The Chinese University of Hong Kong
Michigan State University

DIMENSION THEORY OF ITERATED FUNCTION SYSTEMS

DE-JUN FENG AND HUYI HU

ABSTRACT. Let $\{S_i\}_{i=1}^{\ell}$ be an iterated function system (IFS) on $\mathbb{R}^d$ with attractor K. Let (Σ, σ) denote the one-sided full shift over the alphabet $\{1, \ldots, \ell\}$. We define the projection entropy function h_π on the space of invariant measures on Σ associated with the coding map $\pi : \Sigma \to K$, and develop some basic ergodic properties about it. This concept turns out to be crucial in the study of dimensional properties of invariant measures on K. We show that for any conformal IFS (resp., the direct product of finitely many conformal IFS), without any separation condition, the projection of an ergodic measure under π is always exactly dimensional and, its Hausdorff dimension can be represented as the ratio of its projection entropy to its Lyapunov exponent (resp., the linear combination of projection entropies associated with several coding maps). Furthermore, for any conformal IFS and certain affine IFS, we prove a variational principle between the Hausdorff dimension of the attractor and that of projections of ergodic measures.

CONTENTS

1. Introduction

Let $\{S_i : X \to X\}_{i=1}^{\ell}$ be a family of contractive maps on a nonempty closed set $X \subset \mathbb{R}^d$. Following Barnsley [2], we say that $\Phi = \{S_i\}_{i=1}^{\ell}$ is an *iterated function system* (IFS) on X. Hutchinson [28] showed that there is a unique nonempty compact set $K \subset X$, called the *attractor* of $\{S_i\}_{i=1}^{\ell}$, such that $K = \bigcup_{i=1}^{\ell} S_i(K)$. A probability measure μ on $\mathbb{R}^d$ is said to be *exactly dimensional* if there is a constant C such that the *local dimension*

$$d(\mu, x) = \lim_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}$$

exists and equals C for μ-a.e. $x \in \mathbb{R}^d$, where $B(x,r)$ denotes the closed ball of radius r centered at x. It was shown by Young [65] that in such case, the Hausdorff dimension of μ is equal to C. (See also [14, 43, 51].)

The motivation of the paper is to study the Hausdorff dimension of an invariant measure μ (see Section 2 for precise meaning) for conformal and affine IFS with overlaps. To deal with overlaps, we regard such a system as the image of a natural projection π from the one-sided full shift space over ℓ symbols. Hence we obtain a dynamical system. We introduce a notion *projection entropy*, which plays the similar role as the classical entropy for IFS satisfying the open set condition, and it becomes the classical entropy if the projection is finite to one. The concept of projection entropy turns out to be crucial in the study of dimensional properties of invariant measures on attractors of either conformal IFS with overlaps or affine IFS.

We develop some basic properties about projection entropy (Theorem 2.2, 2.3). We prove that for conformal IFS with overlaps, every ergodic measure μ is exactly dimensional and $d(\mu, x)$ is equal to the projection entropy divided by the Lyapunov exponent (Theorem 2.8). Furthermore, if Φ is a direct product of conformal IFS (see Definition 2.10 for precise meaning), then for every ergodic measure on K the local dimension can be expressed by a Ledrappier-Young type formula in terms of projection entropies and Lyapunov exponents (Theorem 2.11). We also prove variational results about Hausdorff dimension for conformal IFS and certain affine IFS (Theorem 2.13 and 2.15), which says that the Hausdorff dimension of the attractor K is equal to the supremum of Hausdorff dimension of μ taking over all ergodic measures. The results we obtain cover some interesting cases such as $S_i(x) = \text{diag}(\rho_1, \ldots, \rho_d)x + a_i$, where $i = 1, \ldots, \ell$ and ρ_i^{-1} are *Pisot or Salem numbers* and $a_i \in \mathbb{Z}^d$.

The problem whether a given measure is exactly dimensional, and whether the Hausdorff dimension of an attractor can be assumed or approximated by that of an invariant measure have been well studied in the literature for $C^{1+\alpha}$ conformal IFS which satisfy the open set condition (cf. [6, 23, 47]). It is well known that in such case, any ergodic measure μ is exactly dimensional with the Hausdorff dimension given by the classic entropy divided by the Lyapunov exponent. Furthermore there is a unique invariant measure μ with $\dim_H(\mu) = \dim_H(K)$, the Hausdorff dimension of K. However the problems become much complicated and intractable without the assumption of the open set condition. Partial results have only been obtained for conformal IFS that satisfy the *finite type condition* (see [45] for the definition). In that case, a Bernoulli measure is exactly dimensional and its Hausdorff dimension may be expressed as the upper Lyapunov exponent of certain random matrices (see e.g. [16, 17, 36, 39, 37]), and furthermore the Hausdorff dimension of K can be computed (see e.g. [35, 54, 45]).

There are some results for certain special non-overlapping affine IFS. McMullen [44] and Bedford [5] independently computed the Hausdorff dimension and the box dimension of the attractor of the following planar affine IFS

$$S_i(x) = \begin{bmatrix} n^{-1} & 0 \\ 0 & k^{-1} \end{bmatrix} x + \begin{bmatrix} a_i/n \\ b_i/k, \end{bmatrix}, \qquad i = 1, \ldots, \ell,$$

where all a_i, b_i are integers, $0 \leq a_i < n$ and $0 \leq b_i < k$. Furthermore they showed that there is a Bernoulli measure of full Hausdorff dimension. This result was extended by Kenyon and Peres [33] to higher dimensional self-affine Sierpinski sponges, for which ergodic measures are proved to be exactly dimensional with Hausdorff dimension given by a Ledrappier-Young type formula. Another extension of McMullen and Bedford's result to a boarder class of planar affine IFS $\{S_i\}_{i=1}^{\ell}$ was given by Gatzouras and Lalley [20], in which S_i map the unit square $(0, 1)^2$ into disjoint rectangles with sides parallel to the axes (where the longer sides are parallel to the x-axis, furthermore once projected onto the x-axis these rectangles are either identical, or disjoint). Further extensions were given recently by Barański [1], Feng and Wang [19], Luzia [41] and Olivier [46]. For other related results, see e.g. [52, 38, 34, 21, 25, 27, 17, 60, 3, 31].

Along another direction, in [11] Falconer gave a variational formula for the Hausdorff and box dimensions for "almost all" self-affine sets under some assumptions. This formula remains true under some weaker conditions [61, 29]. Käenmäki [30] proved that for "almost all" self-affine sets there exists an ergodic measure m so that $m \circ \pi^{-1}$ is of full Hausdorff dimension.

Our arguments use ergodic theory and Rohlin's theory about conditional measures. The proofs of Theorem 2.6 and Theorem 2.11 are based on some ideas from the work of Ledrappier and Young [40] and techniques in analyzing the densities of conditional measures associated with overlapping IFS.

So far we have restricted ourselves on the study of finite contractive IFS. However we point out that part of our results remain valid for certain non-contractive infinite IFS (see Section 10 for details).

The paper is organized as follows. The main results are given in Section 2. In Section 3, we prove some density results about conditional measures. In Section 4, we investigate the properties of projection entropy and prove Theorem 2.2 and 2.3. In Section 5, we give some local geometric properties of a C^1 IFS. In Section 6, we prove a generalized version of Theorem 2.6, which is based on a key proposition (Proposition 6.1) about the densities of conditional measures. In Section 7, we prove Theorem 2.11 and 2.12. In Section 8, we prove Theorem 2.13 and in Section 9, we prove Theorem 2.15. In Section 10 we give a remark regarding certain non-contractive infinite IFS.

2. Statement of main results

Let $\{S_i\}_{i=1}^{\ell}$ be an IFS on a closed set $X \subset \mathbb{R}^d$. Denote by K its attractor. Let $\Sigma = \{1, \ldots, \ell\}^{\mathbb{N}}$ associated with the left shift σ (cf. [9]). Let $\mathcal{M}_\sigma(\Sigma)$ denote the space of σ-invariant measures on Σ, endowed with the weak-star topology. Let $\pi : \Sigma \to K$ be the canonical projection defined by

$$(1) \qquad \{\pi(x)\} = \bigcap_{n=1}^{\infty} S_{x_1} \circ S_{x_2} \circ \cdots \circ S_{x_n}(K), \qquad \text{where } x = (x_i)_{i=1}^{\infty}.$$

A measure μ on K is called *invariant* (resp., *ergodic*) for the IFS if there is an invariant (resp. ergodic) measure ν on Σ such that $\mu = \nu \circ \pi^{-1}$.

Let $(\Omega, \mathcal{F}, \nu)$ be a probability space. For a sub-σ-algebra $\mathcal{A}$ of $\mathcal{F}$ and $f \in L^1(\Omega, \mathcal{F}, \nu)$, we denote by $\mathbf{E}_\nu(f|\mathcal{A})$ the the *conditional expectation of f given $\mathcal{A}$*. For countable $\mathcal{F}$-measurable partition ξ of Ω, we denote by $\mathbf{I}_\nu(\xi|\mathcal{A})$ the *conditional information of ξ given $\mathcal{A}$*, which is given by the formula

$$(2) \qquad \mathbf{I}_\nu(\xi|\mathcal{A}) = -\sum_{A \in \xi} \chi_A \log \mathbf{E}_\nu(\chi_A|\mathcal{A}),$$

where χ_A denotes the characteristic function on A. The *conditional entropy of ξ given $\mathcal{A}$*, written $H_\nu(\xi|\mathcal{A})$, is defined by the formula

$$H_\nu(\xi|\mathcal{A}) = \int \mathbf{I}_\nu(\xi|\mathcal{A})\, d\nu.$$

(See e.g. [48] for more details.) The above information and entropy are unconditional when $\mathcal{A} = \mathcal{N}$, the trivial σ-algebra consisting of sets of measure zero and one, and in this case we write

$$\mathbf{I}_\nu(\xi|\mathcal{N}) =: \mathbf{I}_\nu(\xi) \quad \text{and} \quad H_\nu(\xi|\mathcal{N}) =: H_\nu(\xi).$$

Now we consider the space $(\Sigma, \mathcal{B}(\Sigma), m)$, where $\mathcal{B}(\Sigma)$ is the Borel σ-algebra on Σ and $m \in \mathcal{M}_\sigma(\Sigma)$. Let $\mathcal{P}$ denote the Borel partition

$$(3) \qquad \mathcal{P} = \{[j] : 1 \le j \le \ell\}$$

of Σ, where $[j] = \{(x_i)_{i=1}^{\infty} \in \Sigma : \ x_1 = j\}$. Let $\mathcal{I}$ denote the σ-algebra

$$\mathcal{I} = \{B \in \mathcal{B}(\Sigma) : \ \sigma^{-1}B = B\}.$$

For convenience, we use γ to denote the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^d$.

Definition 2.1. For any $m \in \mathcal{M}_\sigma(\Sigma)$, we call

$$h_\pi(\sigma, m) := H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - H_m(\mathcal{P}|\pi^{-1}\gamma)$$

the *projection entropy of m under π w.r.t. $\{S_i\}_{i=1}^{\ell}$*, and we call

$$h_\pi(\sigma, m, x) := \mathbf{E}_m\left(f \middle| \mathcal{I}\right)(x)$$

the *local projection entropy of m at x under π w.r.t. $\{S_i\}_{i=1}^{\ell}$*, where f denotes the function $\mathbf{I}_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - \mathbf{I}_m(\mathcal{P}|\pi^{-1}\gamma)$.

It is clear that $h_\pi(\sigma, m) = \int h_\pi(\sigma, m, x)\, dm(x)$. Our first result is the following theorem.

Theorem 2.2. *Let $\{S_i\}_{i=1}^{\ell}$ be an IFS. Then*

(i) *For any $m \in \mathcal{M}_\sigma(\Sigma)$, we have $0 \le h_\pi(\sigma, m) \le h(\sigma, m)$, where $h(\sigma, m)$ denotes the classical measure-theoretic entropy of m associated with σ.*

(ii) *The map $m \mapsto h_\pi(\sigma, m)$ is affine on $\mathcal{M}_\sigma(\Sigma)$. Furthermore if $m = \int \nu \, d\mathbb{P}(\nu)$ is the ergodic decomposition of m, we have*

$$h_\pi(\sigma, m) = \int h_\pi(\sigma, \nu) \, d\mathbb{P}(\nu).$$

(iii) *For any $m \in \mathcal{M}_\sigma(\Sigma)$, we have*

$$\lim_{n\to\infty} \frac{1}{n} \mathbf{I}_m(\mathcal{P}_0^{n-1} | \pi^{-1}\gamma)(x) = h(\sigma, m, x) - h_\pi(\sigma, m, x)$$

for m-a.e. $x \in \Sigma$, where $h(\sigma, m, x)$ denotes the local entropy of m at x, that is, $h(\sigma, m, x) = \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma))(x)$.

Part (iii) of the theorem is an analogue of the classical relativized Shannon-McMillan-Breiman theorem (see, e.g. [8, Lemma 4.1]). However, we should notice that the sub σ-algebra $\pi^{-1}\gamma$ in our consideration is not σ-invariant in general (see Remark 4.11).

Part (iii) also implies that if the map $\pi\colon \Sigma \to K$ is finite-to-one, then

$$h_\pi(\sigma, m) = h(\sigma, m)$$

for any $m \in \mathcal{M}_\sigma(\Sigma)$. In Section 4, we will present a sufficient and necessary condition for the equality (see Corollary 4.16). However for general overlapping IFS, the projection entropy can be strictly less than the classical entropy.

In our next theorem, we give a geometric characterization of the projection entropy for certain affine IFS, which will be used later in the proof of our variational results about the Hausdorff and box dimensions of self-affine sets.

Theorem 2.3. *Assume that $\Phi = \{S_i\}_{i=1}^{\ell}$ is an IFS on $\mathbb{R}^d$ of the form*

$$S_i(x) = Ax + c_i \qquad (i = 1, \ldots, \ell),$$

where A is a $d \times d$ non-singular contractive real matrix and $c_i \in \mathbb{R}^d$. Let K denote the attractor of Φ. Let $\mathcal{Q}$ denote the partition $\{[0,1)^d + \alpha : \ \alpha \in \mathbb{Z}^d\}$ of $\mathbb{R}^d$. For $n = 0, 1, \ldots$, and $x \in \mathbb{R}^d$, we set $\mathcal{Q}_n = \{A^n Q : \ Q \in \mathcal{Q}\}$. Then

(i) *For any $m \in \mathcal{M}_\sigma(\Sigma)$, we have*

$$h_\pi(\sigma, m) = \lim_{n\to\infty} \frac{H_m(\pi^{-1}\mathcal{Q}_n)}{n}.$$

(ii) *Moreover,*

$$\lim_{n\to\infty} \frac{\log \#\{Q \in \mathcal{Q} : \ A^n Q \cap K \ne \emptyset\}}{n} = \sup\{h_\pi(\sigma, m) : \ m \in \mathcal{M}_\sigma(\Sigma)\}.$$

To give the applications of projection entropy in dimension theory of IFS, we need some more notation and definitions.

Definition 2.4. $\{S_i : X \to X\}_{i=1}^{\ell}$ is called a C^1 *IFS* on a compact set $X \subset \mathbb{R}^d$ if each S_i extends to a contracting C^1-diffeomorphism $S_i : U \to S_i(U) \subset U$ on an open set $U \supset X$.

For any $d \times d$ real matrix M, we use $\|M\|$ to denote the usual norm of M, and $[\![M]\!]$ the smallest singular value of M, i.e.,

$$\begin{aligned} \|M\| &= \max\{|Mv| : \ v \in \mathbb{R}^d, |v| = 1\} \quad \text{and} \\ [\![M]\!] &= \min\{|Mv| : \ v \in \mathbb{R}^d, |v| = 1\}. \end{aligned} \tag{4}$$

Definition 2.5. Let $\{S_i\}_{i=1}^{\ell}$ be a C^1 IFS. For $x = (x_j)_{j=1}^{\infty} \in \Sigma$, the *upper and lower Lyapunov exponents of* $\{S_i\}_{i=1}^{\ell}$ *at* x are defined respectively by

$$\begin{aligned} \overline{\lambda}(x) &= -\liminf_{n\to\infty} \frac{1}{n} \log [\![S'_{x_1\ldots x_n}(\pi\sigma^n x)]\!], \\ \underline{\lambda}(x) &= -\limsup_{n\to\infty} \frac{1}{n} \log \|S'_{x_1\ldots x_n}(\pi\sigma^n x)\|, \end{aligned}$$

where $S'_{x_1\ldots x_n}(\pi\sigma^n x)$ denotes the differential of $S_{x_1\ldots x_n} := S_{x_1} \circ S_{x_2} \circ \ldots \circ S_{x_n}$ at $\pi\sigma^n x$. When $\overline{\lambda}(x) = \underline{\lambda}(x)$, the common value, denoted as $\lambda(x)$, is called the *Lyapunov exponent of* $\{S_i\}_{i=1}^{\ell}$ *at* x.

It is easy to check that both $\overline{\lambda}$ and $\underline{\lambda}$ are positive-valued σ-invariant functions on Σ (i.e. $\overline{\lambda} = \overline{\lambda} \circ \sigma$ and $\underline{\lambda} = \underline{\lambda} \circ \sigma$). Recall that for a probability measure μ on $\mathbb{R}^d$, the *local upper and lower dimensions* are defined respectively by

$$\overline{d}(\mu, x) = \limsup_{r\to 0} \frac{\log \mu(B(x,r))}{\log r}, \quad \underline{d}(\mu, x) = \liminf_{r\to 0} \frac{\log \mu(B(x,r))}{\log r},$$

where $B(x,r)$ denotes the closed ball of radius r centered at x. If $\overline{d}(\mu, x) = \underline{d}(\mu, x)$, the common value is denoted as $d(\mu, x)$ and is called the *local dimension* of m at x.

The following theorem gives an estimate of local dimensions of invariant measures on the attractor of an arbitrary C^1 IFS, without any separation condition.

Theorem 2.6. *Let* $\{S_i\}_{i=1}^{\ell}$ *be a* C^1 *IFS with attractor* K. *Then for* $\mu = m \circ \pi^{-1}$, *where* $m \in \mathcal{M}_\sigma(\Sigma)$, *we have the following estimates:*

$$\overline{d}(\mu, \pi x) \le \frac{h_\pi(\sigma, m, x)}{\underline{\lambda}(x)} \quad \text{and} \quad \underline{d}(\mu, \pi x) \ge \frac{h_\pi(\sigma, m, x)}{\overline{\lambda}(x)} \quad \text{for } m\text{-a.e. } x \in \Sigma,$$

where $h_\pi(\sigma, m, x)$ *denotes the local projection entropy of* m *at* x *under* π *(see Definition 2.1). In particular, if* m *is ergodic, we have*

$$\frac{h_\pi(\sigma, m)}{\int \overline{\lambda}\, dm} \le \underline{d}(\mu, z) \le \overline{d}(\mu, z) \le \frac{h_\pi(\sigma, m)}{\int \underline{\lambda}\, dm} \quad \text{for } \mu\text{-a.e. } z \in K.$$

Definition 2.7. Let $\{S_i\}_{i=1}^{\ell}$ be a C^1 IFS and $m \in \mathcal{M}_\sigma(\Sigma)$. We say that $\{S_i\}_{i=1}^{\ell}$ is m-*conformal* if $\lambda(x)$ exists (i.e., $\overline{\lambda}(x) = \underline{\lambda}(x)$) for m-a.e. $x \in \Sigma$.

As a direct application of Theorem 2.6, we have

Theorem 2.8. *Assume that $\{S_i\}_{i=1}^{\ell}$ is m-conformal for some $m \in \mathcal{M}_\sigma(\Sigma)$. Let $\mu = m \circ \pi^{-1}$. Then we have*

$$d(\mu, \pi x) = \frac{h_\pi(\sigma, m, x)}{\lambda(x)} \quad \textit{for } m\textit{-a.e. } x \in \Sigma. \tag{5}$$

In particular, if m is ergodic, we have

$$d(\mu, z) = \frac{h_\pi(\sigma, m)}{\int \lambda \, dm} \quad \textit{for } \mu\textit{-a.e. } z \in K. \tag{6}$$

Recall that $S : U \to S(U)$ is a conformal map if $S'(x) : \mathbb{R}^d \to \mathbb{R}^d$ satisfies $\|S'(x)\| \neq 0$ and $|S'(x)y| = \|S'(x)\||y|$ for all $x \in U$ and $y \in \mathbb{R}^d$.

Definition 2.9. A C^1 IFS $\{S_i\}_{i=1}^{\ell}$ is said to be *weakly conformal* if

$$\frac{1}{n}(\log [\![S'_{x_1 \dots x_n}(\pi\sigma^n x)]\!] - \log \|S'_{x_1 \dots x_n}(\pi\sigma^n x)\|)$$

converges to 0 uniformly on Σ as n tends to ∞. We say that $\{S_i\}_{i=1}^{\ell}$ is *conformal* if each S_i extends to a conformal map $S_i : U \to S_i(U) \subset U$ on an open set $U \supset K$, where K is the attractor of $\{S_i\}_{i=1}^{\ell}$.

By definition, a conformal IFS is always weakly conformal. Furthermore, a weakly conformal IFS is m-conformal for each $m \in \mathcal{M}_\sigma(\Sigma)$ (see Proposition 5.6(ii)). There are some natural examples of weakly conformal IFS which are not conformal. For instance, let $S_i(x) = A_i x + a_i$ ($i = 1, \dots, \ell$) such that, for each i, A_i is a contracting linear map with eigenvalues equal to each other in modulus, and $A_i A_j = A_j A_i$ for different i, j. Then such an IFS is always weakly conformal but not necessary to be conformal. The first conclusion follows from the asymptotic behavior

$$\lim_{n\to\infty} [\![A_i^n]\!]^{1/n} = \lim_{n\to\infty} \|A_i^n\|^{1/n} = \rho(A_i) \qquad (i = 1, \dots, \ell),$$

where $\rho(A_i)$ denotes the spectral radius of A_i (cf. [64]).

Theorem 2.8 verifies the existence of local dimensions for invariant measures on the attractor of an arbitrary weakly conformal IFS attractors, without any separation assumption. We point out that the exact dimensionality for overlapping self-similar measures was first claimed by Ledrappier, nevertheless no proof has been written out (cf. [50, p. 1619]). We remark that this property was also conjectured later by Fan, Lau and Rao in [15].

We can extend the above result to a class of non-conformal IFS.

Definition 2.10. Assume for $j = 1, \dots, k$, $\Phi_j := \{S_{i,j}\}_{i=1}^{\ell}$ is a C^1 IFS defined on a compact set $X_j \subset \mathbb{R}^{q_j}$. Let $\Phi := \{S_i\}_{i=1}^{\ell}$ be the IFS on $X_1 \times \dots \times X_k \subset \mathbb{R}^{q_1} \times \dots \times \mathbb{R}^{q_k}$ given by

$$S_i(z_1, \dots, z_k) = (S_{i,1}(z_1), \dots, S_{i,k}(z_k)) \ \ (i = 1, \dots, \ell, \ j = 1, \dots, k, \ z_j \in X_j).$$

We say that Φ is the *direct product* of $\Phi_1, \dots, \Phi_k$, and write $\Phi = \Phi_1 \times \dots \times \Phi_k$.

Theorem 2.11. *Let $\Phi = \{S_i\}_{i=1}^{\ell}$ be the direct product of k C^1 IFS $\Phi_1, \ldots, \Phi_k$. Let $\mu = m \circ \pi^{-1}$, where $m \in \mathcal{M}_\sigma(\Sigma)$. Assume that $\Phi_1, \ldots, \Phi_k$ are m-conformal. Then*

(i) *$d(\mu, z)$ exists for μ-a.e. z.*
(ii) *Assume furthermore that m is ergodic. Then μ is exactly dimensional. Let τ be a permutation on $\{1, \ldots, k\}$ such that*

$$\lambda_{\tau(1)} \leq \lambda_{\tau(2)} \leq \cdots \leq \lambda_{\tau(k)},$$

where $\lambda_j = \int \lambda_j(x)\, dm(x)$, and $\lambda_j(x)$ denotes the Lyapunov exponent of Φ_j at $x \in \Sigma$. Then we have

$$d(\mu, z) = \frac{h_{\pi_1}(\sigma, m)}{\lambda_{\tau(1)}} + \sum_{j=2}^{k} \frac{h_{\pi_j}(\sigma, m) - h_{\pi_{j-1}}(\sigma, m)}{\lambda_{\tau(j)}} \quad \text{for } \mu\text{-a.e. } z, \tag{7}$$

where π_j denotes the canonical projection w.r.t. the IFS $\Phi_{\tau(1)} \times \cdots \times \Phi_{\tau(j)}$, and $h_{\pi_j}(\sigma, m)$ denotes the projection entropy of m under π_j.

We mention that fractals satisfy the conditions of the theorem include many interesting examples such as those studied in [44, 5, 20, 33], etc.

As an application of Theorem 2.11, we have

Theorem 2.12. *Let $\{S_i\}_{i=1}^{\ell}$ be an IFS on $\mathbb{R}^d$ of the form*

$$S_i(x) = A_i x + a_i, \qquad i = 1, \ldots, \ell,$$

such that each A_i is a nonsingular contracting linear map on $\mathbb{R}^d$, and $A_iA_j = A_jA_i$ for any $1 \leq i, j \leq \ell$. Then for any ergodic measure m on Σ, $\mu = m \circ \pi^{-1}$ is exactly dimensional.

Indeed, under the assumption of Theorem 2.12, we can show that there is a nonsingular linear transformation T on $\mathbb{R}^d$ such that the IFS $\{T \circ S_i \circ T^{-1}\}_{i=1}^{\ell}$ is the direct product of some weakly conformal IFS. Hence we can apply Theorem 2.11 in this situation.

We remark that formula (7) provides an analogue of that for the Hausdorff dimension of $C^{1+\alpha}$ hyperbolic measures along the unstable (resp. stable) manifold established by Ledrappier and Young [40].

The problem of the existence of local dimensions has also a long history in smooth dynamical systems. In [65], Young proved that an ergodic hyperbolic measure invariant under a $C^{1+\alpha}$ surface diffeomorphism is always exact dimensional. For a measures μ in high-dimensional $C^{1+\alpha}$ systems, Ledrappier and Young [40] proved the existence of δ^u and δ^s, the local dimensions along stable and unstable local manifolds, respectively, and the upper local dimension of μ is bounded by the sum of δ^u, δ^s, and the multiplicity of 0 as an exponent. Eckmann and Ruelle [10] indicated that it is unknown whether the local dimension of μ is the sum of δ^u and δ^s if μ is a hyperbolic measure. Then the problem was referred as Eckmann-Ruelle conjecture, and affirmatively answered by Barreira, Pesin and Schmeling in [4] seventeen years later. Some partial dimensional results were obtained for measures invariant under hyperbolic endomorphism [58, 59]. Recently, Qian and

Xie [53] proved the exact dimensionality of ergodic measures invariant under C^2 expanding endomorphism on smooth Riemannian manifolds.

In the remaining part of this section, we present some variational results about the Hausdorff dimension and the box dimension of attractors of IFS and that of invariant measures. First we consider conformal IFS.

Theorem 2.13. *Let K be the attractor of a weakly conformal IFS $\{S_i\}_{i=1}^{\ell}$. Then we have*

$$\begin{aligned}
&(8) \quad \dim_H K &&= \dim_B K \\
&(9) &&= \sup\left\{\dim_H \mu : \ \mu = m\circ\pi^{-1}, \ m \in \mathcal{M}_\sigma(\Sigma), \ m \textit{ is ergodic}\right\} \\
& &&= \max\left\{\dim_H \mu : \ \mu = m\circ\pi^{-1}, \ m \in \mathcal{M}_\sigma(\Sigma)\right\} \\
&(10) &&= \sup\left\{\frac{h_\pi(\sigma, m)}{\int \lambda \, dm} : \ m \in \mathcal{M}_\sigma(\Sigma)\right\},
\end{aligned}$$

where $\dim_B K$ denotes the box dimension of K.

Equality (8) was first proved by Falconer [12] for $C^{1+\alpha}$ conformal IFS. It is not known whether the supremum in (9) and (10) can be attained in the general setting of Theorem 2.13. However, this is true if the IFS $\{S_i\}_{i=1}^{\ell}$ satisfies an additional separation condition defined as follows.

Definition 2.14. An IFS $\{S_i\}_{i=1}^{\ell}$ on a compact set $X \subset \mathbb{R}^d$ is said to satisfy the *asymptotically weak separation condition* (AWSC), if

$$\lim_{n\to\infty} \frac{1}{n} \log t_n = 0,$$

where t_n is given by

$$(11) \qquad t_n = \sup_{x\in\mathbb{R}^d} \#\{S_u : \ u \in \{1, \ldots, \ell\}^n, \ x \in S_u(K)\},$$

here K is the attractor of $\{S_i\}_{i=1}^{\ell}$.

The above definition was first introduced in [18] under a slightly different setting. For example, if $1/\rho$ is a Pisot or Salem number, then the IFS $\{\rho x + a_i\}_{i=1}^{\ell}$ on $\mathbb{R}$, with $a_i \in \mathbb{Z}$, satisfies the AWSC (see Proposition 5.3 and Remark 5.5 in [18]). Recall that a real number $\beta > 1$ is said to be a *Salem number* if it is an algebraic integer whose algebraic conjugates all have modulus not greater than 1, with at least one of which on the unit circle. Whilst $\beta > 1$ is called a *Pisot number* if it is an algebraic integer whose algebraic conjugates all have modulus less than 1. For instance, the largest root (≈ 1.72208) of $x^4 - x^3 - x^2 - x + 1$ is a Salem number, and the golden ratio $(\sqrt{5} + 1)/2$ is a Pisot number. One is referred to [57] for more examples and properties about Pisot and Salem numbers. Under the AWSC assumption, we can show that the projection entropy map $m \mapsto h_\pi(\sigma, m)$ is upper semi-continuous on $\mathcal{M}_\sigma(\Sigma)$ (see Proposition 4.20) and, as a consequence, the supremum (9) and (10) can be attained at ergodic measures (see Remark 8.2).

