Consider a dynamical system with asymptotic entropy expansiveness and approximate product property. We show that it has ergodic measures of arbitrary intermediate entropies. A similar conclusion actually holds in every neighborhood of an invariant measure. Moreover, for such systems the ergodic measures of zero entropy forms a residual set in the space of invariant measures.

1. Introduction

The research on systems with certain hyperbolicity is a mainstream direction in the field of dynamical systems, in which various orbit-tracing properties play important roles. In 1971, Bowen introduced the notion of specification in his seminal paper [4] to study periodic points and measures for Axiom A diffeomorphisms. Since then, a number of variations of the specification property have been considered to study broader classes of dynamical systems, which exhibit some weaker hyperbolicity. Based on these specification-like properties, a bunch of results were successfully achieved, including the ones on the structure of the space of invariant measures, growth rate of periodic orbits, large deviations, multifractal analysis, etc. For an overview of the definitions and results of specification-like properties, the readers are referred to the book [9] and the survey [14].

Among the those specification-like properties discussed in [14], approximate product property is the weakest one. It is also weaker than the so-called gluing orbit property discussed in [3] and [29]. However, it is still strong enough for us to derive interesting results. In [22] Pfister and Sullivan introduced this notion. They showed that approximate product property implies entropy denseness and used this fact to prove large deviations for β-shifts.

Entropy denseness was first discussed and proved for \( \mathbb{Z}^d \) shifts with specification property by Eizenberg, Kifer and Weiss [10]. We perceive that this property is closely related to a conjecture of Katok. It is a historic question whether positive topological entropy implies a rich structure of the space of invariant measures. The answer is negative and there are many \( C^0 \) examples of positive topological entropy.
that are uniquely ergodic (e.g. [12] and [1]). However, it seems that we may expect a positive answer for smooth systems as in this case positive topological entropy implies some sort of hyperbolicity. In his seminal paper [13] Katok showed that every $C^{1+\alpha}$ diffeomorphism in dimension 2 has horseshoes of large entropies. This implies that the system has ergodic measures of arbitrary intermediate metric entropy. Katok believed that this holds in any dimension.

**Conjecture** (Katok). For every $C^2$ diffeomorphism $f$ on a Riemannian manifold $X$, the set

$$H(X, f) := \{ h_\mu(f) : \mu \text{ is an ergodic measure for } (X, f) \}$$

includes $[0, h(f))$.

Early progresses on Katok’s conjecture were made by the author [26, 27, 28]. Then the remarkable result [24] Quas and Soo showed that a system is universal, which is stronger than the conclusion of Katok’s conjecture, if $(X, f)$ satisfies asymptotic entropy expansiveness, almost weak specification and small boundary property. In [11] the author and collaborators showed that certain homogeneous systems satisfy almost weak specification hence proved Katok’s conjecture by applying the result of Quas and Soo. Recently, Li, Shi, Wang and Wang [16] proved a flow version of Katok’s conjecture for star flows and Burguet [6] extended the result of Quas and Soo to request only the almost weak specification property.

In this paper, we prove Katok’s conjecture for systems satisfying asymptotic entropy expansiveness and approximate product property. We actually proved a strong result on the structure of the space of invariant measures, which extends a theorem of Sigmund [25], in a setting where periodic points may be absent. Our result might indicate that we can ask for something more than Katok conjectured, at least for a large class of systems. The main part of the proof is a combination of the argument of Pfister and Sullivan [22] to show entropy denseness and the argument developed in [30] by the author to estimate the entropy from above.

For a system $(X, f)$, denote by $\mathcal{M}(X, f)$ the space of invariant measures and by $\mathcal{M}_e(X, f)$ the subset of ergodic ones. The main results of the paper state as follows:

**Theorem 1.1.** Let $(X, f)$ be an asymptotically entropy expansive system with approximate product property. Then for every $\alpha \in [0, h(f))$,

$$\mathcal{M}_e(X, f, \alpha) := \{ \mu \in \mathcal{M}_e(X, f) : h_\mu(f) = \alpha \}$$

is a residual subset in the compact metric subspace

$$\mathcal{M}^\alpha = \mathcal{M}^\alpha(X, f) := \{ \mu \in \mathcal{M}(X, f) : h_\mu(f) \geq \alpha \}.$$  

In particular, $\mathcal{M}_e(X, f, 0)$, the set of ergodic measures of zero entropy, is a residual subset of $\mathcal{M}(X, f)$.

**Corollary 1.2.** Let $(X, f)$ be an asymptotically entropy expansive system with approximate product property. For every $\mu \in \mathcal{M}(X, f)$, every open neighborhood $U$ of $\mu$, denote

$$H(X, f, U) := \{ h_\nu(f) : \nu \in U \cap \mathcal{M}_e(X, f) \}.$$  

Then we have

$$\begin{cases} 
H(X, f, U) \supset [0, h_\mu(f)], & \text{if } h_\mu(f) < h(f); \\
H(X, f, U) \supset [0, h_\mu(f)], & \text{if } h_\mu(f) = h(f).
\end{cases}$$
In particular, \( H(X, f) = [0, h(f)] \).

**Remark.** We remark that it is possible that \( h_{\mu}(f) \notin H(X, f, U) \) when \( h_{\mu}(f) = h(f) \). See Example 8.1. Another remark is that only asymptotic entropy expansiveness and entropy denseness, even along with the fact that \( \mathcal{M}(X, f) \) is a Poulsen simplex, is not sufficient for the conclusion of Theorem 1.1 and Corollary 1.2. See Example 8.2.

Corollary 1.2 covers most known results on Katok’s conjecture, including [28], [24, Section 3.2], [11] and [30, Theorem 1.3]. However, in [24] and [6], a seemingly much stronger conclusion, universality, is proved under the assumption of almost weak (tempered) specification. We remark that in general universality should not be expected for systems with only approximate product property. See example 8.3.

In the proof of Theorem 1.1, asymptotic entropy expansiveness is used in two places: to bound the local entropy \( h_{\text{loc}}(f, \varepsilon) \) and to provide upper semi-continuity of the entropy map. We doubt if this condition is necessary and suspect that only approximate product property might be enough (without asking for measures of maximal entropy).

As a simple corollary, we can show that a system with approximate product property must have zero topological entropy if it is either minimal or uniquely ergodic. In [30] we have shown that for systems with gluing orbit property, zero topological entropy implies minimality and equicontinuity. We do not know if this still holds for systems with approximate product property.

Notions and results in this paper naturally extends to the continuous-time case. The proof can be carried out with a little extra effort, namely a discretization argument as in the proof of [7, Lemma 5.10].

2. Preliminaries

Let \((X, d)\) be a compact metric space. Let \( f : X \to X \) be a continuous map. Then \((X, f)\) is conventionally called a topological dynamical system or just a system.

We shall denote by \( \mathbb{Z}^+ \) the set of all positive integers and by \( \mathbb{N} \) the set of all nonnegative integers, i.e. \( \mathbb{N} = \mathbb{Z}^+ \cup \{0\} \). For \( n \in \mathbb{Z}^+ \), denote

\[
\mathbb{Z}_n := \{0, 1, \cdots, n - 1\} \quad \text{and} \quad \Sigma_n := \{0, 1, \cdots, n - 1\}^{\mathbb{Z}^+}.
\]

2.1. Specification-like properties.

**Definition 2.1.** Let \( \mathcal{C} = \{x_k\}_{k \in \mathbb{Z}^+} \) be a sequence in \( X \). Let \( \mathcal{S} = \{m_k\}_{k \in \mathbb{Z}^+} \) and \( \mathcal{G} = \{t_k\}_{k \in \mathbb{Z}^+} \) be sequences of positive integers. The pair \((\mathcal{C}, \mathcal{S})\) shall be called an orbit sequence while \( \mathcal{G} \) shall be called a gap. For \( \varepsilon > 0 \) and \( z \in X \), we say that \((\mathcal{C}, \mathcal{S}, \mathcal{G})\) is \( \varepsilon \)-traced by \( z \) if for each \( k \in \mathbb{Z}^+ \),

\[
d(f^{s_k + j}(z), f^k(x_k)) \leq \varepsilon \quad \text{for each } j = 0, 1, \cdots, m_j - 1,
\]

where

\[
s_1 = s_1(\mathcal{S}, \mathcal{G}) := 0 \quad \text{and} \quad s_k = s_k(\mathcal{S}, \mathcal{G}) := \sum_{i=1}^{k-1} (m_i + t_i - 1) \quad \text{for } k \geq 2.
\]

**Definition 2.2.** \((X, f)\) is said to have specification property if for every \( \varepsilon > 0 \) there is \( M = M(\varepsilon) > 0 \) such that for any orbit sequence \((\mathcal{C}, \mathcal{S})\) and any gap \( \mathcal{G} \) satisfying \( \min \mathcal{G} \geq M \), there is \( z \in X \) that \( \varepsilon \)-traces \((\mathcal{C}, \mathcal{S}, \mathcal{G})\).
Definition 2.3. \((X, f)\) is said to have gluing orbit property if for every \(\varepsilon > 0\) there is \(M = M(\varepsilon) > 0\) such that for any orbit sequence \((\mathcal{C}, \mathcal{S})\), there is a gap \(\mathcal{G}\) satisfying \(\max \mathcal{G} \leq M\) and \(z \in X\) such that \((\mathcal{C}, \mathcal{S}, \mathcal{G})\) can be \(\varepsilon\)-traced by \(z\).

