METRIC ENTROPY AND DEGENERACY OF SMOOTH INTERVAL MAPS

GANG LIAO\(^1\) AND SHIROU WANG\(^2\)

Abstract. We consider \(C^r\) \((r > 1)\) maps on an interval (or a circle). By introducing the notions of folding and degenerate rate, we investigate the mechanisms of upper semi-continuity of metric entropy. To be specific, we prove that on any subset of measures with uniform folding or degenerate rate, the metric entropy is upper semi-continuous. Moreover, the sharpness of the conditions on the uniformity of folding and degenerate rate are also investigated.

1. Introduction

This paper concerns the study of metric entropy for smooth one-dimensional dynamical systems and aims to establish general criteria for the upper semi-continuity of metric entropy.

Let \(f\) be a \(C^r\) \((r > 1)\) map on an interval (or a circle) \(X\). Denote by \(\mathcal{M}_{\text{inv}}(f)\) the set of \(f\)-invariant Borel probability measures on \(X\). Given \(\mu \in \mathcal{M}_{\text{inv}}(f)\), the metric entropy, written as \(h_\mu(f)\), measures the complexity of \(f\) with respect to \(\mu\). If \(f\) is further non-invertible, from the investigation of backward process, a quantity called folding entropy was derived by Ruelle during the study of entropy production for non-equilibrium statistical mechanics [14]. To be precise, if denoting \(\epsilon\) as the measurable partition of \(X\) into single points, the folding entropy of \(f\) with respect to \(\mu\) is the conditional entropy \(H_\mu(\epsilon| f^{-1}\epsilon)\).

Our first result establishes a formula for smooth interval (or circle) maps that relates the metric entropy to folding entropy.

Theorem 1.1. Let \(f\) be a \(C^r\) \((r > 1)\) map on an interval (or a circle) \(X\). Then for any \(\mu \in \mathcal{M}_{\text{inv}}(f)\), it holds that \(h_\mu(f) = H_\mu(\epsilon| f^{-1}\epsilon)\).

Remark. In [9], a topological entropy formula was established by Misiurewicz and Szlenk for piecewise monotone interval maps:

\[
(1) \quad h_{\text{top}}(f) = \lim_{n \to +\infty} \frac{1}{n} \log c_n,
\]

Key words and phrases. Metric entropy; upper semi-continuity; interval maps; folding rate; degenerate rate.

2010 Mathematics Subject Classification. 37A35, 37C40, 37E05, 37D25, 37D35.

\(^{1}\) School of Mathematical Sciences, Center for Dynamical Systems and Differential Equations, Soochow University, Suzhou 215006, China; \(^{2}\) Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton T6G2G1, Alberta, Canada; Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; Gang Liao was partially supported by NSFC (11701402, 11790274), BK 20170327 and Jiangsu province “Double Plan”; Shirou Wang was partially supported by NSFC (11771026, 11471344) and PIMS PTCS.
where $c_n$ denotes the smallest number of intervals on which $f^n$ is monotone. The folding entropy $H_{\mu}(\|f^{-1}\epsilon\|)$ actually measures the uncertainty coming from the measure-weighted preimages of $f$ under $\mu$. Thus Theorem 1.1 can be viewed as a measure-theoretical analogue of [4], both of which investigate entropy through backward iterations.

According to [9], if $f$ has no degeneracy, the entropy map $\mu \mapsto h_\mu(f)$ is upper semi-continuous on $\mathcal{M}_{inv}(f)$. On the other hand, examples without upper semi-continuity of metric entropy [8, 15] exist when degenerate points are permissible. We note that when dealing with general interval maps, the difficulty of analysis arises from the existence of degenerate points, especially in case that accumulation among them happens. Indeed, if denote $\Sigma(f) = \{x \in X : f'(x) = 0\}$ as the set of degenerate points, then for any single point, the number of preimages far away from $\Sigma(f)$ is finite and uniformly bounded. Instead, infinite oscillations may happen surrounding $\Sigma(f)$, which brings about horseshoes with large entropy in arbitrarily small scales and thus leads to the out of control of metric entropy in the approximation process.

The upper semi-continuity of metric entropy has been studied broadly in the ergodic theory of dynamical systems and it is a key mechanism for the existence of equilibrium states. It is usually accompanied by certain kind of smoothness or hyperbolicity for differentiable dynamical systems. For $C^\infty$ systems, the upper semi-continuity of metric entropy always holds by the results of Yomdin [17] and Newhouse [10]. For any finite $r > 0$, however, there exist examples [8, 3, 15] for which an atomic measure (zero entropy) is accumulated by a sequence of measures with metric entropy uniformly bounded from below by zero, hence the metric entropy fails to be upper semi-continuous. In the context of hyperbolicity, uniformly hyperbolic systems and all diffeomorphisms away from tangencies [6] are entropy expansive, so the upper semi-continuity of metric entropy follows by Bowen [1]. Moreover, with dominated splittings (uniformly hyperbolic in the projective bundles), the upper semi-continuity property also holds true for nonuniformly hyperbolic systems [4].

In this paper, in the one-dimensional setting, we investigate mechanisms, distinct from the smooth and hyperbolic ones, for the upper semi-continuity of metric entropy. We shall give two criteria for the upper semi-continuity of metric entropy through the control of the complexity around $\Sigma(f)$.

Before elaborating the two criteria, we introduce necessary notations first. Let $\mathcal{O}_f$ be the set of sequences of open sets $\mathcal{V} = \{V_m\}_{m \geq 1}$ such that each $V_m$ is an open neighborhood of $\Sigma(f)$ and $d_H(V_m, \Sigma(f))$ decrease to zero as $m \to +\infty$, where $d_H$ denotes the Hausdorff distance between two sets. Let $\mathcal{C}$ be the set of bounded sequences $\bar{\eta} = \{\eta_m\}_{m=1}^\infty$ of positive real numbers such that $\eta_m \to 0$ as $m \to +\infty$. In the following, we shall use $\mathcal{V}$ and $\bar{\eta}$ to characterize the complexity arising from degeneracy.

1.1. Folding rate and u.s.c. of metric entropy. By Theorem 1.1 one can reduce the analysis of metric entropy to the folding entropy. Note that the folding entropy captures the complexity when one selects uncertain preimages to be deterministic with respect to an invariant measure through function $\phi(x) = -x \log x$. By the concavity of $\phi$, the folding entropy could be bounded by the average of the
logarithm of the number of preimages over the whole space:

\[ H_\mu(e^{f^{-1}e}) \leq \int_X \log \#\{f^{-1}(x)\}d\mu, \]

where for a set \( A \), \( \#A \) denotes the number of elements of \( A \) when it is finite, and let \( \#A = +\infty \), for otherwise. Note that the right hand side of (2) is well-defined since \( \log \#\{f^{-1}(x)\} \) is non-negative and measurable.

To analyze the level of folding, we introduce the notions of local folding and folding rate, which give the first criterion for the upper semi-continuity of metric entropy.

Let \( \mu \in \mathcal{M}_{inv}(f) \), for \( \bar{\eta} = \{\eta_m\}_{m=1}^\infty \in C \) and \( \mathcal{V} = \{V_m\}_{m \geq 1} \in \mathcal{O}_f \), we say that \( f \) admits folding rate \( \bar{\eta} \) with respect to \( (\mu, \mathcal{V}) \) if

\[ F_m(\mu, \mathcal{V}) := \int_{V_m} \log^+ \#\{f^{-1}(f) \cap (V_m \setminus \Sigma(f))\}d\mu \leq \eta_m, \quad \forall m \geq 1, \]

where \( \log^+ x = \max\{\log x, 0\} \).

We call \( F_m(\mu, \mathcal{V}) \)’s the folding functions of \( \mu \). The folding rate characterizes how folding functions vary when approaching \( \Sigma(f) \). In particular, we consider measures with uniform folding rate and refer to

\[ \mathcal{M}_{\bar{\eta}}^{fol}(\mathcal{V}) =: \{\nu \in \mathcal{M}_{inv}(f) : F_m(\nu, \mathcal{V}) \leq \eta_m, \forall m \geq 1\} \]

as the set of all measures with folding rate \( \bar{\eta} \) relative to \( \mathcal{V} \).

In the consideration of upper semi-continuity of metric entropy, together with the harmonic property of metric entropy, it suffices to analyze invariant measures with all ergodic components \( \eta \) satisfying

\[ \int_X \log |f'(x)|d\nu < +\infty. \]

For otherwise, by the Ruelle inequality and upper semi-continuity of Lyapunov exponent, the metric entropy is obviously upper semi-continuous at all ergodic measures with the left hand side of (3) being infinity. Obviously, (3) implies \( \nu(\Sigma(f)) = 0 \). We note that \( \mathcal{M}_{\bar{\eta}}^{fol}(\mathcal{V}) \) has closed intersection with any closed set of invariant measures \( \mu \) with \( \mu(\Sigma(f)) = 0 \). In fact, let \( \mu_i \in \mathcal{M}_{\bar{\eta}_i}^{fol}(\mathcal{V}) \) such that \( \mu_i \to \mu \) in the weak* topology. Then given any small neighborhood \( U \) of \( \Sigma(f) \) contained in \( V_m \), we have

\[ \int_{V_m} \log^+ \#\{f^{-1}(f) \cap (V_m \setminus U)\}d\mu \leq \liminf_{i \to +\infty} \int_{V_m} \log^+ \#\{f^{-1}(f) \cap (V_m \setminus U)\}d\mu_i \leq \eta_m. \]

By the arbitrariness of \( U \), we have \( F_m(\mu, \mathcal{V}) \leq \eta_m \), i.e., \( \mu \in \mathcal{M}_{\bar{\eta}}^{fol}(\mathcal{V}) \).