Next we consider some affine IFS.

Theorem 2.15. *Let $\Phi = \{S_i\}_{i=1}^{\ell}$ be an affine IFS on $\mathbb{R}^d$ given by*

$$S_i(x_1, \dots, x_d) = (\rho_1 x_1, \cdots, \rho_d x_d) + (a_{i,1}, \dots, a_{i,d}),$$

where $\rho_1 > \rho_2 > \cdots > \rho_d > 0$ and $a_{i,j} \in \mathbb{R}$. Let K denote the attractor of Φ, and write $\lambda_j = \log(1/\rho_j)$ for $j = 1, \dots, d$ and $\lambda_{d+1} = \infty$. View Φ as the direct product of $\Phi_1, \dots, \Phi_d$, where $\Phi_j = \{S_{i,j}(x_j) = \rho_j x_j + a_{i,j}\}_{i=1}^{\ell}$. Assume that $\Phi_1 \times \cdots \times \Phi_j$ satisfies the AWSC for $j = 1, \dots, d$. Then we have

$$\begin{aligned}\dim_H K &= \max\left\{\dim_H \mu:\ \mu = m \circ \pi^{-1},\ m \text{ is ergodic}\right\} \\ &= \max\left\{\sum_{j=1}^{d}\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) h_{\pi_j}(\sigma, m):\ m \text{ is ergodic}\right\},\end{aligned}$$

where π_j is the canonical projection w.r.t. the IFS $\Phi_1 \times \cdots \times \Phi_j$. Furthermore

$$\dim_B K = \sum_{j=1}^{d}\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) H_j,$$

where $H_j := \max\{h_{\pi_j}(\sigma, m):\ m \text{ is ergodic}\}$.

It is direct to check that if Φ_j satisfies the AWSC for each $1 \le j \le d$, then so does $\Phi_1 \times \cdots \times \Phi_j$. Hence for instance, the condition of Theorem 2.15 fulfills when $1/\rho_j$ are Pisot numbers or Salem numbers and $(a_{i,1}, \dots, a_{i,d}) \in \mathbb{Z}^d$. Different from the earlier works on the Hausdorff dimension of deterministic self-affine sets and self-affine measures (see e.g. [44, 5, 33, 20, 27, 1, 46]), our model in Theorem 2.15 admits certain overlaps. The two variational results in Theorem 2.15 provide some new insights in the study of overlapping self-affine IFS. An interesting question is whether the results of Theorem 2.15 remain true without the AWSC assumption. It is related to the open problem whether a non-conformal repeller carries an ergodic measure of full dimension (see [22] for a survey). We remark that in the general case, we do have the following inequality(see Lemma 9.2):

$$\underline{\dim}_B K \ge \sum_{j=1}^{d}\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) \sup\{h_{\pi_j}(\sigma, m):\ m \text{ is ergodic}\}.$$

Furthermore Theorem 2.15 can be extended somewhat (see Remark 9.3 and Theorem 9.4).

3. Density results about conditional measures

We prove some density results about conditional measures in this section. To begin with, we give a brief introduction to Rohlin's theory of Lebesgue spaces, measurable partitions and conditional measures. The reader is referred to [55, 49] for more details.

A probability space $(X, \mathcal{B}, m)$ is called a *Lebesgue space* if it is isomorphic to a probability space which is the union of $[0, s]$ $(0 \le s \le 1)$ with Lebesgue measure and a countable number of atoms. Now let $(X, \mathcal{B}, m)$ be a Lebesgue space. A *measurable partition* η of X is a partition of X such that, up to a set of measure zero, the quotient space X/η is separated by a countable number of measurable sets $\{B_i\}$. The quotient space X/η with its inherit probability space structure, written as $(X_\eta, \mathcal{B}_\eta, m_\eta)$, is again a Lebesgue space. Also, any measurable partition η determine a sub-σ-algebra of $\mathcal{B}$, denoted by $\widehat{\eta}$, whose elements are unions of elements of η. Conversely, any sub-σ-algebra of $\mathcal{B}'$ of $\mathcal{B}$ is also countably generated, say by $\{B_i'\}$, and therefore all the sets of the form $\cap A_i$, where $A_i = B_i'$ or its complement, form a measurable partition. In particular, $\mathcal{B}$ itself is corresponding to a partition into single points. An important property of Lebesgue space and measurable partitions is the following.

Theorem 3.1 (Rohlin [55]). *Let η be a measurable partition of a Lebesgue space $(X, \mathcal{B}, m)$. Then, for every x in a set of full m-measure, there is a probability measure m_x^η defined on $\eta(x)$, the element of η containing x. These measures are uniquely characterized (up to sets of m-measure 0) by the following properties: if $A \subset X$ is a measurable set, then $x \mapsto m_x^\eta(A)$ is $\widehat{\eta}$-measurable and $m(A) = \int m_x^\eta(A) dm(x)$. These properties imply that for any $f \in L^1(X, \mathcal{B}, m)$, $m_x^\eta(f) = \mathbf{E}_m(f|\widehat{\eta})(x)$ for m-a.e. x, and $m(f) = \int \mathbf{E}_m(f|\widehat{\eta}) dm$.*

The family of measures $\{m_x^\eta\}$ in the above theorem is called the *canonical system of conditional measures associated with η*.

Throughout the remaining part of this section, we assume that $(X, \mathcal{B}, m)$ is a Lebesgue space. Let η be a measurable partition of X, and let $\{m_x^\eta\}$ denote the corresponding canonical system of conditional measures. Suppose that $\pi : X \to \mathbb{R}^d$ is a $\mathcal{B}$-measurable map. Denote $\gamma := \mathcal{B}(\mathbb{R}^d)$, the Borel-$\sigma$-algebra on $\mathbb{R}^d$. For $y \in \mathbb{R}^d$, we use $B(y, r)$ to denote the closed ball in $\mathbb{R}^d$ of radius r centered at y. Also, we denote for $x \in X$,

$$B^\pi(x, r) = \pi^{-1} B(\pi x, r). \tag{12}$$

Lemma 3.2. *Let $A \in \mathcal{B}$.*

(i) *The map $x \mapsto m_x^\eta(B^\pi(x, r) \cap A)$ is $\hat{\eta} \vee \pi^{-1}\gamma$-measurable for each $r > 0$, where $\hat{\eta} \vee \pi^{-1}\gamma$ denotes the smallest sub-σ-algebra of $\mathcal{B}$ containing $\hat{\eta}$ and $\pi^{-1}\gamma$.*

(ii) *The following functions*

$$\liminf_{r \to 0} \frac{m_x^\eta(B^\pi(x, r) \cap A)}{m_x^\eta(B^\pi(x, r))}, \quad \limsup_{r \to 0} \frac{m_x^\eta(B^\pi(x, r) \cap A)}{m_x^\eta(B^\pi(x, r))}$$

and

$$\inf_{r > 0} \frac{m_x^\eta(B^\pi(x, r) \cap A)}{m_x^\eta(B^\pi(x, r))}$$

are $\hat{\eta} \vee \pi^{-1}\gamma$-measurable, where we interpret $0/0 = 0$.

Proof. We first prove (i). Let $A \in \mathcal{B}$ and $r > 0$. For $n \in \mathbb{N}$, let $\mathcal{D}_n$ denote the collection

$$\mathcal{D}_n = \{[0, 2^{-n})^d + \alpha : \ \alpha \in 2^{-n}\mathbb{Z}^d\}.$$

For $y \in \mathbb{R}^d$, denote

$$W_n(y) = \bigcup_{Q \in \mathcal{D}_n :\, Q \cap B(y, r) \neq \emptyset} Q.$$

Write $\mathcal{W}_n := \{W_n(y) : y \in \mathbb{R}^d\}$. It is clear that $\mathcal{W}_n$ is countable for each $n \in \mathbb{N}$. Furthermore, we have $W_n(y) \downarrow B(y,r)$ for each $y \in \mathbb{R}^d$ as $n \to \infty$, that is, $W_{n+1}(y) \subset W_n(y)$ and $\bigcap_{n=1}^{\infty} W_n(y) = B(y,r)$. As a consequence, we have $\pi^{-1}W_n(\pi x) \downarrow B^\pi(x,r)$ and hence

$$m_x^\eta(B^\pi(x,r)\cap A) = \lim_{n\to\infty} m_x^\eta(\pi^{-1}W_n(\pi x)\cap A) \qquad (x \in X).$$

Therefore to show that $x \mapsto m_x^\eta(B^\pi(x,r)\cap A)$ is $\hat{\eta}\vee\pi^{-1}\gamma$-measurable, it suffices to show that $x \mapsto m_x^\eta(\pi^{-1}W_n(\pi x)\cap A)$ is $\hat{\eta}\vee\pi^{-1}\gamma$-measurable for each $n \in \mathbb{N}$.

Fix $n \in \mathbb{N}$. For $F \in \mathcal{W}_n$, let $\Gamma_n(F) = \{x \in X : W_n(\pi x) = F\}$. Then $\Gamma_n(F) \in \pi^{-1}\gamma$. By Theorem 3.1, $m_x^\eta(\pi^{-1}F\cap A)$ is an $\hat{\eta}$-measurable function of x for each $F \in \mathcal{W}_n$. However

$$m_x^\eta(\pi^{-1}W_n(\pi x)\cap A) = \sum_{F\in\mathcal{W}_n} \chi_{\Gamma_n(F)}(x) m_x^\eta(\pi^{-1}F\cap A).$$

Hence $m_x^\eta(\pi^{-1}W_n(\pi x)\cap A)$ is $\hat{\eta}\vee\pi^{-1}\gamma$-measurable, so is $m_x^\eta(B^\pi(x,r)\cap A)$.

To see (ii), note that for $x \in \Sigma$ and $r > 0$ satisfying $m_x^\eta(B^\pi(x,r)) > 0$, we have

$$\frac{m_x^\eta(B^\pi(x,r)\cap A)}{m_x^\eta(B^\pi(x,r))} = \lim_{q\downarrow r:\, q\in\mathbb{Q}^+} \frac{m_x^\eta(B^\pi(x,q)\cap A)}{m_x^\eta(B^\pi(x,q))}.$$

Hence for the three limits in (ii), we can restrict r to be positive rationals. It together with (i) yields the desired measurability. □

Lemma 3.3. *Let $A \in \mathcal{B}$. Then for m-a.e. $x \in X$,*

$$\lim_{r\to 0} \frac{m_x^\eta(B^\pi(x,r)\cap A)}{m_x^\eta(B^\pi(x,r))} = \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)(x). \tag{13}$$

Proof. Let $\overline{f}(x)$ and $\underline{f}(x)$ be the values obtained by taking the upper and lower limits in the left hand side of (13). By Lemma 3.2, both $\overline{f}$ and $\underline{f}$ are $\hat{\eta}\vee\pi^{-1}\gamma$-measurable. In the following we only show that $\overline{f}(x) = \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)(x)$ for m-a.e. x. The proof for $\underline{f}(x) = \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)(x)$ is similar.

We first prove that

$$\int_{B\cap\pi^{-1}D} \overline{f}\, dm = \int_{B\cap\pi^{-1}D} \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)\, dm \qquad (B \in \hat{\eta},\ D \in \gamma). \tag{14}$$

By Theorem 3.1, for any given $C \in \eta$, m_x^η $(x \in C)$ represent the same measure supported on C, which we rewrite as m_C. Fix $C \in \eta$. We define measures μ_C and ν_C on $\mathbb{R}^d$ by $\mu_C(E) = m_C(\pi^{-1}E\cap A)$ and $\nu_C(E) = m_C(\pi^{-1}E)$ for all $E \in \gamma$. It is clear that $\mu_C \ll \nu_C$. Define

$$g_C(z) = \limsup_{r\to 0} \frac{\mu_C(B(z,r))}{\nu_C(B(z,r))} \qquad (z \in \mathbb{R}^d).$$

Then $\overline{f}(x) = g_{\eta(x)}(\pi x)$ for all $x \in \Sigma$. According to the differentiation theory of measures on $\mathbb{R}^d$ (see, e.g., [43, Theorem 2.12]), $g_C = \frac{d\mu_C}{d\nu_C}$ ν_C-a.e. Hence for each $D \in \gamma$, we have $\int_D g_C(z)\, d\nu_C(z) = \mu_C(D)$, i.e., $\int_{\pi^{-1}D} g_C(\pi y)\, dm_C(y) = \mu_C(D) = m_C(\pi^{-1}D \cap A)$. That is,

$$\int_{\pi^{-1}D} \overline{f}\, dm_x^\eta = m_x^\eta(\pi^{-1}D \cap A) \qquad (x \in X). \tag{15}$$

To see (14), let $B \in \hat{\eta}$. Then

$$\begin{aligned}
\int_{B\cap\pi^{-1}D} \overline{f}\, dm &= \int \chi_B \chi_{\pi^{-1}D} \overline{f}\, dm = \int \mathbf{E}_m\left(\chi_B \chi_{\pi^{-1}D}\overline{f} | \hat{\eta}\right)\, dm \\
&= \int \chi_B \mathbf{E}_m\left(\chi_{\pi^{-1}D}\overline{f}|\hat{\eta}\right)\, dm \\
&= \int_B \left(\int_{\pi^{-1}D} \overline{f}\, dm_x^\eta\right) dm(x) \qquad \text{(by Theorem 3.1)} \\
&= \int_B m_x^\eta(\pi^{-1}D \cap A) dm(x) \qquad \text{(by (15))} \\
&= \int \chi_B(x) \mathbf{E}_m\left(\chi_{\pi^{-1}D\cap A}|\hat{\eta}\right)(x)\, dm(x) \qquad \text{(by Theorem 3.1)}.
\end{aligned}$$

Thus we have

$$\begin{aligned}
\int_{B\cap\pi^{-1}D} \overline{f}\, dm &= \int \mathbf{E}_m\left(\chi_B\chi_{\pi^{-1}D\cap A}|\hat{\eta}\right)(x)\, dm(x) \\
&= \int \chi_B \chi_{\pi^{-1}D\cap A} dm = m(B\cap\pi^{-1}D\cap A) \\
&= \int \mathbf{E}_m(\chi_{B\cap\pi^{-1}D}\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)\, dm \\
&= \int \chi_{B\cap\pi^{-1}D}\mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)\, dm \\
&= \int_{B\cap\pi^{-1}D} \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)\, dm.
\end{aligned}$$

This establishes (14).

Let $R = \overline{f} - \mathbf{E}_m(\chi_A|\hat{\eta}\vee\pi^{-1}\gamma)$. Then R is $\hat{\eta}\vee\pi^{-1}\gamma$-measurable and

$$\int_{B\cap\pi^{-1}(D)} R\, dm = 0 \qquad (B \in \hat{\eta},\ D \in \pi^{-1}\gamma).$$

Denote $\mathcal{F} = \{B\cap\pi^{-1}(D) :\ B \in \hat{\eta},\ D \in \pi^{-1}\gamma\}$ and let

$$\mathcal{F}' = \left\{\bigcup_{i=1}^k F_i :\ k \in \mathbb{N},\ F_1, \ldots, F_k \in \mathcal{F} \text{ are disjoint}\right\}.$$

It is clear that $\int_F R\, dm = 0$ for all $F \in \mathcal{F}'$. Moreover it is a routine to check that $\mathcal{F}'$ is an algebra which contains $\hat{\eta}$ and $\pi^{-1}\gamma$, and hence $\mathcal{F}'$ generates the σ-algebra $\hat{\eta}\vee\pi^{-1}\gamma$.

We claim that $R = 0$ m-a.e. Assume this is not true. Then there exists $\epsilon > 0$ such that the set $\{R > \epsilon\}$, or $\{R < -\epsilon\}$, has positive m-measure. Without loss of generality, we assume that $m\{R > \epsilon\} > 0$. Since $\mathcal{F}'$ is an algebra which generates $\hat{\eta} \vee \pi^{-1}\gamma$, there exists a sequence $F_i \in \mathcal{F}'$ such that $m(F_i \triangle \{R > \epsilon\})$ tends to 0 as $i \to \infty$ (cf. [63, Theorem 0.7]). We conclude that $\int_{F_i} R\, dm$ tends to $\int_{\{R>\epsilon\}} R\, dm > 0$ as $i \to \infty$, which contradicts the fact $\int_{F_i} R\, dm = 0$. $\square$

Remark 3.4. (i) Letting $\eta = \mathcal{N}$ be the trivial partition of X in the above lemma, we obtain $\lim_{r\to 0} \dfrac{m(B^\pi(x,r)\cap A)}{m(B^\pi(x,r))} = \mathbf{E}_m(\chi_A|\pi^{-1}\gamma)(x)$ m-a.e.

(ii) In general, $\mathbf{E}_{m_x^\eta}(\chi_A|\pi^{-1}\gamma)(x) = \mathbf{E}_m(\chi_A|\hat{\eta} \vee \pi^{-1}\gamma)(x)$ m-a.e., both of them equal $\lim_{r\to 0} \dfrac{m_x^\eta(B^\pi(x,r)\cap A)}{m_x^\eta(B^\pi(x,r))}$ m-a.e. by (i).

Proposition 3.5. *Let ξ be a countable measurable partition of X. Then for m-a.e. $x \in X$,*

$$\lim_{r\to 0} \log \frac{m_x^\eta\left(B^\pi(x,r)\cap \xi(x)\right)}{m_x^\eta\left(B^\pi(x,r)\right)} = -\mathbf{I}_m\left(\xi|\hat{\eta}\vee\pi^{-1}\gamma\right)(x), \tag{16}$$

where $\mathbf{I}_m(\cdot|\cdot)$ denotes the conditional information (see (2) for the definition). Furthermore, set

$$g(x) = -\inf_{r>0} \log \frac{m_x^\eta\left(B^\pi(x,r)\cap \xi(x)\right)}{m_x^\eta\left(B^\pi(x,r)\right)} \tag{17}$$

and assume $H_m(\xi) < \infty$. Then $g \geq 0$ and $g \in L^1(X, \mathcal{B}, m)$.

Proof. (16) follows directly from Lemma 3.3 and the following equality

$$\lim_{r\to 0} \log \frac{m_x^\eta\left(B^\pi(x,r)\cap \xi(x)\right)}{m_x^\eta\left(B^\pi(x,r)\right)} = \sum_{A\in\xi} \chi_A(x) \lim_{r\to 0} \log \frac{m_x^\eta\left(B^\pi(x,r)\cap A\right)}{m_x^\eta\left(B^\pi(x,r)\right)}.$$

Now we turn to the proof of (17). It is clear that g is non-negative. By Lemma 3.2, g is measurable. In the following we show that $g \in L^1(X, \mathcal{B}, m)$.

Let $C \in \eta$ and $A \in \xi$ be given. As in the proof of Lemma 3.3, we define measures μ_C and ν_C on $\mathbb{R}^d$ by $\mu_C(E) = m_C(\pi^{-1}E \cap A)$ and $\nu_C(E) = m_C(\pi^{-1}E)$ for all $E \in \gamma$. By Theorem 7.4 in [56], we have

$$\mu_C\left\{z \in \mathbb{R}^d : \inf_{r>0} \frac{\mu_C(B(z,r))}{\nu_C(B(z,r))} < \lambda\right\} \leq 3^d\lambda \qquad (\lambda > 0).$$

Hence for any $\lambda > 0$,

$$m_C\left(\left\{x \in X : \inf_{r>0} \frac{m_C\left(B^\pi(x,r)\cap A\right)}{m_C\left(B^\pi(x,r)\right)} < \lambda\right\} \cap A\right) \leq 3^d\lambda.$$

Integrating C with respect to m_η, we obtain

$$m\left(\left\{x \in X : \inf_{r>0} \frac{m_x^\eta\left(B^\pi(x,r)\cap A\right)}{m_x^\eta\left(B^\pi(x,r)\right)} < \lambda\right\} \cap A\right) \leq 3^d\lambda.$$

Denote $g^A(x) = \inf_{r>0} \dfrac{m_x^\eta\left(B^\pi(x,r)\cap A\right)}{m_x^\eta\left(B^\pi(x,r)\right)}$. Then the above inequality can be rewritten as

$$m(A \cap \{g^A < \lambda\}) \leq 3^d\lambda.$$

Note that by (17), $g(x) = -\sum_{A\in\xi} \chi_A(x) \log g^A(x)$. Since g is non-negative, we have

$$\begin{aligned}
\int g\,dm &= \int_0^\infty m\{g > t\}\,dt = \int_0^\infty \sum_{A\in\xi} m(A \cap \{g^A < e^{-t}\})\,dt \\
&\leq \sum_{A\in\xi} \int_0^\infty \min\{m(A), 3^d e^{-t}\}\,dt \\
&\leq \sum_{A\in\xi} \Big(-m(A)\log m(A) + m(A) + m(A)\log 3^d\Big) \\
&= H_m(\xi) + 1 + \log 3^d.
\end{aligned}$$

This finishes the proof of the proposition. □

Remark 3.6. Consider the case $X = \Sigma$ and $\xi = \mathcal{P}$, where $\mathcal{P}$ is defined as in (3). Suppose that $\{S_i\}_{i=1}^{\ell}$ is a family of mappings such that $S_i\colon \pi(\Sigma) \to S_i(\pi(\Sigma)) \subset \mathbb{R}^d$ is homeomorphic for each i. Then in (16) and (17), we can change the terms $B^\pi(x,r)$ to $\pi^{-1}R_{r,x}(\pi x)$, where $R_{r,x}(z) := S_{x_1}^{-1}B(S_{x_1}(z), r)$. To see it, fix i and define $\pi' = S_i \circ \pi$. Then we have

$$\begin{aligned}
\lim_{r\to 0} \frac{m_x^\eta\left(\pi^{-1}R_{r,x}(\pi x) \cap [i]\right)}{m_x^\eta\left(\pi^{-1}R_{r,x}(\pi x)\right)} &= \lim_{r\to 0} \frac{m_x^\eta\left(B^{\pi'}(x,r) \cap [i]\right)}{m_x^\eta\left(B^{\pi'}(x,r)\right)} \\
&= \mathbf{E}_m(\chi_{[i]}|\hat{\eta} \vee (\pi')^{-1}\gamma)(x).
\end{aligned}$$

However, $(\pi')^{-1}\gamma = \pi^{-1}\gamma$ due to the assumption on S_i. Hence the last term in the above formula equals $\mathbf{E}_m(\chi_{[i]}|\hat{\eta} \vee \pi^{-1}\gamma)(x)$. Thus we can replace the terms $B^\pi(x,r)$ by $\pi^{-1}R_{r,x}(\pi x)$ in (16). For the change in (17), we may use a similar argument.

Lemma 3.7. *Let $\pi\colon X \to \mathbb{R}^d$ and $\phi\colon X \to \mathbb{R}^k$ be two $\mathcal{B}$-measurable maps. Let η be the partition of X given by $\eta = \{\pi^{-1}(z)\colon z \in \mathbb{R}^d\}$. Let $A \in \mathcal{B}$ and $t > 0$. Then for m-a.e. $x \in X$, we have*

$$m_x^\eta(B^\phi(x,t) \cap A) \geq \limsup_{r\to 0} \frac{m\left(B^\phi(x,t) \cap A \cap B^\pi(x,r)\right)}{m\left(B^\pi(x,r)\right)} \tag{18}$$

and

$$m_x^\eta(U^\phi(x,t) \cap A) \leq \liminf_{r\to 0} \frac{m\left(U^\phi(x,t) \cap A \cap B^\pi(x,r)\right)}{m\left(B^\pi(x,r)\right)}, \tag{19}$$

where $B^\phi(x,t) := \phi^{-1}B(\phi x, t)$, $U^\phi(x,t) := \phi^{-1}U(\phi x, t)$, here $U(z,t)$ denotes the open ball in $\mathbb{R}^d$ centered at z of radius t.

Proof. Fix $A \in \mathcal{B}$ and $t > 0$. Similar to the proof of Lemma 3.2, for $n \in \mathbb{N}$, let $\mathcal{D}_n$ denote the collection

$$\mathcal{D}_n = \{[0, 2^{-n})^k + \alpha :\ \alpha \in 2^{-n}\mathbb{Z}^k\}.$$

For $y \in \mathbb{R}^k$, denote

$$W_n(y) = \bigcup_{Q \in \mathcal{D}_n\colon Q \cap B(y,t) \neq \emptyset} Q, \qquad \widehat{W}_n(y) = \bigcup_{Q \in \mathcal{D}_n\colon Q \subset U(y,t)} Q.$$

Write $\mathcal{W}_n := \{W_n(y) : \ y \in \mathbb{R}^k\}$ and $\widehat{\mathcal{W}_n} := \{\widehat{W}_n(y) : \ y \in \mathbb{R}^k\}$. It is clear that both $\mathcal{W}_n$ and $\widehat{\mathcal{W}_n}$ are countable for each $n \in \mathbb{N}$. Furthermore, we have $W_n(y) \downarrow B(y,t)$ and $\widehat{W}_n(y) \uparrow U(y,t)$ for each $y \in \mathbb{R}^k$ as $n \to \infty$. As a consequence, we have $\phi^{-1}W_n(\phi x) \downarrow B^\phi(x,t)$ and $\phi^{-1}\widehat{W}_n(\phi x) \uparrow U^\phi(x,t)$ for $x \in X$. Therefore

$$m_x^\eta(B^\phi(x,t) \cap A) = \lim_{n\to\infty} m_x^\eta(\phi^{-1}W_n(\phi x) \cap A)$$

and

$$m_x^\eta(U^\phi(x,t) \cap A) = \lim_{n\to\infty} m_x^\eta(\phi^{-1}\widehat{W}_n(\phi x) \cap A)$$

for each $x \in X$.

In the following we only prove (18). The proof of (19) is essentially identical. For $n \in \mathbb{N}$ and $F \in \mathcal{W}_n$, let $\Gamma_n(F) = \{x \in X : \ W_n(\phi x) = F\}$. Then for m-a.e. x and all $n \in \mathbb{N}$, we have

$$\begin{aligned}
m_x^\eta(\pi^{-1}W_n(\phi x) \cap A) &= \sum_{F \in \mathcal{W}_n} \chi_{\Gamma_n(F)}(x) m_x^\eta(\phi^{-1}F \cap A) \\
&= \sum_{F \in \mathcal{W}_n} \chi_{\Gamma_n(F)}(x) \mathbf{E}_m(\chi_{\phi^{-1}F\cap A}|\hat{\eta})(x) \\
&= \sum_{F \in \mathcal{W}_n} \chi_{\Gamma_n(F)}(x) \mathbf{E}_m(\chi_{\phi^{-1}F\cap A}|\pi^{-1}\gamma)(x) \\
&= \sum_{F \in \mathcal{W}_n} \chi_{\Gamma_n(F)}(x) \lim_{r\to 0} \frac{m\left(\phi^{-1}F \cap A \cap B^\pi(x,r)\right)}{m\left(B^\pi(x,r)\right)} \\
&\qquad (\text{ by Lemma 3.3}) \\
&= \lim_{r\to 0} \frac{m\left(\phi^{-1}W_n(\phi x) \cap A \cap B^\pi(x,r)\right)}{m\left(B^\pi(x,r)\right)} \\
&\geq \limsup_{r\to 0} \frac{m\left(B^\phi(x,t) \cap A \cap B^\pi(x,r)\right)}{m\left(B^\pi(x,r)\right)}.
\end{aligned}$$

Letting $n \to \infty$, we obtain (18). □

Remark 3.8. Under the condition of Lemma 3.7, assume that

$$g\colon \pi(X) \to g(\pi(X)) \subset \mathbb{R}^d$$

is a homeomorphism. Then we may replace the terms $B^\pi(x,r)$ in (18) and (19) by $B^{g\pi}(x,r)$. To see it, let $\pi' = g \circ \pi$. It is easy to see the partition η is just the same as $\{(\pi')^{-1}(z)\colon z \in \mathbb{R}^d\}$.

Proposition 3.9. *Let $T\colon X \to X$ be a measure-preserving transformation on $(X, \mathcal{B}, m)$, and let η be a measurable partition of X. Suppose that $\pi\colon X \to \mathbb{R}^d$ is* *a bounded $\mathcal{B}$-measurable function. Then for any $r > 0$,*

$$\lim_{n\to\infty} \frac{1}{n} \log m^{\eta}_{T^n x}(B^{\pi}(T^n x, r)) = 0 \quad \textit{for } m\textit{-a.e. } x \in X.$$

Proof. Fix $r > 0$ and $t > 0$. Since $\pi(X)$ is a bounded subset of $\mathbb{R}^d$, we can cover it by ℓ balls $B(\pi x_i, r/2)$ of radius $r/2$, where $x_i \in X$ and $i = 1, \dots, \ell$. Define

$$A_n = \{x \in X\colon\ m^{\eta}_x(B^{\pi}(x, r)) \le e^{-nt}\}, \qquad n \in \mathbb{N}.$$

If a ball $B^{\pi}(x_i, r/2)$ intersects A_n, then for any $y \in A_n \cap B^{\pi}(x_i, r/2)$, we have $B^{\pi}(x_i, r/2) \subset B^{\pi}(y, r)$ because $B(\pi x_i, r/2) \subset B(\pi y, r)$ by the triangle inequality. So the definition of A_n gives $m^{\eta}_y(A_n \cap B^{\pi}(x_i, r/2)) \le m^{\eta}_y(B^{\pi}(y, r)) \le e^{-nt}$. Hence

$$m(A_n \cap B^{\pi}(x_i, r/2)) = \int m^{\eta}_y(A_n \cap B^{\pi}(x_i, r/2))\, dm(y) \le e^{-nt}$$

and $m(A_n) \le \ell e^{-nt}$.

