T_1^{-sM-t-j}\mathscr{A}.$$

We have

$$\{sM + t + j | 0 \leq s \leq q, 0 \leq j \leq M-1\} = \{t, t+1, \ldots, t + qM + M - 1\}$$

without multiplicity. Therefore,

$$\begin{aligned}H_{\sigma_N}(\mathscr{A}_N) &\leq H_{\sigma_N}\bigg(\bigvee_{k=0}^{t+(q+1)M-1} T_1^{-k}\mathscr{A}\bigg) \quad \text{by } N < t + (q+1)M \\ &\leq \sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) + \sum_{k=0}^{t-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}) \quad \text{by Lemma 3.5.}\end{aligned}$$

This implies

$$\begin{aligned}\sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) &\geq H_{\sigma_N}(\mathscr{A}_N) - \sum_{k=0}^{t-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}) \\ &\geq H_{\sigma_N}(\mathscr{A}_N) - M \log |\mathscr{A}| \quad \text{by } t < M.\end{aligned}$$

Now, we sum over t and obtain

$$\sum_{t=1}^{M-1} \sum_{s=0}^{q} H_{\sigma_N}(T_1^{-sM-t}\mathscr{A}_M) \geq M H_{\sigma_N}(\mathscr{A}_N) - M^2 \log |\mathscr{A}|.$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Combining with equation (4.2), this implies

$$\sum_{k=0}^{N-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}_M) \geq M H_{\sigma_N}(\mathscr{A}_N) - 2M^2 \log|\mathscr{A}|.$$

It follows that

$$\frac{1}{M} H_{\mu_N}(\mathscr{A}_N) \geq \frac{1}{MN} \sum_{k=0}^{N-1} H_{\sigma_N}(T_1^{-k}\mathscr{A}_M) \geq \frac{1}{N} H_{\sigma_N}(\mathscr{A}_N) - \frac{2M \log|\mathscr{A}|}{N}. \qquad \square$$

This completes the proof of Theorem 4.1. □

5. *Proof of* $P^{\boldsymbol{a}}_{\mathrm{var}}(f) \leq P^{\boldsymbol{a}}(f)$

It seems difficult to implement the zero-dimensional trick to prove $P^{\boldsymbol{a}}_{\mathrm{var}}(f) \leq P^{\boldsymbol{a}}(f)$. Hence, the proof is more complicated.

THEOREM 5.1. *Suppose that* (X_i, T_i) $(i = 1, 2, \ldots, r)$ *are dynamical systems and* $\pi_i : X_i \to X_{i+1}$ $(i = 1, 2, \ldots, r-1)$ *are factor maps. Then we have*

$$P^{\boldsymbol{a}}_{\mathrm{var}}(f) \leq P^{\boldsymbol{a}}(f)$$

for any continuous function $f : X_1 \to \mathbb{R}$.

Proof. Take and fix $\mu \in \mathscr{M}^{T_1}(X_1)$. Let $\mu_i = \pi^{(i-1)}{}_*\mu$. We need to prove

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}).$$

However, the following argument assures that giving an evaluation up to a constant is sufficient: suppose there is a positive number C which does not depend on f nor $(T_i)_i$ satisfying

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}) + C. \tag{5.1}$$

Applying this to $S_m f$ and $\boldsymbol{T}^m = ({T_i}^m)_i$ for $m \in \mathbb{N}$ yields

$$\sum_{i=1}^{r} w_i h_{\mu_i}({T_i}^m) + w_1 \int_{X_1} S_m f \, d\mu \leq P^{\boldsymbol{a}}(S_m f, \boldsymbol{T}^m) + C.$$

We employ Lemma 3.3 and get

$$m \sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + m w_1 \int_{X_1} f \, d\mu \leq m P^{\boldsymbol{a}}(f, \boldsymbol{T}) + C.$$

Divide by m and let $m \to \infty$. We obtain the desired inequality

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}).$$

Therefore, we only need to prove inequality (5.1).

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Let $\mathscr{A}^{(i)} = \{A_1^{(i)}, A_2^{(i)}, \ldots, A_{m_i}^{(i)}\}$ be an arbitrary partition of X_i for each i. We will prove

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i, \mathscr{A}^{(i)}) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}) + C.$$

We start by approximating elements of $\mathscr{A}^{(i)}$ with compact sets using backward induction. For $1 \leq i \leq r$, let

$$\Lambda_i^0 = \{0, 1, \ldots, m_r\} \times \{0, 1, \ldots, m_{r-1}\} \times \cdots \times \{0, 1, \ldots, m_{i+1}\} \times \{0, 1, \ldots, m_i\},$$
$$\Lambda_i = \{0, 1, \ldots, m_r\} \times \{0, 1, \ldots, m_{r-1}\} \times \cdots \times \{0, 1, \ldots, m_{i+1}\} \times \{1, 2, \ldots, m_i\}.$$

We will denote an element $(j_r, j_{r-1}, \ldots, j_i)$ in Λ_i^0 or Λ_i by $j_r j_{r-1} \cdots j_i$. For each $A_j^{(r)} \in \mathscr{A}^{(r)}$, take a compact set $C_j^{(r)} \subset A_j^{(r)}$ such that

$$\log m_r \cdot \sum_{j=1}^{m_r} \mu_r\big(A_j^{(r)} \setminus C_j^{(r)}\big) < 1.$$

Define $C_0^{(r)}$ as the remainder of X_r, which may not be compact:

$$C_0^{(r)} = \bigcup_{j=1}^{m_r} A_j^{(r)} \setminus C_j^{(r)} = X_r \setminus \bigcup_{j=1}^{m_r} C_j^{(r)}.$$

Then $\mathscr{C}^{(r)} := \{C_0^{(r)}, C_1^{(r)}, \ldots, C_{m_r}^{(r)}\}$ is a measurable partition of X_r.