Let \(\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty}\) be two sequences of integers. We write
\[
\{a_n\}_{n=1}^{\infty} \leq \{b_n\}_{n=1}^{\infty} \text{ if } a_n \leq b_n \text{ for each } n \in \mathbb{Z}^+.
\]

For a sequence \(\mathcal{S} = \{a_n\}_{n=1}^{\infty}\) of positive integers and a function \(L : \mathbb{Z}^+ \to \mathbb{Z}^+\), we write
\[
L(\mathcal{S}) := \{L(a_n)\}_{n=1}^{\infty}.
\]

We say that the function \(L : \mathbb{Z}^+ \to \mathbb{Z}^+\) is tempered if \(L\) is nondecreasing and
\[
\lim_{n \to \infty} \frac{L(n)}{n} = 0.
\]

Denote by \(\sigma\) the shift operator on sequences, i.e.
\[
\sigma(\{a_n\}_{n=1}^{\infty}) = \{a_{n+1}\}_{n=1}^{\infty}.
\]

Definition 2.4. \((X, f)\) is said to have almost weak specification property (as in [24]), or tempered specification property (the name we suggest to avoid ambiguity) if for every \(\varepsilon > 0\) there is a tempered function \(L_{\varepsilon} : \mathbb{Z}^+ \to \mathbb{Z}^+\) such that for any orbit sequence \((\mathcal{C}, \mathcal{S})\) and any gap \(\mathcal{G}\) satisfying \(\mathcal{G} \geq L_{\varepsilon}(\sigma(\mathcal{S}))\), there is \(z \in X\) that \(\varepsilon\)-traces \((\mathcal{C}, \mathcal{S}, \mathcal{G})\).

Definition 2.5. \((X, f)\) is said to have tempered gluing orbit property if for every \(\varepsilon > 0\) there is a tempered function \(L_{\varepsilon} : \mathbb{Z}^+ \to \mathbb{Z}^+\) such that for any orbit sequence \((\mathcal{C}, \mathcal{S})\), there is a gap \(\mathcal{G}\) satisfying \(\mathcal{G} \leq L_{\varepsilon}(\sigma(\mathcal{S}))\) and \(z \in X\) such that \((\mathcal{C}, \mathcal{S}, \mathcal{G})\) can be \(\varepsilon\)-traced by \(z\).

Remark. Definition 2.2–2.5 are equivalent to their analogs respectively, where the tracing property holds for any finite orbit sequences. A proof of the equivalence for gluing orbit property can be found in [29, Lemma 2.10]. The proof for tempered cases is analogous. The properties are called periodic if for any finite orbit sequence we require the tracing point \(z\) to be a periodic point with the specified period (cf. [14] and [31]).

The notion of tempered specification was first introduced in [20] without a name. It was called almost weak specification in some references such as [24] and suggested to be called weak specification in [14]. The author suggests the name tempered specification to avoid any possible ambiguity with other specification-like properties. Hence its variation should be called tempered gluing orbit.

Readers are referred to [14] for a survey of specification-like properties. The references [2], [3] and [29] contain the original definition of gluing orbit property as well as a bunch of interesting examples.

It is clear that specification implies tempered specification, gluing orbit implies tempered gluing orbit and (tempered) specification implies (tempered) gluing orbit. There are systems with gluing orbit properties that does not have tempered specification (e.g. irrational rotation), and vice versa (e.g. quasi-hyperbolic toral automorphisms that are central-skew [17, 18, 20]).
2.2. Topological entropy and entropy expansiveness.

**Definition 2.6.** Let $K$ be a subset of $X$. For $n \in \mathbb{Z}^+$ and $\varepsilon > 0$, a subset $E \subset K$ is called an $(n, \varepsilon)$-separated set in $K$ if for any distinct points $x, y$ in $E$, there is $k \in \{0, \cdots, n-1\}$ such that
\[ d(f^k(x), f^k(y)) > \varepsilon. \]
Denote by $s(K, n, \varepsilon)$ the maximal cardinality of $(n, \varepsilon)$-separated subsets of $K$. Let
\[ h(K, f, \varepsilon) := \limsup_{n \to \infty} \frac{\ln s(K, n, \varepsilon)}{n}. \]
Then the topological entropy of $f$ on $K$ is defined as
\[ h(K, f) := \lim_{\varepsilon \to 0} h(K, f, \varepsilon). \]
In particular, $h(f) := h(X, f)$ is the topological entropy of the system $(X, f)$.

**Remark.** Note that $h(K, f, \varepsilon)$ grows as $\varepsilon$ tends to 0. So we actually have
\[ h(K, f) = \sup\{h(K, f, \varepsilon) : \varepsilon > 0\}. \]

**Proposition 2.7.** A set of the form
\[ B_n(x, \varepsilon) = \{ y \in X : d(f^k(y), f^k(x)) < \varepsilon, k = 0, 1, \cdots, n-1 \} \]
is called an $(n, \varepsilon)$-ball of $(X, f)$. A subset $E$ of $X$ is called an $(n, \varepsilon)$-spanning set if
\[ X = \bigcup_{x \in E} B_n(x, \varepsilon). \]
Denote by $r(n, \varepsilon)$ the minimal cardinality of $(n, \varepsilon)$-spanning subset of $X$. In particular, denote $r(\varepsilon) := r(1, \varepsilon)$. Let
\[ h^*(f, \varepsilon) := \limsup_{n \to \infty} \frac{\ln r(n, \varepsilon)}{n}. \]
Then
\[ h(f) = \lim_{\varepsilon \to 0} h^*(f, \varepsilon) = \sup\{h^*(f, \varepsilon) : \varepsilon > 0\}. \]

**Definition 2.8.** For $\varepsilon > 0$ and $x \in X$, denote
\[ \Gamma_\varepsilon(x) := \{ y \in X : d(f^n(x), f^n(y)) < \varepsilon \text{ for every } n \in \mathbb{N} \}. \]
Let
\[ h_{\text{loc}}(f, \varepsilon) := \sup\{h(\Gamma_\varepsilon(x), f) : x \in X\}. \]
(1) We say that $(X, f)$ is entropy expansive if there is $\varepsilon_0 > 0$ such that
\[ h_{\text{loc}}(f, \varepsilon_0) = 0. \]
(2) We say that $(X, f)$ is asymptotically entropy expansive if
\[ \lim_{\varepsilon \to 0} h_{\text{loc}}(f, \varepsilon) = 0. \]

**Proposition 2.9** ([5, Corollary 2.4]). For every subset $K$ and every $\varepsilon > 0$,
\[ h(K, f) \leq h(K, f, \varepsilon) + h_{\text{loc}}(f, \varepsilon). \]

**Corollary 2.10.** If $(X, f)$ is asymptotically entropy expansive, then $h(f) < \infty$.

**Proposition 2.11.** Let $(X, f)$ be asymptotically entropy expansive. Then the followings hold:
(1) $h(f) < \infty$. 

---

**APPENDIX**

**Remark.**
(2) If \( \Lambda \) is a compact \( f \)-invariant subset of \( X \), then \( (\Lambda, f|_{\Lambda}) \) is also asymptotically entropy expansive.

2.3. Measures and metric entropy. Denote by \( \mathcal{M}(X) \) the space of probability measures on \( X \), by \( \mathcal{M}(X, f) \) the subspace of invariant measures of \( (X, f) \) and by \( \mathcal{M}_e(X, f) \) the subset of the ergodic ones. As \( X \) is compact, \( \mathcal{M}(X) \) is a compact metrizable space under the weak-* topology. We shall denote by \( D^* \) the diameter of \( \mathcal{M}(X) \).