This estimate gives directly that $g(x) := \log m^{\eta}_x(B^{\pi}(x, r)) \in L^1(X, \mathcal{B}, m)$. Note that $g(T^n x) = \sum_{i=1}^{n} g(T^i x) - \sum_{i=1}^{n-1} g(T^i x)$. By the Birkhoff ergodic theorem we can get $\lim_{n\to\infty} \frac{1}{n} g(T^n x) = 0$ for m-a.e. $x \in X$, which is the desired result. □

Lemma 3.10. *Let $\mathcal{A}$ be a sub-σ-algebra of $\mathcal{B}$. Let $A \in \mathcal{B}$ with $m(A) > 0$. Then*

$$\mathbf{E}_m(\chi_A | \mathcal{A})(x) > 0$$

for m-a.e. $x \in A$.

Proof. Let $W := \{\mathbf{E}_m(\chi_A|\mathcal{A}) \le 0\}$. Then $W \in \mathcal{A}$. Hence

$$0 \ge \int_W \mathbf{E}_m(\chi_A|\mathcal{A})\, dm = \int_W \chi_A\, dm(x) = m(A \cap W),$$

which implies $m(A \cap W) = 0$. This finishes the proof. □

4. Projection measure-theoretic entropies associated with IFS

Throughout this section, let $\{S_i\}_{i=1}^{\ell}$ be an IFS on a closed set $X \subset \mathbb{R}^d$, and (Σ, σ) the one-sided full shift over $\{1, \dots, \ell\}$. Let $\mathcal{M}_\sigma(\Sigma)$ denote the collection of all σ-invariant Borel probability measures on Σ. Let $\pi\colon \Sigma \to \mathbb{R}^d$ be defined as in (1), and $h_\pi(\sigma, \cdot)$ as in Definition 2.1.

4.1. Some basic properties.

In this subsection, we present some basic properties of projection measure-theoretic entropy. Our first result is the following.

Proposition 4.1. (i) *$0 \le h_\pi(\sigma, m) \le h(\sigma, m)$ for every $m \in \mathcal{M}_\sigma(\Sigma)$, where $h(\sigma, m)$ denotes the classical measure-theoretic entropy of m.*

(ii) *The projection entropy function is affine on $\mathcal{M}_\sigma(\Sigma)$, i.e., for any $m_1, m_2 \in \mathcal{M}_\sigma(\Sigma)$ and any $0 \le p \le 1$, we have*

$$h_\pi(\sigma, pm_1 + (1-p)m_2) = ph_\pi(\sigma, m_1) + (1-p)h_\pi(\sigma, m_2). \tag{20}$$

The proof of the above proposition will be given later. Now let us recall some notation. If ξ is a partition of Σ, then $\widehat{\xi}$ denotes the σ-algebra generated by ξ. If $\xi_1, \ldots, \xi_n$ are countable partitions of Σ, then $\bigvee_{i=1}^n \xi_i$ denotes the partition consisting of sets $A_1 \cap \cdots \cap A_n$ with $A_i \in \xi_i$. Similarly for σ-algebras $\mathcal{A}_1, \mathcal{A}_2, \ldots$, $\bigvee_n \mathcal{A}_n$ denotes the σ-algebra generated by $\bigcup_n \mathcal{A}_n$.

Let $\mathcal{P}$ be the partition of Σ defined as in (3). Write $\mathcal{P}_0^n = \bigvee_{i=0}^n \sigma^{-i}\mathcal{P}$ for $n \geq 0$. Let γ denote the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$ on $\mathbb{R}^d$. Similar to Definition 2.1, we give the following definition.

Definition 4.2. Let $k \in \mathbb{N}$ and $\nu \in \mathcal{M}_{\sigma^k}(\Sigma)$. Define

$$h_\pi(\sigma^k, \nu) := H_\nu\left(\mathcal{P}_0^{k-1} \middle| \sigma^{-k}\pi^{-1}\gamma\right) - H_\nu\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma\right).$$

The term $h_\pi(\sigma^k, \nu)$ can be viewed as the projection measure-theoretic entropy of ν w.r.t. the IFS $\{S_{i_1} \circ \cdots \circ S_{i_k} : \ 1 \leq i_j \leq \ell \text{ for } 1 \leq j \leq k\}$. The following proposition exploits the connection between $h_\pi(\sigma^k, \nu)$ and $h_\pi(\sigma, m)$, where $m = \frac{1}{k}\sum_{i=0}^{k-1} \nu \circ \sigma^{-i}$.

Proposition 4.3. *Let $k \in \mathbb{N}$ and $\nu \in \mathcal{M}_{\sigma^k}(\Sigma)$. Set $m = \frac{1}{k}\sum_{i=0}^{k-1} \nu \circ \sigma^{-i}$. Then m is σ-invariant, and $h_\pi(\sigma, m) = \frac{1}{k} h_\pi(\sigma^k, \nu)$.*

To prove Propositions 4.1 and 4.3, we first give some lemmas about the (conditional) information and entropy (see §2 for the definitions).

Lemma 4.4 (cf. [48]). *Let m be a Borel probability measure on Σ. Let ξ, η be two countable Borel partitions of Σ with $H_m(\xi) < \infty$, $H_m(\eta) < \infty$, and $\mathcal{A}$ a sub-σ-algebra of $\mathcal{B}(\Sigma)$. Then we have*

(i) $\mathbf{I}_{m \circ \sigma^{-1}}(\xi|\mathcal{A}) \circ \sigma = \mathbf{I}_m(\sigma^{-1}\xi|\sigma^{-1}\mathcal{A})$.
(ii) $\mathbf{I}_m(\xi \vee \eta|\mathcal{A}) = \mathbf{I}_m(\xi|\mathcal{A}) + \mathbf{I}_m(\eta|\widehat{\xi} \vee \mathcal{A})$.
(iii) $H_m(\xi \vee \eta|\mathcal{A}) = H_m(\xi|\mathcal{A}) + H(\eta|\widehat{\xi} \vee \mathcal{A})$.
(iv) *If $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \cdots$ is an increasing sequence of sub-σ-algebras with $\mathcal{A}_n \uparrow \mathcal{A}$, then $\mathbf{I}_m(\xi|\mathcal{A}_n)$ converges almost everywhere and in L^1 to $\mathbf{I}_m(\xi|\mathcal{A})$. In particular,* $\lim_{n\to\infty} H_m(\xi|\mathcal{A}_n) = H_m(\xi|\mathcal{A})$.

Lemma 4.5. *Denote $g(x) = -x\log x$ for $x \geq 0$. For any integer $k \geq 2$ and $x_1, \ldots, x_k \geq 0$, we have $\frac{1}{k}\sum_{i=1}^k g(x_i) \leq g\left(\frac{1}{k}\sum_{i=1}^k x_i\right) \leq \sum_{i=1}^k g(x_i/k)$ and*

$$\sum_{i=1}^k g(x_i) - (x_1 + \ldots + x_k)\log k \leq g(x_1 + \ldots + x_k) \leq \sum_{i=1}^k g(x_i). \tag{21}$$

Moreover for any $p_1, p_2 \geq 0$ with $p_1 + p_2 = 1$,

$$\sum_{j=1}^2 p_j g(x_j) \leq g\left(\sum_{j=1}^2 p_j x_j\right) \leq \sum_{j=1}^2 p_j g(x_j) + g(p_j)x_j. \tag{22}$$

Proof. Standard. □

Lemma 4.6. *Let m be a Borel probability measure on Σ. Assume ξ and η are two countable Borel partitions of Σ such that each member in ξ intersects at most k members of η. Then $H_m(\xi) \geq H_m(\xi \vee \eta) - \log k$.*

Proof. Although the result is standard, we give a short proof for the convenience of the reader. Denote $g(x) = -x \log x$ for $x \in [0,1]$. Then

$$\begin{aligned}
H_m(\xi) &= \sum_{A\in\xi} g(m(A)) = \sum_{A\in\xi} g\left(\sum_{B\in\eta,\ B\cap A\neq\emptyset} m(A\cap B)\right)\\
&\geq \sum_{A\in\xi}\left[\left(\sum_{B\in\eta,\ B\cap A\neq\emptyset} g(m(A\cap B))\right) - m(A)\log k\right] \quad \text{(by (21))}\\
&\geq \left(\sum_{A\in\xi}\sum_{B\in\eta} g(m(A\cap B))\right) - \log k\\
&= H_m(\xi\vee\eta) - \log k.
\end{aligned}$$

This finishes the proof. □

The following simple lemma plays an important role in our analysis.

Lemma 4.7. $\widehat{\mathcal{P}} \vee \sigma^{-1}\pi^{-1}\gamma = \widehat{\mathcal{P}} \vee \pi^{-1}\gamma$.

Proof. We only prove $\widehat{\mathcal{P}}\vee\sigma^{-1}\pi^{-1}\gamma \subseteq \widehat{\mathcal{P}}\vee\pi^{-1}\gamma$. The other direction can be proved by an essentially identical argument. Note that each member in $\widehat{\mathcal{P}} \vee \sigma^{-1}\pi^{-1}\gamma$ can be written as

$$\bigcup_{j=1}^{\ell} [j] \cap \sigma^{-1}\pi^{-1}A_j$$

with $A_j \in \gamma$. However, it is direct to check that

$$[j] \cap \sigma^{-1}\pi^{-1}A_j = [j] \cap \pi^{-1}(S_j(A_j)).$$

Since S_j is injective and contractive (thus continuous), we have $S_j(A_j) \in \gamma$. Therefore $\bigcup_{j=1}^{\ell}[j] \cap \sigma^{-1}\pi^{-1}A_j \in \widehat{\mathcal{P}} \vee \pi^{-1}\gamma$. □

Lemma 4.8. *Let m be a Borel probability measure on Σ and $k \in \mathbb{N}$. We have*

$$\begin{aligned}
&H_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-k}\pi^{-1}\gamma\right) - H_m\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right)\\
&\quad= \sum_{j=0}^{k-1} H_{m\circ\sigma^{-j}}(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - H_{m\circ\sigma^{-j}}(\mathcal{P}|\pi^{-1}\gamma).
\end{aligned}$$

Moreover if $m \in \mathcal{M}_\sigma(\Sigma)$, then

$$H_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-k}\pi^{-1}\gamma\right) - H_m\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right) = kh_\pi(\sigma, m).$$

Proof. For $j=0,1,\dots,k-1$, we have

$$\begin{aligned}
& \mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-j}\pi^{-1}\gamma\right)-\mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-(j+1)}\pi^{-1}\gamma\right)\\
= & \mathbf{I}_m\left(\sigma^{-j}\mathcal{P}\middle|\sigma^{-j}\pi^{-1}\gamma\right)+\mathbf{I}_m\left(\bigvee_{0\le i\le k-1,\, i\neq j}\sigma^{-i}\mathcal{P}\middle|\sigma^{-j}\widehat{\mathcal{P}}\vee\sigma^{-j}\pi^{-1}\gamma\right)\\
& -\mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-(j+1)}\pi^{-1}\gamma\right) \qquad \text{(by Lemma 4.4(ii))}\\
= & \mathbf{I}_m\left(\sigma^{-j}\mathcal{P}\middle|\sigma^{-j}\pi^{-1}\gamma\right)+\mathbf{I}_m\left(\bigvee_{0\le i\le k-1,\, i\neq j}\sigma^{-i}\mathcal{P}\middle|\sigma^{-j}\widehat{\mathcal{P}}\vee\sigma^{-(j+1)}\pi^{-1}\gamma\right)\\
& -\mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-(j+1)}\pi^{-1}\gamma\right) \qquad \text{(by Lemma 4.7)}\\
= & \mathbf{I}_m\left(\sigma^{-j}\mathcal{P}\middle|\sigma^{-j}\pi^{-1}\gamma\right)-\mathbf{I}_m\left(\sigma^{-j}\mathcal{P}\middle|\sigma^{-(j+1)}\pi^{-1}\gamma\right) \qquad \text{(by Lemma 4.4(ii))}\\
= & \mathbf{I}_{m\circ\sigma^{-j}}\left(\mathcal{P}\middle|\pi^{-1}\gamma\right)\circ\sigma^j-\mathbf{I}_{m\circ\sigma^{-j}}\left(\mathcal{P}\middle|\sigma^{-1}\pi^{-1}\gamma\right)\circ\sigma^j \qquad \text{(by Lemma 4.4(i))}.
\end{aligned}$$

Summing j over $\{0,\dots,k-1\}$ yields

$$\begin{aligned}
& \mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right)-\mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\sigma^{-k}\pi^{-1}\gamma\right)\\
& =\sum_{j=0}^{k-1}\left(\mathbf{I}_{m\circ\sigma^{-j}}\left(\mathcal{P}\middle|\pi^{-1}\gamma\right)\circ\sigma^j-\mathbf{I}_{m\circ\sigma^{-j}}\left(\mathcal{P}\middle|\sigma^{-1}\pi^{-1}\gamma\right)\circ\sigma^j\right).
\end{aligned} \tag{23}$$

Taking integration, we obtain the desired formula. □

For any $n\in\mathbb{N}$, let $\mathcal{D}_n$ be the partition of $\mathbb{R}^d$ given by

$$\mathcal{D}_n=\{[0,2^{-n})^d+\alpha:\ \alpha\in 2^{-n}\mathbb{Z}^d\}. \tag{24}$$

Lemma 4.9. *Let $m\in\mathcal{M}_\sigma(\Sigma)$. For each $n\in\mathbb{N}$, we have*

$$H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n})-H_m(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n})\ge -d\log(\sqrt{d}+1).$$

Proof. Since m is σ-invariant, by Lemma 4.4(iii), we have

$$\begin{aligned}
& H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n})-H_m(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n})\\
& =H_m(\mathcal{P}\vee\sigma^{-1}\pi^{-1}\mathcal{D}_n)-H_m(\sigma^{-1}\pi^{-1}\mathcal{D}_n)\\
& \quad -H_m(\mathcal{P}\vee\pi^{-1}\mathcal{D}_n)+H_m(\pi^{-1}\mathcal{D}_n)\\
& =H_m(\mathcal{P}\vee\sigma^{-1}\pi^{-1}\mathcal{D}_n)-H_m(\mathcal{P}\vee\pi^{-1}\mathcal{D}_n).
\end{aligned} \tag{25}$$

Observe that for each $1\le j\le\ell$ and $Q\in\mathcal{D}_n$,

$$[j]\cap\sigma^{-1}\pi^{-1}(Q)=[j]\cap\pi^{-1}(S_j(Q)).$$

Since S_j is contractive, $\operatorname{diam}(S_j(Q))\le 2^{-n}\sqrt{d}$ and thus $S_j(Q)$ intersects at most $(\sqrt{d}+1)^d$ members in $\mathcal{D}_n$. It deduces that $[j]\cap\sigma^{-1}\pi^{-1}(Q)$ intersects at most

$(\sqrt{d}+1)^d$ members in $\mathcal{P}\vee\pi^{-1}\mathcal{D}_n$. By Lemma 4.6, we have

$$\begin{aligned}H_m(\mathcal{P}\vee\sigma^{-1}\pi^{-1}\mathcal{D}_n) &\geq H_m(\mathcal{P}\vee\sigma^{-1}\pi^{-1}\mathcal{D}_n\vee\pi^{-1}\mathcal{D}_n)\\ &\qquad -d\log(\sqrt{d}+1)\\ &\geq H_m(\mathcal{P}\vee\pi^{-1}\mathcal{D}_n)-d\log(\sqrt{d}+1).\end{aligned}\tag{26}$$

Combining it with (25) yields the desired inequality. □

Proof of Proposition 4.1. We first prove part (i) of the proposition, i.e.,

$$0\leq h_\pi(\sigma,m)\leq h(\sigma,m).$$

Since $\widehat{\mathcal{D}_n}\uparrow\gamma$ as n tends to ∞, by Lemma 4.4(iv), we have

$$\lim_{n\to\infty}H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\mathcal{D}_n)-H_m(\mathcal{P}|\pi^{-1}\mathcal{D}_n)=H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma)-H_m(\mathcal{P}|\pi^{-1}\gamma).$$

It together with Lemma 4.9 yields

$$H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma)-H_m(\mathcal{P}|\pi^{-1}\gamma)\geq -d\log(\sqrt{d}+1).$$

Using the same argument to the IFS $\{S_{i_1\ldots i_k}:1\leq i_j\leq\ell,1\leq j\leq k\}$, we have

$$H_m\left(\mathcal{P}_0^{k-1}|\sigma^{-k}\pi^{-1}\gamma\right)-H_m\left(\mathcal{P}_0^{k-1}|\pi^{-1}\gamma\right)\geq -d\log(\sqrt{d}+1).$$

It together with Lemma 4.8 yields $h_\pi(\sigma,m)\geq -d\log(\sqrt{d}+1)/k$. Since k is arbitrary, we have $h_\pi(\sigma,m)\geq 0$. To see $h_\pi(\sigma,m)\leq h(\sigma,m)$, it suffices to observe that

$$\begin{aligned}kh_\pi(\sigma,m) &= H_m\left(\mathcal{P}_0^{k-1}|\sigma^{-k}\pi^{-1}\gamma\right)-H_m\left(\mathcal{P}_0^{k-1}|\pi^{-1}\gamma\right)\\ &\leq H_m\left(\mathcal{P}_0^{k-1}|\sigma^{-k}\pi^{-1}\gamma\right)\leq H_m\left(\mathcal{P}_0^{k-1}\right).\end{aligned}$$

Now we turn to the proof of part (ii). Let $m_1,m_2\in\mathcal{M}_\sigma(\Sigma)$ and $m=pm_1+(1-p)m_2$ for some $p\in[0,1]$. Using (22), for any finite or countable Borel partition ξ we have

$$|H_m(\xi)-pH_{m_1}(\xi)-(1-p)H_{m_2}(\xi)|\leq g(p)+g(1-p)\leq\log 2.\tag{27}$$

Let $k\in\mathbb{N}$. By Lemma 4.8, Lemma 4.4(iv), and (25), we have

$$\begin{aligned}h_\pi(\sigma,m) &= \frac{1}{k}\left(H_m\left(\mathcal{P}_0^{k-1}|\sigma^{-k}\pi^{-1}\gamma\right)-H_m\left(\mathcal{P}_0^{k-1}|\pi^{-1}\gamma\right)\right)\\ &= \frac{1}{k}\lim_{n\to\infty}\left(H_m\left(\mathcal{P}_0^{k-1}|\sigma^{-k}\pi^{-1}\widehat{\mathcal{D}_n}\right)-H_m\left(\mathcal{P}_0^{k-1}|\pi^{-1}\widehat{\mathcal{D}_n}\right)\right)\\ &= \frac{1}{k}\lim_{n\to\infty}\left(H_m\left(\mathcal{P}_0^{k-1}\vee\sigma^{-k}\pi^{-1}\mathcal{D}_n\right)-H_m\left(\mathcal{P}_0^{k-1}\vee\pi^{-1}\mathcal{D}_n\right)\right).\end{aligned}\tag{28}$$

The above statement is true when m is replaced by m_1 and m_2. However by (27),

$$H_m\left(\mathcal{P}_0^{k-1}\vee\sigma^{-k}\pi^{-1}\mathcal{D}_n\right)-H_m\left(\mathcal{P}_0^{k-1}\vee\pi^{-1}\mathcal{D}_n\right)$$

differs from

$$\sum_{j=1}^{2} p_j \left[H_{m_j}\left(\mathcal{P}_0^{k-1} \vee \sigma^{-k}\pi^{-1}\mathcal{D}_n\right) - H_{m_j}\left(\mathcal{P}_0^{k-1} \vee \pi^{-1}\mathcal{D}_n\right)\right]$$

at most $2\log 2$, where $p_1 = p$ and $p_2 = 1-p$. This together with (28) yields (20). $\square$

Proof of Proposition 4.3. Let $k \geq 2$ and $\nu \in \mathcal{M}_{\sigma^k}(\Sigma)$. We claim that $h_\pi(\sigma^k, \nu \circ \sigma^{-j}) = h_\pi(\sigma^k, \nu)$ for any $1 \leq j \leq k-1$. To prove the claim, it suffices to prove $h_\pi(\sigma^k, \nu\circ\sigma^{-1}) = h_\pi(\sigma^k, \nu)$. Note that both ν and $\nu \circ \sigma^{-1}$ are σ^k-invariant. By Lemma 4.8, we have

$$\begin{aligned} h_\pi(\sigma^k, \nu) &= H_\nu\left(\mathcal{P}_0^{k-1} \middle| \sigma^{-k}\pi^{-1}\gamma\right) - H_\nu\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right) \\ &= \sum_{j=0}^{k-1}\left(H_{\nu\circ\sigma^{-j}}\left(\mathcal{P}\middle|\sigma^{-1}\pi^{-1}\gamma\right) - H_{\nu\circ\sigma^{-j}}\left(\mathcal{P}\middle|\pi^{-1}\gamma\right)\right), \end{aligned}$$

whilst

$$\begin{aligned} h_\pi(\sigma^k, \nu\circ\sigma^{-1}) &= H_{\nu\circ\sigma^{-1}}\left(\mathcal{P}_0^{k-1} \middle| \sigma^{-k}\pi^{-1}\gamma\right) - H_{\nu\circ\sigma^{-1}}\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right) \\ &= \sum_{j=0}^{k-1}\left(H_{\nu\circ\sigma^{-j-1}}\left(\mathcal{P}\middle|\sigma^{-1}\pi^{-1}\gamma\right) - H_{\nu\circ\sigma^{-j-1}}\left(\mathcal{P}\middle|\pi^{-1}\gamma\right)\right). \end{aligned}$$

Since ν is σ^k-invariant, we obtain $h_\pi(\sigma^k, \nu\circ\sigma^{-1}) = h_\pi(\sigma^k, \nu)$. This finishes the proof of the claim. To complete the proof of the proposition, let $m = \frac{1}{k}\sum_{i=0}^{k-1} \nu \circ \sigma^{-i}$. It is clear that m is σ-invariant. By Proposition 4.1(ii), $h_\pi(\sigma^k, \cdot)$ is affine on $\mathcal{M}_{\sigma^k}(\Sigma)$. Hence

$$h_\pi(\sigma^k, m) = \frac{1}{k}\sum_{i=0}^{k-1} h_\pi(\sigma^k, \nu\circ\sigma^{-i}) = h_\pi(\sigma^k, \nu).$$

Combining it with Lemma 4.8 yields the equality $h_\pi(\sigma, m) = \frac{1}{k}h_\pi(\sigma^k, \nu)$. $\square$

4.2. **A version of Shannon-McMillan-Breiman Theorem associated with IFS.** In this subsection, we prove the following Shannon-McMillan-Breiman type theorem associated with IFS, which is needed in the proof of Theorem 2.11. It is also of independent interest.

Proposition 4.10. *Let $\{S_i\}_{i=1}^{\ell}$ be an IFS and $m \in \mathcal{M}_\sigma(\Sigma)$. Then*

$$\lim_{k\to\infty}\frac{1}{k}\mathbf{I}_m\left(\mathcal{P}_0^{k-1}\middle|\pi^{-1}\gamma\right)(x) = \mathbf{E}_m(f|\mathcal{I})(x) = h(\sigma, m, x) - h_\pi(\sigma, m, x). \tag{29}$$

almost everywhere and in L^1, where

$$f := \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma)) + \mathbf{I}_m(\mathcal{P}|\pi^{-1}\gamma) - \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma),$$

$\mathcal{I} = \{B \in \mathcal{B}(\Sigma) :\ \sigma^{-1}B = B\}$, and $h(\sigma, m, x)$, $h_\pi(\sigma, m, x)$ denote the classical local entropy and the local projection entropy of m at x (see Definition 2.1), respectively. Moreover if m is ergodic, then the limit in (29) equals $h(\sigma, m) - h_\pi(\sigma, m)$ for m-a.e. $x \in \Sigma$.

Remark 4.11. If ξ is a countable Borel partition of Σ, and $\mathcal{A} \subset \mathcal{B}(\Sigma)$ is a sub-σ-algebra with $\sigma^{-1}\mathcal{A} = \mathcal{A}$, then the relativized Shannon-McMillan-Breiman Theorem states that

$$\lim_{k\to\infty} \frac{1}{k}\mathbf{I}_m\left(\xi_0^{k-1}|\mathcal{A}\right)(x) = \mathbf{E}_m(g|\mathcal{I})(x) \qquad \text{for } m\text{-a.e. } x \in \Sigma,$$

where $g = \mathbf{I}_m\left(\xi|\mathcal{A} \vee \xi_1^\infty\right)$ (see, e.g., [8, Lemma 4.1]). However under the setting of Proposition 4.10, the sub-σ-algebra $\pi^{-1}\gamma$ is not invariant in general.

In the following we present a generalized version of Proposition 4.10.

Proposition 4.12. *Let ξ be a countable Borel partition of Σ with $H_m(\xi) < \infty$, and let $\mathcal{A} \subset \mathcal{B}(\Sigma)$ be a sub-σ-algebra so that $\widehat{\xi} \vee \sigma^{-1}\mathcal{A} = \widehat{\xi} \vee \mathcal{A}$. Let $m \in \mathcal{M}_\sigma(\Sigma)$. Then*

$$\lim_{k\to\infty} \frac{1}{k}\mathbf{I}_m\left(\xi_0^{k-1}|\mathcal{A}\right)(x) = \mathbf{E}_m(f|\mathcal{I})(x) \tag{30}$$

almost everywhere and in L^1, where

$$f := \mathbf{I}_m\left(\xi|\sigma^{-1}\mathcal{A} \vee \bigvee_{i=1}^{\infty} \sigma^{-i}\widehat{\xi}\right) + \mathbf{I}_m(\xi|\mathcal{A}) - \mathbf{I}_m(\xi|\sigma^{-1}\mathcal{A}),$$

and $\mathcal{I} = \{B \in \mathcal{B}(\Sigma) :\ \sigma^{-1}B = B\}$.

To prove Proposition 4.12, we need the following lemma.

Lemma 4.13 ([42], Corollary 1.6, p. 96)**.** *Let $m \in \mathcal{M}_\sigma(\Sigma)$. Let $F_k \in L^1(\Sigma, m)$ be a sequence that converges almost everywhere and in L^1 to $F \in L^1(\Sigma, m)$. Then*

$$\lim_{k\to\infty} \frac{1}{k}\sum_{j=0}^{k-1} F_{k-j}(\sigma^j(x)) = \mathbf{E}_m(F|\mathcal{I})(x)$$

almost everywhere and in L^1.

Proof of Proposition 4.12. For $k \geq 2$ and $x \in \Sigma$, we write

$$g_k(x) = \mathbf{I}_m\left(\xi_0^{k-1}|\mathcal{A}\right)(x) - \mathbf{I}_m\left(\xi_0^{k-2}|\mathcal{A}\right)(\sigma x).$$

Then

$$\mathbf{I}_m\left(\xi_0^{k-1}|\mathcal{A}\right)(x) = \mathbf{I}_m(\xi|\mathcal{A})(\sigma^{k-1}x) + \sum_{j=0}^{k-2} g_{k-j}(\sigma^j x). \tag{31}$$

We claim that

$$g_k(x) = \mathbf{I}_m\left(\xi\big|\sigma^{-1}\mathcal{A} \vee \bigvee_{i=1}^{k-1} \sigma^{-i}\widehat{\xi}\right)(x) + \mathbf{I}_m\left(\xi\big|\mathcal{A}\right)(x) - \mathbf{I}_m\left(\xi\big|\sigma^{-1}\mathcal{A}\right)(x). \tag{32}$$

By the claim and Lemma 4.4(iv), g_k converges almost everywhere and in L^1 to f. It together with (31) and Lemma 4.13 yields (30).

Now we turn to the proof of (32). Let $k \geq 2$. We have

(33)
$$\begin{aligned}\mathbf{I}_m\left(\xi_0^{k-1}\middle|\sigma^{-1}\mathcal{A}\right)(x) &= \mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\right)(x) + \mathbf{I}_m\left(\bigvee_{i=1}^{k-1}\sigma^{-i}\xi\middle|\sigma^{-1}\mathcal{A}\vee\widehat{\xi}\right)(x)\\ &= \mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\right)(x) + \mathbf{I}_m\left(\bigvee_{i=1}^{k-1}\sigma^{-i}\xi\middle|\mathcal{A}\vee\widehat{\xi}\right)(x),\end{aligned}$$

using the property $\sigma^{-1}\mathcal{A}\vee\widehat{\xi} = \mathcal{A}\vee\widehat{\xi}$. Meanwhile, we have

$$\begin{aligned}&\mathbf{I}_m\left(\xi_0^{k-1}\middle|\sigma^{-1}A\right)(x)\\ &= \mathbf{I}_m\left(\bigvee_{i=1}^{k-1}\sigma^{-i}\xi\middle|\sigma^{-1}\mathcal{A}\right)(x) + \mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\vee\bigvee_{i=1}^{k-1}\sigma^{-i}\widehat{\xi}\right)(x)\\ &= \mathbf{I}_m\left(\xi_0^{k-2}\middle|\mathcal{A}\right)(\sigma x) + \mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\vee\bigvee_{i=1}^{k-1}\sigma^{-i}\widehat{\xi}\right)(x).\end{aligned}\tag{34}$$

Combining (33) with (34) yields

$$\begin{aligned}&\mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\right)(x) + \mathbf{I}_m\left(\bigvee_{i=1}^{k-1}\sigma^{-i}\xi\middle|\mathcal{A}\vee\widehat{\xi}\right)(x)\\ &= \mathbf{I}_m\left(\xi_0^{k-2}\middle|\mathcal{A}\right)(\sigma x) + \mathbf{I}_m\left(\xi\middle|\sigma^{-1}\mathcal{A}\vee\bigvee_{i=1}^{k-1}\sigma^{-i}\widehat{\xi}\right)(x).\end{aligned}\tag{35}$$

However

$$\mathbf{I}_m\left(\xi_0^{k-1}\middle|\mathcal{A}\right)(x) = \mathbf{I}_m\left(\xi\middle|\mathcal{A}\right)(x) + \mathbf{I}_m\left(\bigvee_{i=1}^{k-1}\sigma^{-i}\xi\middle|\mathcal{A}\vee\widehat{\xi}\right)(x).\tag{36}$$

Combining (35) with (36) yields (32). This finishes the proof of Proposition 4.12. □

We remark that Proposition 4.10 can be stated in terms of conditional measures. To see it, let

$$\eta = \{\pi^{-1}(z) : z \in \mathbb{R}^d\}$$

be the measurable partition of Σ generated by the canonical projection π associated with $\{S_i\}_{i=1}^{\ell}$. For $m \in \mathcal{M}_\sigma(\Sigma)$, let $\{m_x^\eta\}_{x\in\Sigma}$ denote the canonical system of conditional measures w.r.t. η. For $x \in \Sigma$ and $k \in \mathbb{N}$, let $\mathcal{P}_0^k(x)$ denote the element in the partition $\mathcal{P}_0^k$ containing x. Then Proposition 4.10 can be restated as the following.