Next, consider the partition $\pi_{r-1}^{-1}(\mathscr{C}^{(r)}) \vee \mathscr{A}^{(r-1)}$ of X_{r-1}. For $j_r j_{r-1} \in \Lambda_{r-1}$, let

$$B_{j_r j_{r-1}}^{(r-1)} = \pi_{r-1}^{-1}(C_{j_r}^{(r)}) \cap A_{j_{r-1}}^{(r-1)}.$$

Then

$$\pi_{r-1}^{-1}(\mathscr{C}^{(r)}) \vee \mathscr{A}^{(r-1)} = \{B_{j_r j_{r-1}}^{(r-1)} \mid j_r j_{r-1} \in \Lambda_{r-1}\},$$

and for each $j_r \in \Lambda_r^0$,

$$\bigcup_{j_{r-1}=1}^{m_{r-1}} B_{j_r j_{r-1}}^{(r-1)} = \pi_{r-1}^{-1}(C_{j_r}^{(r-1)}).$$

For each $j_r j_{r-1} \in \Lambda_{r-1}$, take a compact set $C_{j_r j_{r-1}}^{(r-1)} \subset B_{j_r j_{r-1}}^{(r-1)}$ (which could be empty) such that

$$\log |\Lambda_{r-1}| \cdot \sum_{j_r=0}^{m_r} \sum_{j_{r-1}=1}^{m_{r-1}} \mu_{r-1}\big(B_{j_r j_{r-1}}^{(r-1)} \setminus C_{j_r j_{r-1}}^{(r-1)}\big) < 1.$$

Define $C_{j_r 0}^{(r-1)}$ as the remainder of $\pi_{r-1}^{-1}(C_{j_r}^{(r)})$:

$$C_{j_r 0}^{(r-1)} = \pi_{r-1}^{-1}(C_{j_r}^{(r)}) \setminus \bigcup_{j_{r-1}=1}^{m_{r-1}} C_{j_r j_{r-1}}^{(r-1)}.$$

Then $\mathscr{C}^{(r-1)} = \{C_{j_r j_{r-1}}^{(r-1)} \mid j_r j_{r-1} \in \Lambda_{r-1}^0\}$ is a measurable partition of X_{r-1}.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Continue in this manner, and suppose we have obtained the partition $\mathscr{C}^{(k)} = \{C_J^{(k)} | J \in \Lambda_k^0\}$ of X_k for $k = i+1, i+2, \ldots, r$. We will define $\mathscr{C}^{(i)}$. Each element in $\pi_i^{-1}(\mathscr{C}^{(i+1)}) \vee \mathscr{A}^{(i)}$ can be expressed using $J' \in \Lambda_{i+1}^0$ and $j_i \in \{1, 2, \ldots, m_i\}$ by

$$B_{J'j_i}^{(i)} = \pi_i^{-1}(C_{J'}^{(i+1)}) \cap A_{j_i}^{(i)}.$$

Choose a compact set $C_J^{(i)} \subset B_J^{(i)}$ for each $J \in \Lambda_i$ so that

$$\log|\Lambda_i| \cdot \sum_{J' \in \Lambda_{i+1}^0} \sum_{j_i=1}^{m_i} \mu_i\big(B_{J'j_i}^{(i)} \setminus C_{J'j_i}^{(i)}\big) < 1.$$

Finally, for $J' \in \Lambda_{j+1}^0$, let

$$C_{J'0}^{(i)} = \pi_i^{-1}(C_{J'}^{(i+1)}) \setminus \bigcup_{j_i=1}^{m_i} C_{J'j_i}^{(i)}.$$

Set $\mathscr{C}^{(i)} = \{C_J^{(i)} | J \in \Lambda_i^0\}$. This is a partition of X_i.

LEMMA 5.2. *For $\mathscr{C}^{(i)}$ constructed above, we have*

$$h_{\mu_i}(T_i, \mathscr{A}^{(i)}) \le h_{\mu_i}(T_i, \mathscr{C}^{(i)}) + 1.$$

Proof. By Lemma 3.5,

$$\begin{aligned} h_{\mu_i}(T_i, \mathscr{A}^{(i)}) &\le h_{\mu_i}(T_i, \mathscr{A}^{(i)} \vee \pi_i^{-1}(\mathscr{C}^{(i+1)})) \\ &\le h_{\mu_i}(T_i, \mathscr{C}^{(i)}) + H_{\mu_i}(\mathscr{A}^{(i)} \vee \pi_i^{-1}(\mathscr{C}^{(i+1)}) | \mathscr{C}^{(i)}). \end{aligned}$$