For two sequences of functions \( \mathcal{F}^{(1)} =: \{F_m^{(1)}(\cdot)\}_{m \geq 1} \) and \( \mathcal{F}^{(2)} =: \{F_m^{(2)}(\cdot)\}_{m \geq 1} \), we say that \( \mathcal{F}^{(2)} \) uniformly dominates \( \mathcal{F}^{(1)} \), denoted as \( \mathcal{F}^{(1)} \leq \mathcal{F}^{(2)} \), if for any \( \gamma > 0 \) and any index \( m_2 \), we have \( F_m^{(1)}(\cdot) \leq F_m^{(2)}(\cdot) + \gamma \) for all \( m_1 \) sufficiently large. We say that \( \mathcal{F}^{(1)} \) and \( \mathcal{F}^{(2)} \) are uniformly equivalent if both \( \mathcal{F}^{(1)} \leq \mathcal{F}^{(2)} \) and \( \mathcal{F}^{(2)} \leq \mathcal{F}^{(1)} \) hold.

Note that for any two sequences of neighborhoods of \( \Sigma(f) \), \( \mathcal{V}^{(i)} \in \mathcal{O}_f, i = 1, 2, \mathcal{F}(\cdot, \mathcal{V}^{(1)}) =: \{F_m(\cdot, \mathcal{V}^{(1)})\}_{m \geq 1} \) and \( \mathcal{F}(\cdot, \mathcal{V}^{(2)}) =: \{F_m(\cdot, \mathcal{V}^{(2)})\}_{m \geq 1} \) are uniformly equivalent on \( \mathcal{M}_{inv}(f) \). Thus for any sequence \( \mathcal{V} \in \mathcal{O}_f \), through the variations of
\(\tilde{\eta} = \{\eta_m\}_{m=1}^{\infty} \in C\), we can exhaust all settings of uniform folding rates. So from now on, we fix an sequence \(V = \{V_m\}_{m \geq 1} \in O_f\) and omit the notation \(V\) for simplicity.

Now we establish the upper semi-continuity of metric entropy for \(C^r (r > 1)\) interval (or circle) maps from the perspective of folding rate.

**Theorem 1.2.** Let \(f\) be a \(C^r (r > 1)\) map on an interval (or a circle) \(X\). Then for any \(\tilde{\eta} \in C\), the metric entropy is upper semi-continuous on \(M^\text{fol}_{\tilde{\eta}}(f)\).

**Remark.** In the setting of \(C^{1+\alpha}\) non-uniformly hyperbolic systems, Newhouse [10] introduced the notion of hyperbolic rate and prove that the metric entropy is upper semi-continuous on the set of measures with uniform hyperbolic rate. The (uniform) hyperbolic rate has a similar flavour as the (uniform) hyperbolic rate in the sense that both of them control the process for creating defect of upper semi-continuity: the former avoids the possibility of large entropy in the accumulation of small horseshoes near degenerate set, while the latter achieves this by restricting the non-uniform hyperbolic behaviors.

The notion of folding rate is natural from the perspective of folding entropy formula in Theorem [10], but it may happen that \(f\) does not admit any folding rate at some measure \(\mu\) since he sequence \(F_m(\mu)\) does not necessarily converge to zero and even is not finite. Besides, while the number of pre-images is controlled by the reciprocal of \(|f'|\) near \(\Sigma(f)\), we do not exactly know whether the integrability \(\int_X \log|f'(x)|d\mu| < +\infty\) implies \(\mu \in M^\text{fol}_{\tilde{\eta}}(f)\) for some \(\tilde{\eta} \in C\). In other words, there can be some measures not falling into the consideration of folding rate. To remedy this deficiency, we develop the second mechanism directly from the integrability of \(\log|f'|\), rather than the logarithm of pre-image numbers.

### 1.2. Degenerate rate and u.s.c. of metric entropy.

For \(V = \{V_m\}_{m \geq 1} \in O_f\), \(\mu \in M_{inv}(f)\) and \(\tilde{\eta} = \{\eta_m\}_{m=1}^{\infty} \in C\), we say that \(f\) admits degenerate rate \(\tilde{\eta}\) with respect to \(\{\mu, V\}\) if

\[
D_m(\mu, V) = \left| \int_{V_m} \log|f'(x)|d\mu \right| \leq \eta_m, \ \forall m \geq 1.
\]

We call \(D_m(\mu, V)\) the degenerate functions of \(f\) with respect to \(\mu\). The degenerate rate characterizes how the degenerate functions diminish when approaching \(\Sigma(f)\).

We consider measures with uniform degenerate rate and refer to

\[
M^\text{deg}_{\tilde{\eta}}(f, V) = \left\{ \nu \in M_{inv}(f) : D_m(\nu, V) \leq \eta_m, \ \forall m \geq 1 \right\}
\]

as the set of all measures with degenerate rate \(\tilde{\eta}\) relative to \(V\).

It is obvious that \(D_m(\mu, V)\) is finite for any \(m \geq 1\), as long as \(\int_X \log|f'(x)|\) is \(\mu\)-integrable. Moreover, it further holds that

\[
\{D_m(\mu, V)\}_{m=1}^{\infty} \in C.
\]

Thus, all invariant measures with \(\log|f'(x)|\) integrable are included in some set of uniform degenerate rate.

By definition, \(M^\text{deg}_{\tilde{\eta}}(f, V)\) is a closed subset of \(M_{inv}(f)\). In fact, let \(\{\mu_i\}_{i=1}^{\infty} \subset M^\text{deg}_{\tilde{\eta}}(f, V)\) be a sequence satisfying \(\lim_{i \to \infty} \mu_i = \mu\). For one thing, for any \(t > 1\), there exists small neighborhood \(V_m\) such that \(|f'(x)| > t, \ \forall x \in V_m\), so \(\mu_i(V_m) < \eta_m/\log t, \ \forall i \in \mathbb{N}\), which implies \(\mu(V_m) \leq \liminf_i \mu_i(V_m) \leq \eta_m/\log t \leq 1/\log t\).
By the arbitrariness of $t$, one has $\mu(\Sigma_f) = 0$. For another, for any $s > m$, 
\[ \left| \int_{V_m \setminus V_s} \log |f'(x)|d\mu \right| \leq \eta_m, \]
so
\[ \lim_{i \to +\infty} \left| \int_{V_m \setminus V_s} \log |f'(x)|d\mu \right| \leq \eta_m, \]
which implies $\left| \int_{V_m \setminus \Sigma_f} \log |f'(x)|d\mu \right| \leq \eta_m$. Putting together, one has 
\[ \left| \int_{V_m} \log |f'(x)|d\mu \right| \leq \eta_m, \quad \forall m \in \mathbb{N}. \]

Similar to the folding rate, for the fixed $\mathcal{V} = \{V_m\}_{m \geq 1} \in \mathcal{O}_f$, we can exhaust all settings of uniform degenerate rates through the variation of $\hat{\eta} \in \mathcal{C}$, and we omit the notation $\mathcal{V}$ if no confusion occurs.

The following theorem establishes the upper semi-continuity of metric entropy when measures with uniform degenerate rate are considered.

**Theorem 1.3.** Let $f$ be a $C^r (r > 1)$ map on an interval (or a circle) $X$. Then for any sequence $\hat{\eta} \in \mathcal{C}$, the metric entropy is upper semi-continuous at all ergodic measures in $\mathcal{M}_{\text{deg} \hat{\eta}}(f)$.

1.3. **Sharpness of sufficient criteria.** By constructing interval maps for which the upper semi-continuity of metric entropy fails, we study the sharpness of the conditions of uniform folding and degenerate rate in Theorem 1.2 and Theorem 1.3, respectively. In current known examples ([8], [3], [15]), the entropy of non upper semi-continuity points of metric entropy are all equal to zero. Therefore, a further question is:

**Question** ([2]) Is the metric entropy of a $C^r (r > 1)$ interval map upper semi-continuous at ergodic measures with positive entropy?

The positivity of metric entropy of an ergodic measure implies the hyperbolicity of it in one-dimensional case by the Ruelle inequality. We give an answer to the above question by constructing a new example in which the defect of upper semi-continuity at an ergodic measure with positive entropy follows from a homoclinic tangency whose image is a generic point of that measure. It turns out that there exists a sequence of measures approximating the ergodic measure which do not admit uniform folding and degenerate rate. This shows that the condition of uniform folding and degenerate rate in Theorem 1.2 and Theorem 1.3 are crucial and cannot be removed.

**Theorem 1.4.** For any $1 < r < \infty$, there exist $C^r$ interval maps which admit ergodic measures with positive metric entropy as non upper semi-continuity points of metric entropy.

The paper is organized as follows. In section 2, we recall some basic facts of entropy theory and prove the entropy formula in Theorem 1.1. Theorem 1.2 and Theorem 1.3 are proved in section 3 and section 4, respectively. In section 5, we prove Theorem 1.4 by constructing interval maps showing the sharpness of conditions of both uniform folding and degenerate rate for the upper semi-continuity of metric entropy.
2. **An Entropy Formula for Interval Maps**

In this section, we prove Theorem 1.1. Before going any further, we recall basic concepts and results of entropy theory that will be used here and the following sections.