Proposition 4.14. *For $m \in \mathcal{M}_\sigma(\Sigma)$, we have*

$$-\lim_{k\to\infty} \frac{1}{k} \log m_x^\eta(\mathcal{P}_0^k(x)) = \mathbf{E}_m(f|\mathcal{I})(x) \qquad \text{for } m\text{-a.e. } x \in \Sigma, \tag{37}$$

where $f := \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma)) + \mathbf{I}_m(\mathcal{P}|\pi^{-1}\gamma) - \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma)$. Moreover if m is ergodic, then the limit in (37) equals $h(\sigma, m) - h_\pi(\sigma, m)$ for m-a.e. $x \in \Sigma$.

Proof. It suffices to show that for each $k \in \mathbb{N}$,

$$\log m_x^\eta(\mathcal{P}_0^k(x)) = -\mathbf{I}_m(\mathcal{P}_0^k|\pi^{-1}\gamma)(x) \text{ almost everywhere}.$$

To see this, by Theorem 3.1 we have

$$\sum_{A\in\mathcal{P}_0^k} \chi_A(x) m_x^\eta(A) = \sum_{A\in\mathcal{P}_0^k} \chi_A(x)\mathbf{E}_m(\chi_A|\pi^{-1}\gamma)(x) \quad \text{ for } m\text{-a.e. } x \in \Sigma.$$

Taking logarithm yields the desired result. □

Remark 4.15. In Proposition 4.14, for m-a.e. $x \in \Sigma$, we have

$$\lim_{k\to\infty} -\frac{1}{k} \log m_x^\eta(\mathcal{P}_0^k(y)) = \mathbf{E}_m(f|\mathcal{I})(y) \quad \text{ for } m_x^\eta\text{-a.e. } y \in \eta(x).$$

To see this, denote

$$R = \left\{ y \in \Sigma : \ -\lim_{k\to\infty} \frac{1}{k} \log m_y^\eta(\mathcal{P}_0^k(y)) = \mathbf{E}_m(f|\mathcal{I})(y) \right\}.$$

Then $1 = m(R) = \int m_x^\eta(R \cap \eta(x))\, dm(x)$. Hence $m_x^\eta(R \cap \eta(x)) = 1$ m-a.e. For $y \in R \cap \eta(x)$, we have

$$\lim_{k\to\infty} -\frac{1}{k} \log m_x^\eta(\mathcal{P}_0^k(y)) = \lim_{k\to\infty} -\frac{1}{k} \log m_y^\eta(\mathcal{P}_0^k(y)) = \mathbf{E}_m(f|\mathcal{I})(y).$$

As a corollary of Proposition 4.14, we have

Corollary 4.16. *Let $m \in \mathcal{M}_\sigma(\Sigma)$. Then*

$$\begin{aligned} h_\pi(\sigma, m) = h(\sigma, m) \quad &\Longleftrightarrow \quad \lim_{k\to\infty} \frac{1}{k} \log m_x^\eta(\mathcal{P}_0^k(x)) = 0 \ \ m\text{-a.e.} \\ &\Longleftrightarrow \quad \dim_H m_x^\eta = 0 \ \ m\text{-a.e.} \end{aligned}$$

In particular, if $\dim_H \pi^{-1}(z) = 0$ for each $z \in \mathbb{R}^d$, then $h_\pi(\sigma, m) = h(\sigma, m)$. Here $\dim_H$ denotes the Hausdorff dimension.

Proof. Let f be defined as in Proposition 4.14. Then

$$\int \mathbf{E}_m(f|\mathcal{I})\, dm = \int f\, dm = h(\sigma, m) - h_\pi(\sigma, m).$$

By (37), $\mathbf{E}_m(f|\mathcal{I})(x) \geq 0$ for m-a.e. $x \in \Sigma$. Hence we have

$$\begin{aligned} h(\sigma, m) = h_\pi(\sigma, m) \quad &\Longleftrightarrow \quad \mathbf{E}_m(f|\mathcal{I}) = 0 \ \ m\text{-a.e.} \\ &\Longleftrightarrow \quad \lim_{k\to\infty} \frac{1}{k} \log m_x^\eta(\mathcal{P}_0^k(x)) = 0 \ \ m\text{-a.e.} \end{aligned}$$

Using dimension theory of measures (see, e.g., [14]), we have

$$\dim_H m_x^\eta = \operatorname{ess\,sup}_{y\in\eta(x)} \liminf_{k\to\infty} \frac{\log m_x^\eta(\mathcal{P}_0^k(y))}{\log \ell^{-k}}.$$

It together with Remark 4.15 yields

$$\mathbf{E}_m(f|\mathcal{I}) = 0 \ \ m\text{-a.e.} \Longleftrightarrow \dim_H m_x^\eta = 0 \ \ m\text{-a.e.}$$

This finishes the proof of the first part of the corollary.

To complete the proof, assume that $\dim_H \pi^{-1}(z) = 0$ for each $z \in \mathbb{R}^d$. Then for each $x \in \Sigma$, $\dim_H \eta(x) = 0$ and hence $\dim_H m_x^\eta = 0$. Thus $h_\pi(\sigma, m) = h(\sigma, m)$. $\square$

4.3. **Projection entropy under the ergodic decomposition.** In this subsection, we first prove the following result.

Proposition 4.17. *Let $\{S_i\}_{i=1}^\ell$ be an IFS and $m \in \mathcal{M}_\sigma(\Sigma)$. Assume that $m = \int \nu \, d\mathbb{P}(\nu)$ is the ergodic decomposition of m. Then*

$$h_\pi(\sigma, m) = \int h_\pi(\sigma, \nu) \, d\mathbb{P}(\nu).$$

Proof. Let $\mathcal{I}$ denote the σ-algebra $\{B \in \mathcal{B}(\Sigma) : \ \sigma^{-1}B = B\}$, and let $m \in \mathcal{M}_\sigma(\Sigma)$. Then there exists an m-measurable partition ε of Σ such that $\widehat{\varepsilon} = \mathcal{I}$ modulo sets of zero m-measure (see [49, pp. 37-38]). Let $\{m_x^\varepsilon\}$ denote the conditional measures of m associated with the partition ε. Then $m = \int m_x^\varepsilon \, dm(x)$ is just the ergodic decomposition of m (see e.g., [32, Theorem 2.3.3]). Hence to prove the proposition, we need to show that

$$h_\pi(\sigma, m) = \int h_\pi(\sigma, m_x^\varepsilon) \, dm(x). \tag{38}$$

We first show the direction "$\leq$" in (38). Note that $\mathcal{I}$ is σ-invariant and $\widehat{\mathcal{P}} \vee \sigma^{-1}\pi^{-1}\gamma = \widehat{\mathcal{P}} \vee \pi^{-1}\gamma$. Hence we have $\widehat{\mathcal{P}} \vee \sigma^{-1}\pi^{-1}\gamma \vee \mathcal{I} = \widehat{\mathcal{P}} \vee \pi^{-1}\gamma \vee \mathcal{I}$. Taking $\xi = \mathcal{P}$ and $\mathcal{A} = \pi^{-1}\gamma \vee \mathcal{I}$ in Proposition 4.12 yields

$$\lim_{k\to\infty} \frac{1}{k} \mathbf{I}_m \left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma \vee \mathcal{I}\right)(x) = \mathbf{E}_m(f|\mathcal{I})(x) \tag{39}$$

almost everywhere and in L^1, where

$$f := \mathbf{I}_m\left(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma)\right) + \mathbf{I}_m(\mathcal{P}|\pi^{-1}\gamma \vee \mathcal{I}) - \mathbf{I}_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma \vee \mathcal{I}).$$

By Remark 3.4(ii), we have

$$\mathbf{I}_{m_x^\varepsilon}\left(\mathcal{P}_0^{k-1}|\pi^{-1}\gamma\right)(x) = \mathbf{I}_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma \vee \mathcal{I}\right)(x).$$

Hence according to the ergodicity of m_x^ε and Proposition 4.10, we have

$$\begin{aligned} h(\sigma, m_x^\varepsilon) - h_\pi(\sigma, m_x^\varepsilon) &= \lim_{k\to\infty} \frac{1}{k} \mathbf{I}_{m_x^\varepsilon}\left(\mathcal{P}_0^{k-1}|\pi^{-1}\gamma\right)(x) \\ &= \lim_{k\to\infty} \frac{1}{k} \mathbf{I}_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma \vee \mathcal{I}\right)(x) \end{aligned}$$

almost everywhere and

$$\int h(\sigma, m_x^\varepsilon) - h_\pi(\sigma, m_x^\varepsilon)\, dm(x) = \lim_{k\to\infty} \frac{1}{k} H_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma \vee \mathcal{I}\right). \tag{40}$$

Using Proposition 4.10 again we have

$$h(\sigma, m) - h_\pi(\sigma, m) = \lim_{k\to\infty} \frac{1}{k} H_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma\right). \tag{41}$$

However, $H_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma \vee \mathcal{I}\right) \leq H_m\left(\mathcal{P}_0^{k-1} \middle| \pi^{-1}\gamma\right)$ (see e.g. [63, Theorem 4.3 (v)]). By (40), (41) and the above inequality, we have

$$\int h(\sigma, m_x^\varepsilon) - h_\pi(\sigma, m_x^\varepsilon)\, dm(x) \leq h(\sigma, m) - h_\pi(\sigma, m).$$

It is well known (see [63, Theorem 8.4]) that $\int h(\sigma, m_x^\varepsilon)\, dm(x) = h(\sigma, m)$. Hence we obtain the inequality $h_\pi(\sigma, m) \leq \int h_\pi(\sigma, m_x^\varepsilon)\, dm(x)$.

Now we prove the direction "$\geq$ " in (38). For any $n \in \mathbb{N}$, let $\mathcal{D}_n$ be defined as in (24). Since $\widehat{\mathcal{D}_n} \uparrow \gamma$, we have

$$h_\pi(\sigma, m) = \lim_{n\to\infty} H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n}) - H_m(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n}). \tag{42}$$

Now fix $n \in \mathbb{N}$ and denote $A(m) = H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n}) - H_m(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n})$ and

$$\begin{aligned} B(m) &= H_m(\sigma^{-1}\pi^{-1}\mathcal{D}_n|\mathcal{P} \vee \pi^{-1}\widehat{\mathcal{D}_n}) \\ &= H_m(\mathcal{P} \vee \sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n} \vee \pi^{-1}\widehat{\mathcal{D}_n}) - H_m(\mathcal{P} \vee \pi^{-1}\widehat{\mathcal{D}_n}). \end{aligned}$$

Then by (25) and (26), we have

$$B(m) - c \leq A(m) \leq B(m), \tag{43}$$

where $c = d\log(\sqrt{d}+1)$. As a conditional entropy function, $B(m)$ is concave on $\mathcal{M}_\sigma(\Sigma)$ (see, e.g., [26, Lemma 3.3 (1)]). Hence by Jensen's inequality and (43), we have

$$A(m) \geq B(m) - c \geq \int B(m_x^\varepsilon)\, dm(x) - c \geq \int A(m_x^\varepsilon)\, dm(x) - c.$$

That is,

$$\begin{aligned} &H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n}) - H_m(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n}) \\ &\quad\geq \int H_{m_x^\varepsilon}(\mathcal{P}|\sigma^{-1}\pi^{-1}\widehat{\mathcal{D}_n}) - H_{m_x^\varepsilon}(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n})\, dm(x) - c. \end{aligned}$$

Letting $n \to \infty$, using (42) and Lebesgue dominated convergence theorem, we have

$$h_\pi(\sigma, m) \geq \int h_\pi(\sigma, m_x^\varepsilon)\, dm(x) - c.$$

Replacing σ by σ^k we have

$$h_\pi(\sigma^k, m) \geq \int h_\pi(\sigma^k, m_x^{\varepsilon_k})\, dm(x) - c, \tag{44}$$

where ε_k denotes a measurable partition of Σ such that

$$\widehat{\varepsilon_k} = \{B \in \mathcal{B}(\Sigma) :\ \sigma^{-k}B = B\}$$

modulo sets of zero m-measure. Note that $m = \int m_x^{\varepsilon_k}\ dm(x)$ is the ergodic decomposition of m with respect to σ^k. Hence $m = \int (1/k) \sum_{i=0}^{k-1} m_x^{\varepsilon_k} \circ \sigma^{-i}\ dm(x)$ is the ergodic decomposition of m with respect to σ. It follows that

$$\frac{1}{k}\sum_{i=0}^{k-1} m_x^{\varepsilon_k} \circ \sigma^{-i} = m_x^{\varepsilon} \quad m\text{-a.e.} \tag{45}$$

By (44), Proposition 4.3 and (45), we have

$$\begin{aligned} h_\pi(\sigma^k, m) &= \frac{1}{k}\sum_{i=0}^{k-1} h_\pi(\sigma^k, m \circ \sigma^{-i}) \\ &\geq \frac{1}{k}\sum_{i=0}^{k-1} \int h_\pi(\sigma^k, m_x^{\varepsilon_k} \circ \sigma^{-i})\ dm(x) - c \\ &= \int h_\pi\Big(\sigma^k, \frac{1}{k}\sum_{i=0}^{k-1} m_x^{\varepsilon_k} \circ \sigma^{-i}\Big)\ dm(x) - c \\ &= \int h_\pi(\sigma^k, m_x^{\varepsilon})\ dm(x) - c. \end{aligned}$$

Using Proposition 4.3 again yields

$$h_\pi(\sigma, m) \geq \int h_\pi(\sigma, m_x^{\varepsilon})\ dm(x) - c/k \quad \text{for any } k \in \mathbb{N}.$$

Hence we have $h_\pi(\sigma, m) \geq \int h_\pi(\sigma, m_x^{\varepsilon})\ dm(x)$, as desired. □

Proof of Theorem 2.2. It follows directly from Propositions 4.1, 4.10 and 4.17. □

4.4. **The projection entropy for certain affine IFS and the proof of Theorem 2.3.** In this subsection, we assume that $\Phi = \{S_i\}_{i=1}^{\ell}$ is an IFS on $\mathbb{R}^d$ of the form

$$S_i(x) = Ax + c_i \qquad (i = 1, \ldots, \ell),$$

where A is a $d \times d$ non-singular real matrix with $\|A\| < 1$ and $c_i \in \mathbb{R}^d$. Let K denote the attractor of Φ.

Let $\mathcal{Q}$ denote the partition $\{[0,1)^d + \alpha :\ \alpha \in \mathbb{Z}^d\}$ of $\mathbb{R}^d$. For $n = 0, 1, \ldots$, and $x \in \mathbb{R}^d$, we set

$$\mathcal{Q}_n = \{A^n Q :\ Q \in \mathcal{Q}\}, \quad \mathcal{Q}_n + x = \{A^n Q + x :\ Q \in \mathcal{Q}\}.$$

We have the following geometric characterization of h_π for the IFS Φ (i.e., Theorem 2.3).

Proposition 4.18. (i) *Let $m \in \mathcal{M}_\sigma(\Sigma)$. Then*

$$h_\pi(\sigma, m) = \lim_{n \to \infty} \frac{H_m(\pi^{-1}\mathcal{Q}_n)}{n}. \tag{46}$$

(ii)

$$\lim_{n\to\infty} \frac{\log \#\{Q \in \mathcal{Q}:\ A^n Q \cap K \neq \emptyset\}}{n} = \sup\{h_\pi(\sigma, m):\ m \in \mathcal{M}_\sigma(\Sigma)\}.$$

To prove the above proposition, we need the following lemma.

Lemma 4.19. *Assume that Ω is a subset of $\{1,\dots,\ell\}$ such that $S_i(K)\cap S_j(K) = \emptyset$ for all $i, j \in \Omega$ with $i \neq j$. Suppose that ν is an invariant measure on Σ supported on $\Omega^{\mathbb{N}}$, i.e., $\nu([j]) = 0$ for all $j \in \{1,\dots,\ell\}\backslash\Omega$. Then $h_\pi(\sigma,\nu) = h(\sigma,\nu)$.*

Proof. It suffices to prove that $h_\pi(\sigma,\nu) \geq h(\sigma,\nu)$. Recall that

$$h_\pi(\sigma,\nu) = H_\nu(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - H_\nu(\mathcal{P}|\pi^{-1}\gamma)$$

and $H_\nu(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) \geq H_\nu(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma)) = h(\sigma,\nu)$. Hence we only need to show $H_\nu(\mathcal{P}|\pi^{-1}\gamma) = 0$. To do this, denote

$$\delta = \min\{d(S_i(K), S_j(K)):\ i, j \in \Omega, i \neq j\}.$$

Then $\delta > 0$. Let ξ be an arbitrary finite Borel partition of K so that $\mathrm{diam}(A) < \delta/2$ for $A \in \xi$. Set $\mathcal{W} = \{[i]:\ i \in \Omega\}$. Since ν is supported on $\Omega^{\mathbb{N}}$, we have

$$H_\nu(\mathcal{P}|\pi^{-1}\widehat{\xi}) = H_\nu(\mathcal{P}\vee\pi^{-1}\xi) - H_\nu(\pi^{-1}\xi) = H_\nu(\mathcal{W}\vee\pi^{-1}\xi) - H_\nu(\pi^{-1}\xi).$$

However for each $A \in \xi$, there is at most one $i \in \Omega$ such that $S_i(K)\cap A \neq \emptyset$, i.e., $[i]\cap\pi^{-1}A \neq \emptyset$. This forces that $H_\nu(\mathcal{W}\vee\pi^{-1}\xi) = H_\nu(\pi^{-1}\xi)$. Hence

$$H_\nu(\mathcal{P}|\pi^{-1}\widehat{\xi}) = 0.$$

By the arbitrariness of ξ and Lemma 4.4(iv), we have $H_\nu(\mathcal{P}|\pi^{-1}\gamma) = 0$. $\square$

Proof of Proposition 4.18. We first prove (i). Let $m \in \mathcal{M}_\sigma(\Sigma)$. Denote $\gamma = \mathcal{B}(\mathbb{R}^d)$. According to Proposition 4.3, we have

$$H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\gamma) - H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\gamma) = p h_\pi(\sigma, m) \qquad (p \in \mathbb{N}).$$

Now fix p. Since $\widehat{\mathcal{Q}_n} \uparrow \gamma$, by Lemma 4.4(iv), there exists k_0 such that for $k \geq k_0$,

$$\begin{aligned}
&|H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\gamma) - H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\widehat{\mathcal{Q}_{kp}})| \leq 1, \quad \text{and}\\
&|H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\gamma) - H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\widehat{\mathcal{Q}_{(k+1)p}})| \leq 1.
\end{aligned}$$

It follows that for $k \geq k_0$,

$$\begin{aligned}
p h_\pi(\sigma, m) - 2 &\leq H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\widehat{\mathcal{Q}_{kp}}) - H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\widehat{\mathcal{Q}_{(k+1)p}})\\
&\leq p h_\pi(\sigma, m) + 2.
\end{aligned} \tag{47}$$

Now we estimate the difference of conditional entropies in (47). Note that

$$\begin{aligned}
H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\widehat{\mathcal{Q}_{kp}}) &= H_m(\mathcal{P}_0^{p-1}\vee\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp}) - H_m(\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp})\\
&= H_m(\mathcal{P}_0^{p-1}\vee\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp}) - H_m(\pi^{-1}\mathcal{Q}_{kp})
\end{aligned}$$

and

$$H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\widehat{\mathcal{Q}_{(k+1)p}}) = H_m(\mathcal{P}_0^{p-1}\vee\pi^{-1}\mathcal{Q}_{(k+1)p}) - H_m(\pi^{-1}\mathcal{Q}_{(k+1)p}).$$

Hence we have

$$
\begin{aligned}
&H_m(\mathcal{P}_0^{p-1}|\sigma^{-p}\pi^{-1}\widehat{\mathcal{Q}_{kp}}) - H_m(\mathcal{P}_0^{p-1}|\pi^{-1}\widehat{\mathcal{Q}_{(k+1)p}}) \\
&= H_m(\mathcal{P}_0^{p-1}\vee\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp}) - H_m(\mathcal{P}_0^{p-1}\vee\pi^{-1}\mathcal{Q}_{(k+1)p}) \\
&\quad + H_m(\pi^{-1}\mathcal{Q}_{(k+1)p}) - H_m(\pi^{-1}\mathcal{Q}_{kp}).
\end{aligned} \tag{48}
$$

Observe that for each $[u]\in\mathcal{P}_0^{p-1}$ and any $Q\in\mathcal{Q}$,

$$[u]\cap\sigma^{-p}\pi^{-1}A^{kp}Q = [u]\cap\pi^{-1}S_uA^{kp}Q.$$

Since the linear part of S_u is A^p, the set $S_uA^{kp}Q$ intersects at most 2^d elements of $\mathcal{Q}_{(k+1)p}$. Therefore each element of $\mathcal{P}_0^{p-1}\vee\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp}$ intersects at most 2^d elements of $\mathcal{P}_0^{p-1}\vee\pi^{-1}\mathcal{Q}_{(k+1)p}$. Similarly, the statement is also true if the two partitions are interchanged. Therefore by Lemma 4.6, we have

$$|H_m(\mathcal{P}_0^{p-1}\vee\sigma^{-p}\pi^{-1}\mathcal{Q}_{kp}) - H_m(\mathcal{P}_0^{p-1}\vee\pi^{-1}\mathcal{Q}_{(k+1)p})| \le d\log 2.$$

It together with (47) and (48) yields

$$
\begin{aligned}
ph_\pi(\sigma,m) - 2 - d\log 2 &\le H_m(\pi^{-1}\mathcal{Q}_{(k+1)p}) - H_m(\pi^{-1}\mathcal{Q}_{kp}) \\
&\le ph_\pi(\sigma,m) + 2 + d\log 2
\end{aligned}
$$

for $k\ge k_0$. Hence we have

$$
\begin{aligned}
&\limsup_{k\to\infty}\frac{H_m(\pi^{-1}\mathcal{Q}_{kp})}{kp} \le h_\pi(\sigma,m) + \frac{2+d\log 2}{p} \quad \text{and} \\
&\liminf_{k\to\infty}\frac{H_m(\pi^{-1}\mathcal{Q}_{kp})}{kp} \ge h_\pi(\sigma,m) - \frac{2+d\log 2}{p}.
\end{aligned}
$$

By a volume argument, there is a large integer N (N depends on A, d, p; and it is independent of k) such that for any $i=0,\dots,p-1$, each element of $\mathcal{Q}_{kp+i}$ intersects at most N elements of $\mathcal{Q}_{kp}$, and vice versa. Hence by Lemma 4.6, $|H_m(\pi^{-1}\mathcal{Q}_{kp}) - H_m(\pi^{-1}\mathcal{Q}_{kp+i})| < \log N$ for $0\le i\le p-1$. It follows that

$$
\begin{aligned}
&\limsup_{k\to\infty} H_m(\pi^{-1}\mathcal{Q}_{kp})/(kp) = \limsup_{n\to\infty} H_m(\pi^{-1}\mathcal{Q}_n)/n \text{ and} \\
&\liminf_{k\to\infty} H_m(\pi^{-1}\mathcal{Q}_{kp})/(kp) = \liminf_{n\to\infty} H_m(\pi^{-1}\mathcal{Q}_n)/n.
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
h_\pi(\sigma,m) - \frac{2+d\log 2}{p} &\le \liminf_{n\to\infty}\frac{H_m(\pi^{-1}\mathcal{Q}_n)}{n} \le \limsup_{n\to\infty}\frac{H_m(\pi^{-1}\mathcal{Q}_n)}{n} \\
&\le h_\pi(\sigma,m) + \frac{2+d\log 2}{p}.
\end{aligned}
$$

Letting p tend to infinity, we obtain (46).

To show (ii), we assume $K\subset[0,1)^d$, without loss of generality. Note that the number of (non-empty) elements in the partition $\pi^{-1}\mathcal{Q}_n$ is just equal to

$$N_n := \#\{Q\in\mathcal{Q}:\ A^nQ\cap K\ne\emptyset\}.$$

Hence by (21), we have

$$H_m(\pi^{-1}\mathcal{Q}_n) \le \log N_n, \quad \forall\, m \in \mathcal{M}_\sigma(\Sigma).$$

This together with (i) proves

$$\liminf_{n\to\infty} \frac{\log N_n}{n} \ge \sup\{h_\pi(\sigma, m) :\ m \in \mathcal{M}_\sigma(\Sigma)\}.$$

To prove (ii), we still need to show

$$\limsup_{n\to\infty} \log N_n/n \le \sup\{h_\pi(\sigma, m) :\ m \in \mathcal{M}_\sigma(\Sigma)\}. \tag{49}$$

We may assume that $\limsup_{n\to\infty} \log N_n/n > 0$, otherwise there is nothing to prove. Let n be a large integer so that $N_n > 7^d$. Choose a subset Γ of

$$\{Q :\ A^n Q \cap K \ne \emptyset, Q \in \mathcal{Q}\}$$

such that $\#\Gamma > 7^{-d} N_n$, and

$$2Q \cap 2\widetilde{Q} = \emptyset \qquad \text{for different } Q, \widetilde{Q} \in \Gamma, \tag{50}$$

where $2Q := \bigcup_{P\in\mathcal{Q}:\ \overline{P}\cap\overline{Q}\ne\emptyset} P$, and $\overline{P}$ denotes the closure of P. For each $Q \in \Gamma$, since $A^n Q \cap K \ne \emptyset$, we can pick a word $u = u(Q) \in \Sigma_n$ such that $S_u K \cap A^n Q \ne \emptyset$. Consider the collection $W = \{u(Q) : Q \in \Gamma\}$. The separation condition (50) for elements in Γ guarantees that

$$S_{u(Q)}(K) \cap S_{u(\widetilde{Q})}(K) = \emptyset \quad \text{for all } Q, \widetilde{Q} \in \Gamma \text{ with } Q \ne \widetilde{Q}.$$

Define a Bernoulli measure ν on $W^{\mathbb{N}}$ by

$$\nu([w_1 \dots w_k]) = (\#\Gamma)^{-k} \qquad (k \in \mathbb{N},\ w_1, \dots, w_k \in W).$$

Then ν can be viewed as a σ^n-invariant measure on Σ (by viewing $W^{\mathbb{N}}$ as a subset of Σ). By Lemma 4.19, we have $h_\pi(\sigma^n, \nu) = h(\sigma^n, \nu) = \log \#\Gamma$. Define $\mu = \frac{1}{n}\sum_{i=0}^{n-1} \nu \circ \sigma^{-i}$. Then $\mu \in \mathcal{M}_\sigma(\Sigma)$, and by Proposition 4.3,

$$h_\pi(\sigma, \mu) = \frac{h_\pi(\sigma^n, \nu)}{n} = \frac{\log \#\Gamma}{n} \ge \frac{\log(7^{-d} N_n)}{n},$$

from which (49) follows. □

4.5. **Upper semi-continuity of $h_\pi(\sigma, \cdot)$ under the AWSC.** In this subsection, we prove the following proposition.

Proposition 4.20. *Assume that $\{S_i\}_{i=1}^{\ell}$ is an IFS which satisfies the AWSC (see Definition 2.14). Then the map $m \mapsto h_\pi(\sigma, m)$ is upper semi-continuous on $\mathcal{M}_\sigma(\Sigma)$.*

We first prove a lemma.

Lemma 4.21. *Let $\{S_i\}_{i=1}^{\ell}$ be an IFS with attractor $K \subset \mathbb{R}^d$. Assume that*

$$\#\{1 \le i \le \ell :\ x \in S_i(K)\} \le k$$

for some $k \in \mathbb{N}$ and each $x \in \mathbb{R}^d$. Then $H_\nu(\mathcal{P}|\pi^{-1}\gamma) \le \log k$ for any Borel probability measure ν on Σ.

Proof. A compactness argument shows that there is $r_0 > 0$ such that

$$\#\{1 \le i \le \ell : \ B(x, r_0) \cap S_i(K) \ne \emptyset\} \le k$$

for each $x \in \mathbb{R}^d$. Let $n \in \mathbb{N}$ so that $2^{-n}\sqrt{d} < r_0$. Then for each $Q \in \mathcal{D}_n$, where $\mathcal{D}_n$ is defined as in (24), there are at most k different $i \in \{1, \ldots, \ell\}$ such that $S_i(K) \cap Q \ne \emptyset$. It follows that each member in $\pi^{-1}\mathcal{D}_n$ intersects at most k members of $\mathcal{P} \vee \pi^{-1}\mathcal{D}_n$. By Lemma 4.6, we have

$$H_\nu(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n}) = H_\nu(\mathcal{P} \vee \pi^{-1}\mathcal{D}_n) - H_\nu(\pi^{-1}\mathcal{D}_n) \le \log k.$$

Note that $\pi^{-1}\widehat{\mathcal{D}_n} \uparrow \pi^{-1}\gamma$. Applying Lemma 4.4(iv), we obtain

$$H_\nu(\mathcal{P}|\pi^{-1}\gamma) = \lim_{n\to\infty} H_\nu(\mathcal{P}|\pi^{-1}\widehat{\mathcal{D}_n}) \le \log k.$$

□

As a corollary, we have

Corollary 4.22. *Under the condition of Lemma 4.21, we have*

$$h_\pi(\sigma, m) \ge h(\sigma, m) - \log k$$

for any $m \in \mathcal{M}_\sigma(\Sigma)$.