**Lemma 2.12** (cf. [32, Theorem 6.4]). There is a metric \( D \) on \( \mathcal{M}(X) \) such that \( D \) induces the weak-* topology on \( \mathcal{M}(X) \) and

\[
D\left(\sum_{k=1}^{n} a_k \mu_k, \sum_{k=1}^{n} a_k \nu_k\right) \leq \sum_{k=1}^{n} a_k D(\mu_k, \nu_k)
\]

for any \( \mu_1, \ldots, \mu_n, \nu_1, \ldots, \nu_n \in \mathcal{M}(X) \) and any \( a_1, \ldots, a_n > 0 \) satisfying \( \sum_{k=1}^{n} a_k = 1 \).

By [32], \( \mathcal{M}_e(X, f) \) is exactly the set of extreme points of \( \mathcal{M}(X, f) \) and \( \mathcal{M}(X, f) \) is a Choquet simplex, i.e. every \( \mu \in \mathcal{M}(X, f) \) is the barycenter of a unique probability measure supported on the set of its extreme points (i.e. \( \mathcal{M}_e(X, f) \)). As a corollary, \( \mathcal{M}_e(X, f) \) is a \( G_{\delta} \) subset of \( \mathcal{M}(X, f) \). See [23] for more details on Choquet simplices.

**Definition 2.13.** Let \( \mu \) be an invariant probability measure for \( (X, f) \). Fix \( \delta \in (0, 1) \). Denote

\[
r_\mu(n, \varepsilon, \delta) := \min\{|U| : U \text{ is a collection of } (n, \varepsilon) \text{-balls such that } \mu(\bigcup_{U \in \mathcal{U}} U) > 1 - \delta\}.
\]

Then the metric entropy of \( (X, f) \) with respect to \( \mu \) is defined as

\[
h_\mu(f) := \lim_{\varepsilon \to 0} \limsup_{n \to \infty} \frac{\ln r_\mu(n, \varepsilon, \delta)}{n}.
\]

**Proposition 2.14** (cf. [32, Theorem 8.1]). For any \( \mu, \nu \in \mathcal{M}(X, f) \) and \( \lambda \in [0, 1] \),

\[
h_{\lambda \mu + (1 - \lambda) \nu}(f) = \lambda h_\mu(f) + (1 - \lambda) h_\nu(f).
\]

**Proposition 2.15** (Variational Principle). For any system \( (X, f) \), we have

\[
h(f) = \sup\{h_\mu(f) : \mu \in \mathcal{M}(X, f)\}.
\]

**Proposition 2.16** ([21, Corollary 4.1]). If \( (X, f) \) is asymptotically entropy expansive, then the map \( \mu \mapsto h_\mu(f) \) is upper semi-continuous with respect to the weak-* topology on \( \mathcal{M}(X, f) \). As a corollary, there is a \( \mu_M \in \mathcal{M}_e(X, f) \) such that \( h_{\mu_M}(f) = h(f) \).

3. Approximate Product Property

**Definition 3.1.** Let \( \mathcal{G} = \{x_k\}_{k \in \mathbb{Z}^+} \) be a sequence in \( X \) and \( \mathcal{G} = \{t_k\}_{k \in \mathbb{Z}^+} \) be an increasing sequence of nonnegative integers. For \( n \in \mathbb{Z}^+ \), \( \delta_1, \delta_2, \varepsilon > 0 \) and \( z \in X \), we say that \( \mathcal{G} \) is \( (n, \delta_1, \delta_2, \varepsilon) \)-traced by \( z \) if \( \mathcal{G} \) is \( (n, \delta_1) \)-spaced, i.e.

\[
t_1 = 0 \text{ and } n \leq t_{k+1} - t_k < n(1 + \delta_1) \text{ for each } k \in \mathbb{Z}^+,
\]

and the following tracing property holds:

\[
|\{j \in \mathbb{Z}_n : d(f^{t_k+j}(z), f^j(x_k)) > \varepsilon\}| < \delta_2n \text{ for each } k \in \mathbb{Z}^+.
\]
Definition 3.2. The system $(X, f)$ is said to have approximate product property, if for every $\delta_1, \delta_2, \epsilon > 0$, there is $M = M(\delta_1, \delta_2, \epsilon) > 0$ such that for every $n > M$ and every sequence $\mathcal{C}$ in $X$, there are an $(n, \delta_1)$-spaced sequence $\mathcal{G}$ and $z \in X$ such that $\mathcal{C}$ is $(n, \delta_1, \mathcal{G}, \epsilon)$-traced by $z$.

Approximate product property is almost the weakest specification-like property. It is weaker than all other specification-like properties discussed in [14], as well as gluing orbit property and tempered gluing orbit property, and seems independent with the decomposition introduced by Climenhaga and Thompson [8].

Lemma 3.3. Suppose that $(X, f)$ has tempered gluing orbit property. Then $(X, f)$ has approximate product property.

Proof. Suppose that we are given $\delta_1, \delta_2, \epsilon > 0$ and $(X, f)$ has tempered gluing orbit property. There is a tempered function $L_\epsilon : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ such that for any orbit sequence $(\mathcal{C}, S)$, there is a gap $\mathcal{G}$ satisfying $\mathcal{G} \leq L_\epsilon(\sigma(S))$ and $z \in X$ such that $(\mathcal{C}, S, \mathcal{G})$ can be $\epsilon$-traced by $z$. Then there is $M$ such that

$$\frac{L_\epsilon(n)}{n} < \delta_1$$

for every $n > M$.

For every $n > M$ and every sequence $\mathcal{C} = \{x_k\}_{k \in \mathbb{Z}^+}$ in $X$, assume that $(\mathcal{C}, \{n\}_{n \in \mathbb{Z}^+}, \{t_k\}_{k=1}^\infty)$ is $\epsilon$-traced by $z$ and $t_k \leq L_\epsilon(n)$ for each $k$. Denote

$$s_k := \sum_{j=1}^{k-1}(n + t_k - 1).$$

Then

$$s_1 = 0, n \leq s_{k+1} - s_k = t_k - 1 < L_\epsilon(n) < \delta_1 n$$

for each $k$ and

$$\left|\{j \in \mathbb{Z}_n : d(f^{s_k+j}(z), f^j(x_k)) > \epsilon\}\right| = 0 < \delta_2 n$$

for each $k$.

Hence $\mathcal{C}$ is $(n, \delta_1, \delta_2, \{s_k\}_{k=1}^\infty, \epsilon)$-traced by $z$. This implies that $(X, f)$ has approximate product property. \qed

Analogous to [2, Example 3.4], it is not difficult to show the following fact.

Proposition 3.4. Suppose that $(X, f)$ has approximate product property and $(Y, g)$ has tempered specification property, then their direct product $(X \times Y, f \times g)$ has approximate product property.

In [22], Pfister and Sullivan has shown that approximate product property implies entropy denseness.

Proposition 3.5 ([22]). Suppose that $(X, f)$ has approximate product property. Then for every $\mu \in \mathcal{M}(X, f)$, every $\eta > 0$, there is $\nu \in \mathcal{M}_e(X, f)$ such that

$$D(\mu, \nu) < \eta \text{ and } h_\nu(f) > h_\mu(f) - \eta.$$  

Recall that $\mathcal{M}(X, f)$ is a Choquet simplex. Entropy denseness implies that $\mathcal{M}_e(X, f)$, which is exactly the set of extreme points of $\mathcal{M}(X, f)$, is dense in $\mathcal{M}(X, f)$. In this case we know that Then $\mathcal{M}(X, f)$ is a Poulsen simplex if it is non-trivial, i.e. it is not a singleton. The structure of a Poulsen simplex has been studied in [19]. Hence we have the following corollary.

Corollary 3.6. Suppose that $(X, f)$ has approximate product property. Then the following hold.
(1) As Poulsen simplex is unique up to affine homeomorphisms, we have that $M(X, f)$ is affinely homeomorphic to $M(\{0, 1\}^\mathbb{Z}, \sigma)$, where $\{0, 1\}^\mathbb{Z}$ is the two-sided full shift.
(2) $M_b(X, f)$ is homeomorphic to the Hilbert space $l^2$.
(3) $M_e(X, f)$ is arcwise connected by simple arcs.

4. Empirical Measures

In this section we investigate some facts on empirical measures to prepare ourselves for the proof of the main result. Our proofs of these lemmas mainly follow [7, Section 5.3]. We shall also need some notions and lemmas from the work of Pfister and Sullivan. We list them at the end of this section.

For $x \in X$ and $n \in \mathbb{N}$, we define the empirical measure $E(x, n)$ such that

$$\int \phi dE(x, n) := \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x))$$
for every $\phi \in C(X)$.