Since $C_J^{(i)} \subset B_J^{(i)}$ for $J \in \Lambda_i$,

$$\begin{aligned} &H_{\mu_i}(\mathscr{A}^{(i)} \vee \pi_i^{-1}(\mathscr{C}^{(i+1)}) | \mathscr{C}^{(i)}) \\ &= -\sum_{\substack{J \in \Lambda_i^0 \\ \mu_i(C_J^{(i)}) \neq 0}} \mu_i(C_J^{(i)}) \sum_{K \in \Lambda_i} \frac{\mu_i(B_K^{(i)} \cap C_J^{(i)})}{\mu_i(C_J^{(i)})} \log\left(\frac{\mu_i(B_K^{(i)} \cap C_J^{(i)})}{\mu_i(C_J^{(i)})}\right) \\ &= -\sum_{\substack{J' \in \Lambda_{i+1}^0 \\ \mu_i(C_{J'0}^{(i)}) \neq 0}} \mu_i(C_{J'0}^{(i)}) \sum_{j_i=1}^{m_i} \frac{\mu_i(B_{J'j_i}^{(i)} \cap C_{J'0}^{(i)})}{\mu_i(C_{J'0}^{(i)})} \log\left(\frac{\mu_i(B_{J'j_i}^{(i)} \cap C_{J'0}^{(i)})}{\mu_i(C_{J'0}^{(i)})}\right). \end{aligned}$$

By Lemma 3.4, we have

$$-\sum_{j_i=1}^{m_i} \frac{\mu_i(B_{J'j_i}^{(i)} \cap C_{J'0}^{(i)})}{\mu_i(C_{J'0}^{(i)})} \log\left(\frac{\mu_i(B_{J'j_i}^{(i)} \cap C_{J'0}^{(i)})}{\mu_i(C_{J'0}^{(i)})}\right) \le \log|\Lambda_i|.$$

Therefore,

$$H_{\mu_i}(\mathscr{A}^{(i)} \vee \pi_i^{-1}(\mathscr{C}^{(i+1)}) | \mathscr{C}^{(i)}) \le \log|\Lambda_i| \sum_{J' \in \Lambda_{i+1}^0} \mu_i\left(\pi_i^{-1}(C_{J'}^{(i+1)}) \setminus \bigcup_{j_i=1}^{m_i} C_{J'j_i}^{(i)}\right) < 1. \qquad \square$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Recall the definition of $\boldsymbol{w}$ in equation (2.1). We have

$$\begin{aligned}\sum_{i=1}^{r} w_i h_{\mu_i}(T_i, \mathscr{C}^{(i)}) + w_1 \int_{X_1} f\, d\mu &= \lim_{N\to\infty} \frac{1}{N}\bigg\{ H_{\mu_r}(\mathscr{C}_N^{(r)}) + a_1 a_2 \cdots a_{r-1} N \int_{X_1} f\, d\mu \\ &\quad + \sum_{i=1}^{r-1} a_i a_{i+1} \cdots a_{r-1}\Big(H_{\mu_i}(\mathscr{C}_N^{(i)}) - H_{\mu_{i+1}}(\mathscr{C}_N^{(i+1)})\Big)\bigg\} \\ &= \lim_{N\to\infty} \frac{1}{N}\bigg\{ H_{\mu_r}(\mathscr{C}_N^{(r)}) + a_1 a_2 \cdots a_{r-1} \int_{X_1} S_N f\, d\mu \\ &\quad + \sum_{i=1}^{r-1} a_i a_{i+1} \cdots a_{r-1} H_{\mu_i}(\mathscr{C}_N^{(i)} | \pi_i^{-1}(\mathscr{C}_N^{(i+1)}))\bigg\}.\end{aligned}$$

Here, we used the relation

$$\begin{aligned} H_{\mu_i}(\mathscr{C}_N^{(i)}) - H_{\mu_{i+1}}(\mathscr{C}_N^{(i+1)}) &= H_{\mu_i}(\mathscr{C}_N^{(i)}) - H_{\mu_i}(\pi_i^{-1}(\mathscr{C}_N^{(i+1)})) \\ &= H_{\mu_i}(\mathscr{C}_N^{(i)} | \pi_i^{-1}(\mathscr{C}_N^{(i+1)})).\end{aligned}$$

We fix N and evaluate from above the following terms using backward induction:

$$H_{\mu_r}(\mathscr{C}_N^{(r)}) + a_1 a_2 \cdots a_{r-1} \int_{X_1} S_N f\, d\mu + \sum_{i=1}^{r-1} a_i a_{i+1} \cdots a_{r-1} H_{\mu_i}(\mathscr{C}_N^{(i)} | \pi_i^{-1}(\mathscr{C}_N^{(i+1)})). \tag{5.2}$$

First, consider the term

$$a_1 a_2 \cdots a_{r-1}\bigg(H_\mu(\mathscr{C}_N^{(1)} | \pi_1^{-1}(\mathscr{C}_N^{(2)})) + \int_{X_1} S_N f\, d\mu\bigg).$$

For $C \in \mathscr{C}_N^{(i+1)}$, let $\mathscr{C}_N^{(i)}(C) = \{D \in \mathscr{C}_N^{(i)} | \pi_i(D) \subset C\}$, then by Lemma 3.4,

$$\begin{aligned} H_\mu(\mathscr{C}_N^{(1)} | \pi_1^{-1}(\mathscr{C}_N^{(2)})) + \int_{X_1} S_N f\, d\mu &\le \sum_{\substack{C \in \mathscr{C}_N^{(2)} \\ \mu_2(C) \neq 0}} \mu_2(C) \bigg\{ \sum_{D \in \mathscr{C}_N^{(1)}(C)} \bigg(-\frac{\mu(D)}{\mu_2(C)} \log \frac{\mu(D)}{\mu_2(C)} + \frac{\mu(D)}{\mu_2(C)} \sup_D S_N f\bigg)\bigg\} \\ &\le \sum_{C \in \mathscr{C}_N^{(2)}} \mu_2(C) \log \sum_{D \in \mathscr{C}_N^{(1)}(C)} e^{\sup_D S_N f}.\end{aligned}$$