Proof. By the definition of $h_\pi(\sigma, m)$ and Lemma 4.21, we have

$$h_\pi(\sigma, m) = H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - H_m(\mathcal{P}|\pi^{-1}\gamma) \ge H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) - \log k.$$

However, $H_m(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma) \ge H_m(\mathcal{P}|\sigma^{-1}\mathcal{B}(\Sigma)) = h(\sigma, m)$. This implies the desired result. □

To prove Proposition 4.20, we need the following lemma.

Lemma 4.23. *Let* $\{S_i\}_{i=1}^{\ell}$ *be an IFS with attractor* K. *Suppose that* Ω *is a subset of* $\{1, \ldots, \ell\}$ *such that there is a map* $g\colon \{1, \ldots, \ell\} \to \Omega$ *so that*

$$S_i = S_{g(i)} \qquad (i = 1, \ldots, \ell).$$

Let $(\Omega^{\mathbb{N}}, \widetilde{\sigma})$ *denote the one-sided full shift over* Ω. *Define* $G : \Sigma \to \Omega^{\mathbb{N}}$ *by* $(x_j)_{j=1}^\infty \mapsto (g(x_j))_{j=1}^\infty$. *Then*

- (i) K *is also the attractor of* $\{S_i\}_{i\in\Omega}$. *Moreover if we let* $\widetilde{\pi}\colon \Omega^{\mathbb{N}} \to K$ *denote the canonical projection w.r.t.* $\{S_i\}_{i\in\Omega}$, *then we have* $\pi = \widetilde{\pi} \circ G$.
- (ii) *Let* $m \in \mathcal{M}_\sigma(\Sigma)$. *Then* $\nu = m \circ G^{-1} \in \mathcal{M}_{\widetilde{\sigma}}(\Omega^{\mathbb{N}})$. *Furthermore,* $h_\pi(\sigma, m) = h_{\widetilde{\pi}}(\widetilde{\sigma}, \nu)$. *In particular,* $h_\pi(\sigma, m) \le \log(\#\Omega)$.

Proof. (i) is obvious. To see (ii), let $m \in \mathcal{M}_\sigma(\Sigma)$. It is easily seen that the following diagram commutes:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\ \sigma\ } & \Sigma \\ {\scriptstyle G}\downarrow & & \downarrow{\scriptstyle G} \\ \Omega^{\mathbb{N}} & \xrightarrow{\ \widetilde{\sigma}\ } & \Omega^{\mathbb{N}}. \end{array}$$

That is, $\widetilde{\sigma} \circ G = G \circ \sigma$. Hence $\nu = m \circ G^{-1} \in \mathcal{M}_{\widetilde{\sigma}}(\Omega^{\mathbb{N}})$. To show that $h_\pi(\sigma, m) = h_{\widetilde{\pi}}(\widetilde{\sigma}, \nu)$, let $\mathcal{Q} = \{[i] : i \in \Omega\}$ be the canonical partition of $\Omega^{\mathbb{N}}$. Then

$$\begin{aligned} h_{\widetilde{\pi}}(\widetilde{\sigma}, \nu) &= H_{m\circ G^{-1}}(\mathcal{Q}|\widetilde{\sigma}^{-1}\widetilde{\pi}^{-1}\gamma) - H_{m\circ G^{-1}}(\mathcal{Q}|\widetilde{\pi}^{-1}\gamma) \\ &= H_m(G^{-1}(\mathcal{Q})|G^{-1}\widetilde{\sigma}^{-1}\widetilde{\pi}^{-1}\gamma) - H_m(G^{-1}(\mathcal{Q})|G^{-1}\widetilde{\pi}^{-1}\gamma) \\ &= H_m(G^{-1}(\mathcal{Q})|\sigma^{-1}\pi^{-1}\gamma) - H_\nu(G^{-1}(\mathcal{Q})|\pi^{-1}\gamma), \end{aligned}$$

using the facts $G \circ \sigma = \widetilde{\sigma} \circ G$ and $\widetilde{\pi} \circ G = \pi$. Since $\mathcal{P} \vee G^{-1}(\mathcal{Q}) = \mathcal{P}$, we have

$$\begin{aligned} & h_\pi(\sigma, m) - h_{\widetilde{\pi}}(\widetilde{\sigma}, m \circ G^{-1}) \\ &= \left(H_m\left(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma\right) - H_m\left(\mathcal{P}|\pi^{-1}\gamma\right)\right) \\ &\quad - \left(H_m(G^{-1}(\mathcal{Q})|\sigma^{-1}\pi^{-1}\gamma) - H_m(G^{-1}(\mathcal{Q})|\pi^{-1}\gamma)\right) \\ &= \left(H_m\left(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma\right) - H_m(G^{-1}(\mathcal{Q})|\sigma^{-1}\pi^{-1}\gamma)\right) \\ &\quad - \left(H_m\left(\mathcal{P}|\pi^{-1}\gamma\right) - H_m(G^{-1}(\mathcal{Q})|\pi^{-1}\gamma)\right) \\ &= H_m\left(\mathcal{P}|\sigma^{-1}\pi^{-1}\gamma \vee G^{-1}(\widehat{\mathcal{Q}})\right) - H_m\left(\mathcal{P}|\pi^{-1}\gamma \vee G^{-1}(\widehat{\mathcal{Q}})\right). \end{aligned}$$

An argument similar to the proof of Lemma 4.7 shows that

$$\sigma^{-1}\pi^{-1}\gamma \vee G^{-1}(\widehat{\mathcal{Q}}) = \pi^{-1}\gamma \vee G^{-1}(\widehat{\mathcal{Q}}).$$

Hence we have $h_\pi(\sigma, m) = h_{\widetilde{\pi}}(\widetilde{\sigma}, m \circ G^{-1})$. □

Proof of Proposition 4.20. Let (ν_n) be a sequence in $\mathcal{M}_\sigma(\Sigma)$ converging to m in the weak-star topology. We need to show that $\limsup_{n\to\infty} h_\pi(\sigma, \nu_n) \le h_\pi(\sigma, m)$. To see this, it suffices to show that

$$\limsup_{n\to\infty} h_\pi(\sigma, \nu_n) \le h_\pi(\sigma, m) + \frac{1}{k}\log t_k \tag{51}$$

for each $k \in \mathbb{N}$, where t_k is given as in Definition 2.14.

To prove (51), we fix $k \in \mathbb{N}$. Define an equivalence relation $\sim$ on $\{1, \ldots, \ell\}^k$ by $u \sim v$ if $S_u = S_v$. Let $\underline{u}$ denotes the equivalence class containing u. Denote $S_{\underline{u}} = S_u$. Set $\mathcal{J} = \{\underline{u} : u \in \{1, \ldots, \ell\}^k\}$. Let $(\mathcal{J}^{\mathbb{N}}, T)$ denote the one-sided full shift space over the alphabet $\mathcal{J}$. Let $G : \Sigma \to \mathcal{J}^N$ be defined by

$$(x_i)_{i=1}^\infty \mapsto \left(\underline{x_{jk+1}\cdots x_{(j+1)k}}\right)_{j=0}^\infty.$$

It is clear that the following diagram commutes:

$$\begin{array}{ccc} \Sigma & \xrightarrow{\sigma^k} & \Sigma \\ G\big\downarrow & & \big\downarrow G \\ \mathcal{J}^N & \xrightarrow{T} & \mathcal{J}^N \end{array}$$

That is, $T \circ G = G \circ \sigma^k$. It implies that $\nu_n \circ G^{-1}, m \circ G^{-1} \in \mathcal{M}_T(\mathcal{J}^{\mathbb{N}})$ and

$$\lim_{n\to\infty} \nu_n \circ G^{-1} = m \circ G^{-1}.$$

Hence we have

$$h(T, m \circ G^{-1}) \geq \limsup_{n\to\infty} h(T, \nu_n \circ G^{-1}), \tag{52}$$

where we use the upper semi-continuity of the classical measure-theoretic entropy map on $(\mathcal{J}^{\mathbb{N}}, T)$. Define $\widetilde{\pi} : \mathcal{J}^{\mathbb{N}} \to K$ by

$$\widetilde{\pi}\left((\underline{u_i})_{i=1}^{\infty}\right) = \lim_{n\to\infty} S_{\underline{u_1}} \circ \cdots \circ S_{\underline{u_n}}(K).$$

Then $\widetilde{\pi} \circ G = \pi$. By the assumption of AWSC (11) and Corollary 4.22 (considering the IFS $\{S_{\underline{u}} : \underline{u} \in \mathcal{J}\}$), we have

$$\begin{aligned} h_{\widetilde{\pi}}(T, m \circ G^{-1}) &\geq h(T, m \circ G^{-1}) - \log t_k \\ &\geq \limsup_{n\to\infty} h(T, \nu_n \circ G^{-1}) - \log t_k \qquad (\text{ by } (52)) \\ &\geq \limsup_{n\to\infty} h_{\widetilde{\pi}}(T, \nu_n \circ G^{-1}) - \log t_k, \end{aligned}$$

where the last inequality follows from Proposition 4.1(i). Then (51) follows from the above inequality, together with Proposition 4.3 and the following claim:

$$h_{\widetilde{\pi}}(T, \nu \circ G^{-1}) = h_{\pi}(\sigma^k, \nu) \qquad (\nu \in \mathcal{M}_{\sigma}(\Sigma)). \tag{53}$$

However, (53) just comes from Lemma 4.23, where we consider the IFS $\{S_u :\ u \in \{1, \ldots, \ell\}^k\}$ rather than $\{S_i\}_{i=1}^{\ell}$. ☐

5. Some geometric properties of C^1 IFS

In this section we give some geometric properties of C^1 IFS.

Lemma 5.1. *Let $S : U \to S(U) \subset \mathbb{R}^d$ be a C^1 diffeomorphism on an open set $U \subset \mathbb{R}^d$, and X a compact subset of U. Let $c > 1$. Then there exists $r_0 > 0$ such that*

$$c^{-1} [\![S'(x)]\!] \cdot |x - y| \leq |S(x) - S(y)| \leq c \|S'(x)\| \cdot |x - y| \tag{54}$$

for all $x \in X$, $y \in U$ with $|x - y| \leq r_0$, where $S'(x)$ denotes the differential of S at x, $[\![\cdot]\!]$ and $\| \cdot \|$ are defined as in (4). As a consequence,

$$B(S(x), c^{-1} [\![S'(x)]\!] r) \subset S\left(B(x, r)\right) \subset B(S(x), c\|S'(x)\| r) \tag{55}$$

for all $x \in X$ and $0 < r \leq r_0$.

Proof. Let $c > 1$. We only prove (54), for it is not hard to derive (55) from (54). Assume on the contrary that (54) is not true. Then there exist two sequences $(x_n) \subset X$, $(y_n) \subset U$ such that $x_n \neq y_n$, $\lim_{n\to\infty} |x_n - y_n| = 0$ and for each $n \geq 1$,

$$\begin{aligned} &\text{either } |S(x_n) - S(y_n)| \geq c \|S'(x_n)\| \cdot |x_n - y_n|, \\ &\text{or} \quad\ |S(x_n) - S(y_n)| \leq c^{-1} [\![S'(x_n)]\!] \cdot |x_n - y_n|. \end{aligned} \tag{56}$$

Since X is compact, without lost of generality, we assume that

$$\lim_{n\to\infty} x_n = x = \lim_{n\to\infty} y_n.$$

Write $S = (f_1, f_2, \dots, f_d)^t$. Then each component f_j of S is a C^1 real-valued function defined on U. Choose a small $\epsilon > 0$ such that

$$\{z \in \mathbb{R}^d : |z - x| \le \epsilon \text{ for some } x \in X\} \subset U.$$

Take $N \in \mathbb{N}$ such that $|x_n - y_n| < \epsilon$ for $n \ge N$. By the mean value theorem, for each $n \ge N$ and $1 \le j \le d$, there exists $z_{n,j}$ on the segment L_{x_n,y_n} connecting x_n and y_n such that

$$f_j(x_n) - f_j(y_n) = \nabla f_j(z_{n,j}) \cdot (x_n - y_n),$$

where ∇f_j denote the gradient of f_j. Therefore $|S(x_n) - S(y_n)| = |M_n(x_n - y_n)|$ with $M_n := (\nabla f_1(z_{n,1}), \dots, \nabla f_d(z_{n,d}))^t$. It follows

$$\text{(57)} \qquad [\![M_n]\!] \cdot |x_n - y_n| \le |S(x_n) - S(y_n)| \le \|M_n\| \cdot |x_n - y_n|.$$

Since S is C^1, M_n tends to $S'(x)$ as $n \to \infty$. Thus we have $[\![M_n]\!] \to [\![S'(x)]\!]$ and $\|M_n\| \to \|S'(x)\|$. Meanwhile, $[\![S'(x_n)]\!] \to [\![S'(x)]\!]$ and $\|S'(x_n)\| \to \|S'(x)\|$. These limits together (57) lead to a contradiction with (56). □

Let $\{S_1, \dots, S_\ell\}$ be a C^1 IFS on a compact set $X \subset \mathbb{R}^d$. Let $\pi : \Sigma \to \mathbb{R}^d$ be defined as in (1). By Lemma 5.1, we have directly

Lemma 5.2. *Let $c > 1$. Then there exists $r_0 > 0$ such that for any $1 \le i \le \ell$, $x \in \Sigma$ and $0 < r < r_0$,*

$$B(S_i(\pi x), c^{-1}[\![S_i'(\pi x)]\!] r) \subset S_i\left(B(\pi x, r)\right) \subset B(S_i(\pi x), c\|S_i'(\pi x)\| r).$$

Let $\overline{\rho}, \underline{\rho} : \Sigma \to \mathbb{R}$ be defined by

$$\text{(58)} \qquad \overline{\rho}(x) = \|S_{x_1}'(\pi\sigma x)\|, \quad \underline{\rho}(x) = [\![S_{x_1}'(\pi\sigma x)]\!] \qquad (x = (x_i)_{i=1}^\infty \in \Sigma).$$

Let $\mathcal{P}$ be the partition of Σ defined as in (3). For $x \in \Sigma$, let $\mathcal{P}(x)$ denote the element in $\mathcal{P}$ which contains x. Then we have

Lemma 5.3. *Let $c > 1$. Then there exists $r_0 > 0$ such that for any $z \in \Sigma$ and $0 < r < r_0$,*

$$B^\pi(z, c^{-1}\underline{\rho}(z) r) \cap \mathcal{P}(z) \subset B^{\pi\sigma}(z, r) \cap \mathcal{P}(z) \subset B^\pi(z, c\overline{\rho}(z) r) \cap \mathcal{P}(z),$$

where $B^\pi(z, r)$ is defined as in (12).

Proof. Let $z = (z_j)_{j=1}^\infty \in \Sigma$. Taking $i = z_1$ and $x = \sigma z$ in Lemma 5.2 we obtain

$$B(S_{z_1}(\pi\sigma z), c^{-1}[\![S_{z_1}'(\pi\sigma z)]\!] r) \subset S_{z_1}(B(\pi\sigma z, r)) \subset B(S_{z_1}(\pi\sigma z), c\|S_{z_1}'(\pi\sigma z)\| r).$$

That is,

$$B(\pi z, c^{-1}\underline{\rho}(z) r) \subset S_{z_1} B(\pi\sigma z, r) \subset B(\pi z, c\overline{\rho}(z) r),$$

where we use the fact $S_{z_1}(\pi\sigma z) = \pi z$, which can be checked directly from the definition of π. Thus we have

$$B^\pi(\pi z, c^{-1}\underline{\rho}(z) r) \cap \mathcal{P}(z) \subset \pi^{-1}\left(S_{z_1}\left(B(\pi\sigma z, r)\right)\right) \cap \mathcal{P}(z) \subset B^\pi(z, c\overline{\rho}(z) r) \cap \mathcal{P}(z).$$

At last we show that $\pi^{-1}\left(S_{z_1}\left(B(\pi\sigma z, r)\right)\right) \cap \mathcal{P}(z) = B^{\pi\sigma}(z, r) \cap \mathcal{P}(z)$. To see this, let $y = (y_j)_{j=1}^{\infty} \in \Sigma$. Then we have the following equivalent implications.

$$\begin{aligned}
& y \in \pi^{-1}\left(S_{z_1}\left(B(\pi\sigma z, r)\right)\right) \cap \mathcal{P}(z) \\
& \quad \Longleftrightarrow y_1 = z_1, \quad \pi y \in S_{z_1}\left(B(\pi\sigma z, r)\right) \\
& \quad \Longleftrightarrow y_1 = z_1, \quad S_{y_1}(\pi\sigma y) \in S_{z_1}\left(B(\pi\sigma z, r)\right) \\
& \quad \Longleftrightarrow y_1 = z_1, \quad \pi\sigma y \in B(\pi\sigma z, r) \\
& \quad \Longleftrightarrow y_1 = z_1, \quad y \in B^{\pi\sigma}(z, r) \\
& \quad \Longleftrightarrow y \in B^{\pi\sigma}(z, r) \cap \mathcal{P}(z).
\end{aligned}$$

This finishes the proof of the lemma. $\square$

Lemma 5.4. *Assume that $\{S_i\}_{i=1}^{\ell}$ is a weakly conformal IFS with attractor K. Then for any $c > 1$, there exists $D > 0$ such that for any $n \in \mathbb{N}$, $u \in \{1, \ldots, \ell\}^n$, and $x, y \in K$ we have*

$$D^{-1}c^{-n}\|S_u'(x)\| \cdot |x - y| \leq |S_u(x) - S_u(y)| \leq Dc^n\|S_u'(x)\| \cdot |x - y|.$$

and

$$D^{-1}c^{-n}\|S_u'(x)\| \leq \operatorname{diam}(S_u(K)) \leq Dc^n\|S_u'(x)\|. \tag{59}$$

Proof. The results were proved in the conformal case in [18, Lemma 3.5 and Corollary 3.6]. A slight modification of that proof works for the weakly conformal case. $\square$

As a corollary, we have

Corollary 5.5. *Under the assumption of Lemma 5.4, for $\alpha > 0$, there is $r_0 > 0$ such that for any $0 < r < r_0$ and $z \in K$, there exist $n \in \mathbb{N}$ and $u \in \{1, \ldots, \ell\}^n$ such that $S_u(K) \subset B(z, r)$ and*

$$|S_u(x) - S_u(y)| \geq r^{1+\alpha}|x - y| \qquad (x, y \in K). \tag{60}$$

Proof. Denote $a = \inf\{\|S_i'(x)\| :\ x \in K, 1 \leq i \leq \ell\}$ and $b = \sup\{\|S_i'(x)\| :\ x \in K, 1 \leq i \leq \ell\}$. Then $0 < a \leq b < 1$. Choose c so that

$$1 < c < b^{\frac{-\alpha}{3(2+\alpha)}}. \tag{61}$$

Let D be the constant in Lemma 5.4 corresponding to c. Take $n_0 \in \mathbb{N}$ and $r_0 > 0$ such that

$$\left(c^3 b^{\alpha/(2+\alpha)}\right)^{n_0} < D^{-3}ab^{\alpha/(2+\alpha)}, \quad (1 + \alpha/2) \cdot \frac{\log r_0}{\log a} = n_0. \tag{62}$$

Now fix $z \in K$ and $0 < r < r_0$. We shall show that there exist $n \in \mathbb{N}$ and $u \in \{1, \ldots, \ell\}^n$ such that $S_u(K) \subset B(z, r)$ and (60) holds. To see this, take $\omega = (\omega_i)_{i=1}^{\infty} \in \Sigma$ such that $z = \pi\omega$, where π is defined as in (1). Let n be the unique integer such that

$$\|S'_{\omega_1\cdots\omega_n}(\pi\sigma^n\omega)\| < r^{1+\alpha/2} \leq \|S'_{\omega_1\cdots\omega_{n-1}}(\pi\sigma^{n-1}\omega)\|. \tag{63}$$

It follows $a^n < r^{1+\alpha/2} \le b^{n-1}$, which together with (62) forces that

$$n > n_0 \quad \text{and} \quad c^{3n} < D^{-3} a r^{-\alpha/2}. \tag{64}$$

To see (64), we first assume on the contrary that $n \le n_0$. Then

$$a^n \ge a^{n_0} = a^{(1+\alpha/2)\log r_0/\log \alpha} = r_0^{1+\alpha/2} > r^{1+\alpha/2},$$

which contradicts the fact $a^n < r^{1+\alpha/2}$. Hence $n > n_0$. To see $c^{3n} < D^{-3} a r^{-\alpha/2}$, note that

$$\begin{aligned} c^{3n} r^{\alpha/2} &\le c^{3n} b^{(n-1)\alpha/(2+\alpha)} \qquad (\text{ using } r^{1+\alpha/2} \le b^{n-1} \text{ }) \\ &\le \left(c^3 b^{\alpha/(2+\alpha)}\right)^n b^{-\alpha/(2+\alpha)} \\ &\le \left(c^3 b^{\alpha/(2+\alpha)}\right)^{n_0} b^{-\alpha/(2+\alpha)} \qquad (\text{ using } n > n_0 \text{ and (61) }) \\ &\le D^{-3} a \qquad (\text{by (62)}). \end{aligned}$$

This completes the proof of (64). By (59), we have

$$\text{diam} S_{\omega_1 \cdots \omega_n}(K) \le D c^n \|S'_{\omega_1 \cdots \omega_n}(\pi \sigma^n \omega)\| \le D c^n r^{1+\alpha/2} < r.$$

Since $z \in S_{\omega_1 \cdots \omega_n}(K)$, the above inequality implies $S_{\omega_1 \cdots \omega_n}(K) \subset B(z, r)$. By (59) again, we have

$$\|S'_u(x)\| \ge D^{-2} c^{-2n} \|S'_u(y)\|, \quad \forall\, u \in \{1, \ldots, \ell\}^n, \ \forall\, x, y \in K. \tag{65}$$

By Lemma 5.4, we have for $x, y \in K$,

$$\begin{aligned} &|S_{\omega_1 \cdots \omega_n}(x) - S_{\omega_1 \cdots \omega_n}(y)| \\ &\quad \ge D^{-1} c^{-n} \|S'_{\omega_1 \cdots \omega_n}(x)\| \cdot |x - y| \\ &\quad \ge D^{-3} c^{-3n} \|S'_{\omega_1 \cdots \omega_n}(\pi \sigma^n \omega)\| \cdot |x - y| \qquad (\text{by (65)}) \\ &\quad \ge D^{-3} c^{-3n} \|S'_{\omega_1 \cdots \omega_{n-1}}(\pi \sigma^{n-1} \omega)\| [\![S'_{\omega_n}(\pi \sigma^n \omega)]\!] \cdot |x - y| \\ &\quad \ge D^{-3} c^{-3n} a r^{1+\alpha/2} |x - y| \qquad (\text{by (63)}) \\ &\quad \ge r^{1+\alpha} |x - y| \qquad (\text{by (64)}). \end{aligned}$$

Hence the corollary follows by taking $u = \omega_1 \cdots \omega_n$. $\square$

Proposition 5.6. *Let $\{S_i\}_{i=1}^{\ell}$ be a C^1 IFS with attractor K. Assume that K is not a singleton. Then*

(i) *for any $m \in \mathcal{M}_\sigma(\Sigma)$, we have for m-a.e. $x = (x_i)_{i=1}^\infty \in \Sigma$,*

$$\liminf_{n \to \infty} \frac{\log \text{diam} S_{x_1 \ldots x_n}(K)}{n} \ge -\overline{\lambda}(x),$$

$$\limsup_{n \to \infty} \frac{\log \text{diam} S_{x_1 \ldots x_n}(K)}{n} \le -\underline{\lambda}(x),$$

where $\underline{\lambda}, \overline{\lambda}$ are defined as in Definition 2.5. In particular, if $\{S_i\}_{i=1}^{\ell}$ is m-conformal, then for m-a.e. $x = (x_i)_{i=1}^\infty \in \Sigma$,

$$\lim_{n \to \infty} \frac{\log \text{diam} S_{x_1 \ldots x_n}(K)}{n} = -\lambda(x).$$

(ii) *If* $\{S_i\}_{i=1}^{\ell}$ *is weakly conformal, then it is* m*-conformal for each* $m \in \mathcal{M}_\sigma(\Sigma)$.

Proof. We first prove (i). Take $c > 1$ small enough so that $c \sup_{x\in\Sigma} \overline{\rho}(x) < 1$. Let $r_0 > 0$ be given as in Lemma 5.2. Let $x = (x_i)_{i=1}^{\infty} \in \Sigma$. Applying Lemma 5.2 repeatedly, we have

$$S_{x_1\cdots x_n}(B(\pi\sigma^n x, r_0)) \subset B(\pi x, c^n \overline{\rho}(x) \cdots \overline{\rho}(\sigma^{n-1}x) r_0). \tag{66}$$

Since $\{S_i\}_{i=1}^{\ell}$ is contractive, there is a constant k such that

$$S_{x_{n+1}\cdots x_{n+k}}(K) \subset B(\pi\sigma^n x, r_0).$$

This together with (66) yields

$$\mathrm{diam} S_{x_1\ldots x_{n+k}}(K) \le \mathrm{diam} S_{x_1\cdots x_n}(B(\pi\sigma^n x, r_0)) \le c^n \overline{\rho}(x) \ldots \overline{\rho}(\sigma^{n-1}x) r_0. \tag{67}$$

Since K is not a singleton, there exists $0 < r_1 < r_0$ such that for each $z \in K$, there exists $w \in K$ such that $r_1 \le |z - w| \le r_0$. Indeed, to obtain r_1, one chooses an integer n_0 large enough such that $\sup_{u\in\Sigma_{n_0}} \mathrm{diam} S_u(K) \le r_0$, then set

$$r_1 = (1/2) \inf_{u\in\Sigma_{n_0}} \mathrm{diam} S_u(K).$$

For each such pair (z, w), applying (54) repeatedly yields

$$\mathrm{diam} S_{x_1\ldots x_n}(K) \ge |S_{x_1\ldots x_n}(z) - S_{x_1\ldots x_n}(w)| \ge r_1 c^{-n} \prod_{j=1}^{n} \llbracket S'_{x_j}(S_{x_{j+1}\ldots x_n}(z) \rrbracket.$$

Hence by taking $z = \pi\sigma^n x$, we have

$$\mathrm{diam} S_{x_1\ldots x_n}(K) \ge r_1 c^{-n} \underline{\rho}(x) \ldots \underline{\rho}(\sigma^{n-1}x). \tag{68}$$

Denote

$$g_*(x) = \liminf_{n\to\infty} \frac{\log \mathrm{diam} S_{x_1\ldots x_n}(K)}{n} \quad \text{and}$$
$$g^*(x) = \limsup_{n\to\infty} \frac{\log \mathrm{diam} S_{x_1\ldots x_n}(K)}{n}.$$

It is clear that $g_*(x) = g_*(\sigma x)$ and $g^*(x) = g^*(\sigma x)$. Let $\mathcal{I}$ denote the σ-algebra $\{B \in \mathcal{B}(\Sigma) : \sigma^{-1}B = B\}$. Then by (68), the Birkhoff ergodic theorem, and Theorem 34.2 in [7], we have for m-a.e. $x \in \Sigma$,

$$\begin{aligned} g_*(x) = \mathbf{E}_m(g_*|\mathcal{I})(x) &\ge \mathbf{E}_m\left(\lim_{n\to\infty} \frac{-n\log c + \sum_{i=0}^{n-1} \log \underline{\rho}\circ\sigma^{-i}}{n} \big| \mathcal{I}\right)(x) \\ &= -\log c + \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathbf{E}_m(\log \underline{\rho}\circ\sigma^{-i}|\mathcal{I})(x) \\ &= -\log c + \mathbf{E}_m(\log \underline{\rho}|\mathcal{I})(x) \end{aligned} \tag{69}$$

and similarly by (67),

$$g^*(x) \le \log c + \mathbf{E}_m(\log \overline{\rho}|\mathcal{I})(x). \tag{70}$$

For $p \in \mathbb{N}$, write $A_p(x) = \log [\![S'_{x_1 \cdots x_p}(\pi\sigma^p x)]\!]$ and $A^*_p(x) = \log \|S'_{x_1 \cdots x_p}(\pi\sigma^p x)\|$. Consider the IFS $\{S_{i_1 \ldots i_p} : 1 \le i_j \le \ell, 1 \le j \le p\}$ rather than $\{S_i\}_{i=1}^{\ell}$. Then (69) and (70) can be replaced by

$$g_*(x) \ge -\log c + \frac{1}{p}\mathbf{E}_m(A_p|\mathcal{I}_p)(x), \qquad g^*(x) \le \log c + \frac{1}{p}\mathbf{E}_m(A^*_p|\mathcal{I}_p)(x),$$

where $\mathcal{I}_p := \{B \in \mathcal{B}(\Sigma) : \sigma^{-p}B = B\}$. Taking the conditional expectation with respect to $\mathcal{I}$ in the above inequalities and noting that g_*, g^* are σ-invariant, we obtain

$$g_*(x) \ge -\log c + \frac{1}{p}\mathbf{E}_m(A_p|\mathcal{I})(x), \qquad g^*(x) \le \log c + \frac{1}{p}\mathbf{E}_m(A^*_p|\mathcal{I})(x). \tag{71}$$

Since $A_p(x)$ is sup-additive (i.e., $A_{p+q}(x) \ge A_p(x) + A_q(\sigma^p x)$) and $A^*_p(x)$ is sub-additive (i.e., $A^*_{p+q}(x) \le A^*_p(x) + A^*_q(\sigma^p x)$), by Kingman's sub-additive ergodic theorem (cf. [63]), we have

$$\lim_{p\to\infty} A_p(x)/p = -\overline{\lambda}(x), \qquad \lim_{p\to\infty} A^*_p(x)/p = -\underline{\lambda}(x) \tag{72}$$

almost everywhere and in L^1. Hence letting $c \to 1$ and $p \to \infty$ in (71) and using Theorem 34.2 in [7], we obtain that $g_*(x) \ge -\overline{\lambda}(x)$ and $g^*(x) \le -\underline{\lambda}(x)$ almost everywhere. This finishes the proof of (i).