Given a set $U \subset M(X, T)$, let

$$E(U, n) := \{x \in X : E(x, n) \in U\}.$$

Let $\delta > 0$ and $B_\delta = B_\delta(\mu) := B(\mu, \delta)$. For $N \in \mathbb{Z}^+$, denote

$$Z_{N,\delta} = Z_{N,\delta}(\mu) := \{x \in X : f^k(x) \in E(B_\delta, N) \text{ for every } k \in \mathbb{N}\} = \{x \in X : f^k(x) \in E(B_\delta, N) \text{ for every } k \in \mathbb{N}\}. \quad (2)$$

Then $f(Z_{N,\delta}) \subset Z_{N,\delta}$. By [32], the map $x \mapsto E(x, N)$ is continuous. Hence the set $Z_{N,\delta}$ is closed. For $\varepsilon > 0$, denote

$$\text{Var}(\varepsilon) := \max\{D(E(x, 1), E(y, 1)) : d(x, y) \leq \varepsilon\}.$$ 
By continuity of the map $x \mapsto E(x, 1)$, we have

$$\lim_{\varepsilon \to 0} \text{Var}(\varepsilon) = 0. \quad (3)$$

**Lemma 4.1.** For any $N \in \mathbb{Z}^+$ and any $\nu \in M(Z_{N,\delta}, f|_{Z_{N,\delta}})$, we have $D(\nu, \mu) \leq \delta$.

*Proof.* Assume that $\nu \in M(Z_{N,\delta}, f|_{Z_{N,\delta}})$ is ergodic. There is a generic point $x \in Z_{N,\delta}$ such that $E(x, n)$ converges to $\nu$ as $n \to \infty$.

Write $n = kN + l$ such that $k \in \mathbb{N}$ and $0 \leq l < N$. Note that

$$nE(x, n) = \sum_{j=0}^{k-1} (NE(f^jN(x), N)) + lE(f^kN(x), l).$$

For each $j \in \mathbb{N}$, as $x \in Z_{N,\delta}$, by (2), we have $E(f^jN(x), N) \in B_\delta$ and hence

$$D \left( E \left( f^jN(x), N \right), \mu \right) \leq \delta.$$
Recall that $D^*$ is the diameter of the set of probability measures on $X$. Then for $n > \frac{NB^*}{\eta}$, we have

$$D (\mathcal{E}(x, n), \mu) = D \left( \sum_{j=0}^{k-1} \frac{N}{n} \mathcal{E} \left( f^{jN}(x), N \right) + \frac{1}{n} \mathcal{E} \left( f^{kN}(x), l \right), \mu \right)$$

$$\leq \sum_{j=0}^{k-1} \frac{N}{n} D \left( \mathcal{E} \left( f^{jN}(x), N \right), \mu \right) + \frac{1}{n} D \left( \mathcal{E} \left( f^{kN}(x), l \right), \mu \right)$$

$$< \delta + \frac{ND^*}{n}.$$ 

This implies that $D(\nu, \mu) \leq \delta$ as $\mathcal{E}(x, n) \to \nu$.

When $\nu$ is not ergodic, the result follows from ergodic decomposition.

\[ \Box \]

**Lemma 4.2.** Assume that

$$\frac{2D^*}{\eta} < T \leq \frac{1}{\delta_1} \text{ and } \text{Var}(\varepsilon) + (\delta_1 + \delta_2)D^* < \eta. \quad (4)$$

Let $\mathcal{C}$ be a sequence in $E(B(x, \eta), M)$ that is $(\mathcal{M}, \delta_1, \delta_2, \mathcal{G}, \varepsilon)$-traced by $z$. Then $z \in Z_{T \mathcal{M}, 3\eta}$.

*Proof.* Given any $n \in \mathbb{N}$, we need to show that $D (\mathcal{E}(f^n(z), T \mathcal{M}), \mu) < 3\eta$.

Denote $\mathcal{C} = \{x_k\}$ and $\mathcal{G} = \{t_k\}_{k=1}^\infty$. There is unique $k$ such that $t_k < n \leq t_{k+1}$. Denote

$$s := (t_{k+1} - n) + \sum_{j=1}^{T-2} (t_{k+j+1} - t_{k+j}) = t_{k+T-1} - n.$$

By (4), we have

$$T \mathcal{M} > (T - 1)M(1 + \delta_1) > s.$$

Then

$$\mathcal{E}(f^n(z), T \mathcal{M}) = \frac{t_{k+1} - n}{T \mathcal{M}} \mathcal{E}(f^n(y), t_{k+1} - n)$$

$$+ \sum_{j=1}^{T-2} \frac{t_{k+j+1} - t_{k+j}}{T \mathcal{M}} \mathcal{E}(f^{t_{k+j}}(y), t_{k+j+1} - t_{k+j})$$

$$+ \frac{T \mathcal{M} - s}{T \mathcal{M}} \mathcal{E}(f^{t_{k+T-1}}(y), T \mathcal{M} - s).$$

For each $j$, denote

$$r_j := |\{l \in \mathbb{Z}_M : d(f^{t_{k+j} + l}(z), l^j(x_{k+j})) > \varepsilon\}| < \delta_2 M.$$

Then the tracing property (1) yields that

$$D (\mathcal{E}(f^{t_{k+j}}(z), t_{k+j+1} - t_{k+j}), \mu)$$

$$\leq \frac{M}{t_{k+j+1} - t_{k+j}} D (\mathcal{E}(f^{t_{k+j}}(z), M), \mu) + \frac{(t_{k+j+1} - t_{k+j} - M)D^*}{t_{k+j+1} - t_{k+j}}$$

$$< D (\mathcal{E}(f^{t_{k+j}}(z), M), \mathcal{E}(x_{k+j}, M)) + D (\mathcal{E}(x_{k+j}, M), \mu) + \delta_1 D^*$$

$$< \frac{M - r_j}{M} \text{Var}(\varepsilon) + \frac{r_j}{M} D^* + \eta + \delta_1 D^*$$

$$< 2\eta.$$
Hence
\[
D(\mathcal{E}(f^n(z), TM), \mu) \leq \frac{t_{k+1} - n}{TM} D^* + \frac{t_{k+T-1} - t_{k+1}}{TM} \cdot 2\eta + \frac{TM - s}{TM} D^* \\
\leq \frac{2M}{TM} D^* + 2\eta \\
< 3\eta.
\]

\[
\square
\]

**Definition 4.3.** Let \( S \) be a subset of \( X \). For \( n \in \mathbb{Z}^+ \), \( \delta > 0 \) and \( \varepsilon > 0 \), we say that \( S \) is \((n, \delta, \varepsilon)\)-separated if for any distinct points \( x, y \in S \), we have
\[
|\{k \in \mathbb{Z}_M : d(f^k(x), f^k(y)) > \varepsilon\}| > \delta M.
\]

It is easy to see that if \( 0 < \delta < \delta' \), then every \((n, \delta', \varepsilon)\)-separated set is also \((n, \delta, \varepsilon)\)-separated.

**Lemma 4.4** ([22, Proposition 2.1]). Let \( \nu \) be an ergodic measure and \( h < h_\nu(f) \). Then there are \( \delta > 0 \) and \( \gamma > 0 \) such that for any neighborhood \( U \) of \( \nu \), there is \( N^* = N^*(h, \delta, \gamma, U) > 0 \) such that for any \( n \geq N^* \) there is an \((n, \delta, \gamma)\)-separated set \( \Gamma_n \subset E(U, n) \) with \( |\Gamma_n| \geq e^{nh} \).

**Lemma 4.5** ([22, Lemma 2.1]). For \( n \in \mathbb{Z}^+ \) and \( \delta \in (0, \frac{1}{2}) \), denote
\[
Q(n, \delta) := |\{A \in \mathbb{Z}_n : |A| > (1 - \delta)n\}|.
\]

Then
\[
\frac{\ln Q(n, \delta)}{n} \leq -\delta \ln \delta - (1 - \delta) \ln(1 - \delta).
\]

5. **Local Denseness**

The following proposition is the crucial part of the paper. Compared to the result of Pfister and Sullivan (Theorem 3.5), it contains an upper estimate of the entropy of \( \nu \). This is carried out by applying an argument developed in [30] by the author.

**Proposition 5.1.** Let \((X, f)\) be an asymptotically entropy expansive system with approximate product property. Then for every ergodic measure \( \mu \in \mathcal{M}_e(X, f) \), every \( \eta_0 > 0 \), every \( h_0 \in (0, h_\mu(f)) \) and every \( \beta_0 > 0 \), there is an ergodic measure \( \nu \in \mathcal{M}_e(X, f) \) such that
\[
D(\nu, \mu) < \eta_0 \text{ and } |h_\nu(f) - h_0| < \beta_0.
\]

The conclusion of proposition 5.1 can be directly proved for every invariant measure \( \mu \), with an argument similar to the one in [7, Section 5.3]. However, we choose to use the result of entropy denseness that is already proved in [22] to make our exposition more concise.