Applying this inequality to equation (5.2), the following term appears:

$$a_2 a_3 \cdots a_{r-1}\bigg(H_{\mu_2}(\mathscr{C}_N^{(2)} | \pi_2^{-1}(\mathscr{C}_N^{(3)})) + a_1 \sum_{C \in \mathscr{C}_N^{(2)}} \mu_2(C) \log \sum_{D \in \mathscr{C}_N^{(1)}(C)} e^{\sup_D S_N f}\bigg). \tag{5.3}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

This can be evaluated similarly using Lemma 3.4 as

$$H_{\mu_2}(\mathscr{C}_N^{(2)}|\pi_2^{-1}(\mathscr{C}_N^{(3)})) + a_1 \sum_{C\in\mathscr{C}_N^{(2)}} \mu_2(C) \log \sum_{D\in\mathscr{C}_N^{(1)}(C)} e^{\sup_D S_N f}$$
$$= \sum_{\substack{C\in\mathscr{C}_N^{(3)}\\ \mu_3(C)\neq 0}} \mu_3(C)\bigg\{ \sum_{D\in\mathscr{C}_N^{(2)}(C)} \bigg(-\frac{\mu_2(D)}{\mu_3(C)} \log \frac{\mu_2(D)}{\mu_3(C)} + \frac{\mu_2(D)}{\mu_3(C)} \log\bigg(\sum_{E\in\mathscr{C}_N^{(1)}(D)} e^{\sup_E S_N f}\bigg)^{a_1}\bigg)\bigg\}$$
$$\leq \sum_{C\in\mathscr{C}_N^{(3)}} \mu_3(C) \log \sum_{D\in\mathscr{C}_N^{(2)}(C)} \bigg(\sum_{E\in\mathscr{C}_N^{(1)}(D)} e^{\sup_E S_N f}\bigg)^{a_1}.$$

Continue likewise and obtain the following upper bound for equation (5.2):

$$\log \sum_{C^{(r)}\in\mathscr{C}_N^{(r)}} \bigg(\sum_{C^{(r-1)}\in\mathscr{C}_N^{(r-1)}(C^{(r)})} \bigg(\cdots \bigg(\sum_{C^{(1)}\in\mathscr{C}_N^{(1)}(C^{(2)})} e^{\sup_{C^{(1)}} S_N f}\bigg)^{a_1} \cdots \bigg)^{a_{r-2}}\bigg)^{a_{r-1}}. \tag{5.4}$$

For $1 \leq i \leq r$, let $\mathscr{C}_c^{(i)} = \{C \in \mathscr{C}^{(i)} | C \text{ is compact}\}$. There is a positive number ε_i such that $d^{(i)}(y_1, y_2) > \varepsilon_i$ for any $C_1, C_2 \in \mathscr{C}_c^{(i)}$ and $y_1 \in C_1$, $y_2 \in C_2$. Fix a positive number ε with

$$\varepsilon < \min_{1\leq i\leq r} \varepsilon_i. \tag{5.5}$$

Let $\mathscr{F}^{(i)}$ be a chain of open (N, ε)-covers of X_i (see Definition 3.1). Consider

$$\log \mathscr{P}^{\boldsymbol{a}}(f, N, \varepsilon, (\mathscr{F}^{(i)})_i)$$
$$= \log \sum_{U^{(r)}\in\mathscr{F}^{(r)}} \bigg(\sum_{U^{(r-1)}\in\mathscr{F}^{(r-1)}(U^{(r)})} \bigg(\cdots \bigg(\sum_{U^{(1)}\in\mathscr{F}^{(1)}(U^{(2)})} e^{\sup_{U^{(1)}} S_N f}\bigg)^{a_1} \cdots \bigg)^{a_{r-2}}\bigg)^{a_{r-1}}. \tag{5.6}$$

We will evaluate equation (5.4) from above by equation (5.6) up to a constant. We need the next lemma.

LEMMA 5.3

(1) *For any* $V \subset X_r$ *with* $\operatorname{diam}(V, d_N^{(r)}) < \varepsilon$,

$$|\{D \in \mathscr{C}_N^{(r)} | D \cap V \neq \varnothing\}| \leq 2^N.$$

(2) *Let* $1 \leq i \leq r-1$ *and* $C \in \mathscr{C}_N^{(i+1)}$. *For any* $V \subset X_i$ *with* $\operatorname{diam}(V, d_N^{(i)}) < \varepsilon$,

$$|\{D \in \mathscr{C}_N^{(i)}(C) | D \cap V \neq \varnothing\}| \leq 2^N.$$

Proof. (1) $D \in \mathscr{C}_N^{(r)}$ can be expressed using $C_{k_s}^{(r)} \in \mathscr{C}^{(r)}$ $(s = 0, 1, \ldots, N-1)$ as

$$D = C_{k_0}^{(r)} \cap T_r^{-1} C_{k_1}^{(r)} \cap T_r^{-2} C_{k_2}^{(r)} \cap \cdots \cap T_r^{-N+1} C_{k_{N-1}}^{(r)}.$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

If $D \cap V \neq \varnothing$, we have $T_r^{-s}(C_{k_s}^{(r)}) \cap V \neq \varnothing$ for every $0 \leq s \leq N-1$. Then for each s,

$$\varnothing \neq T_r^s\Big(T_r^{-s}(C_{k_s}^{(r)}) \cap V\Big) \subset C_{k_s}^{(r)} \cap T_r^s(V).$$

By equation (5.5), each k_s is either 0 or one of the elements in $\{1, 2, \ldots, m_r\}$. Therefore, there are at most 2^N such sets.