Let $T$ be a measurable transformation on a Lebesgue space $(X, B, \mu)$. (Note that in our one-dimensional setting, for a Borel probability measure $\mu$, if $B_\mu(X)$ denotes the completion of Borel $\sigma$-algebra with respect to $\mu$, then $(X, B_\mu(X), \mu)$ constitutes a Lebesgue space.) A partition $\xi$ of $X$ is measurable if there exist measurable subsets $E_1, E_2, \cdots$ of $X$ such that

$$\xi = \{E_1, X \setminus E_1\} \cup \{E_2, X \setminus E_2\} \cup \cdots \mod 0,$$

where $\{A_{\beta_1} : \beta_1 \in J_1\} \cup \{B_{\beta_2} : \beta_2 \in J_2\} = \{A_{\beta_1} \cap B_{\beta_2} : \beta_1 \in J_1, \beta_2 \in J_2\}$. In other words, there exists a full $\mu$-measure subset $F_0 \subset X$ such that for any atom $P$ of $\xi$, we have

$$P = E_1^* \cap E_2^* \cap \cdots \cap F_0,$$

where $E_i^*$ is either $E_i$ or $X \setminus E_i$ for $i \geq 1$. For any measurable partition $\xi$ of $X$ and $n \in \mathbb{Z}^+$, denote $T^{-n}\xi = \{T^{-n}B : B \in \xi\}$ which is still a measurable partition of $X$. For $A \in B$, let

\begin{align*}
(4) & \quad \xi|_A = \{B \cap A : B \in \xi\} \cup \{X \setminus A\}, \\
(5) & \quad \xi \cap A = \{B \cap A : B \in \xi\}.
\end{align*}

Then $\xi|_A$ and $\xi \cap A$ are measurable partitions of $X$ and $A$, respectively.

For $x \in X$, define

$$p_\xi(x) = \xi(x),$$

where $\xi(x)$ denotes the element of $\xi$ containing $x$. The factor space $X/\xi$ is the measure space whose points are the elements of $\xi$ with measurable structure as follows: a subset $C$ of $X/\xi$ is measurable if and only if $p_\xi^{-1}(C) \in B$. An associated measure $\mu_\xi$ on $X/\xi$ is defined as $\mu_\xi(C) = \mu(p_\xi^{-1}(C))$

Recall that every measurable partition $\xi$ enjoys a unique system of conditional measures $\{\mu_C\}_{C \in \xi} (\mathbb{P})$: for any $g \in L^1(X, B, \mu)$,

$$\int_X g d\mu = \int_{X/\xi} \left( \int_C g_C d\mu_C \right) d\mu_\xi,$$

where $g_C(x) = g(x), x \in C$. For the measurable partition $\xi$ and the probability measure $\mu$, let $H_\mu(\xi) = \int \log \mu(\xi(x)) d\mu$ be the entropy of $\xi$ with respect to $\mu$. Furthermore, for any two measurable partitions $\xi$ and $\eta$,

\begin{align*}
(7) & \quad H_\mu(\xi|\eta) = \int_{X/\eta} H_{\mu_\eta}(\xi|_B) d\mu_\eta = \int_X -\log \mu_\eta(x)(\xi(x) \cap \eta(x)) d\mu,
\end{align*}

is the conditional entropy of $\xi$ with respect to $\eta$.

It is widely known that the metric entropy of $T$ with respect to $\mu$ satisfies

\begin{align*}
(8) & \quad h_\mu(T) = \sup_{\xi} \{ H_\mu(\xi|) \vee_{n \geq 1} T^{-n}\xi \},
\end{align*}

where the supremum is taken over all measurable partitions of $X$. 

\[ \text{\textcopyright} 2019 Springer Nature Switzerland AG. All rights reserved. \]
To derive the entropy formula, we need two inequalities of metric entropy for a $C^r (r > 1)$ interval (or circle) map $f$:

\[ h_\mu(f) \leq \int_X \max\{\lambda(f,x), 0\} d\mu(x) \tag{9} \]

\[ h_\mu(f) \leq H_\mu(\epsilon|f^{-1}\epsilon) - \int_X \min\{\lambda(f,x), 0\} d\mu(x) \tag{10} \]

where $\lambda(f,x)$ denotes the Lyapunov exponent of $f$ at $x$ which exists for $\mu$-$a.e.$, $x$ by Oseledets theorem [11]. The inequality (9) is the classical Ruelle (or Margulis-Ruelle) inequality [13]. The second inequality (10) is a special case of inequality (2) of [5] in the one-dimensional setting which was first suggested by Ruelle [14] when considering entropy production for non-invertible differentiable maps. It was proved by Liu [7] for $C^{1+\alpha}$ maps under some conditions on degenerate sets and was addressed for all $C^{1+\alpha}$ maps in [5]. In the present paper, for one-dimensional setting, we further deduce the equality as in Theorem 1.1.

**Proof of Theorem 1.1.** Since $\epsilon$ is the partition into single points which implies that $\bigvee_{n \geq 1} f^{-n}\epsilon = f^{-1}\epsilon$, from [3] it always holds that

\[ h_\mu(f) \geq H_\mu(\epsilon|f^{-1}\epsilon). \tag{11} \]

Thus we only need to show that

\[ h_\mu(f) \leq H_\mu(\epsilon|f^{-1}\epsilon). \]

Denote

\[ A_1 = \{x : \lambda(f,x) \leq 0\} \quad \text{and} \quad A_2 = \{x : \lambda(f,x) > 0\}. \]

Without loss of generality, we assume $A_1$ and $A_2$ are both positive $\mu$-measured. Otherwise, the situations would be easier. Let $\mu^{(i)} = \mu(A_i)$, $i = 1, 2$. Then $\mu = \sum_{i=1,2} \mu(A_i) \mu^{(i)}$.

For $\mu^{(1)}$, by (9), we have $h_{\mu^{(1)}}(f) = 0$. Combined with (11), one further gets

\[ H_{\mu^{(1)}}(\epsilon|f^{-1}\epsilon) = h_{\mu^{(1)}}(f) = 0. \]

For $\mu^{(2)}$, directly from (10), one gets

\[ h_{\mu^{(2)}}(f) \leq H_{\mu^{(2)}}(\epsilon|f^{-1}\epsilon). \]

Observing that $\mu^{(1)}(f^{-1}(A_2)) = \mu^{(2)}(f^{-1}(A_1)) = 0$, we have

\[
H_\mu(\epsilon|f^{-1}\epsilon) = -\int \log \mu_{f^{-1}\epsilon}(\epsilon(x)) d\mu(x) \\
= -\sum_{i=1,2} \mu^{(i)}(A_i) \int \log \mu_{f^{-1}\epsilon}(\epsilon(x)) d\mu^{(i)}(x) \\
= \sum_{i=1,2} \mu^{(i)}(A_i) H_{\mu^{(i)}}(\epsilon|f^{-1}\epsilon).
\]

Therefore,

\[ h_\mu(f) = \sum_{i=1,2} \mu^{(i)}(A_i) h_{\mu^{(i)}}(f) \leq \sum_{i=1,2} \mu^{(i)}(A_i) H_{\mu^{(i)}}(\epsilon|f^{-1}\epsilon) = H_\mu(\epsilon|f^{-1}\epsilon). \]

\[ \square \]
3. Upper semi-continuity and folding rate—Proof of Theorem 1.2

In this section, we prove Theorem 1.2. By Theorem 1.1, we reduce our proof to the following statement: the map

\[ \mu \mapsto H_\mu(\epsilon|f^{-1}\epsilon) \]

is upper semi-continuous on \( \mathcal{M}_{\text{fol}}^f \).

Throughout this section, we fix any \( \mu \in \mathcal{M}_{\text{fol}}^f \). Without loss of generality, suppose \( X = [0, 1] \). We start with an increasing sequence of partitions \( \{\Gamma_k\}_{k \in \mathbb{N}} \) of \( X \):

\[ \Gamma_k = \left\{ \left( \frac{q}{2^k}, \frac{q + 1}{2^k} \right) : 0 \leq q \leq 2^k - 1, q \in \mathbb{Z} \right\}. \]

By taking a translation, we assume \( \mu(\partial \Gamma_k) = 0, \forall k \in \mathbb{N} \). For certain subscript \( k_0 \in \mathbb{N} \) which will be determined later, we take a sequence of finite partitions \( \{\mathcal{P}_n\}_{n \geq 0} \) from \( \{\Gamma_k\}_{k \in \mathbb{N}} \) as follows:

\[ \mathcal{P}_0 = \Gamma_{k_0}, \quad \mathcal{P}_n = \Gamma_{k_0 + n}. \]

Obviously, \( \text{diam}(\mathcal{P}_n) \to 0 \) as \( n \to +\infty \).

We define the pre-image partition from \( \mathcal{P}_0 \) as follows: for each element \( P \in \mathcal{P}_0 \), denote by \( P^{-1,\epsilon} \) and \( \tilde{P}^{-1,\epsilon} \) the set of connected components of \( f^{-1}P \) and the set of connected components of \( f^{-1}P \) having nonempty intersection with \( M \setminus V_m \), respectively. Then

\[ \tilde{D} = \bigcup_{P \in \mathcal{P}_0} \tilde{P}^{-1,\epsilon} \cup \{ \tilde{V}_m \}, \]

is a finite measurable partition of \( X \), where \( \tilde{V}_m = M \setminus \bigcup_{P \in \mathcal{P}_0} \tilde{P}^{-1,\epsilon} \). Note that \( \tilde{D} \subset V_m \) is an open neighborhood of \( \Sigma(f) \) for \( k_0 \) large enough.