To see (ii), assume that $\{S_i\}_{i=1}^{\ell}$ is weakly conformal and $m \in \mathcal{M}_\sigma(\Sigma)$. Then $|A_p(x) - A^*_p(x)|/p$ converges to 0 uniformly as p tends to infinity. This together with (72) yields $\overline{\lambda}(x) = \underline{\lambda}(x)$ for m-a.e. $x \in \Sigma$. This proves (ii). ☐

6. Estimates for local dimensions of invariant measures for C^1 IFS

In this section, we prove a general version of Theorem 2.6, which is also needed in the proof of Theorem 2.11. Let $\{T_i\}_{i=1}^{\ell}$ be a C^1 IFS on $\mathbb{R}^d$, and $\{S_i\}_{i=1}^{\ell}$ a C^1 IFS on $\mathbb{R}^k$. Let $\phi : \Sigma \to \mathbb{R}^d$ and $\pi : \Sigma \to \mathbb{R}^k$ denote the canonical projections associated with $\{T_i\}_{i=1}^{\ell}$ and $\{S_i\}_{i=1}^{\ell}$ respectively. Let η and ξ be two partitions of Σ defined respectively by

$$\eta = \{\phi^{-1}(z) : z \in \mathbb{R}^d\}, \qquad \xi = \sigma^{-1}\eta.$$

Let $\mathcal{P}$ be the partition of Σ given as in (3) and let $\overline{\rho}(x), \underline{\rho}(x)$ be defined as in (58). Applying Lemma 5.3 to the IFS $\{S_i\}_{i=1}^{\ell}$, we have for any $c > 1$ there exist $0 < \delta < c - 1$ and $r_0 > 0$ such that for any $r \in (0, r_0)$ and $x \in \Sigma$,

$$B^\pi(x, (c-\delta)^{-1}\underline{\rho}(x)r) \cap \mathcal{P}(x) \subset B^{\pi\sigma}(x, r) \cap \mathcal{P}(x) \subset B^\pi(x, (c-\delta)\overline{\rho}(x)r) \cap \mathcal{P}(x). \tag{73}$$

The following technical proposition is substantial in our proof.

Proposition 6.1. *Let $m \in \mathcal{M}_\sigma(\Sigma)$ and $c > 1$. Let δ, r_0 be given as above. Then there exists $\Lambda \subset \Sigma$ with $m(\Lambda) = 1$ such that for all $x \in \Lambda$ and $r \in (0, r_0)$,*

$$\frac{m^\eta_x(B^\pi(x, c\overline{\rho}(x)r) \cap \mathcal{P}(x))}{m^\eta_{\sigma x}(B^\pi(\sigma x, r))} \ge f(x) \cdot \frac{m^\xi_x(B^{\pi\sigma}(x, r) \cap \mathcal{P}(x))}{m^\xi_x(B^{\pi\sigma}(x, r))} \tag{74}$$

and

(75)
$$\frac{m_x^\eta(B^\pi(x,c^{-1}\underline{\rho}(x)r)\cap\mathcal{P}(x))}{m_{\sigma x}^\eta(B^\pi(\sigma x,r))}\le f(x)\cdot\frac{m_x^\xi(B^{\pi\sigma}(x,(1-c\delta/2)r)\cap\mathcal{P}(x))}{m_x^\xi(B^{\pi\sigma}(x,(1-c\delta/2)r))},$$

where $f:=\sum_{A\in\mathcal{P}}\chi_A\frac{\mathbf{E}_m(\chi_A|\phi^{-1}\gamma)}{\mathbf{E}_m(\chi_A|\sigma^{-1}\phi^{-1}\gamma)}$, $\gamma=\mathcal{B}(\mathbb{R}^d)$.

Proof. Write $R_{t,x}(z)=T_{x_1}^{-1}B(T_{x_1}z,t)$ for $t>0$, $x=(x_i)_{i=1}^\infty\in\Sigma$ and $z\in\mathbb{R}^d$. It is direct to check that

$$\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\cap\mathcal{P}(x)=B^\phi(x,t)\cap\mathcal{P}(x). \tag{76}$$

Hence for m-a.e. x,

$$\begin{aligned}\frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x,t))} &= \frac{m(B^\phi(x,t)\cap\mathcal{P}(x))}{m(B^\phi(x,t))}\cdot\frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x,t)\cap\mathcal{P}(x))}\\ &= \frac{m(B^\phi(x,t)\cap\mathcal{P}(x))}{m(B^\phi(x,t))}\cdot\frac{m(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x))}{m(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\cap\mathcal{P}(x))}.\end{aligned}$$

Letting $t\to0$ and applying Proposition 3.5 and Remark 3.6, we have

$$\lim_{t\to0}\frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x,t))}=\sum_{A\in\mathcal{P}}\chi_A(x)\frac{\mathbf{E}_m(\chi_A|\phi^{-1}\gamma)(x)}{\mathbf{E}_m(\chi_A|\sigma^{-1}\phi^{-1}\gamma)(x)}=:f(x). \tag{77}$$

for m-a.e. x. Let $\widetilde{\Lambda}$ denote the set of $x\in\Sigma$ such that the following properties (1)-(4) hold:

(1) $\displaystyle\lim_{t\to0}\frac{m(B^\phi(x,t)\cap\mathcal{P}(x))}{m(B^\phi(x,t))}=\sum_{A\in\mathcal{P}}\chi_A\mathbf{E}_m(\chi_A|\phi^{-1}\gamma)(x)>0.$

(2) $\displaystyle\lim_{t\to0}\frac{m(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\cap\mathcal{P}(x))}{m(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x))}=\sum_{A\in\mathcal{P}}\chi_A\mathbf{E}_m(\chi_A|\sigma^{-1}\phi^{-1}\gamma)(x)>0.$

(3) For all $q\in\mathbb{Q}^+$,

$$m_x^\eta(B^\pi(x,q)\cap\mathcal{P}(x))\ge\limsup_{t\to0}\frac{m\left(B^\pi(x,q)\cap\mathcal{P}(x)\cap B^\phi(x,t)\right)}{m\left(B^\phi(x,t)\right)},$$
$$m_x^\eta(U^\pi(x,q)\cap\mathcal{P}(x))\le\liminf_{t\to0}\frac{m\left(B^\pi(x,q)\cap\mathcal{P}(x)\cap B^\phi(x,t)\right)}{m\left(B^\phi(x,t)\right)},$$
$$m_x^\xi(B^{\pi\sigma}(x,q)\cap\mathcal{P}(x))\ge\limsup_{t\to0}\frac{m\left(B^{\pi\sigma}(x,q)\cap\mathcal{P}(x)\cap\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\right)}{m\left(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\right)},$$
$$m_x^\xi(U^{\pi\sigma}(x,q)\cap\mathcal{P}(x))\le\liminf_{t\to0}\frac{m\left(B^{\pi\sigma}(x,q)\cap\mathcal{P}(x)\cap\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\right)}{m\left(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)\right)},$$

where $U^\pi(x,q):=\pi^{-1}U(\pi x,q)$, $U^{\pi\sigma}(x,q):=\sigma^{-1}\pi^{-1}U(\pi\sigma x,q)$ and $U(z,q)$ denotes the open ball in $\mathbb{R}^k$ of radius q centered at z.

(4) $\displaystyle\lim_{t\to0}\frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x,t))}=f(x).$

Then we have $m(\widetilde{\Lambda})=1$ by Proposition 3.5, Lemma 3.7, Remarks 3.6, 3.8 and (77).

Now let $\Lambda = \widetilde{\Lambda} \cap \sigma^{-1}\widetilde{\Lambda}$. Then $m(\Lambda) = 1$. Fix $x \in \Lambda$ and $r \in (0, r_0)$. Let $q_1 \in \mathbb{Q}^+ \cap (r, cr/(c-\delta))$. Choose $q_2, q_3 \in \mathbb{Q}^+$ such that $q_1 < q_2 < cr/(c-\delta)$ and $q_2(c-\delta)\overline{\rho}(x) < q_3 < c\overline{\rho}(x)r$. By (73), we have $B^\pi(x, q_3) \cap \mathcal{P}(x) \supset B^{\pi\sigma}(x, q_2) \cap \mathcal{P}(x)$. It together with (76) yields

$$B^\pi(x, q_3) \cap \mathcal{P}(x) \cap B^\phi(x, t)) \supset B^{\pi\sigma}(x, q_2) \cap \mathcal{P}(x) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)). \tag{78}$$

Hence we have

$$\begin{aligned}
& \frac{m_x^\eta(B^\pi(x, c\overline{\rho}(x)r) \cap \mathcal{P}(x))}{m_{\sigma x}^\eta(B^\pi(\sigma x, r))} \\
\geq\ & \frac{m_x^\eta(B^\pi(x, q_3) \cap \mathcal{P}(x))}{m_{\sigma x}^\eta(U^{\pi\sigma}(x, q_1))} \\
\geq\ & \frac{\limsup_{t\to 0} m(B^\pi(x, q_3) \cap \mathcal{P}(x) \cap B^\phi(x, t))/m(B^\phi(x, t))}{\liminf_{t\to 0} m(B^\pi(\sigma x, q_1) \cap \phi^{-1}R_{t,x}(\phi\sigma x))/m(\phi^{-1}R_{t,x}(\phi\sigma x))} \\
& \quad \text{(by Lemma 3.7 and Remark 3.8)} \\
\geq\ & \lim_{t\to 0} \frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x, t))} \cdot \limsup_{t\to 0} \frac{m(B^\pi(x, q_3) \cap \mathcal{P}(x) \cap B^\phi(x, t))}{m(\sigma^{-1}B^\pi(\sigma x, q_1) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x))} \\
=\ & \lim_{t\to 0} \frac{m(\phi^{-1}R_{t,x}(\phi\sigma x))}{m(B^\phi(x, t))} \cdot \limsup_{t\to 0} \frac{m(B^\pi(x, q_3) \cap \mathcal{P}(x) \cap B^\phi(x, t))}{m(B^{\pi\sigma}(x, q_1) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x))}.
\end{aligned}$$

Denote

$$\begin{aligned}
X_t &:= m(B^{\pi\sigma}(x, q_2) \cap \mathcal{P}(x) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)), \\
Y_t &:= m(B^{\pi\sigma}(x, q_1) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)), \\
Z_t &:= m(\sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x)).
\end{aligned}$$

Using the property (4), we have

$$\begin{aligned}
& \frac{m_x^\eta(B^\pi(x, c\overline{\rho}(x)r) \cap \mathcal{P}(x))}{m_{\sigma x}^\eta(B^\pi(\sigma x, r))} \\
\geq\ & f(x) \cdot \limsup_{t\to 0} \frac{m(B^\pi(x, q_3) \cap \mathcal{P}(x) \cap B^\phi(x, t))}{m(B^{\pi\sigma}(x, q_1) \cap \sigma^{-1}\phi^{-1}R_{t,x}(\phi\sigma x))} \\
\geq\ & f(x) \cdot \limsup_{t\to 0} X_t/Y_t \qquad \text{(by (78))} \\
\geq\ & f(x) \cdot \limsup_{t\to 0} \frac{X_t/Z_t}{Y_t/Z_t} \geq f(x) \cdot \frac{\liminf_{t\to 0} X_t/Z_t}{\limsup_{t\to 0} Y_t/Z_t} \\
\geq\ & f(x) \cdot \frac{m_x^\xi(U^{\pi\sigma}(x, q_1) \cap \mathcal{P}(x))}{m_x^\xi(B^{\pi\sigma}(x, q_1))} \qquad \text{(by Lemma 3.7 and Remark 3.8)} \\
\geq\ & f(x) \cdot \frac{m_x^\xi(B^{\pi\sigma}(x, r) \cap \mathcal{P}(x))}{m_x^\xi(B^{\pi\sigma}(x, q_1))}.
\end{aligned}$$

Letting $q_1 \downarrow r$, we obtain (74). (75) follows from an analogous argument. □

Let (ϕ,π) denote the map $\Sigma\to\mathbb{R}^d\times\mathbb{R}^k$, $x\mapsto(\phi x,\pi x)$. It is easy to see that (ϕ,π) is the canonical projection w.r.t. the direct product of $\{T_i\}_{i=1}^{\ell}$ and $\{S_i\}_{i=1}^{\ell}$. In the following we give a general version of Theorem 2.6.

Theorem 6.2. *Let $m\in\mathcal{M}_\sigma(\Sigma)$. Then for m-a.e. $x\in\Sigma$, we have*

$$\limsup_{r\to 0}\frac{\log m_x^\eta(B^\pi(x,r))}{\log r}\le\frac{\mathbf{E}_m(g|\mathcal{I})(x)}{-\underline{\lambda}(x)}\quad\textit{and}\tag{79}$$

$$\liminf_{r\to 0}\frac{\log m_x^\eta(B^\pi(x,r))}{\log r}\ge\frac{\mathbf{E}_m(g|\mathcal{I})(x)}{-\overline{\lambda}(x)},\tag{80}$$

where

$$\begin{aligned}g\ :=\ &\mathbf{I}_m(\mathcal{P}|\sigma^{-1}\phi^{-1}\mathcal{B}(\mathbb{R}^d))-\mathbf{I}_m(\mathcal{P}|\phi^{-1}\mathcal{B}(\mathbb{R}^d))\\&+\mathbf{I}_m(\mathcal{P}|(\phi,\sigma)^{-1}\mathcal{B}(\mathbb{R}^d\times\mathbb{R}^k))-\mathbf{I}_m(\mathcal{P}|\sigma^{-1}(\phi,\pi)^{-1}\mathcal{B}(\mathbb{R}^d\times\mathbb{R}^k)),\end{aligned}$$

and $\overline{\lambda}(x)$, $\underline{\lambda}(x)$ denote the upper and lower Lyapunov exponents of $\{S_i\}_{i=1}^{\ell}$ at x (see Definition 2.5). In particular, if $\{S_i\}_{i=1}^{\ell}$ is m-conformal, we have

$$\lim_{r\to 0}\frac{\log m_x^\eta(B^\pi(x,r))}{\log r}=\frac{h_{(\phi,\pi)}(\sigma,m,x)-h_\phi(\sigma,m,x)}{\lambda(x)}.$$

Proof. It suffices to prove (79) and (80). For short we only prove (79). The proof of (80) is analogous.

We first prove the following inequality

$$\limsup_{r\to 0}\frac{\log m_x^\eta(B^\pi(x,r))}{\log r}\le\frac{\mathbf{E}_m(g|\mathcal{I})(x)}{\mathbf{E}_m(\log\overline{\rho}|\mathcal{I})(x)}\qquad m\text{-a.e.},\tag{81}$$

where $\overline{\rho}(x)=\|S'_{x_1}(\sigma x)\|$ for $x=(x_i)_{i=1}^\infty$. To see it, let $c>1$ so that

$$c\sup_{x\in\Sigma}\overline{\rho}(x)<1.$$

Let r_0 and f be given as in Proposition 6.1. For $n\in\mathbb{N}$ and $x\in\Sigma$, define

$$\overline{\rho}_n(x)=\overline{\rho}(x)\overline{\rho}(\sigma x)\cdots\overline{\rho}(\sigma^{n-1}x).$$

Write

$$\begin{aligned}H_n(x)&:=\log\frac{m_x^\eta\left(B^\pi(x,c^n\overline{\rho}_n(x)r_0)\right)}{m_{\sigma x}^\eta\left(B^\pi(\sigma x,c^{n-1}\overline{\rho}_{n-1}(\sigma x)r_0)\right)},\\G_n(x)&:=\log\frac{m_x^\eta\left(B^\pi(x,c^n\overline{\rho}_n(x)r_0)\cap\mathcal{P}(x)\right)}{m_x^\eta\left(B^\pi(x,c^n\overline{\rho}_n(x)r_0)\right)},\\W_n(x)&:=\log\frac{m_x^\xi\left(B^{\pi\sigma}(x,c^{n-1}\overline{\rho}_{n-1}(\sigma x)r_0)\cap\mathcal{P}(x)\right)}{m_x^\xi\left(B^{\pi\sigma}(x,c^{n-1}\overline{\rho}_{n-1}(\sigma x)r_0)\right)}.\end{aligned}$$

Then by Proposition 6.1 we have for m-a.e. x, $H_n(x)+G_n(x)\ge\log f(x)+W_n(x)$, that is,

$$H_n(x)\ge\log f(x)-G_n(x)+W_n(x).$$

However

$$\log m_x^\eta\left(B^\pi(x, c^n\overline{\rho}_n(x)r_0)\right) = \sum_{j=0}^{n-1} H_{n-j}(\sigma^j x) + \log m_{\sigma^n x}^\eta\left(B^\pi(\sigma^n x, r_0)\right).$$

Hence for m-a.e. x,

$$\begin{aligned}\frac{\log m_x^\eta\left(B^\pi(x, c^n\overline{\rho}_n(x)r_0)\right)}{n} &\geq \frac{1}{n}\sum_{j=0}^{n-1}\left[\log f(\sigma^j x) - G_{n-j}(\sigma^j x) + W_{n-j}(\sigma^j x)\right] \\ &\quad + \frac{1}{n}\log m_{\sigma^n x}^\eta\left(B^\pi(\sigma^n x, r_0)\right).\end{aligned}$$

Note that by Proposition 3.5,

$$\begin{aligned}&G_n \to G := -\mathbf{I}_m(\mathcal{P}|\hat{\eta} \vee \pi^{-1}\mathcal{B}(\mathbb{R}^k)),\\ &W_n \to W := -\mathbf{I}_m(\mathcal{P}|\sigma^{-1}\hat{\eta} \vee \sigma^{-1}\pi^{-1}\mathcal{B}(\mathbb{R}^k))\end{aligned}$$

pointwise and in L^1. By Lemma 4.13 and Proposition 3.9, we have for m-a.e. x,

$$\begin{aligned}\liminf_{n\to\infty} \frac{\log m_x^\eta\left(B^\pi(x, c^n\overline{\rho}_n(x)r_0)\right)}{n} &\geq \mathbf{E}_m((\log f - G + W)|\mathcal{I})(x) \\ &= \mathbf{E}_m(g|\mathcal{I})(x).\end{aligned}$$

In the meantime, by Birkhoff ergodic Theorem, we have

$$\lim_{n\to\infty} \frac{1}{n}\log(c^n\overline{\rho}_n(x)r_0) = \log c + \mathbf{E}_m(\log\overline{\rho}|\mathcal{I})(x) \quad m\text{-a.e.}$$

Hence we have

$$\begin{aligned}\limsup_{r\to 0} \frac{\log m_x^\eta\left(B^\pi(x,r)\right)}{\log r} &= \limsup_{n\to\infty} \frac{\log m_x^\eta\left(B^\pi(x, c^n\overline{\rho}_n(x)r_0)\right)}{\log(c^n\overline{\rho}_n(x)r_0)} \\ &\leq \frac{\mathbf{E}_m(g|\mathcal{I})(x)}{\log c + \mathbf{E}_m(\log\overline{\rho}|\mathcal{I})(x)}.\end{aligned}$$

Taking $c \to 1$, we obtain (81).

Let $q \in \mathbb{N}$. Considering the IFS $\{T_{i_1\dots i_q} :\ 1 \leq i_j \leq \ell,\ 1 \leq j \leq q\}$ and $\{S_{i_1\dots i_q} :\ 1 \leq i_j \leq \ell,\ 1 \leq j \leq q\}$, analogous to (81) we have

$$\limsup_{r\to 0} \frac{\log m_x^\eta(B^\pi(x,r))}{\log r} \leq \frac{\mathbf{E}_m(g_q|\mathcal{I})(x)}{\mathbf{E}_m(\log h_q|\mathcal{I})(x)}, \tag{82}$$

where

$$\begin{aligned}g_q \ := \ &\mathbf{I}_m(\mathcal{P}_0^{q-1}|\sigma^{-q}\phi^{-1}\mathcal{B}(\mathbb{R}^d)) - \mathbf{I}_m(\mathcal{P}_0^{q-1}|\phi^{-1}\mathcal{B}(\mathbb{R}^d)) \\ &+\mathbf{I}_m(\mathcal{P}_0^{q-1}|(\phi,\pi)^{-1}\mathcal{B}(\mathbb{R}^d\times\mathbb{R}^k)) - \mathbf{I}_m(\mathcal{P}_0^{q-1}|\sigma^{-q}(\phi,\pi)^{-1}\mathcal{B}(\mathbb{R}^d\times\mathbb{R}^k))\end{aligned}$$

and $h_q(x) := \|S'_{x_1\dots x_q}(\sigma^q x)\|$ for $x = (x_i)_{i=1}^\infty$.

Due to (23), we have $\mathbf{E}_m(g_q|\mathcal{I})(x) = q\mathbf{E}_m(g|\mathcal{I})(x)$. It is easily seen that $h_q(x)$ is sub-multiplicative in the sense that $h_{p+q}(x) \leq h_p(x)h_q(\sigma^p x)$. Thus by Kingman sub-additive ergodic theorem (cf. [63]), we have

$$\lim_{q\to\infty} \frac{1}{q}\mathbf{E}_m(\log h_q|\mathcal{I})(x) = -\underline{\lambda}(x) \quad \text{for } m\text{-a.e. } x.$$

Hence letting $q \to \infty$ in (82) we obtain (79). This finishes the proof of Theorem 6.2. $\square$

Proof of Theorem 2.6. In Theorem 6.2, we take $T_i(x) = x/2$ for all $1 \leq i \leq \ell$ to obtain Theorem 2.6. To see it, we know that the attractor of $\{T_i\}_{i=1}^{\ell}$ is just the singleton $\{0\}$. Hence η is the trivial partition $\{\Sigma, \emptyset\}$ of Σ, and thus we have $m_x^{\eta} \equiv m$. $\square$

7. Proofs of Theorem 2.11 and Theorem 2.12

7.1. **Proof of Theorem 2.11.** Let $\Phi = \{S_i\}_{i=1}^{\ell}$ be the direct product of k C^1 IFS $\Phi_1, \ldots, \Phi_k$, which are defined respectively on compact $X_i \subset \mathbb{R}^{q_i}$ ($i = 1, \ldots, k$). For each i, let Γ_i denote the canonical projection w.r.t. Φ_i, and let $\lambda_i(x)$ denote the Lyapunov exponent of Φ_i at x provided it exists.

Let $m \in \mathcal{M}_\sigma(\Sigma)$. Assume that $\Phi_1, \ldots, \Phi_k$ are m-conformal. Let Ω denote the collection of all permutations of $\{1, \ldots, k\}$. For $\tau \in \Omega$, we denote

$$\Lambda_\tau := \left\{x \in \Sigma : \ \lambda_i(x) \text{ exists for all } i, \ \lambda_{\tau(1)}(x) \leq \lambda_{\tau(2)}(x) \leq \cdots \leq \lambda_{\tau(k)}(x)\right\}.$$

Then $m\left(\bigcup_{\tau \in \Omega} \Lambda_\tau\right) = 1$. Let π denote the canonical projection associated with the IFS Φ. In the following we show that the local dimension $d(m \circ \pi^{-1}, \pi x)$ exists for m-a.e. $x \in \Sigma$.

Without loss of generality we only show that $d(m \circ \pi^{-1}, \pi x)$ exists for m-a.e. $x \in \Lambda_e$, where e denotes the identity in Ω. Here we may assume $m(\Lambda_e) > 0$. For other Λ_τ's, the proof is essentially identical under a change of coordinates.

For $i = 1, \ldots, k$, let π_i denote the canonical projection w.r.t. $\Phi_1 \times \cdots \times \Phi_i$. It is clear that $\pi = \pi_k$. Bear in mind that

$$\lambda_1(x) \leq \lambda_2(x) \leq \cdots \leq \lambda_k(x) \qquad (x \in \Lambda_e).$$

For $i = 1, \ldots, k$, we use $\{m_x^i\}$ to denote the family of conditional measures $\{m_x^{\eta_i}\}$ of m associated with the partition

$$\eta_i = \left\{\pi_i^{-1}(z) : \ z \in \prod_{t=1}^{i} \mathbb{R}^{q_t}\right\}.$$

For convenience, we use $\{m_x^0\}$ denote the family of conditional measures of m with the trivial partition $\{\Sigma, \emptyset\}$. It is clear that $m_x^0 = m$ for all $x \in \Sigma$.

For $i = 1, \ldots, k$, we give a metric d_i on $\prod_{t=1}^{i} \mathbb{R}^{q_t}$ by

$$d_i((z_1, \ldots, z_i), (w_1, \ldots, w_i)) = \sup_{1 \leq t \leq i} |z_t - w_t|_{\mathbb{R}^{q_t}}.$$

and define $d = d_k$. We claim that for any $x \in \Lambda_e$ and $\epsilon > 0$,

$$\eta_i(x) \cap \mathcal{P}_0^n(x) \subset B^\pi(x, e^{-n(\lambda_{i+1}(x) - \epsilon)}) \tag{83}$$

when n is large enough. Here $B^\pi(x, r)$ is defined as in (12). To see the claim, let $x \in \Lambda_e$ and $y \in \eta_i(x)$. Then $\pi_i y = \pi_i x$. Thus

$$d(\pi y, \pi x) = \sup_{1 \leq t \leq k} |\Gamma_t y, \Gamma_t x|_{\mathbb{R}^{q_t}} = \sup_{i+1 \leq t \leq k} |\Gamma_t y, \Gamma_t x|_{\mathbb{R}^{q_t}}.$$

Since $y \in \mathcal{P}_0^n(x)$ and $\lambda_{i+1}(x) \le \ldots \le \lambda_k(x)$, by Proposition 5.6, we have

$$d(\pi y, \pi x) \le e^{-n(\lambda_{i+1}(x)-\epsilon)}$$

when n is large enough, and (83) follows.

For $i = 0, 1, \ldots, k$ and $x \in \Sigma$, denote

$$h_i(x) = \lim_{n\to\infty} \frac{-\log m_x^i(\mathcal{P}_0^n(x))}{n+1}$$

provided that the limit exists. By Proposition 4.14,

$$h_i(x) = h(\sigma, m, x) - h_{\pi_i}(\sigma, m, x) \quad \text{for } m\text{-a.e. } x \in \Sigma. \tag{84}$$

For $i = 0, 1, \ldots, k-1$ and $x \in \Sigma$, denote

$$\vartheta_i(x) = \liminf_{r\to 0} \frac{\log m_x^i(B^{\Gamma_{i+1}}(x,r))}{\log r}.$$

By Theorem 6.2 and (84), we have

$$\vartheta_i(x) = \frac{h_{\pi_{i+1}}(\sigma, m, x) - h_{\pi_i}(\sigma, m, x)}{\lambda_{i+1}(x)} = \frac{h_i(x) - h_{i+1}(x)}{\lambda_{i+1}(x)} \tag{85}$$

for m-a.e. $x \in \Sigma$.

For $i = 0, 1, \ldots, k$ and $x \in \Sigma$, define

$$\overline{\delta}_i(x) = \limsup_{r\to 0} \frac{\log m_x^i(B^\pi(x,r))}{\log r}, \quad \underline{\delta}_i(x) = \liminf_{r\to 0} \frac{\log m_x^i(B^\pi(x,r))}{\log r}.$$

We claim that

(C1) $\overline{\delta}_k(x) = \underline{\delta}_k(x) = 0$ for all $x \in \Sigma$.
(C2) $h_i(x) - h_{i+1}(x) \ge \lambda_{i+1}(\overline{\delta}_i(x) - \overline{\delta}_{i+1}(x))$ for m-a.e. $x \in \Lambda_e$ and $i = 0, 1 \ldots, k-1$;
(C3) $\underline{\delta}_{i+1}(x) + \vartheta_i(x) \le \underline{\delta}_i(x)$ for m-a.e. $x \in \Lambda_e$ and $i = 0, 1 \ldots, k-1$;

It is easy to see that (C1)-(C3) together with (84)-(85) force that for m-a.e. $x \in \Lambda_e$, $\underline{\delta}_i(x) = \overline{\delta}_i(x)$ (we denoted the common value as $\delta_i(x)$) for $i = 0, \ldots, k$ and, furthermore

$$d(m \circ \pi^{-1}, \pi x) = \delta_0(x) = \sum_{i=0}^{k-1} \vartheta_i(x) = \sum_{i=0}^{k-1} \frac{h_i(x) - h_{i+1}(x)}{\lambda_{i+1}(x)}. \tag{86}$$

which is the desired result in Theorem 2.11. In the following we prove (C1)-(C3) respectively.