(2) The proof works in the same way as in item (1). C can be written using $J_k \in \Lambda_{i+1}^0$ $(k = 0, 1, \ldots, N-1)$ as

$$C = C_{J_0}^{(i+1)} \cap T_{i+1}^{-1} C_{J_1}^{(i+1)} \cap T_{i+1}^{-2} C_{J_2}^{(i+1)} \cap \cdots \cap T_{i+1}^{-N+1} C_{J_{N-1}}^{(i+1)}.$$

Then any $D \in \mathscr{C}_N^{(i)}(C)$ is of the form

$$D = C_{J_0 k_0}^{(i)} \cap T_i^{-1} C_{J_1 k_1}^{(i)} \cap T_i^{-2} C_{J_2 k_2}^{(i)} \cap \cdots \cap T_i^{-N+1} C_{J_{N-1} k_{N-1}}^{(i)}$$

with $0 \leq k_l \leq m_i$ $(l = 1, 2, \ldots, N-1)$. If $D \cap V \neq \varnothing$, then each k_l is either 0 or one of the elements in $\{1, 2, \ldots, m_i\}$. Therefore, there are at most 2^N such sets. □

For any $C^{(1)} \in \mathscr{C}_N^{(1)}$, there is $V \in \mathscr{F}^{(1)}$ with $V \cap C^{(1)} \neq \varnothing$ and

$$e^{\sup_{C^{(1)}} S_N f} \leq e^{\sup_V S_N f}.$$

Let $C^{(2)} \in \mathscr{C}_N^{(2)}$, then by Lemma 5.3,

$$\sum_{C^{(1)} \in \mathscr{C}_N^{(1)}(C^{(2)})} e^{\sup_{C^{(1)}} S_N f} \leq \sum_{\substack{U \in \mathscr{F}^{(2)} \\ U \cap C^{(2)} \neq \varnothing}} 2^N \sum_{V \in \mathscr{F}^{(1)}(U)} e^{\sup_V S_N f}.$$

By Lemma 3.4,

$$\bigg(\sum_{C^{(1)} \in \mathscr{C}_N^{(1)}(C^{(2)})} e^{\sup_{C^{(1)}} S_N f}\bigg)^{a_1} \leq 2^{a_1 N} \sum_{\substack{U \in \mathscr{F}^{(2)} \\ U \cap C^{(2)} \neq \varnothing}} \bigg(\sum_{V \in \mathscr{F}^{(1)}(U)} e^{\sup_V S_N f}\bigg)^{a_1}.$$

For $C^{(3)} \in \mathscr{C}_N^{(3)}$, we apply Lemmas 5.3 and 3.4 similarly and obtain

$$\begin{aligned} &\bigg(\sum_{C^{(2)} \in \mathscr{C}_N^{(2)}(C^{(3)})} \bigg(\sum_{C^{(1)} \in \mathscr{C}_N^{(1)}(C^{(2)})} e^{\sup_{C^{(1)}} S_N f}\bigg)^{a_1}\bigg)^{a_2} \\ &\quad \leq 2^{a_1 a_2 N} 2^{a_2 N} \sum_{\substack{O \in \mathscr{F}^{(3)} \\ O \cap C^{(3)} \neq \varnothing}} \bigg(\sum_{U \in \mathscr{F}^{(2)}(O)} \bigg(\sum_{V \in \mathscr{F}^{(1)}(U)} e^{\sup_V S_N f}\bigg)^{a_1}\bigg)^{a_2}. \end{aligned}$$

We continue this reasoning and get

$$\begin{aligned} &\sum_{C^{(r)} \in \mathscr{C}_N^{(r)}} \bigg(\sum_{C^{(r-1)} \in \mathscr{C}_N^{(r-1)}(C^{(r)})} \bigg(\cdots \bigg(\sum_{C^{(1)} \in \mathscr{C}_N^{(1)}(C^{(2)})} e^{\sup_{C^{(1)}} S_N f}\bigg)^{a_1} \cdots\bigg)^{a_{r-2}}\bigg)^{a_{r-1}} \\ &\quad \leq 2^{\alpha N} \sum_{U^{(r)} \in \mathscr{F}^{(r)}} \bigg(\sum_{U^{(r-1)} \in \mathscr{F}^{(r-1)}(U^{(r)})} \bigg(\cdots \bigg(\sum_{U^{(1)} \in \mathscr{F}^{(1)}(U^{(2)})} e^{\sup_{U^{(1)}} S_N f}\bigg)^{a_1} \cdots\bigg)^{a_{r-2}}\bigg)^{a_{r-1}}. \end{aligned}$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Here, α stands for $\sum_{i=1}^{r-1} a_i a_{i+1} \cdots a_{r-1}$. We take the logarithm of both sides; the left-hand side equals equation (5.4), which is an upper bound for equation (5.2). Furthermore, consider the infimum over the chain of open (N, ε)-covers $(\mathscr{F}^{(i)})_i$ on the right-hand side. By Remark 3.2, this yields

$$H_{\mu_r}(\mathscr{C}_N^{(r)}) + a_1 a_2 \cdots a_{r-1} \int_{X_1} S_N f \, d\mu + \sum_{i=1}^{r-1} a_i a_{i+1} \cdots a_{r-1} H_{\mu_i}(\mathscr{C}_N^{(i)} | \pi_i^{-1}(\mathscr{C}_N^{(i+1)}))$$
$$\leq \log P_r^{\boldsymbol{a}}(X_r, f, N, \varepsilon) + \alpha N \log 2.$$