We argue that by taking \( k_0 \) large depending on \( m \), for any \( D \in \mathcal{D} \setminus \{ \tilde{V}_m \} \), \( f|_D \) is a diffeomorphism. For otherwise, there exists at least one point in \( D \) at which \( |f'| \) vanishes. This is impossible since \( D \cap \Sigma(f) = \emptyset \).

Thus, for each \( m \geq 1 \), we take \( k_0 = k_0(m) \) sufficiently large and hence obtain a finite partition \( \xi_m = \{ \tilde{V}_m, X \setminus \tilde{V}_m \} \). The following lemma says that one can approximate the folding entropy by refining the pre-image partition \( f^{-1}\epsilon \) through \( \xi_m \). Recall that the weak* topology on \( \mathcal{M}_{\text{inv}}(f) \) can be induced by a metric \( D \) ((14)), and we denote \( B_\rho(\mu) = \{ \nu \in \mathcal{M}_{\text{inv}}(f) : D(\nu, \mu) \leq \rho \} \).

**Lemma 3.1.** For any \( \epsilon > 0 \), there exists \( m_0 > 0 \) satisfying that for any \( m \geq m_0 \), there exist \( m' > m \) and \( \rho > 0 \) such that for any \( \nu \in B_\rho(\mu) \cap \mathcal{M}_{\text{inv}}(f) \), it holds that

\[ H_\nu(\epsilon|f^{-1}\epsilon) - H_\nu((\epsilon|X \setminus V_m')|f^{-1}\epsilon \vee \xi_m) \leq \frac{\epsilon}{2} \]

where \( \epsilon|X \setminus V_m' \) is defined as in (4).

**Proof.** We split

\[ H_\nu(\epsilon|f^{-1}\epsilon) - H_\nu((\epsilon|X \setminus V_m')|f^{-1}\epsilon \vee \xi_m) \]

into the sum of two parts:

\[ H_\nu(\epsilon|f^{-1}\epsilon) - H_\nu(\epsilon|f^{-1}\epsilon \vee \xi_m), \quad H_\nu(\epsilon|f^{-1}\epsilon \vee \xi_m) - H_\nu((\epsilon|X \setminus V_m')|f^{-1}\epsilon \vee \xi_m). \]

For the first one, observe that

\[ H_\nu(\epsilon|f^{-1}\epsilon) - H_\nu(\epsilon|f^{-1}\epsilon \vee \xi_m) = H_\nu(\xi_m|f^{-1}\epsilon) \leq H_\nu(\xi_m) = \psi_0(\nu(\tilde{V}_m)), \]
where \( \psi_0(t) = -t \log t - (1 - t) \log(1 - t) \) which is increasing near zero. Thus, choose \( m_0 \) large such that \( \mu(\tilde{V}_m) \) is small for all \( m \geq m_0 \) and, since \( \mu(\partial \tilde{V}_m) = 0 \), there exists \( \rho > 0 \) such that for any \( \nu \in B_\rho(\mu) \cap \mathcal{M}_{inv}(f) \),

\[
H_\nu(\epsilon|f^{-1}\epsilon) - H_\nu(\epsilon|f^{-1}\epsilon \lor \xi_m) < \frac{\varepsilon}{4}.
\]

For the second one, observe that for \( m' > m \),

\[
H_\nu(\epsilon|f^{-1}\epsilon \lor \xi_m) - H_\nu((\epsilon|X_{\tilde{V}_{m'}})|(f^{-1}\epsilon \lor \xi_m)) \\
= H_\nu(\epsilon | (\epsilon|X_{\tilde{V}_{m'}}) \lor (f^{-1}\epsilon \lor \xi_m)) \\
\leq H_\nu(\epsilon | (\epsilon|X_{\tilde{V}_{m}})).
\]

Since \( \epsilon|X_{\tilde{V}_{m'}} \nrightarrow \infty \) as \( m' \to \infty \), we have

\[
\lim_{m' \to +\infty} H_\nu((\epsilon|X_{\tilde{V}_{m'}})|(f^{-1}\epsilon \lor \xi_m)) = H_\nu(\epsilon|f^{-1}\epsilon \lor \xi_m),
\]

for which the limit is continuous at \( \mu \) since \( \mu(\partial \tilde{V}_{m'}) = 0 \).

\[ \Box \]

Now we fix an \( m \geq m_0 \) as in Lemma \( \text{(3.1)} \). According to \( \text{(13)} \) we only need to deal with \( H_\nu(\epsilon|X_{\tilde{V}_{m'}})|(f^{-1}\epsilon \lor \xi_m) \), which we split into two parts:

\[
H_\nu(\epsilon|X_{\tilde{V}_{m'}})|(f^{-1}\epsilon \lor \xi_m)
\]

\[
= \int_{X/(f^{-1}\epsilon \lor \xi_m)} H_{\nu_C}(\epsilon|X_{\tilde{V}_{m'}}) d\nu_{f^{-1}\epsilon \lor \xi_m}
\]

\[
= \int_{X/(f^{-1}\epsilon \lor \xi_m)} H_{\nu_C}(\epsilon|X_{\tilde{V}_{m'}}) d\nu_{f^{-1}\epsilon \lor \xi_m}
\]

\[
\int_{X/(f^{-1}\epsilon \lor \xi_m)} H_{\nu_C}(\epsilon|X_{\tilde{V}_{m'}}) d\nu_{f^{-1}\epsilon \lor \xi_m}
\]

\[
= : I_m^{(1)}(\nu) + I_m^{(2)}(\nu),
\]

where

\[
(14) \quad \left(X/(f^{-1}\epsilon \lor \xi_m)\right)_{\tilde{V}_{m'}} = \{C \in X/(f^{-1}\epsilon \lor \xi_m) : p_{f^{-1}\epsilon \lor \xi_m}(C) \subseteq X_{\tilde{V}_{m'}}\},
\]

\[
(15) \quad \left(X/(f^{-1}\epsilon \lor \xi_m)\right)_{\tilde{V}_{m'}} = \{C \in X/(f^{-1}\epsilon \lor \xi_m) : p_{f^{-1}\epsilon \lor \xi_m}(C) \subseteq \tilde{V}_{m}\},
\]

and recall that \( p_{f^{-1}\epsilon \lor \xi_m} \) is defined as in \( \text{(3)} \).

Intuitively, \( \text{(14)} \) and \( \text{(15)} \) correspond to the preimage set of any single point lying outside and inside of \( \tilde{V}_{m} \), respectively. Note that \( I_m^{(1)} \) deals with approximations of folding entropy away from \( \Sigma(f) \) and \( I_m^{(2)} \) concerns the dynamics around \( \Sigma(f) \).

In this way, to deal with folding entropy we only need to discuss \( I_m^{(1)} \) and \( I_m^{(2)} \) separately.
3.1. Upper semi-continguity of $I^{(1)}_{m,n}$. Note that
\[ H_{\nu^C}(\epsilon|X \setminus V_m) = H_{\nu^C}(\epsilon), \quad C \in (X/(f^{-1} \epsilon \lor \xi_m)|X \setminus \tilde{V}_m) . \]
Thus,
\[ I^{(1)}_{m,n} = \int_{(X/(f^{-1} \epsilon \lor \xi_m)|X \setminus \tilde{V}_m)} H_{\nu^C}(\epsilon) d\nu_{f^{-1} \epsilon \lor \xi_m} . \]
Let
\[ P^{-1}_n = f^{-1} P_n|X \setminus \tilde{V}_m \lor \tilde{D}, \]
where $\tilde{D}$ is defined in (12). Then $P^{-1}_n$ is the pre-image partition of $P_n$ away from $\Sigma(f)$. Obviously, $P^{-1}_n \subseteq f^{-1} P_n|X \setminus \tilde{V}_m, n \in \mathbb{N}$. Given any $n \in \mathbb{N}$ and $Q \in P^{-1}_n$, let $\tau_n(Q)$ be the element of $f^{-1} P_n|X \setminus \tilde{V}_m$ satisfying $Q \subseteq \tau_n(Q)$. Then all partitions \{P^{-1}_n\}_{n \geq 1} and \{f^{-1} P_n|X \setminus \tilde{V}_m\}_{n \geq 1} contain a common element $\tilde{V}_m$, which we denote as $Q_0$. One can see that $\tau_n(Q_0) = Q_0$ and
\[ H_{\nu}(P^{-1}_n|f^{-1} P_n|X \setminus \tilde{V}_m) = \sum_{Q \in P^{-1}_n, Q \neq Q_0} -\nu(Q) \log \left( \frac{\nu(Q)}{\nu(\tau_n(Q))} \right) =: I^{(1)}_{m,n}(\nu). \]

**Lemma 3.2.** $\lim_{n \to +\infty} I^{(1)}_{m,n}(\nu) = I^{(1)}_{m}(\nu), \quad \nu \in \mathcal{M}_{inv}(f)$.

**Proof.** Define $\phi(t) = -t \log t$ for $t \in (0, 1]$ and $\phi(0) = 0$. Then
\[ I^{(1)}_{m,n}(\nu) = \sum_{Q \in P^{-1}_n, Q \neq Q_0} \nu(\tau_n(Q)) \phi\left( \frac{\nu(Q)}{\nu(\tau_n(Q))} \right) \]
\[ = \sum_{D \in \tilde{D}, D \neq \tilde{V}_m} \sum_{P \in P_n, P \neq \tilde{V}_m} \nu(f^{-1} P \cap (X \setminus \tilde{V}_m)) \phi\left( \frac{\nu(f^{-1} P \cap D)}{\nu(f^{-1} P \cap (X \setminus \tilde{V}_m))} \right) \]
\[ = \sum_{D \in \tilde{D}, D \neq \tilde{V}_m} \int_{X \setminus \tilde{V}_m} \phi\left( \frac{\nu(f^{-1} P \cap D)}{\nu(f^{-1} P \cap (X \setminus \tilde{V}_m))} \right) d\nu(x) \]
\[ = \int_{X \setminus \tilde{V}_m} \sum_{D \in \tilde{D}, D \neq \tilde{V}_m} \phi\left( E(\chi_D|f^{-1} P_n|X \setminus \tilde{V}_m) \right) d\nu(x), \]
where $E_\nu(\cdot | \cdot)$ denotes conditional expectation.