We prove Proposition 5.1 in the rest of this section and split the proof into two subsections.
5.1. **Construction.** Suppose that we are given \( \mu \in \mathcal{M}(X, f), \eta_0, \delta_0 > 0 \) and \( h_0 \in (0, h_\mu(f)) \). We fix
\[
\eta := \frac{\eta_0}{4}, \quad \beta := \frac{\delta_0}{6} \quad \text{and} \quad T > \frac{2D^*}{\eta}.
\]
By Lemma 4.4, there are \( \delta_0 > 0, \gamma_0 > 0 \) and \( N^* = N^*(h_0, \delta_1, \gamma_0, B(\mu, \eta)) \) such that for any \( n \geq N^* \) there is an \((n, \delta_0, \gamma_0)\)-separated set \( \Gamma_n \subset E(B(\mu, \eta), n) \) with
\[
e^{n h_0} \leq |\Gamma_n| < e^{n(h_0 + \beta)}.
\]
By (3), we can fix \( \varepsilon > 0 \) such that
\[
h_{loc}(f, 2\varepsilon) < \beta, \quad \Var(\varepsilon) < \frac{1}{4} \eta \quad \text{and} \quad \varepsilon < \frac{1}{3} \gamma_0.
\]
As \( h(f) < \infty \), we can fix \( \delta_1 > 0 \) such that
\[
\delta_1 < \frac{1}{2T} \quad \text{and} \quad \delta_1(h(f) + \beta) < \beta.
\]
By Lemma 4.5, we can fix \( \delta_2 > 0 \) such that
\[
\delta_2 < \min\left\{ \frac{\delta_0}{2}, \frac{1}{T} \right\} \quad \text{and} \quad -\delta_2 \ln \delta_2 - (1 - \delta_2) \ln(1 - \delta_2) < \beta.
\]
As
\[
\limsup_{n \to \infty} \frac{\ln r(n, \varepsilon)}{n} = h^*(f, \varepsilon) \leq h(f),
\]
there is \( N_0 > 0 \) such that
\[
\ln r(n, \varepsilon) \leq n(h(f) + \beta) \quad \text{for every} \quad n > N_0.
\]
We fix \( M > M(\varepsilon, \delta_1, \delta_2) \) as in Definition 3.2 such that
\[
M > \max\left\{ N^*, \frac{N_0}{\delta_1} \right\} \quad \text{and} \quad \frac{\ln(\delta_1 M)}{M} < \beta.
\]
By Lemma 4.4, there is a fixed \((M, \delta_0, \gamma_0)\)-separated set \( \Gamma_M \subset E(B(\mu, \eta), n) \) with
\[
e^{M h_0} \leq |\Gamma_M| < e^{M(h_0 + \beta)}.
\]
Denote \( M_1 := \lfloor \delta_1 M \rfloor, \Sigma := \Sigma_{M_1} \) and \( \Gamma := (\Gamma_M)^{\mathbb{Z}^+} \). For each \( \xi = \{\xi(k)\}_{k=1}^\infty \in \Sigma \), denote
\[
t_k(\xi) := \sum_{j=1}^{k-1} (M + \xi(j)) \quad \text{for each} \quad k \quad \text{and} \quad \mathcal{G}_\xi := \{t_k(\xi)\}_{k=1}^\infty.
\]
For each \( \xi \in \Sigma \) and each sequence \( \mathcal{G} \in (\Gamma_M)^{\mathbb{Z}^+} \), denote
\[
Y_{\mathcal{G}, \xi} := \{y \in X: \mathcal{G} \text{ is } (M, \delta_1, \mathcal{G}_\xi, \delta_2, \varepsilon)-\text{traced by } y\}.
\]
Let
\[
Y := \bigcup_{\mathcal{G} \in \Gamma, \xi \in \Sigma} Y_{\mathcal{G}, \xi}.
\]
Note that by (5), (6) and (7), we have
\[
\Var(\varepsilon) + (\delta_1 + \delta_2)D^* < \frac{1}{4} \eta + \frac{3D^*}{2T} < \eta.
\]
Hence (4) holds. By Lemma 4.2, we have \( Y \subset Z_{\Gamma M, 3\eta} \).

Denote by \( \sigma_\Gamma \) and \( \sigma_\Sigma \) the shift maps on \( \Gamma \) and \( \Sigma \), respectively. It is clear that the following invariance holds.
Lemma 5.2. For every $\mathcal{C} \in \Gamma$ and $\xi \in \Sigma$, we have
\[ f^{t_2}(\xi)(Y_{\mathcal{C},\xi}) \subset Y_{\delta_r(\mathcal{C}),\sigma_\Sigma(\xi)}. \]

Lemma 5.3. Suppose that $y \in Y_{\mathcal{C},\xi}$ and $y' \in Y_{\mathcal{C}',\xi'}$ such that
\[ t_n(\xi) = t_n(\xi') \quad \text{and} \quad x_n(\mathcal{C}) \neq x_n(\mathcal{C}'). \]
Then $y, y'$ are $(nM(1 + \delta_1), \varepsilon)$-separated.

Proof. Denote $t := t_n(\xi) = t_n(\xi')$. Denote
\[ A := \{ j \in \mathbb{Z}_M : d(f^{t+j}(y), f^j(x_n(\mathcal{C}))) \leq \varepsilon \} \quad \text{and} \quad A' := \{ j \in \mathbb{Z}_M : d(f^{t+j}(y'), f^j(x_n(\mathcal{C}'))) \leq \varepsilon \}. \]
By (10) and the tracing property, we have $|A|, |A'| \geq (1 - \delta_2)M$, hence
\[ |A \cap A'| \geq (1 - 2\delta_2)M > (1 - \delta_0)M. \]

But $x_n(\mathcal{C})$ and $x_n(\mathcal{C}')$ are distinct elements in $\Gamma_M$. They must be $(M, \delta_0, \gamma_0)$-separated. Then there must be $\tau \in A \cap A'$ such that
\[ d(f^\tau(x_n(\mathcal{C})), f^\tau(x_n(\mathcal{C}'))) > \gamma_0 > 3\varepsilon. \]
Hence
\[
d(f^{t+j}(y), f^{t+j}(y')) \geq d(f^\tau(x_n(\mathcal{C})), f^\tau(x_n(\mathcal{C}'))) \]
\[ - d(f^{t+j}(y), f^\tau(x_n(\mathcal{C}))) \]
\[ - d(f^{t+j}(y'), f^\tau(x_n(\mathcal{C}'))) \]
\[ > \varepsilon. \]
Moreover, we have
\[ t + \tau \leq \sum_{k=1}^n (t_{k+1}(\xi) - t_k(\xi)) \leq nM(1 + \delta_1). \]

Denote by
\[ C^\Gamma_{p_1 \cdots p_n} = \{ \mathcal{C} \in \Gamma : x_j(\mathcal{C}) = p_j \quad \text{for each} \quad j = 1, \cdots, n \} \]
a cylinder of rank $n$ in $\Gamma$ and
\[ C^\Sigma_{w_1 \cdots w_n} = \{ \xi \in \Sigma : \xi(j) = w_j \quad \text{for each} \quad j = 1, \cdots, n \} \]
a cylinder of rank $n$ in $\Sigma$. For each cylinder $C^\Gamma$ and $C^\Sigma$, denote
\[ Y_{C^\Gamma,C^\Sigma} = \bigcup_{\mathcal{C} \in C^\Gamma, \xi \in C^\Sigma} Y_{\mathcal{C},\xi}. \]

Corollary 5.4. Suppose that $y_i \in Y_{C_i^\Gamma,C_i^\Sigma}$ for $i = 1, 2$ such that $C_1^\Gamma, C_2^\Gamma$ are distinct cylinders of rank $n$ in $\Gamma$ and $C_1^\Sigma$ is a cylinder of rank $n - 1$ in $\Sigma$. Then $y_1, y_2$ are $(nM(1 + \delta_1), \varepsilon)$-separated.