Divide by N, then let $N \to \infty$ and $\varepsilon \to 0$. We obtain

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i, \mathscr{C}^{(i)}) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}) + \alpha \log 2.$$

Lemma 5.2 yields

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i, \mathscr{A}^{(i)}) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}) + \alpha \log 2 + r.$$

We take the supremum over the partitions $(\mathscr{A}^{(i)})_i$:

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}) + \alpha \log 2 + r.$$

By the argument at the beginning of this proof, we conclude that

$$\sum_{i=1}^{r} w_i h_{\mu_i}(T_i) + w_1 \int_{X_1} f \, d\mu \leq P^{\boldsymbol{a}}(f, \boldsymbol{T}).$$ □

6. *Example: sofic sets*

Kenyon and Peres [**KP96-2**] calculated the Hausdorff dimension of sofic sets in $\mathbb{T}^2$. In this section, we will see that we can calculate the Hausdorff dimension of certain sofic sets in $\mathbb{T}^d$ with arbitrary d. We give an example for the case $d = 3$.

6.1. *Definition of sofic sets.*

This subsection is referred to [**KP96-2**]. Weiss [**We82**] defined *sofic systems* as subshifts which are factors of shifts of finite type. Boyle, Kitchens, and Marcus proved in [**BKM85**] that this is equivalent to the following definition.

Definition 6.1. [**KP96-2**, Proposition 3.6] Consider a finite directed graph $G = \langle V, E \rangle$ in which loops and multiple edges are allowed. Suppose its edges are colored in l colors in a 'right-resolving' fashion: every two edges emanating from the same vertex have different colors. Then the set of color sequences that arise from infinite paths in G is called the *sofic system*.

Let $m_1 \leq m_2 \leq \cdots \leq m_r$ be natural numbers, T an endomorphism on $\mathbb{T}^r = \mathbb{R}^r / \mathbb{Z}^r$ represented by the diagonal matrix $A = \mathrm{diag}(m_1, m_2, \ldots, m_r)$, and

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

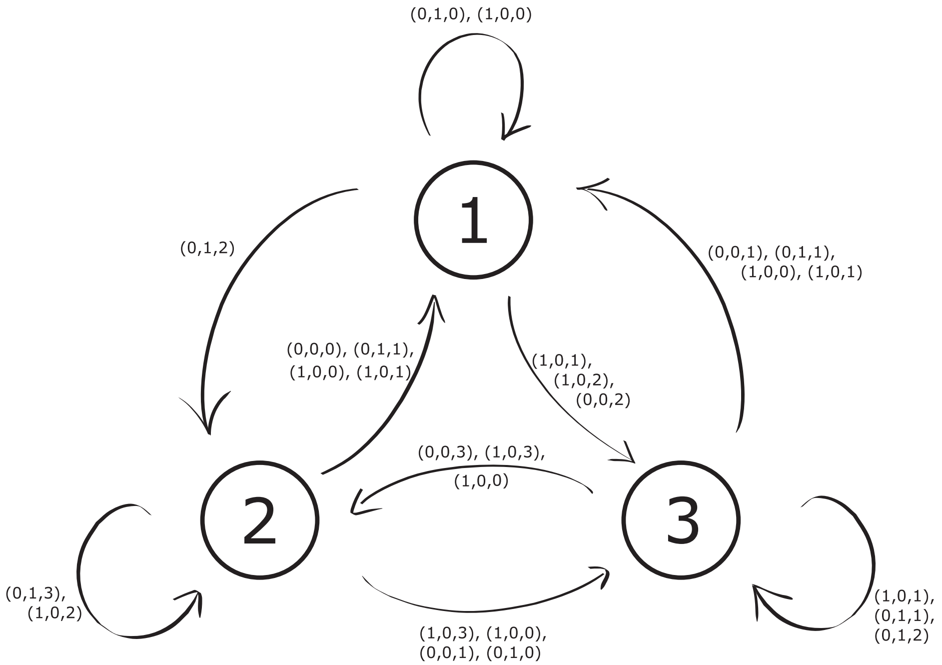


FIGURE 2. Directed graph G.

$D = \prod_{i=1}^{r}\{0, 1, \ldots, m_i - 1\}$. Define a map $R_r : D^{\mathbb{N}} \to \mathbb{T}^r$ by

$$R_r((e^{(n)})_{n=1}^{\infty}) = \left(\sum_{k=0}^{\infty} \frac{e_1^{(k)}}{{m_1}^k}, \ldots, \sum_{k=0}^{\infty} \frac{e_r^{(k)}}{{m_r}^k} \right),$$

where $e^{(k)} = (e_1^{(k)}, \ldots, e_r^{(k)}) \in D$ for each k. Suppose the edges in some finite directed graph are labeled by the elements in D in the right-resolving fashion, and let $S \subset D^{\mathbb{N}}$ be the resulting sofic system. The image of S under R_r is called a *sofic set*.