Since $|f'|$ is uniformly bounded from below on $X \setminus \tilde{V}_m$, $f^{-1} P_n|X \setminus \tilde{V}_m$ increasingly converges to $f^{-1} \epsilon|X \setminus \tilde{V}_m$. Then by martingale convergence theorem, we have
\[ \lim_{n \to +\infty} I^{(1)}_{m,n}(\nu) = \int_{X \setminus \tilde{V}_m} E(\chi_D|f^{-1} \epsilon|X \setminus \tilde{V}_m) \right) d\nu(x) . \]
Note that for almost every $C \in (X/(f^{-1} \epsilon \lor \xi_m)|X \setminus \tilde{V}_m$, and any $x \in C$,
\[ \sum_{D \in \tilde{D}, D \neq \tilde{V}_m} \phi\left( E(\chi_D|f^{-1} \epsilon|X \setminus \tilde{V}_m\right)(x) = H_{\nu^C}(\epsilon) . \]
Thus,
\[
\lim_{n \to +\infty} I_{m,n}^{(1)}(\nu) = \int \left( X/(f^{-1}\epsilon \cup \xi_m) \right)_{X \setminus V_m} \int_D \sum_{D \in \mathcal{D}, D \neq V_m} \phi\left( E(\chi_D \mid f^{-1}\epsilon \setminus X \setminus V_m) \right) d\nu d\nu f^{-1}\epsilon \cup \xi_m
\]
\[= \int \left( X/(f^{-1}\epsilon \cup \xi_m) \right)_{X \setminus V_m} H_{\nu_C}(\epsilon) d\nu f^{-1}\epsilon \cup \xi_m
\]
\[= I_{m}^{(1)}(\nu).
\]
\[\square\]

As the arguments in [14] shows, for any fixed \(m\), \(I_{m,n}^{(1)}(\cdot)\) varies in a monotone way with respect to \(n\).

**Lemma 3.3.** Given any \(m \in \mathbb{N}\), \(I_{m,n}^{(1)}(\cdot)\) is non-increasing on \(\mathcal{M}_{inv}(f)\) as \(n\) increases.

**Proof.** Given any \(\tilde{Q} \in \mathcal{P}_{n+1}^{-1}\) such that \(\tilde{Q} \neq Q_0\),
\[\begin{align*}
    \sum_{Q \in \mathcal{P}_{n+1}^{-1}, Q \subseteq \tilde{Q}} -\nu(Q) \log \frac{\nu(Q)}{\nu(\tau_{n+1}(Q))} &= \nu(\tau_n(\tilde{Q})) \sum_{Q \in \mathcal{P}_{n+1}^{-1}, Q \subseteq \tilde{Q}} \frac{\nu(\tau_{n+1}(Q))}{\nu(\tau_n(\tilde{Q}))} \cdot \phi\left( \frac{\nu(Q)}{\nu(\tau_{n+1}(Q))} \right) \\
    \text{Since } f|_{\tilde{Q}} \text{ is a diffeomorphism, it is not hard to show that } \\
    \sum_{Q \in \mathcal{P}_{n+1}^{-1}, Q \subseteq \tilde{Q}} \nu(\tau_{n+1}(Q)) &= \nu(\tau_n(\tilde{Q})).
\end{align*}
\]

Thus by the concavity of \(\phi\),
\[\begin{align*}
    16 \leq \nu(\tau_n(\tilde{Q})) \phi\left( \sum_{Q \in \mathcal{P}_{n+1}^{-1}, Q \subseteq \tilde{Q}} \frac{\nu(\tau_{n+1}(Q))}{\nu(\tau_n(\tilde{Q}))} \cdot \frac{\nu(Q)}{\nu(\tau_{n+1}(Q))} \right) \\
    &= -\nu(\tilde{Q}) \log \frac{\nu(\tilde{Q})}{\nu(\tau_n(\tilde{Q}))}.
\end{align*}\]

Therefore,
\[
I_{m,n+1}^{(1)}(\nu) = \sum_{Q \in \mathcal{P}_{n+1}^{-1}, Q \neq Q_0} -\nu(Q) \log \frac{\nu(Q)}{\nu(\tau_{n+1}(Q))} \\
= \sum_{\tilde{Q} \in \mathcal{P}_n^{-1}, Q \in \mathcal{P}_{n+1}^{-1}, \tilde{Q} \neq Q_0} -\nu(Q) \log \frac{\nu(Q)}{\nu(\tau_{n+1}(Q))} \\
\leq \sum_{\tilde{Q} \in \mathcal{P}_n^{-1}, \tilde{Q} \neq Q_0} -\nu(\tilde{Q}) \log \frac{\nu(\tilde{Q})}{\nu(\tau_n(\tilde{Q}))} = I_{m,n}^{(1)}(\nu).
\]
\[\square\]
Combine Lemma 3.2 and Lemma 3.3 we get that
\[ I^{(1)}_m(\nu) = \lim_{n \to +\infty} I^{(1)}_{m,n}(\nu) = \inf_{n \geq 1} I^{(1)}_{m,n}(\nu), \quad \nu \in \mathcal{M}_{\text{inv}}(f). \]

Now we can obtain the upper semi-continuity property for \( I^{(1)}_m \).

**Proposition 3.4.** \( I^{(1)}_m \) is upper semi-continuous on \( \mathcal{M}_{\text{inv}}(f) \) at \( \mu \).

*Proof.* Since we get \( \{P_n\}_{n \geq 1} \) from \( \{\Gamma_k\}_{k \geq 1} \) with \( \mu(\partial\Gamma_k) = 0, k \geq 1 \), we have \( \mu(\partial P^{-1}_n) = 0, \forall n \geq 1 \). Hence for any \( Q \in P^{-1}_n, \nu \mapsto \nu(Q) \) is continuous at \( \mu \) in the weak* topology. Thus, for any \( n \in \mathbb{N} \), the finite sum
\[ I^{(1)}_{m,n} = \sum_{Q \in P^{-1}_n, Q \neq \tilde{Q}_0} -\nu(Q) \log \frac{\nu(Q)}{\nu(\tau_n(Q))} \]
is continuous at \( \mu \). It follows that \( I^{(1)}_m = \inf_{n \geq 1} I^{(1)}_{m,n} \) being the infimum of a family of functions continuous at \( \mu \) is upper semi-continuous at \( \mu \). \( \square \)

### 3.2. Estimates on \( I^{(2)}_m \)

As mentioned before, \( I^{(2)}_m \) concerns dynamics around \( \Sigma(f) \) where complicated phenomena such as infinite oscillations may happen, and we shall use folding rate to control it.

By concavity of the function \( \phi(t) = -t \log t \), for any \( C \in (X/(f^{-1}\varepsilon \lor \xi_m))_{\tilde{V}_m} \), we have
\[ H_{\forall\nu}(\epsilon|X\setminus V_m) \leq \log^+ \{f^{-1}(fx) \cap (\tilde{V}_m \setminus V_m') + 1\} \]
\[ \leq \log^+ \{f^{-1}(fx) \cap (\tilde{V}_m \setminus V_m')\} + 1, \quad \forall x \in p^{-1}(C), \]
which implies that
\[ I^{(2)}_m(\nu) \leq \int_{(X/(f^{-1}\varepsilon \lor \xi_m))_{\tilde{V}_m}} \int_{C} \log^+ \{f^{-1}(fx) \cap (\tilde{V}_m \setminus V_m') + 1\} \nu f^{-1}\varepsilon \lor \xi_m \]
\[ \leq \int_{\tilde{V}_m} \log^+ \{f^{-1}(fx) \cap (\tilde{V}_m \setminus V_m')\} \nu(x) + \nu(\tilde{V}_m) \]
\[ \leq \int_{V_m} \log^+ \{f^{-1}(fx) \cap (V_m \setminus \Sigma(f))\} \nu(x) + \nu(V_m). \]

Therefore, for all \( \nu \in \mathcal{M}_{\eta}^{\text{fold}}(f) \) with folding rate \( \tilde{\eta} = \{\eta_m\}_{m \geq 1} \), we have
\[ I^{(2)}_m(\nu) \leq \eta_m + \nu(V_m). \]