Lemma 5.5. Let $\{ y_n \}$ be a sequence in $Y$ such that $y_n \to \tilde{y}$ in $X$. Then there are $\mathcal{C} \in \Gamma$ and $\xi \in \Sigma$ such that $\tilde{y} \in Y_{\mathcal{C},\xi}$. Hence $Y$ is compact.
Hence \( \pi \) is finite. For each \( k \in \mathbb{Z}^+ \) and \( y \in Y_{\tilde{\mathcal{C}}, \tilde{\xi}} \), denote
\[
A_k(y) := \{ \tau \in \mathbb{Z}_\mathcal{M} : d(f^{t_k(\xi)} + \tau)(y), f^\tau(x_k(\mathcal{C})) \leq \varepsilon \} \in \mathcal{P}.
\]
Assume that \( y_n \in Y_{\mathcal{C}_n, \xi_n} \) for each \( n \). Note that \( \Gamma, \Sigma \) and \( \mathcal{P}^{\mathbb{Z}^+} \) are compact metric space as symbolic spaces. We can find a subsequence \( \{ n_j \}_{j=1}^{\infty} \) such that
\[
\mathcal{C}_{n_j} \rightarrow \tilde{\mathcal{C}}, \xi_{n_j} \rightarrow \tilde{\xi} \quad \text{and} \quad \{ A_k(y_{n_j}) \}_{k=1}^{\infty} \rightarrow \{ A_k \}_{k=1}^{\infty}.
\]
For each \( k \in \mathbb{Z}^+ \), there is \( N_k \) such that for every \( n_j > N_k \), we have
\[
x_k(\mathcal{C}_{n_j}) = x_k(\tilde{\mathcal{C}}), t_k(\xi_{n_j}) = t_k(\tilde{\xi}) \quad \text{and} \quad A_k(y_{n_j}) = A_k.
\]
For each \( \tau \in A_k \), we have
\[
d(f^{t_k(\tilde{\xi}) + \tau}(\tilde{y}), f^\tau(x_k(\tilde{\mathcal{C}}))) = \lim_{n_j \rightarrow \infty} d(f^{t_k(\xi_{n_j}) + \tau}(y_{n_j}), f^\tau(x_k(\mathcal{C}_{n_j}))) \leq \varepsilon.
\]
This implies that \( \tilde{y} \in Y_{\mathcal{C}, \tilde{\xi}} \). \( \square \)

5.2. Entropy estimate.

**Lemma 5.6.** For every \( n \in \mathbb{Z}^+ \), every cylinder \( C^\Gamma = C^\Gamma_{P_1 \cdots P_n} \) and every cylinder \( C^\Sigma = C^\Sigma_{w_1 \cdots w_n} \) in \( \Sigma \), there are at most
\[
(Q(M, \delta_2) r(\varepsilon)^{d_2 M r(M_1, \varepsilon)})^{n+2}
\]
points in \( Y_{C^\Gamma, C^\Sigma} \) that are \((nM, 2\varepsilon)\)-separated.

**Proof.** Let \( S(M_1, \varepsilon) \) be a fixed \((M_1, \varepsilon)\)-spanning subset of \( X \) with the minimal cardinality. Then \( |S(M_1, \varepsilon)| = r(M_1, \varepsilon) \). Let \( S_1 \) be a fixed \((1, \varepsilon)\) spanning subset of \( X \) with the minimal cardinality \( r(\varepsilon) \). Let \( \mathcal{A} := \{ A_1, \cdots, A_n \} \) be an \( n \)-tuple in \( \mathcal{P}^n \). Fix any \( \xi \in C \). Denote
\[
Y_{C^\mathcal{A}, C^\infty}(\mathcal{A}) := \{ y \in Y_{C^\mathcal{A}, C^\infty} : d(f^{t_k(\xi) + j}(y), f^j(p_k)) \leq \varepsilon
\]
for every \( j \in A_k, k = 1, 2, \cdots, n \}\}

Denote
\[
\Omega(\mathcal{A}) := \prod_{j=0}^{t_{n+1}(\xi)-1} \Omega_j(\mathcal{A}),
\]
where for \( t_k(\xi) \leq j < t_{k+1}(\xi), k = 1, 2, \cdots, n, \)
\[
\Omega_j(\mathcal{A}) := \begin{cases} \{ f^{j-t_k(\xi)}(p_k) \}, & \text{if } j - t_k(\xi) \in A_k; \\ S_1, & \text{if } j - t_k(\xi) \in \mathbb{Z}_M \setminus A_k; \\ f^{j-t_k(\xi)-(M-1)}(S(M_1, \varepsilon)), & \text{otherwise.} \end{cases}
\]
Let \( S \) be an \((t_{n+1}(\xi), 2\varepsilon)\)-separated set in \( Y_{C^\mathcal{A}, C^\infty}(\mathcal{A}) \). Then there is an injection \( \pi : S \rightarrow \Omega(\mathcal{A}) \) such that for every \( y \in S \),
\[
d(f^j(y), f^j(\pi(y))) \leq \varepsilon \quad \text{for each } j = 0, 1, \cdots, t_{n+1}(\xi) - 1.
\]
Hence
\[
|S| \leq |\Omega(\mathcal{A})| \leq (|S_1|^d_2 M |S(M_1, \varepsilon)|)^n = (r(\varepsilon)^{d_2 M} r(M_1, \varepsilon))^n.
\]
By Lemma 4.5, there are at most \( Q(M, \delta_2)^n \) \( n \)-tuples in \( (P)^n \). So the maximal cardinality of an \( (t_{n+1}(\xi), 2\varepsilon) \)-separated set in \( Y_{\varepsilon, \xi} \) is at most

\[
(Q(M, \delta_2)^r(\varepsilon)^{\delta_2 M r(M_1, \varepsilon)})^n.
\]

The result follows since \( t_{n+1}(\xi) \geq nM \).

**Corollary 5.7.**

\[
s(Y, nM, 2\varepsilon) < (e^{M(h_0 + \beta)} M_1 Q(M, \delta_2)^r(\varepsilon)^{\delta_2 M r(M_1, \varepsilon)})^n.
\]

**Proof.** Just note that the numbers of distinct cylinders of rank \( n \) in \( \Gamma \) and \( \Sigma \) are \( |\Gamma_M|^n \) and \( M_1^n \), respectively.

**Corollary 5.8.** For each \( \tau \in \{0, \cdots, M + M_1 - 1\} \) and every \( n \in \mathbb{Z}^+ \), we have

\[
s(f^\tau(Y), nM, 2\varepsilon) \leq (e^{M(h_0 + \beta)} M_1 Q(M, \delta_2)^r(\varepsilon)^{\delta_2 M r(M_1, \varepsilon)})^{n+2}
\]

**Proof.** Note that if \( S \) is an \( (nM, 2\varepsilon) \)-separated subset of \( f^\tau(Y) \), then \( f^{-\tau}(S) \) includes an \( (nM + \tau, 2\varepsilon) \)-separated subset of \( f^\tau(Y) \). Moreover, we have

\[
nM + \tau < (n + 2)M
\]

for \( \tau \in \{0, \cdots, M + M_1 - 1\} \).

**Lemma 5.9.** Let

\[
\Lambda := \bigcup_{k=0}^{M+M_1-1} f^k(Y).
\]

Then \( \Lambda \) is a compact \( f \)-invariant subset in \( Z_{TM, 3\eta} \).

**Proof.** We have that \( \Lambda \) is compact since \( Y \) is compact. We have \( \Lambda \subset Z_{TM, 3\eta} \) since \( Y \subset Z_{TM, 3\eta} \) and \( f(Z_{TM, 3\eta}) \subset Z_{TM, 3\eta} \).

For every \( z \in \Lambda \), there is \( y \in Y \) and \( \tau \in \{0, \cdots, M + M_1 - 1\} \) such that \( f^\tau(y) = z \).

If \( r < M + M_1 - 1 \), then \( f^\tau(z) = f^{\tau+1}(y) \in \Lambda \).

Assume that \( y \in Y_{\varepsilon, \xi} \). Note that \( M \leq t_2(\xi) \leq M + M_1 \) and by Lemma 5.2, we have

\[
f^{t_2(\xi)}(y) \in Y.
\]

Hence if \( \tau = M + M_1 - 1 \), then

\[
f(x) = f^{\tau+1}(y) = f^{M+M_1-t_2(\xi)}(f^{t_2(\xi)}(y)) \in f^{M+M_1-t_2(\xi)}(Y) \subset \Lambda.
\]

So we can conclude that \( f(\Lambda) \subset \Lambda \).

**Lemma 5.10.** We have

\[
h_0 - \beta_0 < h(\Lambda, f) < h_0 + \beta_0.
\]

Hence \( \Lambda \) supports an ergodic measure \( \nu \) as requested in Proposition 5.1.