6.2. *An example of a sofic set.* Here we will look at an example of a sofic set and calculate its Hausdorff dimension via its weighted topological entropy. Let $D = \{0, 1\} \times \{0, 1, 2\} \times \{0, 1, 2, 3\}$, and consider the directed graph $G = \langle V, E \rangle$ with $V = \{1, 2, 3\}$ and D-labeled edges E in Figure 2.

Let $Y_1 \subset D^{\mathbb{N}}$ be the resulting sofic system. Let $C = \{0, 1\} \times \{0, 1, 2\}$ and $B = \{0, 1\}$. Define $p_1 : D \to C$ and $p_2 : C \to B$ by

$$p_1(i, j, k) = (i, j), \quad p_2(i, j) = i.$$

Let $p_1^{\mathbb{N}} : D^{\mathbb{N}} \to C^{\mathbb{N}}$ and $p_2^{\mathbb{N}} : C^{\mathbb{N}} \to B^{\mathbb{N}}$ be the product map of p_1 and p_2, respectively. Set $Y_2 = p_1^{\mathbb{N}}(Y_1)$ and $Y_3 = p_2^{\mathbb{N}}(Y_2)$. Note that $Y_2 = \{(0, 0), (1, 0), (0, 1)\}^{\mathbb{N}}$ and $Y_3 = \{0, 1\}^{\mathbb{N}}$, meaning they are full shifts.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

The sets $X_i = R_i(Y_i)(i = 1, 2, 3)$ are sofic sets. Define $\pi_1 : X_1 \to X_2$ and $\pi_2 : X_2 \to X_3$ by

$$\pi_1(x, y, z) = (x, y), \quad \pi_2(x, y) = x.$$

Furthermore, let T_1, T_2, and T_3 be the endomorphism on X_1, X_2, and X_3 represented by the matrices diag(2, 3, 4), diag(2, 3), and diag(2), respectively. Then $(X_i, T_i)_i$ and $(\pi_i)_i$ form a sequence of dynamical systems.

For a natural number N, denote by $Y_i|_N$ the restriction of Y_i to its first N coordinates, and let $p_{i,N} : Y_i|_N \to Y_{i+1}|_N$ be the projections for $i = 1, 2$. As in Example 1.4, we have for any exponent $\boldsymbol{a} = (a_1, a_2) \in [0, 1]^2$,

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \lim_{N\to\infty} \frac{1}{N} \log \sum_{u\in\{0,1\}^N} \Bigg(\sum_{v\in {p_{2,N}}^{-1}(u)} |{p_{1,N}}^{-1}(v)|^{a_1} \Bigg)^{a_2}.$$

Now, let us evaluate $|{p_{1,N}}^{-1}(v)|$ using matrix products. This idea of using matrix products is due to Kenyon and Peres [**KP96-2**]. Fix $(a, b) \in \{0, 1\}^2$ and let

$$a_{ij} = |\{e \in E | e \text{ is from } j \text{ to } i \text{ and the first two coordinates of its label are } (a, b)\}|.$$

Define a 3×3 matrix by $A_{(a,b)} = (a_{ij})_{ij}$. Then we have

$$A_{(0,0)} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad A_{(0,1)} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}, \quad A_{(1,0)} = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}, \quad A_{(1,1)} = O.$$

Note that ${A_{(0,0)}}^2 = A_{(0,1)}$ and ${A_{(0,0)}}^3 = A_{(1,0)}$. For $v = (v_1, \ldots, v_N) \in Y_2|_N$, we have

$$|{p_{1,N}}^{-1}(v)| \asymp \|A_{v_1} A_{v_2} \cdots A_{v_N}\|.$$

Here, $A \asymp B$ means there is a constant $c > 0$ independent of N with $c^{-1}B \leq A \leq cB$. For $\alpha = (1 + \sqrt{5})/2$, we have $\alpha^2 = \alpha + 1$ and

$$A_{(0,0)} \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix} = \begin{pmatrix} 1+\alpha \\ \alpha \\ 1+\alpha \end{pmatrix} = \alpha \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix}, \quad A_{(0,1)} \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix} = \alpha^2 \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix}, \quad A_{(1,0)} \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix} = \alpha^3 \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix}.$$

Therefore,

$$\|A_{v_1} A_{v_2} \cdots A_{v_N}\| \asymp \left\| A_{v_1} A_{v_2} \cdots A_{v_N} \begin{pmatrix} \alpha \\ 1 \\ \alpha \end{pmatrix} \right\| \asymp \lambda_{v_1} \lambda_{v_2} \cdots \lambda_{v_N},$$

where $\lambda_{(0,0)} = \alpha$, $\lambda_{(0,1)} = \alpha^2$, $\lambda_{(1,0)} = \alpha^3$.

Take $u \in Y_3 = \{0, 1\}^{\mathbb{N}}$ and suppose there are n numbers of zeros in u. Also, if there are k numbers of (0, 0) terms in $v = (v_1, \ldots, v_N) \in {p_{2,N}}^{-1}(u)$, there are $n - k$ numbers of (0, 1) terms and $N - n$ numbers of (1, 0) terms in v. Then,

$${\lambda_{v_1}}^{a_1} \cdots {\lambda_{v_N}}^{a_1} = \alpha^{a_1 k} \alpha^{2a_1(n-k)} \alpha^{3a_1(N-n)}.$$