Proof of (C1). Since $\eta_k = \left\{\pi^{-1}(z) : z \in \prod_{t=1}^k \mathbb{R}^{q_t}\right\}$, we have

$$m_x^k(B^\pi(x,r)) = m_x^k(\eta_k(x)) = 1$$

for all $x \in \Sigma$. Thus $\overline{\delta}_k(x) = \underline{\delta}_k(x) = 0$ for all $x \in \Sigma$. □

Proof of (C2). We give a proof by contradiction, which is modified from [40, §10.2]. Assume that (C2) is not true. Then there exists $0 \le i \le k$ such that

$$h_i(x) - h_{i+1}(x) < \lambda_{i+1}(x)(\overline{\delta}_i(x) - \overline{\delta}_{i+1}(x))$$

on a subset of Λ_e with positive measure. Hence there exist $\alpha > 0$ and real numbers $h_i, h_{i+1}, \lambda_{i+1}, \overline{\delta}_i, \overline{\delta}_{i+1}$ with $\lambda_{i+1} > 0$ such that

$$h_i - h_{i+1} < \lambda_{i+1}(\overline{\delta}_i - \overline{\delta}_{i+1}) - \alpha \tag{87}$$

and for any $\epsilon > 0$, there exists $B_\epsilon \subset \Lambda_e$ with $m(B_\epsilon) > 0$ so that for $x \in B_\epsilon$,

$$|h_i(x) - h_i| < \epsilon/2, \quad |h_{i+1}(x) - h_{i+1}| < \epsilon/2, \quad |\lambda_{i+1}(x) - \lambda_{i+1}| < \epsilon/2$$

and

$$|\overline{\delta}_i(x) - \overline{\delta}_i| < \epsilon/2, \quad |\overline{\delta}_{i+1}(x) - \overline{\delta}_{i+1}| < \epsilon/2.$$

Fix $\epsilon > 0$. There exists $n_0 \colon B_\epsilon \to \mathbb{N}$ such that for m-a.e. $x \in B_\epsilon$ and $n > n_0(x)$, we have

(1) $\dfrac{\log m_x^{i+1}\left(B^\pi(x, e^{-n(\lambda_{i+1}-2\epsilon)})\right)}{-n(\lambda_{i+1}-2\epsilon)} \le \overline{\delta}_{i+1} + \epsilon$;

(2) $-\dfrac{1}{n}\log m_x^{i+1}(\mathcal{P}_0^n(x)) \ge h_{i+1} - \epsilon$ (by (84));

(3) $\eta_i(x) \cap \mathcal{P}_0^n(x) \subset B^\pi(x, e^{-n(\lambda_{i+1}-2\epsilon)})$ (by (83));

(4) $-\dfrac{1}{n}\log m_x^{i}(\mathcal{P}_0^n(x)) \le h_i + \epsilon$ (by (84)).

Take N_0 such that

$$\Delta := \{x \in B_\epsilon \colon n_0(x) \le N_0\}$$

has the positive measure. By Lemma 3.3 and Lemma 3.10, there exist $c > 0$ and $\Delta' \subset \Delta$ with $m(\Delta') > 0$ such that for $x \in \Delta'$, there exists $n = n(x) \ge N_0$ such that

(5) $\dfrac{m_x^{i+1}(L \cap \Delta)}{m_x^{i+1}(L)} \ge c$, where

$$L := B^\pi(x, e^{-n(\lambda_{i+1}-2\epsilon)});$$

(6) $\dfrac{\log m_x^{i}\left(B^\pi(x, 2e^{-n(\lambda_{i+1}-2\epsilon)})\right)}{-n(\lambda_{i+1}-2\epsilon)} > \overline{\delta}_i - \epsilon$;

(7) $\dfrac{\log(1/c)}{n} < \epsilon$.

Take $x \in \Delta'$ such that (1)–(7) are satisfied with $n = n(x)$. Denote $C = \eta_{i+1}(x)$ and $C' = \eta_i(x)$. Then by (5) and (1),

$$m_x^{i+1}(L \cap \Delta) \ge c m_x^{i+1}(L) \ge c e^{-n(\lambda_{i+1}-2\epsilon)(\overline{\delta}_{i+1}+\epsilon)}.$$

But for each $y \in L \cap \Delta$, we have by (2), $m_y^{i+1}(\mathcal{P}_0^n(y)) \le e^{-n(h_{i+1}-\epsilon)}$. It follows that the number of distinct $\mathcal{P}_0^n$-atoms intersecting $C \cap L \cap \Delta$ is larger than

$$m_x^{i+1}(L \cap \Delta) e^{n(h_{i+1}-\epsilon)}.$$

However each such a $\mathcal{P}_0^n$-atom, say $\mathcal{P}_0^n(y)$, intersects $C' \cap L \cap \Delta$, and this together with (3) guarantees that $C' \cap \mathcal{P}_0^n(y)$ is contained in $C' \cap B^\pi(x, 2e^{-n(\lambda_{i+1}-2\epsilon)})$.

To see this, let $z \in \mathcal{P}_0^n(y) \cap C' \cap L \cap \Delta$. Since $z \in \Delta$, we have $d(\pi z, \pi x) \leq e^{-n(\lambda_{i+1}-2\epsilon)}$. Thus

$$\mathcal{C}' \cap \mathcal{P}_0^n(y) = \eta_i(z) \cap \mathcal{P}_0^n(z) \subset B^\pi(x, e^{-n(\lambda_{i+1}-2\epsilon)}) \subset B^\pi(x, 2e^{-n(\lambda_{i+1}-2\epsilon)}).$$

Meanwhile by (4), $m_x^i(\mathcal{P}_0^n(y)) \geq e^{-n(h_i+\epsilon)}$ (for $w \in \mathcal{P}_0^n(y) \cap C' \cap L$, we have $\eta_i(x) = \eta_i(w)$ and thus $m_x^i(\mathcal{P}_0^n(y)) = m_w^i(\mathcal{P}_0^n(w))$). Hence we have

$$\begin{aligned} m_x^i(B^\pi(x, 2e^{-n(\lambda_{i+1}-2\epsilon)})) &\geq \#\{\mathcal{P}_0^n\text{-atoms intersecting } C' \cap L \cap \Delta\} \cdot e^{-n(h_i+\epsilon)} \\ &\geq m_x^{i+1}(L \cap \Delta) e^{n(h_{i+1}-\epsilon)} e^{-n(h_i+\epsilon)} \\ &\geq c e^{-n(\lambda_{i+1}-2\epsilon)(\overline{\delta}_{i+1}+\epsilon)} e^{n(h_{i+1}-\epsilon)} e^{-n(h_i+\epsilon)}. \end{aligned}$$

Comparing this with (6), we have

$$\begin{aligned} &(\lambda_{i+1} - 2\epsilon)(\overline{\delta}_i - \epsilon) \\ &\quad \leq (\lambda_{i+1} - 2\epsilon)(\overline{\delta}_{i+1} + \epsilon)(\lambda_i - 2\epsilon) + \frac{\log(1/c)}{n} + h_i - h_{i+1} + 2\epsilon \\ &\quad \leq (\lambda_{i+1} - 2\epsilon)(\overline{\delta}_{i+1} + \epsilon)(\lambda_i - 2\epsilon) + h_i - h_{i+1} + 3\epsilon. \end{aligned}$$

Taking $\epsilon \to 0$ yields $h_i - h_{i+1} \geq \lambda_{i+1}(\overline{\delta}_i - \overline{\delta}_{i+1})$, which leads to a contradiction with (87). □

Proof of (C3). Here we give a proof by contradiction, adopting an idea from the proof of [40, Lemma 11.3.1]. Assume that (C3) is not true. Then there exists $0 \leq i \leq k-1$ such that $\underline{\delta}_{i+1}(x) + \vartheta_i(x) > \underline{\delta}_i(x)$ on a subset of Λ_e with positive measure. Hence there exists $\beta > 0$ and real numbers $\underline{\delta}_i, \underline{\delta}_{i+1}, \lambda_i$ such that

$$\underline{\delta}_{i+1} + \vartheta_i > \underline{\delta}_i + \beta, \tag{88}$$

and for any $\epsilon > 0$, there exists $A_\epsilon \subset \Lambda_e$ with $m(A_\epsilon) > 0$ so that for $x \in A_\epsilon$,

$$|\underline{\delta}_i(x) - \underline{\delta}_i| < \epsilon/2, \quad |\underline{\delta}_{i+1}(x) - \underline{\delta}_{i+1}| < \epsilon/2, \quad |\vartheta_i(x) - \vartheta_i| < \epsilon/2. \tag{89}$$

Let $0 < \epsilon < \beta/4$. Find N_1 and a set $A'_\epsilon \subset A_\epsilon$ with $m(A'_\epsilon) > 0$ such that for $x \in A'_\epsilon$ and $n > N_1$,

$$m_x^{i+1}\left(B^\pi(x, 2e^{-n})\right) \leq e^{-n(\underline{\delta}_{i+1}-\epsilon)}. \tag{90}$$

By Lemma 3.3 and Lemma 3.10, we can find $c > 0$ and $A''_\epsilon \subset A'_\epsilon$ with $m(A''_\epsilon) > 0$ and N_2 such that for all $x \in A''_\epsilon$ and $n \geq N_2$,

$$\frac{m_x^i(A'_\epsilon \cap B^\pi(x, e^{-n}))}{m_x^i(B^\pi(x, e^{-n}))} > c.$$

For $x \in A''_\epsilon$ and $n \geq N_2$, we have

$$\begin{aligned} m_x^i(B^\pi(x, e^{-n})) &\leq c^{-1} m_x^i(A'_\epsilon \cap B^\pi(x, e^{-n})) \\ &= c^{-1} \int m_y^{i+1}(A'_\epsilon \cap B^\pi(x, e^{-n}))\, dm_x^i(y) \\ &= c^{-1} \int_{B^{\Gamma_{i+1}}(x, e^{-n})} m_y^{i+1}(A'_\epsilon \cap B^\pi(x, e^{-n}))\, dm_x^i(y). \end{aligned} \tag{91}$$

Let $y \in \eta_i(x)$ such that $\eta_{i+1}(y) \cap A'_\epsilon \cap B^\pi(x, e^{-n}) \neq \emptyset$. Then there exists $w \in A'_\epsilon \cap B^\pi(x, e^{-n})$ such that $\pi_{i+1} y = \pi_{i+1} w$. Hence $A'_\epsilon \cap B^\pi(x, e^{-n}) \subset B^\pi(w, 2e^{-n})$ and by (90)

$$\begin{aligned} m_y^{i+1}(A'_\epsilon \cap B^\pi(w, e^{-n})) &= m_w^{i+1}(A'_\epsilon \cap B^\pi(w, e^{-n})) \\ &\leq m_w^{i+1}(B^\pi(w, 2e^{-n})) \\ &\leq e^{-n(\underline{\delta}_{i+1} - \epsilon)}. \end{aligned}$$

Combining it with (91), we have

$$m_x^i(B^\pi(x, e^{-n})) \leq c^{-1} e^{-n(\underline{\delta}_{i+1} - \sigma)} m_x^i(B^{\Gamma_{i+1}}(x, e^{-n})) \qquad (x \in A''_\epsilon,\ n \geq N_2).$$

Letting $n \to \infty$, we obtain $\underline{\delta}_i(x) \geq \underline{\delta}_{i+1} - \epsilon + \vartheta_i(x)$ for $x \in A''_\epsilon$. Combining it with (89) yields

$$\underline{\delta}_i \geq \underline{\delta}_{i+1} + \vartheta_i - 4\epsilon \geq \underline{\delta}_{i+1} + \vartheta_i - \beta,$$

which contradicts (88). $\square$

7.2. **Proof of Theorem 2.12.**

Definition 7.1. A real square matrix A is called *asymptotically similar* if all the (complex) eigenvalues of A are equal in modulus. Correspondingly, a linear transformation T on a finite-dimensional vector space V is called *asymptotically similar* if its representation matrix (associated with some basis of V) is asymptotically similar.

Lemma 7.2. *Let $(A_1, \ldots, A_\ell)$ be an ℓ-tuple of commuting linear transformations on $\mathbb{R}^d$. Then there are subspaces $V_1, \ldots, V_k$ of $\mathbb{R}^d$ such that*

(i) $\mathbb{R}^d = V_1 \oplus \cdots \oplus V_k$;

(ii) *V_i is A_j-invariant for $1 \leq i \leq k$ and $1 \leq j \leq \ell$;*

(iii) *The restriction of A_j on V_i is asymptotically similar for $1 \leq i \leq k$ and $1 \leq j \leq \ell$.*

Proof. For brevity, we only prove the lemma in the case $\ell = 2$. The reader will see that the idea works for all cases.

Let S, T be two commuting linear transformations on $\mathbb{R}^d$. Let f denote the real minimal polynomial of S. Suppose $f = f_1^{t_1} \cdots f_p^{t_p}$ is the decomposition of f into powers of distinct, real irreducible monic factors f_i. Let W_i denote the null space of $[f_i(S)]^{t_i}$, $i = 1, \ldots, p$. Then W_i's are S-invariant and $\mathbb{R}^d = W_1 \oplus \cdots \oplus W_p$ (cf. [62, Theorem 7.3]). Moreover S_{W_i}, the restriction of S on W_i, is asymptotically similar.

Since $ST = TS$, W_i is also T-invariant for each i. But T_{W_i} may be not asymptotically similar. However, as above, for each i, we can decomposed W_i into $W_i = W_{i,1} \oplus \cdots \oplus W_{i,u_i}$ such that $W_{i,j}$ are the null spaces corresponding to some factors of the minimal polynomial of T_{W_i}. Again, $W_{i,j}$ is T_{W_i}-invariant and S_{W_i}-invariant. Furthermore $T_{W_{i,j}}$ and $S_{W_{i,j}}$ are asymptotically similar. Hence $\mathbb{R}^d = \bigoplus_{i,j} W_{i,j}$ is the desired decomposition for S and T. $\square$

Proof of Theorem 2.12. Let $\{S_i\}_{i=1}^{\ell}$ be the IFS given in the theorem. By Lemma 7.2, there is a non-singular linear transformation Q on $\mathbb{R}^d$ such that $\{QS_iQ^{-1}\}_{i=1}^{\ell}$ is the direct product of k asymptotically conformal IFS. Hence the desired result follows from Theorem 2.11. □

8. A VARIATIONAL PRINCIPLE ABOUT DIMENSIONS OF SELF-CONFORMAL SETS

In this section, we assume that K is the attractor of a C^1 weakly conformal IFS $\Phi = \{S_i\}_{i=1}^{\ell}$ on a compact set $X \subset \mathbb{R}^d$. The main result of this section is the following variational principle.

Theorem 8.1. *Under the above setting, we have*

$$\begin{aligned}
(92)\quad \dim_H K &= \dim_B K \\
(93)\quad &= \sup\left\{\dim_H \mu:\ \mu = m\circ\pi^{-1},\ m\in\mathcal{M}_\sigma(\Sigma),\ m \textit{ is ergodic}\right\} \\
(94)\quad &= \max\left\{\dim_H \mu:\ \mu = m\circ\pi^{-1},\ m\in\mathcal{M}_\sigma(\Sigma)\right\} \\
(95)\quad &= \sup\left\{\frac{h_\pi(\sigma,m)}{\int \lambda\, dm}:\ m\in\mathcal{M}_\sigma(\Sigma)\right\}.
\end{aligned}$$

Proof. Without loss of generality we assume that $\overline{\dim}_B(K) > 0$, where $\overline{\dim}_B$ denotes the upper box-counting dimension (cf. [13]). Let

$$0 < t_3 < t_2 < t_1 < \overline{\dim}_B(K).$$

We first prove that there is an ergodic measure $m \in \mathcal{M}_\sigma(\Sigma)$ such that $\dim_H m\circ\pi^{-1} \geq t_3$. To achieve this, let $\alpha = \frac{t_2}{t_3} - 1$ and let r_0 be given as in Corollary 5.5. Since $\overline{\dim}_B(K) > t_1$, for any $0 < \epsilon < r_0$, there exist $r \in (0,\epsilon)$ and integer $N \geq r^{-t_1}$ such that there are disjoint closed balls $B(z_i, r)$ $(i = 1,\ldots,N)$ with centers $z_i \in K$. By Corollary 5.5, we can find words $w_i \in \Sigma^*$ $(i = 1,\ldots,N)$ such that $S_{w_i}(K) \subset B(z_i,r)$ and

$$(96)\qquad |S_{w_i}(x) - S_{w_i}(y)| \geq r^{1+\alpha}|x-y| \qquad (x,y\in K).$$

This implies $r^{1+\alpha}\mathrm{diam}(K) \leq \mathrm{diam}(S_{w_i}(K)) \leq 2r$. According to this fact and (59), there exist two positive constants A, B (independent of r) such that

$$B\log(1/r) \leq |w_i| \leq A\log(1/r) \text{ for all } 1 \leq i \leq N.$$

Hence by the pigeon hole principle, there is a subset $\mathcal{J}$ of $\{1,\ldots,N\}$ with cardinality

$$\#\mathcal{J} \geq \frac{N}{(A-B)\log(1/r)+1} \geq \frac{r^{-t_1}}{(A-B)\log(1/r)+1} \geq r^{-t_2}$$

such that the words w_i $(i\in\mathcal{J})$ have the same length, say n.

Now we adopt an argument from the proof of [12, Theorem 4]. Let

$$\delta = \min\{d(B(z_i,r),B(z_j,r)):\ i,j\in\mathcal{J}, i\neq j\}.$$

For any positive integer q and distinct sequences $i_1, \ldots, i_q$ and $j_1, \ldots, j_q$ taking values in $\mathcal{J}$, let k be the least integer such that $i_k \neq j_k$. Applying (96) $(k-1)$ times, we have

$$\begin{aligned} & d(S_{w_{i_1}} \circ \cdots \circ S_{w_{i_q}}(K), S_{w_{j_1}} \circ \cdots \circ S_{w_{j_q}}(K)) \\ & \quad \geq r^{(1+\alpha)(k-1)} d(B(z_{i_k}, r), B(z_{j_k}, r)) \geq r^{q(1+\alpha)}\delta. \end{aligned}$$

Define a measure η on the class of finite unions of sets $S_{w_{i_1}} \circ \cdots \circ S_{w_{i_q}}(K)$ by letting $\eta(S_{w_{i_1}} \circ \cdots \circ S_{w_{i_q}}(K)) = (\#\mathcal{J})^{-q}$. This extends to a measure η on the σ-algebra generating by these sets. Let U be any subset of K with $\text{diam}(U) < \delta$ and let q be the least integer such that

$$r^{(q+1)(1+\alpha)}\delta \leq \text{diam}(U) < r^{q(1+\alpha)}\delta.$$

Then U intersects at most one set $S_{w_{i_1}} \circ \cdots \circ S_{w_{i_q}}(K)$, hence

$$\begin{aligned} \eta(U) & \leq (\#\mathcal{J})^{-q} \leq r^{t_2 q} \leq r^{-t_2}\delta^{-t_2/(1+\alpha)}\text{diam}(U)^{t_2/(1+\alpha)} \\ & = r^{-t_2}\delta^{-t_3}\text{diam}(U)^{t_3}. \end{aligned}$$

This implies $\dim_H \eta \geq t_3$.

We point out that the measure η constructed as above is, indeed, the projection of a σ^n-invariant and ergodic measure ν under π. Actually ν is the unique measure on Σ satisfying

$$\nu([w_{i_1} \ldots w_{i_q}]) = (\#\mathcal{J})^{-q} \qquad (q \in \mathbb{N},\ i_1, \ldots, i_q \in \mathcal{J}).$$

Applying Theorem 2.8 to the IFS $\{S_{w_i} :\ i \in \mathcal{J}\}$, We have

$$\dim_H \eta = \dim_H \nu \circ \pi^{-1} = \frac{h_\pi(\sigma^n, \nu)}{-\int \log \|S'_{x_1 \ldots x_n}(\pi\sigma^n x)\| d\nu}.$$

Take $m = \frac{1}{n}\sum_{i=0}^{n-1} \nu \circ \sigma^{-i}$. Then m is σ-invariant and ergodic. Applying Theorem 2.8 and Proposition 4.3, we have

$$\begin{aligned} \dim_H m \circ \pi^{-1} & = \frac{h_\pi(\sigma, m)}{-\int \log \|S'_{x_1}(\pi\sigma x)\| dm} = \frac{h_\pi(\sigma^n, \nu)}{-\int \log \|S'_{x_1 \ldots x_n}(\pi\sigma^n x)\| d\nu} \\ & = \dim_H \eta \geq t_3. \end{aligned}$$

Since $t_3 < \overline{\dim}_B K$ is arbitrarily given, we obtain (92) and (93). To show (94), let (m_i) be a sequence of measures in $\mathcal{M}_\sigma(\Sigma)$ with

$$\lim_{i \to \infty} \dim_H m_i \circ \pi^{-1} = \dim_H K.$$

Take a sequence of positive numbers (a_i) such that $\sum_{i=1}^\infty a_i = 1$. Then $m = \sum_{i=1}^\infty a_i m_i$ is an element in $\mathcal{M}_\sigma(\Sigma)$ with

$$\dim_H m \circ \pi^{-1} = \sup_i \dim_H m_i \circ \pi^{-1} = \dim_H K.$$

To show (95), according to (93), it suffices to show that

$$\dim_H m \circ \pi^{-1} \geq \frac{h_\pi(\sigma, m)}{-\int \log \|S'_{x_1}(\pi\sigma x)\|\ dm(x)} \qquad (m \in \mathcal{M}_\sigma(\Sigma)). \tag{97}$$

Fix m and let $\mu = m \circ \pi^{-1}$. Denote by Λ the righthand side of (97). By Theorem 2.8, $d(\mu, z)$ exists for μ-a.e. $z \in \mathbb{R}^d$. Hence to show (97), we only need to show that for any $\epsilon > 0$, there is a Borel set $E \subset \mathbb{R}^d$ such that $\mu(E) > 0$ and $d(\mu, z) \geq \Lambda - \epsilon$ for $z \in E$. Assume this is false. Then $d(\mu, z) < \Lambda - \epsilon$ for μ-a.e. $z \in \mathbb{R}^d$. Thus by Theorem 2.8 again, we have

$$h_\pi(\sigma, m, x) < \lambda(x)(\Lambda - \epsilon) \qquad \text{for } m\text{-a.e. } x \in \Sigma.$$

Taking integration w.r.t. m on both sides yields

$$h_\pi(\sigma, m) < (\Lambda - \epsilon) \int \lambda \, dm,$$

which leads to a contradiction. □

Remark 8.2. Assume that $\{S_i\}_{i=1}^{\ell}$ is a weakly conformal IFS which satisfies the AWSC (see Definition 2.14). Then the supremum in (93) and (95) can be attained by ergodic measures. To see this, by Proposition 4.20, the map $m \mapsto h_\pi(\sigma, m)$ is upper semi-continuous on $\mathcal{M}_\sigma(\Sigma)$, hence the supremum in (95) is attained at some member, say m_0, in $\mathcal{M}_\sigma(\Sigma)$. Let $m_0 = \int \nu \, d\mathbb{P}(\nu)$ be the ergodic decomposition of m_0. By Theorem 2.2(ii), we have

$$\dim_H K = \frac{h_\pi(\sigma, m_0)}{\int \lambda \, dm_0} = \frac{\int h_\pi(\sigma, \nu) \, d\mathbb{P}(\nu)}{\int\int \lambda \, d\nu \, d\mathbb{P}(\nu)}.$$

Since $\frac{h_\pi(\sigma,\nu)}{\int \lambda \, d\nu} \leq \dim_H K$ for each ν, the above equality implies that $\frac{h_\pi(\sigma,\nu)}{\int \lambda \, d\nu} = \dim_H K$ for $\mathbb{P}$-a.e. ν. Hence the supremum in (95) can be attained at some ergodic measure, so do the supremum in (93).

9. Proof of Theorem 2.15

We first present some lemmas.

Lemma 9.1. *Let $\{S_i\}_{i=1}^{\ell}$ be an IFS with attractor K. For $n \in \mathbb{N}$, write $\Sigma_n = \{1, \ldots, \ell\}^n$ and denote*

$$N_n = \#\{S_u : \; u \in \Sigma_n\}.$$

Then

(i) $\sup\{h_\pi(\sigma, m) : \; m \in \mathcal{M}_\sigma(\Sigma)\} \leq \frac{\log N_n}{n}$.

(ii) *Let $t_n = \sup_{x \in \mathbb{R}^d} \#\{S_u : \; u \in \Sigma_n, x \in S_u(K)\}$. Then*

$$\sup\{h_\pi(\sigma, m) : \; m \in \mathcal{M}_\sigma(\Sigma), \; m \textit{ is ergodic}\} \geq \frac{\log N_n - \log t_n}{n}.$$

Proof. We first show (i). Let $n \in \mathbb{N}$ and $m \in \mathcal{M}_\sigma(\Sigma)$. By the definition of N_n, we can construct a subset Ω of Σ_n with $\#\Omega = N_n$ such that for any $u \in \Sigma_n$, there exists $w \in \Omega$ so that $S_u = S_w$. Hence there is a map $g : \; \Sigma_n \to \Omega$ such that $S_u = S_{g(u)}$ for each $u \in \Sigma_n$. Let $(\Omega^{\mathbb{N}}, T)$ denote the one-sided full shift over Ω. Define $G : \; \Sigma \to \Omega^{\mathbb{N}}$ by

$$G((x_i)_{i=0}^{\infty}) = (w_j)_{j=1}^{\infty} \qquad ((x_i)_{i=1}^{\infty} \in \Sigma),$$

where $w_j = g(x_{(j-1)n+1}x_{(j-1)n+2}\cdots x_{jn})$. Let $\widetilde{\pi}:\ \Omega^{\mathbb{N}} \to \mathbb{R}^d$ denote the canonical projection w.r.t. the IFS $\{S_u:\ u \in \Omega\}$. Then by Lemma 4.23(ii), we have

$$h_\pi(\sigma^n, m) = h_{\widetilde{\pi}}(T, m \circ G^{-1}) \le \log(\#\Omega) = \log N_n.$$

It follows that $h_\pi(\sigma, m) \le \log N_n/n$. This proves (i).

To show (ii), let ν be the Bernoulli measure on $\Omega^{\mathbb{N}}$ with probability weight $(1/N_n, \dots, 1/N_n)$. Then ν can be viewed as a σ^n-invariant measure on Σ. By Lemma 4.23(ii), we have $h_\pi(\sigma^n, \nu) = h_{\widetilde{\pi}}(T, \nu)$. Note that for $x \in \mathbb{R}^d$, there are at most t_n words u in Ω such that $x \in S_u(K)$. By Corollary 4.22, we have

$$h_{\widetilde{\pi}}(T, \nu) \ge h(T, \nu) - \log t_n = \log N_n - \log t_n.$$

Let $\mu = \frac{1}{n}\sum_{i=0}^{n-1} \nu \circ \sigma^{-i}$. Then μ is σ-invariant and ergodic, furthermore

$$h_\pi(\sigma, \mu) = \frac{1}{n} h_\pi(\sigma^n, \nu) = \frac{1}{n} h_{\widetilde{\pi}}(T, \nu) \ge (\log N_n - \log t_n)/n,$$

as desired. $\square$

Lemma 9.2. *Let $\Phi = \{S_i\}_{i=1}^{\ell}$ be an affine IFS on $\mathbb{R}^d$ given by*

$$S_i(x_1, \dots, x_d) = (\rho_1 x_1, \cdots, \rho_d x_d) + (a_{i,1}, \dots, a_{i,d}),$$

where $1 > \rho_1 > \rho_2 > \cdots > \rho_d > 0$ and $a_{i,j} \in \mathbb{R}$. Let K denote the attractor of Φ, and write $\lambda_j = \log(1/\rho_j)$ for $j = 1, \dots, d$ and $\lambda_{d+1} = \infty$. View Φ as the direct product of $\Phi_1, \dots, \Phi_d$, where $\Phi_j = \{S_{i,j}(x_j) = \rho_j x_j + a_{i,j}\}_{i=1}^{\ell}$. Let π_j denote the canonical projection w.r.t. the IFS $\Phi_1 \times \cdots \times \Phi_j$. Then we have

(98)
$$\sum_{j=1}^{d}\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) H_j \le \underline{\dim}_B(K) \le \overline{\dim}_B(K) \le \sum_{j=1}^{d}\left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) \widetilde{H}_j,$$

with $H_j = \sup\left\{h_{\pi_j}(\sigma, m):\ m \in \mathcal{M}_\sigma(\Sigma)\right\}$ and

$$\widetilde{H}_j = \lim_{n\to\infty} \frac{\log \#\left\{S_u^{(j)}:\ u \in \Sigma_n\right\}}{n},$$

where $\left\{S_i^{(j)}\right\}_{i=1}^{\ell}$ is the IFS $\Phi_1 \times \cdots \times \Phi_j$ on $\mathbb{R}^j$.

Proof. Without loss of generality we assume that

$$S_i([0,1]^d) \subset [0,1]^d \qquad (i = 1, \dots, \ell).$$

For $n \in \mathbb{N}$, we write

$$N_n^{(j)} = \#\{S_u^{(j)}:\ u \in \Sigma_n\} \qquad (j = 1, \dots, d),$$

and

$$q_d(n) = n, \quad q_j(n) = \left[\left(\frac{\log \rho_d}{\log \rho_j} - \frac{\log \rho_d}{\log \rho_{j+1}}\right) n\right] \text{ for } 1 \le j \le d-1,$$

where $[x]$ denotes the integral part of x.