**Proof.** By Corollary 5.8, we have for each \( n \in \mathbb{Z}^+ \),

\[
s(\Lambda, nM, 2\varepsilon) \leq \sum_{\tau=0}^{M+M_1-1} s(f^\tau(Y), nM, 2\varepsilon)
\]

\[
< (M + M_1)(e^{M(h_0 + \beta)} M_1 Q(M, \delta_2)^r(\varepsilon)^{\delta_2 M r(M_1, \varepsilon)})^{n+2}.
\]  

(11)
Then for every $t \in \mathbb{Z}^+$, there is $n \in \mathbb{Z}^+$ such that $(n-1)M < t \leq nM$. Hence we have

$$h(\Lambda, f, 2\varepsilon) = \limsup_{t \to \infty} \frac{\ln s(\Lambda, t, 2\varepsilon)}{t} \leq \limsup_{n \to \infty} \frac{\ln s(\Lambda, nM, 2\varepsilon)}{(n-1)M} \leq (h_0 + \beta) + \frac{\ln M + \ln Q(M, \delta_2) + \ln r(M_1, \varepsilon)}{M} + \delta_2 \ln r(\varepsilon).$$

By (8) and (9), we have

$$\ln r(M_1, \varepsilon) \leq \delta_1 M(h(f) + \beta).$$

Then by (5), (6) and (7), we have

$$h(\Lambda, f) \leq h(\Lambda, f, 2\varepsilon) + h_{as}(f, 2\varepsilon) < (h_0 + \beta) + 4\beta + \beta = h_0 + \beta_0.$$

For each $n$, there are $|\Gamma_M|^n$ cylinders of rank $n$ in $\Gamma$ and $M_1^{n-1}$ cylinders of rank $n-1$ in $\Sigma$. There must be a cylinder $C_{\sigma_0}^n$ of rank $n-1$ such that

$$|Q(C_{\sigma_0}^n)| \geq \frac{|\Gamma_M|^n}{M_1^{n-1}} \geq \frac{e^{nMh_0}}{M_1^{n-1}},$$

where

$Q(C_{\sigma_0}^n) := \{C^n : C^n \text{ is a cylinder of rank } n \text{ in } \Gamma \text{ and } Y_{C^n, C_{\sigma_0}^n} \neq \emptyset\}.$

By Corollary 5.4, we have

$$s(\Lambda, nM(1 + \delta_1), \varepsilon) \geq s(Y, nM(1 + \delta_1), \varepsilon) \geq s(Y_{C^n, C_{\sigma_0}^n}, nM(1 + \delta_1), \varepsilon) \geq |Q(C_{\sigma_0}^n)|.$$

Hence

$$h(\Lambda, f) \geq h(\Lambda, f, \varepsilon) \geq \limsup_{n \to \infty} \frac{\ln s(\Lambda, nM(1 + \delta_1), \varepsilon)}{nM(1 + \delta_1)} \geq \limsup_{n \to \infty} \frac{Mh_0 - (n-1)\ln M_1}{nM(1 + \delta_1)} \geq h_0 - \delta_1 h_0 - \frac{\ln(\delta_1 M)}{M} > h_0 - 2\beta.$$

As $(X, f)$ is asymptotically entropy expansive, $\Lambda$ supports an ergodic measure $\nu$ with $h_\nu(f) = h(\Lambda, f)$. By Lemma 4.1, we have $D(\nu, \mu) \leq 3\eta < \eta_0$. □

**Corollary 5.11.** Let $(X, f)$ be a system with approximate product property and positive topological entropy. Then $(X, f)$ is not minimal.

**Proof.** Suppose that $h_0 + \beta_0 < h(f)$. By (11), there is $N$ such that

$$s(\Lambda, nM, 2\varepsilon) < nM(h_0 + \beta_0) \text{ for every } n > N.$$

This implies that $\Lambda$ is a proper subset of $X$ that is compact and $f$-invariant. Hence $(X, f)$ is not minimal. □
6. Proof of the Main Theorem

Lemma 6.1. Let $(X, f)$ be an asymptotically entropy expansive system with approximate product property. For $\alpha' > \alpha$, denote

$$M_{\alpha, \alpha'} := \{ \mu \in M(X, f) : \alpha \leq h_\mu(f) < \alpha' \}.$$ 

Then $M_{\alpha, \alpha'} \cap M_e(X, f)$ is dense in $M^\alpha$.

Proof. Suppose that we are given $\mu \in M^\alpha$ and every $\delta > 0$. Let

$$\tilde{\mu} := \mu + \frac{\delta}{3}D_\ast(\mu_M - \mu).$$

Then

$$D(\tilde{\mu}, \mu) < \frac{\delta}{3} \quad \text{and} \quad h_{\tilde{\mu}}(f) > \alpha.$$ 

By Proposition 3.5, there is $\mu' \in M_e(X, f)$ such that

$$D(\mu', \tilde{\mu}) < \frac{\delta}{3} \quad \text{and} \quad h_{\mu'}(f) > \alpha.$$ 

By Proposition 5.1, there is $\nu \in M_e(X, f)$ such that

$$D(\nu, \mu') < \frac{\delta}{3} \quad \text{and} \quad \alpha \leq h_{\nu}(f) < \alpha'.$$

Then we have $\nu \in (M_{\alpha, \alpha'} \cap M_e) \cap (B(\mu, \delta) \cap M^\alpha)$. Hence $M_{\alpha, \alpha'} \cap M_e(X, f)$ is dense in $M^\alpha$. □

Proof of Theorem 1.1. As $(X, f)$ is asymptotically entropy expansive, the map $\mu \to h_\mu(f)$ is upper semi-continuous. This implies that $M^\alpha(X, f)$ is a compact metric subspace of $M(X, f)$, hence is a Baire space. As $M_e(X, f)$ is a $G_\delta$ set in $M(X, f)$, the set

$$M^\alpha_e := M_e(X, f) \cap M^\alpha$$

is a $G_\delta$ set in $M^\alpha$.

By Lemma 6.1, for every $\alpha' > \alpha$, $M_{\alpha, \alpha'} \cap M^\alpha_e = M_{\alpha, \alpha'} \cap M_e(X, f)$ is dense in $M^\alpha$. Note that $M_{\alpha, \alpha'}$ is relatively open in $M^\alpha(X, f)$. Hence $M_{\alpha, \alpha'}$ is open dense and $M_{\alpha, \alpha'} \cap M^\alpha_e$ is residual in $M^\alpha$. Finally, we have that

$$M_e(X, f, \alpha) = \bigcap_{k=1}^{\infty} (M_{\alpha, \alpha+\frac{1}{k}} \cap M^\alpha_e)$$

is a residual subset of $M^\alpha$.

If $\alpha = 0$, we just note that $M^0(X, f)$ is exactly $M(X, f)$.

□

Proof the Corollary 1.2. Let $\delta > 0$ such that $B(\mu, 2\delta) \subset U$. 

Let $0 \leq \alpha < h_\mu(f)$. Then $B(\mu, \delta) \cap M^\alpha(X, f)$ is open and nonempty in $M^\alpha(X, f)$. By Theorem 1.1, there is an ergodic measure $\nu \in B(\mu, \delta) \cap M^\alpha(X, f)$ such that $h_\nu(f) = \alpha$.

If $h_\mu(f) < h(f)$, then

$$\tilde{\mu} := \mu + \frac{\delta}{D_\ast(\mu_M - \mu)} \in B(\mu, \delta) \text{ and } h_{\tilde{\mu}}(f) > h_\mu(f).$$

There is an ergodic measure $\nu \in B(\tilde{\mu}, \delta) \cap M^{h_\mu(f)}$ such that $h_{\nu}(f) = h_\mu(f)$. Then $\nu \in B(\mu, 2\delta) \subset U$. □
In particular, by asymptotic entropy expansiveness, \((X, f)\) has ergodic measures of the maximal entropy. □

7. Average Lyapunov exponent for asymptotically additive functions

In this section we discuss a result related to entropy denseness of systems with approximate product property.

In [31], Tian, Wang and Wang introduced the notion of the average Lyapunov exponent \(\chi_{\Phi}(\mu)\) for a sequence \(\Phi = \{\phi_n\}_{n=1}^{\infty}\) of asymptotically additive functions for \((X, f)\) with respect to an invariant measure \(\mu \in \mathcal{M}(X, f)\). We refer the readers to their paper [31] for exact definitions. In [31], it is shown that if \((X, f)\) has periodic gluing orbit property, then for each \(a \in \mathcal{L}_{\Phi} := \left( \inf_{\mu \in \mathcal{M}(X, f)} \chi_{\Phi}(\mu), \sup_{\mu \in \mathcal{M}(X, f)} \chi_{\Phi}(\mu) \right)\), there is an ergodic measure \(\nu_a\) of full support such that

\[
\chi_{\Phi}(\nu_a) = a. \tag{12}
\]

We perceive that existence of an ergodic measure satisfying (12), which is called intermediate exponent property after [31], is a corollary of the fact that \(\mathcal{M}_e(X, f)\) is dense in \(\mathcal{M}(X, f)\). However, in general we do not know if such an ergodic measure has full support.