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

Therefore (recall that $Y_2 = \{(0, 0), (1, 0), (0, 1)\}^{\mathbb{N}}$),

$$\sum_{v\in {p_{2,N}}^{-1}(u)} |{p_{1,N}}^{-1}(v)|^{a_1} \asymp \sum_{(v_1,\ldots,v_N)\in {p_{2,N}}^{-1}(u)} {\lambda_{v_1}}^{a_1}\cdots{\lambda_{v_N}}^{a_1}$$
$$= \sum_{k=0}^{n}\binom{n}{k}\alpha^{a_1 k}\alpha^{2a_1(n-k)}\alpha^{3a_1(N-n)} = (\alpha^{a_1}+\alpha^{2a_1})^n\alpha^{3a_1(N-n)}.$$

This implies

$$\sum_{u\in\{0,1\}^N}\bigg(\sum_{v\in {p_{2,N}}^{-1}(u)} |{p_{1,N}}^{-1}(v)|^{a_1}\bigg)^{a_2} \asymp \sum_{n=0}^{N}\binom{N}{n}(\alpha^{a_1}+\alpha^{2a_1})^{a_2 n}\alpha^{3a_1a_2(N-n)}$$
$$= \{(\alpha^{a_1}+\alpha^{2a_1})^{a_2}+\alpha^{3a_1a_2}\}^N.$$

We conclude that

$$h^{\boldsymbol{a}}(\boldsymbol{T}) = \lim_{N\to\infty}\frac{1}{N}\log\{(\alpha^{a_1}+\alpha^{2a_1})^{a_2}+\alpha^{3a_1a_2}\}^N$$
$$= \log\bigg\{\bigg(\bigg(\frac{1+\sqrt{5}}{2}\bigg)^{a_1}+\bigg(\frac{3+\sqrt{5}}{2}\bigg)^{a_1}\bigg)^{a_2}+(2+\sqrt{5})^{a_1a_2}\bigg\}.$$

As in Example 1.4, the Hausdorff dimension of X_1 is obtained by letting $a_1 = \log_4 3$ and $a_2 = \log_3 2$:

$$\dim_H(X_1) = \frac{h^{\boldsymbol{a}}(\boldsymbol{T})}{\log 2} = \log_2\bigg\{\bigg(\bigg(\frac{1+\sqrt{5}}{2}\bigg)^{\log_4 3}+\bigg(\frac{3+\sqrt{5}}{2}\bigg)^{\log_4 3}\bigg)^{\log_3 2}+\sqrt{(2+\sqrt{5})}\bigg\}$$
$$= 2.1061\cdots.$$

Acknowledgements. I am deeply grateful to my mentor, Masaki Tsukamoto, who not only has reviewed this paper several times throughout the writing process but has patiently helped me understand ergodic theory in general with his expertise. I also want to thank my family and friends for their unconditional support and everyone who has participated in my study for their time and willingness to share their knowledge. This work could not have been possible without their help.

References

[Bed84] T. Bedford. Crinkly curves, Markov partitions and box dimension in self-similar sets. *PhD Thesis*, University of Warwick, 1984.
[BF12] J. Barral and D. J. Feng. Weighted thermodynamic formalism on subshifts and applications. *Asian J. Math.* **16** (2012), 319–352.
[BKM85] M. Boyle, B. Kitchens and B. Marcus. A note on minimal covers for sofic systems. *Proc. Amer. Math. Soc.* **95** (1985), 403–411.
[Din70] E. I. Dinaburg. A correlation between topological entropy and metric entropy. *Dokl. Akad. Nauk SSSR* **190** (1970), 19–22.
[Dow11] T. Downarowicz. *Entropy in Dynamical Systems*. Cambridge University Press, Cambridge, MA, 2011.
[Fe11] D. J. Feng. Equilibrium states for factor maps between subshifts. *Adv. Math.* **226** (2011), 2470–2502.
[FH16] D.-J. Feng and W. Huang. Variational principle for weighted topological pressure. *J. Math. Pures Appl. (9)* **106** (2016), 411–452.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press

[Gm71] T. N. T. Goodman. Relating topological entropy and measure entropy. *Bull. Lond. Math. Soc.* **3** (1971), 176–180.
[Gw69] L. W. Goodwyn. Topological entropy bounds measure-theoretic entropy. *Proc. Amer. Math. Soc.* **23** (1969), 679–688.
[KP96] R. Kenyon and Y. Peres. Measures of full dimension on affine-invariant sets. *Ergod. Th. & Dynam. Sys.* **16** (1996), 307–323.
[KP96-2] R. Kenyon and Y. Peres. Hausdorff dimensions of sofic affine-invariant sets. *Israel J. Math.* **94** (1996), 157–178.
[McM84] C. McMullen. The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.* **96** (1984), 1–9.
[Ru73] D. Ruelle. Statistical mechanics on a compact set with Z^v action satisfying expansiveness and specification. *Trans. Amer. Math. Soc.* **187** (1973), 237–251.
[Tsu22] M. Tsukamoto. New approach to weighted topological entropy and pressure. *Ergod. Th. & Dynam. Sys.* **43** (2023), 1004–1034.
[Wal75] P. Walters. A variational principle for the pressure of continuous transformations. *Amer. J. Math.* **97** (1975), 937–971.
[Wal82] P. Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, New York, 1982.
[We82] B Weiss. Subshifts of finite type and sofic systems. *Monatsh. Math.* **77** (1973), 462–474.
[Ya11] Y. Yayama. Applications of a relative variational principle to dimensions of nonconformal expanding maps. *Stoch. Dyn.* **11** (2011), 643–679.

https://doi.org/10.1017/etds.2024.35 Published online by Cambridge University Press