Construct $\Omega_{n,j} \subset \Sigma_{q_j(n)}$ ($j = 1, \ldots, d$) such that $\#\Omega_{n,j} = N^{(j)}_{q_j(n)}$ and for each $u \in \Sigma_{q_j(n)}$, there is $w \in \Omega_{n,j}$ so that $S^{(j)}_u = S^{(j)}_w$. Then the family of following rectangles

$$\prod_{j=1}^{d} S_{w_d w_{d-1} \cdots w_j, j}([0,1]) \qquad (w_1 \in \Omega_{n,1}, \ldots, w_d \in \Omega_{n,d}) \tag{99}$$

is a cover of K. To see it, let $u_j \in \Sigma_{q_j(n)}$ ($j = 1, \ldots, d$). Then we can find $w_j \in \Omega_{n,j}$ ($j = 1, \ldots, d$) such that $S^{(j)}_{u_j} = S^{(j)}_{w_j}$. Hence

$$\begin{aligned} S_{u_d u_{d-1} \ldots u_1}(K) &\subset S_{u_d u_{d-1} \ldots u_1}([0,1]^d) \subset \prod_{j=1}^{d} S_{u_d u_{d-1} \cdots u_1, j}([0,1]) \\ &\subset \prod_{j=1}^{d} S_{u_d u_{d-1} \cdots u_j, j}([0,1]) = \prod_{j=1}^{d} S_{w_d w_{d-1} \cdots w_j, j}([0,1]). \end{aligned}$$

It follows that the family of rectangles in (99) covers K. One can check that each rectangle in (99) is an almost $(\rho_d)^n$-cube. Hence by the definition of box-counting dimension, we have

$$\begin{aligned} \overline{\dim}_B K &\leq \limsup_{n \to \infty} \frac{\prod_{j=1}^{d} \#\Omega_{n,j}}{-\log(\rho_d)^n} = \limsup_{n \to \infty} \frac{\prod_{j=1}^{d} N^{(j)}_{q_j(n)}}{-\log(\rho_d)^n} \\ &= \sum_{j=1}^{d} \left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}} \right) \widetilde{H}_j. \end{aligned}$$

This proves one part of (98).

To see the other part of (98), for $j = 1, \ldots, d$, let $\mathcal{Q}_j$ denote the collection $\{[0,1)^j + \alpha : \ \alpha \in \mathbb{Z}^j\}$, and define

$$M^{(j)}_n = \#\{Q \in \mathcal{Q}_j : \ \mathrm{diag}(\rho_1^n, \ldots, \rho_j^n) Q \cap K_j \neq \emptyset\},$$

where K_j denotes the attractor of $\Phi_1 \times \cdots \times \Phi_j$. Then by Proposition 4.18(ii), we have $H_j = \lim_{n \to \infty} \frac{\log M^{(j)}_n}{n}$. We claim that for $n \in \mathbb{N}$, there exists a subset $\overline{\Omega}_{n,j} \subset \Sigma_n$ with cardinality $\geq 7^{-j} M^{(j)}_n$ such that

$$S^{(j)}_w([0,1]^j) \cap S^{(j)}_{w'}([0,1]^j) = \emptyset \text{ for all } w, w' \in \overline{\Omega}_{n,j} \text{ with } w \neq w'. \tag{100}$$

To show the claim, we construct a finite subset of $\mathcal{Q}_j$, denoted by $W^{(j)}_n$, such that (i) $\#W^{(j)}_n \geq 7^{-j} M^{(j)}_n$; (ii) $\mathrm{diag}(\rho_1^n, \ldots, \rho_j^n) Q \cap K_j \neq \emptyset$ for each $Q \in W^{(j)}_n$; (iii) $2Q \cap 2\widetilde{Q} = \emptyset$ for $Q, \widetilde{Q} \in W^{(j)}_n$ with $Q \neq \widetilde{Q}$, where $2Q := \bigcup_{Q' \in \mathcal{Q}_j : \, Q' \cap Q \neq \emptyset} Q'$. For each $Q \in W^{(j)}_n$, since $\mathrm{diag}(\rho_1^n, \ldots, \rho_j^n) Q \cap K_j \neq \emptyset$, we can pick a word $w(Q) \in \Sigma_n$ such that $\mathrm{diag}(\rho_1^n, \ldots, \rho_j^n) Q \cap S^{(j)}_{w(Q)} K_j \neq \emptyset$ and hence

$$\mathrm{diag}(\rho_1^n, \ldots, \rho_j^n) Q \cap S^{(j)}_{w(Q)}([0,1]^j) \neq \emptyset.$$

Denote $\overline{\Omega}_{n,j} = \{w(Q) : \quad Q \in W_n^{(j)}\}$. The separation condition (iii) for the elements in $W_n^{(j)}$ guarantees (100). This finishes the proof of the claim.

As above, we can construct $\overline{\Omega}_{n,j}$ well for each $j = 1, \ldots, d$ and $n \in \mathbb{N}$. Now fix n and consider the following collection of rectangles:

$$\prod_{j=1}^{d} S_{w_d w_{d-1} \cdots w_j, j}([0,1]) \qquad (w_j \in \overline{\Omega}_{q_j(n),j},\ 1 \leq j \leq d).$$

It is clear that the above rectangles are almost $(\rho_d)^n$-cubes and each of them intersects with K. Furthermore they are disjoint due to (100). Hence by the definition of box-counting dimension, we have

$$\begin{aligned} \underline{\dim}_B(K) &\geq \liminf_{n\to\infty} \frac{\prod_{j=1}^{d} \#\overline{\Omega}_{q_j(n),j}}{-\log(\rho_d)^n} \geq \liminf_{n\to\infty} \frac{\prod_{j=1}^{d} 7^{-j} M_{q_j(n)}^{(j)}}{-\log(\rho_d)^n} \\ &= \sum_{j=1}^{d} \left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) H_j. \end{aligned}$$

This finishes the proof of (98). □

Proof of Theorem 2.15. We divide the proof into two steps:
Step 1. Show the variational principle for $\dim_H K$.

We first give an upper bound for $\dim_H K$. Fix $n \in \mathbb{N}$. Define

$$N_j = \#\{S_u^{(j)} :\ u \in \Sigma_n\} \qquad (j = 1, \ldots, d),$$

where $\{S_i^{(j)}\}_{i=1}^{\ell}$ denotes the IFS $\Phi_1 \times \cdots \times \Phi_j$. Then we can construct

$$\Omega_j \subset \Sigma_n \quad (j = d, d-1, \ldots, 1)$$

such that $\#\Omega_j = N_j$, $\Sigma_n \supset \Omega_d \supset \Omega_{d-1} \supset \cdots \supset \Omega_1$ and furthermore, for each $u \in \Sigma_n$ and $1 \leq j \leq d$, there is $w_j \in \Omega_j$ such that $S_u^{(j)} = S_{w_j}^{(j)}$. Hence there are natural maps $\theta_d, \theta_{d-1}, \ldots, \theta_1$ with

$$\Sigma_n \xrightarrow{\theta_d} \Omega_d \xrightarrow{\theta_{d-1}} \Omega_{d-1} \xrightarrow{\theta_{d-2}} \cdots \xrightarrow{\theta_2} \Omega_2 \xrightarrow{\theta_1} \Omega_1$$

such that $S_u^{(j)} = S_{\theta_j(u)}^{(j)}$ for any $1 \leq j \leq d$ and $u \in \Omega_{j+1}$, with convention $\Omega_{d+1} = \Sigma_n$.

Let $Z_d : \Omega_d \to \mathbb{R}$ be the indicator of Ω_d, i.e., $Z_d(u) = 1$ for all $u \in \Omega_d$. Define

$$Z_{d-1}(w) = \sum_{u \in \theta_{d-1}^{-1}(w)} Z_d(u) \qquad (w \in \Omega_{d-1}).$$

Define inductively

$$Z_j(w) = \sum_{u \in \theta_j^{-1}(w)} Z_{j+1}(u)^{\frac{\log \rho_{j+1}}{\log \rho_{j+2}}} \qquad (w \in \Omega_j,\ j = d-2, \ldots, 1).$$

In particular, define

$$Z_0 = \sum_{u \in \Omega_1} Z_1(u)^{\frac{\log \rho_1}{\log \rho_2}}.$$

Using the technique by Kenyon & Peres [33] (which is an extension of McMullen [44]), we have

$$\dim_H K \le \frac{\log Z_0}{-n \log \rho_1}. \tag{101}$$

More precisely, define a probability vector $(p(u))_{u \in \Omega_d}$ by

$$p(u) = \frac{Z_d(u)}{Z_{d-1}(\theta_{d-1}(u))} \cdot \prod_{j=1}^{d-1} \frac{Z_j(\theta_j \theta_{j+1} \cdots \theta_{d-1}(u))^{\frac{\log \rho_j}{\log \rho_{j+1}}}}{Z_{j-1}(\theta_{j-1} \theta_j \cdots \theta_{d-1}(u))}$$

with convention $Z_0(\theta_0 \ldots \theta_{d-1}(u)) = Z_0$ for any $u \in \Omega_d$. Let ν be the product measure on $(\Omega_d)^{\mathbb{N}}$ by assigning probability $p(u)$ to each digit $u \in \Omega_d$. The measure ν can be viewed as a measure on Σ, which is σ^n-invariant and ergodic. Let $\mu = \nu \circ \pi^{-1}$. Then

$$\liminf_{r \to 0} \frac{\log \mu(B(\pi x, r))}{\log r} \le \frac{\log Z_0}{-n \log \rho_1} \qquad (x \in \Sigma). \tag{102}$$

A detailed proof of (102) was given by Shmerkin (see the proof of (4.3) in [60]) for the case $d = 2$, whilst a slight modification of the proof of [33, Theorem 1.2] provides a proof of (102) for $d \ge 2$. Then (101) follows from (102) and Billingsley's lemma.

Now we want to indicate certain connection between the upper bound $\frac{\log Z_0}{-n \log \rho_1}$ and the projection entropies. First we define the projections $\theta_j^* : \Omega_{j+1}^{\mathbb{N}} \to \Omega_j^{\mathbb{N}}$ $(j = d-1, \ldots, 1)$ by

$$\theta_j^*\left((u_k)_{k=1}^{\infty}\right) = (\theta_j(u_k))_{k=1}^{\infty} \qquad ((u_k)_{k=1}^{\infty} \in \Omega_{j+1}^{\mathbb{N}}).$$

Then it is easy to see that for each $1 \le j \le d-1$, the measure

$$\nu_j := \nu \circ \left(\theta_j^* \circ \theta_{j+1}^* \circ \cdots \circ \theta_{d-1}^*\right)^{-1}$$

is a product measure on $\Omega_j^{\mathbb{N}}$. Let T_j denote the left shift operator on $\Omega_j^{\mathbb{N}}$. By a direct calculation, we have

$$\frac{\log Z_0}{-n \log \rho_1} = \sum_{j=1}^{d} \left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) \frac{h(T_j, \nu_j)}{n}.$$

Thus we have

$$\dim_H K \le \sum_{j=1}^{d} \left(\frac{1}{\lambda_j} - \frac{1}{\lambda_{j+1}}\right) \frac{h(T_j, \nu_j)}{n}. \tag{103}$$

Let $\widetilde{\pi}_j$ $(j = 1, \ldots, d)$ denote the canonical projection from $\Omega_j^{\mathbb{N}}$ to $\mathbb{R}^j$ w.r.t. the IFS $\{S_u^{(j)}\}_{u \in \Omega_j}$ (remember that π_j denotes the canonical projection from Σ to $\mathbb{R}^d$

w.r.t. $\{S_u^{(j)}:\ u\in\Sigma_n\}$). According to Lemma 4.23(ii), we have

$$h_{\widetilde{\pi}_j}(T_j,\nu_j)=h_{\pi_j}(\sigma^n,\nu)\qquad (j=1,\ldots,d). \tag{104}$$

Since $\Phi_1\times\cdots\times\Phi_j$ ($j=1,\ldots,d$) satisfy the AWSC, there is a sequence (t_n) of positive integers with $\lim_n \log t_n/n=0$, such that

$$\sup_{x\in\mathbb{R}^j}\#\{S_u^{(j)}:\ u\in\Omega_j,\ x\in S_u^{(j)}(K_j)\}\leq t_n\qquad (j=1,\ldots,d), \tag{105}$$

where K_j denotes the attractor of $\Phi_1\times\cdots\times\Phi_j$. By Corollary 4.22, we have

$$h_{\widetilde{\pi}_j}(T_j,\nu_j)\geq h(T_j,\nu_j)-\log t_n\geq h(T_j,\nu_j)-\log t_n.$$

It together with (104) yields $h_{\pi_j}(\sigma^n,\nu)\geq h(T_j,\nu_j)-\log t_n$. Now applying Theorem 2.11 to the IFS $\{S_u: u\in\Sigma_n\}$, we have

$$\begin{aligned}
\dim_H \nu\circ\pi^{-1} &= \frac{1}{n}\sum_{j=1}^d\left(\frac{1}{\lambda_j}-\frac{1}{\lambda_{j+1}}\right)h_{\pi_j}(\sigma^n,\nu)\\
&\geq \frac{1}{n}\sum_{j=1}^d\left(\frac{1}{\lambda_j}-\frac{1}{\lambda_{j+1}}\right)(h(T_j,\nu_j)-\log t_n)\\
&\geq \dim_H K-\frac{\log t_n}{n}\cdot\sum_{j=1}^d\left(\frac{1}{\lambda_j}-\frac{1}{\lambda_{j+1}}\right)\qquad \text{(by (103))}.
\end{aligned}$$

Let $m=\frac{1}{n}\sum_{i=1}^n \nu\circ\sigma^{-i}$. Then m is ergodic and $\dim_H m\circ\pi^{-1}=\dim_H \nu\circ\pi^{-1}$. Letting n tend to ∞, we obtain

$$\sup\{\dim_H m\circ\pi^{-1}:\ m\in\mathcal{M}_\sigma(\Sigma),\ m \text{ is ergodic }\}\geq \dim_H K. \tag{106}$$

It is clear the "$\geq$" in above inequality can be replaced by "$=$" since $m\circ\pi^{-1}$ is supported on K. Note that $h_{\pi_j}(\sigma,\cdot)$ ($j=1,\ldots,d$) are upper semi-continuous on $\mathcal{M}_\sigma(\Sigma)$ (see Proposition 4.20 and (105)). By Theorem 2.2(ii) and Theorem 2.11, we see that the supremum in (106) is attained at some ergodic element in $\mathcal{M}_\sigma(\Sigma)$. This finishes the proof of the variational principle for $\dim_H K$.

Step 2. Show the variational principle for $\dim_B K$.

By Lemma 9.2, we only need to show that under the assumption of Theorem 2.15,

$$H_j\geq \widetilde{H}_j\qquad (j=1,\ldots,d), \tag{107}$$

where

$$H_j=\sup\{h_{\pi_j}(\sigma,m):\ m\in\mathcal{M}_\sigma(\Sigma)\},\quad \widetilde{H}_j=\lim_{n\to\infty}\frac{\log\#\{S_u^{(j)}:\ u\in\Sigma_n\}}{n}.$$

To see (107), by (105) and Lemma 9.1, we have

$$H_j\geq\frac{\log\#\left\{S_u^{(j)}:\ u\in\Sigma_n\right\}-\log t_n}{n}\qquad (n\in\mathbb{N}).$$

Letting $n \to \infty$, we obtain (107) by the assumption $\log t_n/n \to 0$. This finishes the proof of the theorem. □

Remark 9.3. With an essentially identical proof, Theorem 2.15 can be extended to the following class of IFS $\Phi = \Phi_1 \times \cdots \times \Phi_k$ on $\mathbb{R}^{q_1} \times \cdots \times \mathbb{R}^{q_k}$, where Φ_j has the form $\{A_j z_j + c_{i,j}\}_{i=1}^{\ell}$ such that A_j is the inverse of an integral matrix and all the eigenvalues of A_j equals ρ_j in modulus, $\rho_1 > \cdots > \rho_k$, $c_{i,j} \in \mathbb{Q}^{q_j}$.

This together with Lemma 7.2 and the proof of Theorem 2.12 yields

Theorem 9.4. *Let $\Phi = \{S_i\}_{i=1}^{\ell}$ be an IFS on $\mathbb{R}^d$ of the form*

$$S_i(x) = Ax + c_i \qquad (i = 1, \ldots, \ell),$$

where A is the inverse of an integral expanding $d \times d$ matrix, $c_i \in \mathbb{Z}^d$. Let K be the attractor of the IFS. Then there is an ergodic measure on K of full Hausdorff dimension.

10. A FINAL REMARK ABOUT INFINITE NON-CONTRACTIVE IFS

In the previous sections, we have made the restriction that an IFS consists of finitely many contractive maps. We remark that part of our results can be extended to certain infinite non-contractive IFS.

Let $\Phi = \{S_i\}_{i=1}^{\infty}$ be a family of maps on $\mathbb{R}^d$ of the form

$$S_i(x) = \rho_i R_i(x) + a_i \qquad (i = 1, 2, \ldots),$$

where $\rho_i > 0$, R_i are orthogonal $d \times d$ matrices, $a_i \in \mathbb{R}^d$.

Let (X, σ) be the left shift over the alphabet $\{i : \ i \in \mathbb{N}\}$, and let m be an ergodic measure on X satisfying $H_m(\mathcal{P}_\infty) < \infty$, where $\mathcal{P}_\infty$ denotes the partition of X given by

$$\mathcal{P}_\infty = \{[i] : \ i \in \mathbb{N}\},$$

where $[i] = \{(x_i)_{i=1}^{\infty} \in X : \ x_1 = i\}$. Assume that Φ is *m-contractive* in the sense that

$$\sum_{i=1}^{\infty} (\log \rho_i) m([i]) < 0, \qquad \sum_{i=1}^{\infty} (\log |a_i|) m([i]) < \infty. \tag{108}$$

Denote

$$\lambda = -\sum_{i=1}^{\infty} (\log \rho_i) m([i]).$$

Let X' denote the set of points $x = (x_i)_{i=1}^{\infty} \in X$ such that

$$\lim_{n\to\infty} (1/n) \log(\rho_{x_1}\rho_{x_2} \ldots \rho_{x_n}) = -\lambda, \quad \lim_{n\to\infty} (1/n) \log |a_{x_n}| = 0.$$

Then X' satisfies $\sigma^{-1}(X') = X'$. Furthermore by Birkhoff's ergodic theorem,

$$m(X') = 1.$$

Define the projection map $\pi : \ X' \to \mathbb{R}^d$ by

$$\pi(x) = \lim_{n\to\infty} S_{x_1} \circ S_{x_2} \circ \cdots \circ S_{x_n}(0) \qquad (x \in X').$$

It is easily checked that π is well defined. Let $\mu = m \circ \pi^{-1}$ be the projection of m under π. We have the following theorem

Theorem 10.1. *Under the above setting, $\mu = m \circ \pi^{-1}$ is exactly dimensional and*

$$\dim_H \mu = \frac{h_\pi(\sigma, m)}{\lambda},$$

where $H_\pi(\sigma, m) = H_m(\mathcal{P}_\infty | \sigma^{-1}\pi^{-1}\gamma) - H_m(\mathcal{P}_\infty | \pi^{-1}\gamma)$, $\gamma = \mathcal{B}(\mathbb{R}^d)$.

We remark that when m is a Bernoulli product measure, $\mu = m \circ \pi^{-1}$ is the stationary measure of certain affine random walk determined by Φ and m, and the decay property of μ at infinity has been extensively studied in the literature (cf. [24] and references therein).

The proof of Theorem 10.1 is essentially identical to that given in Section 6. Indeed we only need to replace Σ in Section 6 by X', and replace 'let $c > 1$ so that $c \sup_{x\in\Sigma} \overline{\rho}(x) < 1$' in the proof of Theorem 6.2 by 'let $1 < c < e^{\lambda}$'.

Acknowledgement.The authors are grateful to François Ledrappier for his encouragement and helpful comments. They are indebted to Eric Olivier for stimulating discussions of the variational principle about the Hausdorff dimension of self-affine sets, and to Wen Huang for the discussions of the entropy theory. They also thank Quansheng Liu, Emile Le Page and Yong-Luo Cao for valuable comments, and Guo-Hua Zhang for critical reading of the manuscript. The first author was partially supported by the RGC grant in CUHK, Fok Ying Tong Education Foundation and NSFC (Grant 10571100). The second author was partially supported by NSF under grants DMS-0240097 and DMS-0503870.

REFERENCES

[1] Barański, K. Hausdorff dimension of the limit sets of some planar geometric constructions. *Adv. Math.* **210** (2007), 215–245.

[2] Barnsley, M. *Fractals everywhere*. Academic Press, Inc., Boston, MA, 1988.

[3] Barral, J. and Mensi, M. Gibbs measures on self-affine Sierpiński carpets and their singularity spectrum. *Ergodic Theory Dynam. Systems* **27** (2007), 1419–1443.

[4] Barreira, L., Pesin, Ya., and Schmeling, J. Dimension and product structure of hyperbolic measures. *Ann. of Math.* **149** (1999), 755–783.

[5] Bedford, T. *Crinkly curves, Markov partitions and box dimension in self-similar sets*, Ph.D. Thesis, University of Warwick, 1984.

[6] Bedford, T. Applications of dynamical systems theory to fractals—a study of cookie-cutter Cantor sets. *Fractal geometry and analysis* (Montreal, PQ, 1989), 1–44, Kluwer Acad. Publ., Dordrecht, 1991.

[7] Billingsley, P. *Probability and measure*. Third edition. John Wiley & Sons, Inc., New York, 1995.

[8] Bogenschütz, T. Entropy, pressure, and a variational principle for random dynamical systems, *Random Comput. Dynam.* **1** (1992/93), 99–116.

[9] Bowen, R. Equilibrium states and the ergodic theory of Anosov diffeomorphisms, *Lecture notes in Math.,* No. **470**, Springer-Verlag, 1975.

[10] Eckmann, J. P. and Ruelle, D. Ergodic theory of chaos and strange attractors. *Rev. Modern Phys.* **57** (1985), 617–656.

[11] Falconer, K. J. The Hausdorff dimension of self-affine fractals. *Math. Proc. Cambridge Philos. Soc.* **103** (1988), 339–350

[12] Falconer, K. J. Dimensions and measures of quasi self-similar sets. *Proc. Amer. Math. Soc.* **106** (1989), 543–554.
[13] Falconer, K. J. *Fractal geometry, mathematical foundations and applications*. Wiley, 1990.
[14] Fan, A. H. Sur les dimensions de mesures. *Studia Math.* **111** (1994), 1–17.
[15] Fan, A. H., Lau, K. S., and Rao, H. Relationships between different dimensions of a measure. *Monatsh. Math.* **135** (2002), 191–201.
[16] Feng, D. J. The smoothness of L^q-spectrum of self-similar measures with overlaps. *J. Lond. Math. Soc.* **68** (2003), 102–118.
[17] Feng, D. J. The limited Rademacher functions and Bernoulli convolutions associated with Pisot numbers. *Adv. Math.* **195** (2005), 24–101.
[18] Feng, D. J. Gibbs properties of self-conformal measures and the multifractal formalism. *Ergodic Theory Dynam. Systems* **27** (2007), 787–812.
[19] Feng, D. J. and Wang, Y. A class of self-affine sets and self-affine measures. *J. Fourier Anal. Appl.* **11** (2005), 107–124.
[20] Gatzouras, D. and Lalley, S. P. Hausdorff and box dimensions of certain self-affine fractals. *Indiana Univ. Math. J.* **41** (1992), 533–568.
[21] Gatzouras, D. and Peres, Y. Invariant measures of full dimension for some expanding maps. *Ergodic Theory Dynam. Systems* **17** (1997), 147–167.
[22] Gatzouras, D. and Peres, Y. The variational principle for Hausdorff dimension: a survey. *Ergodic theory of $\mathbb{Z}^d$ actions* (Warwick, 1993–1994), London Math. Soc. Lecture Note Ser. **228**, 113–125, Cambridge Univ. Press, Cambridge, 1996.
[23] Geronimo, J. S. and Hardin, D. P. An exact formula for the measure dimensions associated with a class of piecewise linear maps. Fractal approximation. *Constr. Approx.* **5** (1989), 89–98.
[24] Guivarc'h, Y. and Le Page, E. On spectral properties of a family of transfer operators and convergence to stable laws for affine random walks. *Ergodic Theory Dynam. Systems* **28** (2008), 423–446.
[25] Hu, H. Dimensions of invariant sets of expanding maps. *Comm. Math. Phys.* **176** (1996), 307–320.
[26] Huang, W., Ye, X. D., and Zhang, G. H. A local variational principle for conditional entropy. *Ergodic Theory Dynam. Systems* **26** (2006), 219–245.
[27] Hueter I. and Lalley, S. P. Falconer's formula for the Hausdorff dimension of a self-affine set in $\mathbb{R}^2$. *Ergodic Theory Dynam. Systems* **15** (1995), 77–97.
[28] Hutchinson, J. E. Fractals and self-similarity. *Indiana Univ. Math. J.* **30** (1981), 713–747.
[29] Jordan, T., Pollicott M., and Simon, K. Hausdorff dimension for randomly perturbed self affine attractors. *Comm. Math. Phys.* **270** (2007), 519–544.
[30] Käenmäki, A. On natural invariant measures on generalised iterated function systems. *Ann. Acad. Sci. Fenn. Math.* **29** (2004), 419–458.
[31] Käenmäki, A. and Shmerkin, P. Overlapping self-affine sets of Kakeya type. *Ergodic Theory Dynam. Systems*. At press.
[32] Keller, G. *Equilibrium states in ergodic theory*. Cambridge University Press, Cambridge, 1998.
[33] Kenyon R. and Peres, Y. Measures of full dimension on affine-invariant sets. *Ergodic Theory Dynam. Systems* **16** (1996), 307–323.
[34] Kenyon R. and Peres, Y. Hausdorff dimensions of sofic affine-invariant sets. *Israel J. Math.* **94** (1996), 157–178.
[35] Lalley, S. P. β-expansions with deleted digits for Pisot numbers β. *Trans. Amer. Math. Soc.* **349** (1997), 4355–4365.
[36] Lalley, S. P. Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution. *J. London Math. Soc.* **57** (1998), 629–654.
[37] Lau, K. S. and Ngai, S. M. Multifractal measure and a weak separation condition. *Adv. Math.* **141** (1999), 45–96.
[38] Ledrappier, F. "On the dimension of some graphs" in *Symbolic dynamics and its applications* (New Haven, CT, 1991), Contemp. Math. **135**, Amer. Math. Soc., Providence, RI, 1992, 285–293.

[39] Ledrappier, F. and Porzio, A. A dimension formula for Bernoulli convolutions. *J. Stat. Phys.* **76** (1994), 1307–1327.

[40] Ledrappier, F. and Young, L.-S. The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula. II. Relations between entropy, exponents and dimension. *Ann. of Math.* **122** (1985), 509–539; 540–574.

[41] Luzia, N. A variational principle for the dimension for a class of non-conformal repellers. *Ergodic Theory Dynam. Systems* **26** (2006), 821–845.

[42] Mañé, R. *Ergodic theory and differentiable dynamics*. Springer-Verlag, Berlin, 1987.

[43] Mattila, P. *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press, 1995.

[44] McMullen, C. The Hausdorff dimension of general Sierpiński carpets. *Nagoya Math. J.* **96** (1984), 1–9.

[45] Ngai, S. M. and Wang, Y. Hausdorff dimension of self-similar sets with overlaps. *J. London Math. Soc.* **63** (2001), 655–672.

[46] Olivier, E. Variational principle for dimension and the uniqueness of the measure with full dimension on (mod 1) Sierpiński carpets. Preprint.

[47] Patzschke, N. Self-conformal multifractal measures. *Adv. in Appl. Math.* **19** (1997), 486–513.

[48] Parry, W. *Topics in ergodic theory*, Cambridge University Press, 1981.

[49] Parry, W. *Entropy and generators in ergodic theory*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[50] Peres, Y. and Solomyak, B. Existence of L^q dimensions and entropy dimension for self-conformal measures. *Indiana Univ. Math. J.* **49** (2000), 1603–1621.

[51] Pesin, Ya. *Dimension theory in dynamical systems: contemporary views and applications*. The University of Chicago Press, 1997.

[52] Przytycki, F. and Urbański, M. On the Hausdorff dimension of some fractal sets. *Studia Math.* **93** (1989), 155–186.

[53] Qian, M. and Xie, J.-S. Entropy formula for endomorphisms: relations between entropy, exponents and dimension. *Discrete Contin. Dyn. Syst.* **21** (2008), 367–392.

[54] Rao, H. and Wen, Z. Y. A class of self-similar fractals with overlap structure. *Adv. in Appl. Math.* **20** (1998), 50–72.

[55] Rohlin, V. A. On the fundamental ideas of measure theory. *Mat. Sbornik N.S.* 25 (67) (1949), 107–150. (see also *Amer. Math. Soc. Translation 1952*, (1952). no. 71.)

[56] Rudin, W. *Real and complex analysis*. Third edition. McGraw-Hill Book Co., New York, 1987.

[57] Salem, R. *Algebraic numbers and Fourier analysis*. D. C. Heath and Co., Boston, 1963.

[58] Schmeling, J. A dimension formula for endomorphisms—the Belykh family. *Ergodic Theory Dynam. Systems* **18** (1998), 1283–1309.

[59] Schmeling, J. and Troubetzkoy, S. Dimension and invertibility of hyperbolic endomorphisms with singularities. *Ergodic Theory Dynam. Systems* **18** (1998), 1257–1282.

[60] Shmerkin, P. Overlapping self-affine sets. *Indiana Univ. Math. J.* **55** (2006), 1291–1331.

[61] Solomyak, B. Measure and dimension for some fractal families. *Math. Proc. Cambridge Philos. Soc.* **124** (1998), 531–546.

[62] Stoll, R. R. *Linear algebra and matrix theory*. McGraw-Hill Company, Inc., New York–Toronto–London, 1952.

[63] Walters, P. *An introduction to ergodic theory*. Springer-Verlag, 1982.

[64] Yamamoto, T. On the extreme values of the roots of matrices. *J. Math. Soc. Japan* **19** (1967), 173–178.

[65] Young, L.-S. Dimension, entropy and Lyapunov exponents. *Ergodic Theory Dynam. Systems* **2** (1982), 109–124.