### 3.3. Proof of Theorem 1.2

Now we are in a position to prove Theorem 1.2. By Lemma 3.1 it suffices to show that for any \( \epsilon > 0 \) and any \( \nu \in B_{\rho}(\mu) \cap \mathcal{M}_{\tilde{\eta}}^{\text{fold}}(f) \),
\[ H_{\rho}(\epsilon|X\setminus V_m, f^{-1}\varepsilon \lor \xi_m) - H_{\mu}(\epsilon|X\setminus V_m, f^{-1}\varepsilon \lor \xi_m) < \frac{\epsilon}{2} \]
holds for \( m \) large enough. Since \( \lim_{m \to +\infty} \eta_m = 0 \), we can choose \( m \) large enough such that \( \eta_m < \frac{\epsilon}{8} \). Shrink \( \rho \) in Lemma 5.1 if necessary such that \( \nu(V_m) < \frac{\epsilon}{8} \) for any \( \nu \in B_{\rho}(\mu) \cap \mathcal{M}_{\text{inv}}(f) \). Then we have
\[ I^{(2)}_m(\nu) - I^{(2)}_m(\mu) \leq \eta_m + \nu(V_m) < \frac{\epsilon}{4}. \]
By Proposition 6.3 shrinking \( \rho \) further if necessary, we have
\[
I_{m}^{(1)}(\nu) - I_{m}^{(1)}(\mu) < \frac{\varepsilon}{4}, \quad \nu \in B_{\rho}(\mu) \cap \mathcal{M}_{inv}(f).
\]
Therefore,
\[
H_{\nu}(\epsilon|_{X\backslash V_{m}}| f^{-1}\epsilon \vee \xi_{m}) - H_{\mu}(\epsilon|_{X\backslash V_{m}}| f^{-1}\epsilon \vee \xi_{m})
= (I_{m}^{(1)}(\nu) - I_{m}^{(1)}(\mu)) + (I_{m}^{(2)}(\nu) - I_{m}^{(2)}(\mu)) < \frac{\varepsilon}{2},
\]
which concludes the proof of Theorem 1.2.

\[\square\]

4. Upper semi-continuity and degenerate rate—Proof of Theorem 1.3

In this section, for a \( C^{1+\alpha} \) interval (or circle) map \( f \), we investigate the relationship between metric entropy and degenerate rate by estimating the variation of entropy with respect to the scale. To achieve this, we will adopt some notations in \cite{5} where a sequence of partitions are introduced to estimate the complexity arising from the degeneracy surrounding \( \Sigma f \). We first briefly explain these notations and introduce necessary estimations from \cite{5}. More emphasis will be put on the different estimations in the present paper.

As in Section 3, we begin with partitions \( \Gamma \) for \( k \in \mathbb{N} \):
\[
\Gamma_{k} = \left\{ \left( \frac{q}{2k}, \frac{q+1}{2k} \right) : 0 \leq q \leq 2^{k} - 1, \ q \in \mathbb{Z} \right\}.
\]
Concerning the bound of derivatives, for any \( \varepsilon > 0 \), let
\[
U_{\varepsilon} = \left\{ x \in X : |f'(x)| < \varepsilon \right\}, \quad G_{\varepsilon} = \left\{ x \in X : |f'(x)| \geq \varepsilon \right\}.
\]
Without loss of generality, it suffices to consider \( V_{m} = U_{1/m}, \forall m \in \mathbb{N} \). By the Hölder continuity of \( f \), if take \( r_{\varepsilon} = (\frac{\varepsilon^{2}}{4K})^{\frac{1}{\alpha'}} \), then we have for any \( x \in G_{\varepsilon} \),
\[
|f'(y)| \geq \frac{\varepsilon}{2}, \quad \forall y \in B(x, r_{\varepsilon}),
\]
which yields
\[
f(B(x, r_{\varepsilon})) \supseteq \Gamma_{k}(f(x)), \quad \text{whenever } \frac{\varepsilon}{2}(\frac{\varepsilon^{2}}{4K})^{\frac{1}{\alpha'}} \geq \frac{1}{2k^{2}}.
\]
Denote \( \alpha' = \frac{\alpha}{2+\alpha} \). Take \( \varepsilon_{k} = \varepsilon_{0} 2^{-k\alpha'} \) with \( \varepsilon_{0} = (2(4K)^{\frac{1}{\alpha'}})^{\alpha'} \). Denote
\[
(f|_{B(x, r_{\varepsilon})})^{-1} = x + [(f|_{B(x, r_{\varepsilon})})^{-1}(f(x))] + R_{x}^{k},
\]
then
\[
|(R_{x}^{k})'(y)| \leq 1, \quad \forall y \in \Gamma_{k}(f(x)).
\]
New partitions \( \mathcal{P}_{k} \) which connect with the influence of degeneracy near \( U_{\varepsilon_{k}} \) are defined as follows:
\[
\mathcal{P}_{k}(x) = \begin{cases} \Gamma_{k}(x), & \text{if } \Gamma_{k}(x) \cap U_{\varepsilon_{k}} = \emptyset; \\ A_{k} := \bigcup_{P \in \Gamma_{k}, P \cap U_{\varepsilon_{k}} \neq \emptyset} P, & \text{otherwise}. \end{cases}
\]
Since \( \text{diam}(\mathcal{P}_{k}) \to 0 \) as \( k \to +\infty \), we have
\[
h_{\nu}(f) = \lim_{k \to \infty} h_{\nu}(f, \mathcal{P}_{k}), \quad \nu \in \mathcal{M}_{inv}(f).
\]
For each $P \in \mathcal{P}_k \setminus \{A_k\}$, if $P^{-1,c}$ denotes the set of connected components of the preimage $f^{-1}P$, then
\[ P^{-1,c}_k =: \{ B_k \} \bigcup_{P \in \mathcal{P}_k \setminus \{A_k\}} \{ A \in P^{-1,c}: A \cap U_{\varepsilon_k} = \emptyset \} \]
is the preimage partition associated with $\mathcal{P}_k$, where
\[ B_k = f^{-1}A_k \bigcup_{P \in \mathcal{P}_k \setminus \{A_k\}} \bigcup_{A \in P^{-1,c}} A. \]

Note that $h_\nu(f, \mathcal{P}_k) \leq H_\nu(\mathcal{P}_k \mid f^{-1}\mathcal{P}_k)$. As in [5], we split $H_\nu(\mathcal{P}_k \mid f^{-1}\mathcal{P}_k)$ into four terms:
\[ \Delta^{(1)}_k = H_\nu(P^{-1,c}_k f^{-1}\mathcal{P}_k), \]
\[ \Delta^{(2)}_k = \sum_{P \in \mathcal{P}_k \setminus \{A_k\}} -\nu(P \cap B_k) \log \frac{\nu(P \cap B_k)}{\nu(B_k)}, \]
\[ \Delta^{(3)}_k = \sum_{P \in \mathcal{P}_k \setminus \{A_k\}, Q \in P^{-1,c} f^{-1}\mathcal{P}_k \setminus \{B_k\}} -\nu(P \cap Q) \log \frac{\nu(P \cap Q)}{\nu(Q)}, \]
\[ \Delta^{(4)}_k = \sum_{Q \in P^{-1,c} f^{-1}\mathcal{P}_k \setminus \{B_k\}} -\nu(A_k \cap Q) \log \frac{\nu(A_k \cap Q)}{\nu(Q)}. \]

We note that $\Delta^{(2)}_k$ and $\Delta^{(4)}_k$ concern the complexity coming from the degeneracy of $f$ around $\Sigma(f)$, while $\Delta^{(1)}_k$ and $\Delta^{(3)}_k$ are approximations of the complexity from the backward process.

Let
\[ \psi(f, x) := \begin{cases} \max \left\{ \frac{1}{|f'(x)|}, 1 \right\}, & x \in X \setminus \Sigma(f), \\ +\infty, & \text{otherwise.} \end{cases} \]

Fix $k_0 \in \mathbb{N}$, and for $k \geq k_0 + 1$, we define
\[ W_k = \{ Q \in P^{-1,c}_k \setminus \{B_k\} : Q \subseteq Q' \text{ for some } Q' \in P^{-1,c}_{k_0} \setminus \{B_{k_0}\} \}. \]

Then $W_k$ is a finer partition of $P^{-1,c}_k$ when restricted away from the degenerate set $\Sigma(f)$.

In [5], for $\Delta^{(i)}_k$, $i = 1, 2, 3, 4$, we have the following estimations:
\[ \Delta^{(1)}_k \leq e^{-1} + 1 + \sum_{Q \in W_k} \nu(f(Q)) \phi \left( \frac{\nu(Q)}{\nu(f(Q))} \right), \]
\[ \Delta^{(2)}_k \leq C \int_{U_{C_2-k_0'}} \log |f'(x)| d\nu(x), \]
\[ \Delta^{(3)}_k \leq C + \int_X \log \psi(f, x) d\nu(x), \]
\[ \Delta^{(4)}_k \leq e^{-1}, \]
where $C$ is a constant independent of $f$ and $k$. By the condition on the uniform degenerate rate, for $k_0$ large, one gets
\[ \Delta^{(2)}_k \leq 1, \quad k \geq k_0, \]
for any \( v \in \mathcal{M}^{\text{deg}}(f) \). Let
\[
I_k(\nu, f) = \sum_{Q \in W_k} \nu(fQ) \phi \left( \frac{\nu(Q)}{\nu(fQ)} \right).
\]
Then \( I_k \) is the part of conditional entropy away from degeneracy.