The crucial part of the proof is also presented in [31].

Lemma 7.1 ([31, Lemma 2.5]). The map \(\chi_{\Phi} : \mathcal{M}(X, f) \to \mathbb{R}\) is continuous with respect to the weak-* topology.

Proposition 7.2. Suppose that \(\mathcal{M}_e(X, f)\) is dense in \(\mathcal{M}(X, f)\). Then \((X, f)\) has intermediate exponent property.

Proof. Suppose that we are given \(a \in \mathcal{L}_{\Phi}\). There are \(\mu_1, \mu_2 \in \mathcal{M}(X, f)\) such that

\[
\chi_{\Phi}(\mu_1) < a < \chi_{\Phi}(\mu_2).
\]

As \(\mathcal{M}_e(X, f)\) is dense, by Lemma 7.1, there are ergodic measures \(\nu_1, \nu_2 \in \mathcal{M}_e(X, f)\) such that

\[
\chi_{\Phi}(\nu_1) < a < \chi_{\Phi}(\nu_2).
\]

Note that denseness of \(\mathcal{M}_e(X, f)\) implies that \(\mathcal{M}(X, f)\) is a Poulsen simplex. By [19], \(\mathcal{M}_e(X, f)\) is arcwise connected. Hence there must be an arc in \(\mathcal{M}_e(X, f)\) that connects \(\nu_1\) and \(\nu_2\), on which there is \(\nu_a\) with \(\chi_{\Phi}(\nu_a) = a\). □

Corollary 7.3. Approximate product property implies intermediate exponent property.

8. Counterexamples

Example 8.1. In [15], Kwietniak, Oprocha and Rams constructed a one-sided shift \((X_2, \sigma)\) that has tempered specification property and multiple but finitely many ergodic measures of maximal entropy. If \(\mu_1, \mu_2\) are two distinct ergodic measures of maximal entropy for \((X_2, \sigma)\), then there is \(\delta > 0\) such that \(B\left(\frac{\mu_1 + \mu_2}{2}, \delta\right)\) contains no ergodic measures of maximal entropy. So Corollary 1.2 is optimal.
Example 8.2. Let $(X_1,f_1)$ be a subshift constructed in [12] which is uniquely ergodic and of positive topological entropy. Let $(X,f)$ be the direct product of $(X_1,f_1)$ and the full shift, then every invariant measure for $(X,f)$ is a direct product of the unique ergodic measure for $(X_1,f_1)$ and an invariant measure for the full shift. $(X,f)$ is asymptotically entropy expansive, entropy dense and $\mathcal{M}(X,f)$ is non-trivial. But it does not have ergodic measures of entropies less than $h(X_1,f_1)$.

Example 8.3. Let $(X,f)$ be the direct product of the irrational rotation and the full shift. Then $(X,f)$ has gluing orbit property. By [24, Section 3.5], this system $(X,f)$ is not universal.

Acknowledgments

The author is supported by National Natural Science Foundation of China (No. 11571387) and CUFE Young Elite Teacher Project (No. QYP1902). The author would like to thank Weisheng Wu and Daniel Thompson for helpful communications.

References

[1] F. Béguin, S. Crovisier, F. Le Roux, Construction of curious minimal uniquely ergodic homeomorphisms on manifolds: the Denjoy-Rees technique. Ann. Sci. École Norm. Sup. (4) 40 (2007), 251–308.
[2] T. Bomfim, M. J. Torres and P. Varandas, Topological features of flows with the reparametrized gluing orbit property. Journal of Differential Equations 262 (2017), 4292–4313.
[3] T. Bomfim and P. Varandas, The gluing orbit property, uniform hyperbolicity and large deviations principles for semiflows. ArXiv:1507.03905.
[4] R. Bowen, Periodic points and measures for Axiom A diffeomorphisms. Trans. Amer. Math. Soc. 154 (1971), 377–397.
[5] R. Bowen, Entropy-expansive maps. Trans. Amer. Math. Soc., 164 (1972), 323–331.
[6] D. Burget. Topological and almost Borel universality for systems with the weak specification property. Ergodic Theory and Dynamical Systems.
[7] D. Constantine, J. Lafont and D. Thompson, The weak specification property for geodesic flows on CAT(-1) spaces. to appear in Groups, Geometry, and Dynamics.
[8] Vaughn Climenhaga and Daniel J. Thompson, Unique equilibrium states for flows and homeomorphisms with non-uniform structure. Adv. Math. 303 (2016), 744–799.
[9] M. Denker, C. Grillenberger and K. Sigmund, Ergodic theory on compact spaces., Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin-New York, 1976.
[10] A. Eizenberg, Y. Kifer, and B. Weiss, Large deviations for $Z^d$-actions, Comm. Math. Phys 164 (1994), no. 3, 433–454.
[11] L. Guan, P. Sun and W. Wu, Measures of Intermediate Entropies and Homogeneous Dynamics, Nonlinearity, 30 (2017), 3349–3361.
[12] F. Hahn and Y. Katznelson, On the entropy of uniquely ergodic transformations, Trans. Amer. Math. Soc. 126 (1967), 335–360.
[13] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, Publ. Math. I.H.E.S. 51 (1980), 137–173.
[14] D. Kwietniak, M. Lacka and P. Oprocha. A panorama of specification-like properties and their consequences, Contemporary Mathematics, 669 (2016), 155–186.
[15] D. Kwietniak, P. Oprocha, and M. Rams, On entropy of dynamical systems with almost specification. Israel Journal of Mathematics, 2016, 213(1), 475–503.
[16] M. Li, Y. Shi, S. Wang and X. Wang. Measures of intermediate entropies for star vector fields. preprint, 2018.
[17] D. Lind. Ergodic group automorphisms and specification. In M. Denker and K. Jacobs, editors, Ergodic Theory, volume 729 of Lecture Notes in Mathematics, pages 93–104. Springer Berlin / Heidelberg, 1979.
[18] D. Lind. Dynamical properties of quasihyperbolic toral automorphisms. Ergodic Theory Dynam. Systems, 2 (1982), 49–68.
[19] J. Lindenstrauss, G. Olsen, and Y. Sternfeld, *The Poulsen simplex*, Ann. Inst. Fourier (Grenoble) 28 (1978), no. 1, vi, 91–114.
[20] B. Marcus. *A note on periodic points for ergodic toral automorphisms*. Monatsh. Math., 89 (1980), 121–129.
[21] M. Misiurewicz, *Topological conditional entropy*. Studia Math., 55 (1976), 175–200.
[22] C-E. Pfister and W.G. Sullivan, *Large deviations estimates for dynamical systems without the specification property*. Application to the $\beta$-shifts, Nonlinearity, 18 (2005), 237–261.
[23] R.R. Phelps, Lectures on Choquet’s theorem, second ed., Lecture Notes in Mathematics, vol. 1757, Springer-Verlag, Berlin, 2001.
[24] A. Quas, and T. Soo, *Ergodic universality of some topological dynamical systems*, Transactions of the American Mathematical Society, 2016, 368(6), 4137–4170.
[25] K. Sigmund, *Generic properties of invariant measures for Axiom A diffeomorphisms*. Invent. Math. 11 (1970), 99–109.
[26] P. Sun, *Zero-entropy invariant measures for skew product diffeomorphisms*, Ergodic Theory and Dynamical Systems, 30 (2010), 923–930.
[27] P. Sun, *Measures of intermediate entropies for skew product diffeomorphisms*, Discrete Contin. Dyn. Syst - A, 2010, 27(3), 1219–1231.
[28] P. Sun, *Density of metric entropies for linear toral automorphisms*, Dynamical Systems, 2012, 27(2), 197–204.
[29] P. Sun, *Minimality and gluing orbit property*. Discrete and Continuous Dynamical Systems - A, 2019, 39(7), 4041-4056.
[30] P. Sun, *Zero-entropy dynamical systems with gluing orbit property*, preprint, 2019.
[31] X. Tian, S. Wang and X. Wang, *Intermediate Lyapunov exponents for system with periodic gluing orbit property*, preprint, 2018.
[32] P. Walters, An Introduction to Ergodic Theory, Springer-Verlag, 1982.

E-mail address: sunpeng@cufe.edu.cn