**Lemma 4.1.** Given any \( k \in \mathbb{N} \), \( I_k(\nu, f) \) is non-increasing on \( \mathcal{M}_{\text{inv}}(f) \) as \( k \) increases. Consequently, \( \lim_{k \to +\infty} I_k(\nu, f) = \inf_{k \geq 1} I_k(\nu, f) \) for any \( \nu \in \mathcal{M}_{\text{inv}}(f) \).

**Proof.**
\[
I_{k+1}(\nu, f) = \sum_{Q' \in W_{k+1}} \nu(fQ') \phi \left( \frac{\nu(Q')}{\nu(fQ')} \right)
\]
\[
= \sum_{Q \in W_k} \nu(fQ) \sum_{Q' \in W_{k+1}, Q \subseteq Q'} \frac{\nu(fQ')}{\nu(fQ)} \phi \left( \frac{\nu(Q')}{\nu(fQ')} \right)
\]
\[
\leq \sum_{Q \in W_k} \nu(fQ) \phi \left( \sum_{Q' \in W_{k+1}, Q \subseteq Q'} \frac{\nu(Q')}{\nu(fQ)} \right)
\]
\[
= \sum_{Q \in W_k} \nu(fQ) \phi \left( \frac{\nu(Q)}{\nu(fQ)} \right) = I_k(\nu, f),
\]
where the inequality comes from the convexity of the function \( \phi \).

Let \( I(\nu, f) = \lim_{k \to +\infty} I_k(\nu, f) \). From (19)–(23), we get an upper bound for \( h_\nu(f) \) :
\[
(24) \quad h_\nu(f) \leq I(\nu, f) + \int_X \log \psi(f, x) d\nu(x) + \tilde{C},
\]
where the constant \( \tilde{C} \) is independent of \( \nu \) and \( k \).

Since for each \( k \), \( W_k \) only contains preimages away from degeneracy, by Theorem 1.1, we have
\[
I(\nu, f) \leq H_\nu(e|f^{-1}e) = h_\nu(f).
\]
Hence, (24) implies that
\[
(25) \quad |I(\nu, f) - h_\nu(f)| \leq \tilde{C} + \int_X \log \psi(f, x) d\nu(x).
\]
For \( x \in X \), let
\[
\lambda(x) = \lim_{n \to +\infty} \frac{1}{n} \log |(f^n)'(x)|
\]
whenever the limit exists. By Oseledets theorem [11], for any \( \mu \in \mathcal{M}_{\text{inv}}(f) \), \( \lambda(x) \) exists for \( \mu \)-a.e., \( x \in X \), which is a constant when \( \mu \) is furthermore ergodic.

Now we can consider the upper semi-continuity of metric entropy. Let \( \{ \mu_i \} \subset \mathcal{M}^{\text{deg}}(f) \) and \( \mu_i \) converge to some ergodic \( \mu \in \mathcal{M}^{\text{deg}}(f) \). Observe that the metric entropy is always upper semi-continuous at \( \mu \) if \( \mu \) admits non-positive Lyapunov exponent (by Ruelle inequality and the upper semi-continuity of Lyapunov exponent). In the following, we only need to deal with the positive case of Lyapunov exponent.
By [25], we shall reduce our proof into two parts: (i) the upper semi-continuity of the regular term $I(\cdot, f)$; (ii) the control of tail term $\int_X \log \psi(f, x) d\mu(x)$.

4.1. Upper semi-continuity of $I(\cdot, f)$. Similar to the Proposition 3.4 with the help of Lemma 4.1, the upper semi-continuity of $I(\cdot, f)$ on $\mathcal{M}_{inv}(f)$ is straightforward.

**Proposition 4.2.** $I(\cdot, f)$ is upper semi-continuous on $\mathcal{M}_{inv}(f)$.

**Proof.** Given any $\mu \in \mathcal{M}_{inv}(f)$, by taking a translation, we assume $\mu(\partial P_k) = 0$ for all $k \in \mathbb{N}$. Then by the $f$-invariance of $\mu$, we have $\mu(\partial Q) = 0$ for any $Q \in P^{-1,c}_{\mu}$. So for any $k \in \mathbb{N},$

$$I_k(\nu, \mu) = \sum_{Q \in W_k} \nu(f(Q)) \phi\left(\frac{\nu(Q)}{\nu(fQ)}\right)$$

is continuous at $\mu$. Thus by Lemma 4.1 $I(\cdot, f) = \inf_{n \geq 1} I_k(\cdot, f)$ is upper semi-continuous at $\mu$. \hfill \Box

4.2. Control of the tail term. We control the tail term $\int_X \log \psi(f, x) d\nu(x)$ by replacing $f$ with $f^N$ for $N$ sufficiently large.

**Lemma 4.3.** Let $\mu \in \mathcal{M}_{inv}(f)$ such that for $\mu$-a.e. $x \in X$, $\lambda(x) > 0$. Then

$$\lim_{N \to +\infty} \frac{1}{N} \int_X \log \psi(f^N, x) d\mu(x) = 0.$$

**Proof.** For any given $\delta > 0$ and $N \in \mathbb{N},$ we split $X$ (mod $\mu$) into three parts:

$$W_1 = \{x \in X : \frac{1}{N} \log |(f^N)'(x)|^{-1} \in (-\lambda(x) - \delta, -\lambda(x) + \delta)\},$$

$$W_2 = \{x \in W_1^c : \frac{1}{N} \log |(f^N)'(x)|^{-1} > 0\},$$

$$W_3 = \{x \in W_1^c : \frac{1}{N} \log |(f^N)'(x)|^{-1} < 0\}.$$

By the positivity of Lyapunov exponent, we have $\psi(f^N, x) = 1$ for $\mu$-a.e. $x \in W_1 \cup W_3$. Also, $\mu(W_2)$ is arbitrarily small for $N$ sufficiently large. Thus, by the integrability property [3],

$$\frac{1}{N} \int_{W_2} \log \psi(f^N, x) d\mu(x) = \frac{1}{N} \int_{W_2} \log |(f^N(x))'|^{-1} d\mu(x)$$

can be arbitrarily small for $N$ sufficiently large and the lemma follows. \hfill \Box

By Lemma 4.3 for any $\gamma > 0$, we take $N$ such that

$$\int_X \log \psi(f^N, x) d\mu(x) < N\gamma.$$

Note that $\Sigma_{fN} = \Sigma_f \cup f^{-1}(\Sigma_f) \cup \cdots \cup f^{-(N-1)}(\Sigma_f)$. Then

$$\tilde{V}_m := V_m \cup f^{-1}(V_m) \cup \cdots \cup f^{-(N-1)}(V_m)$$

is an open neighborhood of $\Sigma_{fN}$. Let

$$\tilde{\mathcal{V}} = \{\tilde{V}_m\}_{m \geq 1}.$$

Obviously, $\tilde{\mathcal{V}} \in \mathcal{O}_{fN}$. 
Lemma 4.4. For any $\nu \in \mathcal{M}^{deg}_\eta(f,V)$, $f^N$ admits degenerate rate $\{N^2\eta_m\}_{m=1}^\infty$ with respect to $(\nu, \mathcal{V})$.

Proof. Note that for any $\nu \in \mathcal{M}_{inv}(f)$, values $\nu(f^{-i}(V_m))$ are the same for all $i \in \mathbb{N}$. Thus for any $m \in \mathbb{N}$,

$$\left| \int_{\tilde{V}_m} \log |(f^N)'(x)|d\nu \right| = \left| \int_{\bigcup_{0 \leq i \leq N-1} f^{-i}(V_m)} \sum_{0 \leq i \leq N-1} \log |f'(x)|d\nu \right|$$

$$\leq \sum_{0 \leq i \leq N-1} \left| \int_{\bigcup_{0 \leq i \leq N-1} f^{-i}(V_m)} \log |f'(x)|d\nu \right|$$

$$\leq \sum_{0 \leq i \leq N-1} N \left| \int_{f^{-i}(V_m)} \log |f'(x)|d\nu \right|$$

$$= N^2 \left| \int_{V_m} \log |f'(x)|d\nu \right| \leq N^2\eta_m.$$  \(\square\)

Now for the sequence of measures $\{\mu_i\} \subset \mathcal{M}^{deg}_\eta(f)$ such that $\mu_i$ converge to some ergodic $\mu$ in $\mathcal{M}^{deg}_\eta(f)$, we could estimate the tail term $\int_X \log \psi(f^N, x)d\mu_i(x)$ for $i$ sufficiently large. For any $\gamma' > 0$ and any large $i$, we have

$$\int_X \log \psi(f^N, x)d\mu_i(x)$$

$$= \int_{X \setminus \tilde{V}_m} \log \psi(f^N, x)d\mu_i(x) + \int_{\tilde{V}_m} \log \psi(f^N, x)d\mu_i(x)$$

$$\leq \int_{X \setminus \tilde{V}_m} \log \psi(f^N, x)d\mu_i(x) + \gamma' + N^2\eta_m$$

$$\leq N\gamma + \gamma' + N^2\eta_m,$$

where the last inequality comes from (26) and (27).

4.3. Proof of Theorem 1.3. Applying (25) for $f^N$ and then divided by $N$, we have for $\nu = \mu_i$ or $\mu \in \mathcal{M}^{deg}_\eta(f, \mathcal{V})$,

$$\left| \frac{1}{N} I(\nu, f^N) - h_\nu(f) \right| \leq \frac{1}{N} (\gamma' + \bar{C}) + \gamma + N\eta_m.$$   

For $f^N$ with sufficiently large $N$, by Proposition 4.2 together with the arbitrariness of $\gamma, \gamma'$ and $m$, we get the upper semi-continuity of $h_\mu(f)$ on $\mathcal{M}^{deg}_\eta(f)$.

5. An example of interval maps with positive metric entropy

In this section, for each $r \in (1, +\infty)$, we construct a $C^r$ interval map $f$ which admits an ergodic measure with positive entropy at which the metric entropy fails to be upper semi-continuous.

Before describing the example, we recall an important fact about one-sided shift which will be used several times in our discussions. Given a one-sided full shift with topological entropy $h$, any real number in $[0, h]$ can be achieved by an ergodic
invariant measure. Actually, this can be deduced directly from the following basic fact.

**Proposition 5.1** (see Theorem 4.26 and its Remark in [16]). Consider the $k$-full shift $(\Sigma_k, \sigma)$. For any probability vector $(p_0, \cdots, p_{k-1})$, the Markov measure associated with the $(p_0, \cdots, p_{k-1})$-shift is ergodic with entropy equals $-\sum_{i=0}^{k-1} p_i \log p_i$. In particularly, the metric entropy of $(\frac{1}{k}, \cdots, \frac{1}{k})$-shift achieves the topological entropy of the full shift $(\Sigma_k, \sigma)$ which equals to $\log k$.

\[ \Lambda = \bigcap_{i=0}^{+\infty} f^{-i}(I_1 \cup I_2). \]

Then $(f, \Lambda)$ is a uniformly expanding set which conjugates to a one-sided full shift of two symbols, thus the topological entropy $h_{\text{top}}(f|_\Lambda) = \log 2$.

In the following, for any subinterval $J$ of $I$, by $|J|$ we mean the length of $J$. Any $y \in I$ stands for either a point in $I$ or a real number when $I$ is considered as a subinterval of $\mathbb{R}$.

Let $x_\star$ and $x'_\star$ be the left end point of $I_1$ and the right end point of $I_2$, respectively. We further assume that $x_\star$ is a fix point of $f$ and $f(x'_\star) = f(x_\star) = x_\star$. For simplicity,

**Figure 1.** accumulation of small horeshoes

5.1. **Description of the example.** In this subsection, we describe in detail how the interval map $f : I \to I$ is constructed for $I = [0, 1]$. It is qualitatively depicted in Figure 1.

In Figure 1, $I_1$ and $I_2$ are two subintervals of $I$ on which $f$ act linearly with slope $\lambda$ such that

\[ f(I_i) \supseteq I_1 \cup I_2, \quad i = 1, 2. \]

Let

\[ \Lambda = \bigcap_{i=0}^{+\infty} f^{-i}(I_1 \cup I_2). \]

Then $(f, \Lambda)$ is a uniformly expanding set which conjugates to a one-sided full shift of two symbols, thus the topological entropy $h_{\text{top}}(f|_\Lambda) = \log 2$.

In the following, for any subinterval $J$ of $I$, by $|J|$ we mean the length of $J$. Any $y \in I$ stands for either a point in $I$ or a real number when $I$ is considered as a subinterval of $\mathbb{R}$.

Let $x_\star$ and $x'_\star$ be the left end point of $I_1$ and the right end point of $I_2$, respectively. We further assume that $x_\star$ is a fix point of $f$ and $f(x'_\star) = f(x_\star) = x_\star$. For simplicity,
we set \(a = x_* = |I_1| = |I_2|\) and \(d_H(I_1, I_2) = 3a\). Let \(\lambda\) be large enough so that

\[I_1 \cup I_2 \subset [x_*, x_* + a\lambda / 2].\]

Since \((f, \Lambda)\) conjugates to a one-sided 2-full shift, by Proposition 5.1 for any \(c \in (0, \min\{\log 2, -\log \lambda\})\), we can find an ergodic measure \(\mu\) with \(h_\mu(f) = c\) and having \(\Lambda\) as its support. Take \(\delta_0 = a / 2\lambda\). Observing that \(x_* \in \Lambda\), we can choose \(x_0 \in \Lambda \cap [x_*, \delta_0 / 2]\) to be a generic point of \(\mu\) in the sense that

\[\frac{1}{n} \sum_{i=0}^{n-1} \delta_{f^i(x_0)} \to \mu, \quad \text{as} \quad n \to +\infty\]

in the weak* topology.

As denoted in Figure 4, we arrange \(z_0\) to be one preimage of \(x_0\) lying on the right side of \(x_*\) such that \([|x_*^0, z_0|] = a\). Since \(d_H(I_1, I_2) = 3a\), we suppose that in the \(\delta_0\)-neighborhood of \(I_1 \cup I_2\), \(f\) is also linear with slope \(\lambda\).

**Lemma 5.2.** Let \(\delta_1 = x_0 - x_*\). Then there exist \(N_1, N_2 \in \mathbb{N}\) such that

\[f^{N_1}([x_0 - \delta_1, x_0]) \supset [z_0 - a, z_0 + 3a], \quad f^{N_2}([x_0, x_0 + \delta_1]) \supset [z_0 - a, z_0 + 3a].\]

**Proof.** First, since \(x_*\) is a fixed point of \(f\), the expanding property of \(f\) on \(I_1\) gives some time \(N_1 \in \mathbb{N}\) satisfying

\[f^{N_1}([x_0 - \delta_1, x_0]) \supset f(I_1) \supset [z_0 - a, z_0 + 3a].\]

For the second, note that there exists \(\delta \in (0, \delta_1)\) such that

\[f^{N_1}([x_0 - \delta_1 + \delta, x_0]) \supset [z_0 - a, z_0 + 3a].\]

Recall that \(x_0\) is a generic point of \(\mu\). Since \(\mu([x_0 - \delta_1, x_0 - \delta_1 + \delta]) > 0\), there exists \(t \in \mathbb{N}\) such that

\[f^t(x_0) \in [x_0 - \delta_1, x_0 - \delta_1 + \delta].\]

Next we consider the \(f\)-iterations of \([x_0, x_0 + \delta_1]\) cut by \([x_* - \delta_0, x_* + \delta_0]\), successively. There exists \(t_0 > 0\) such that for each \(i > t_0\), \(f^i([x_0, x_0 + \delta_1])\) contains an interval \(R_i\) of length \(\delta_0\) with \(f^i(x_0)\) being one of its end points. In particular, we require \(t > t_0\). If \(f^t(x_0)\) is the left end point of \(R_t\), then

\[R_t \supset [x_0 - \delta_1 + \delta, x_0]\]

which implies

\[f^{t+N_1}([x_0, x_0 + \delta_1]) \supset f^{N_t}(R_t) \supset [z_0 - a, z_0 + 3a],\]

and so let \(N_2 = t + N_1\). If \(f^t(x_0)\) is the right end point of \(R_t\), then

\[R_t \supset [x_0 - \delta_1, f^t(x_0)].\]

Once more, using the expanding property of \(f\) on \(I_1\), there exists \(t' \in \mathbb{N}\), such that

\[f^{t'}([x_0 - \delta_1, f^t(x_0)]) \supset [z_0 - a, z_0 + 3a].\]

In this case, let \(N_2 = t + t'\). We finish the proof of the Lemma. \(\square\)
By Lemma [5.2] and the continuity of $f$, there exists $\eta > 0$ such that for any $x \in [x_0 - \eta, x_0 + \eta]$, it holds that
$$f^{N_1}(x - \delta_1, x) \supset [z_0, z_0 + 2a], \quad f^{N_2}(x, x + \delta_1) \supset [z_0, z_0 + 2a].$$
Since $x_0$ is a generic point and $\mu([x_0 - \eta, x_0 + \eta]) > 0$, there exist $n_1 < n_2 < \cdots < n_k < \cdots$ such that $f^{n_k}(x_0) \in [x_0 - \eta, x_0 + \eta]$.

Let $\{J_n\}_{n \geq 1}$ be a sequence of disjoint subintervals of $I$ accumulating to $z_0$ from the right with the following properties:

(i) On each interval $J_k$, $f|_{I_k} = A_k^c \cos \omega_k(x - c_k) + (x_0 + A_k^e)$, where $\{c_k\}_{k \geq 1}$ is a sequence of real numbers with $c_k \searrow z_0$, and
$$A_k = \left(\frac{\delta_0}{2} \lambda^{-n_k}\right)^{\frac{1}{k}}, \quad \omega_k = \frac{L}{A_k},$$
where $L \geq \lambda$ is chosen such that all of our constructions are restricted to satisfy
$$\sup_{x \in I} |f'(x)| \leq L.$$
(ii) On each interval $J_k$, $f$ oscillates $M_k$ times, i.e., the number of period is $M_k$, where
$$M_k = \frac{L \gamma_0}{2 \pi k^2 \left(\frac{2}{\delta_0} \lambda^{n_k}\right)^{\frac{1}{k}}}$$
and $\gamma_0$ is a sufficiently small real number.

From (i) we can see $x_0 \in f(J_k)$. From (i) and (ii), we have $|J_k| = 2 \pi M_k \omega_k = \gamma_0 k^2$. Thus, for a small $\gamma_0$, we can have
$$\sum_{k=1}^{+\infty} |J_k| = \gamma_0 \sum_{k=1}^{+\infty} \frac{1}{k^2} < a,$$
and make
$$\bigcup_{k \geq 1} J_k \subset [z_0, z_0 + 2a].$$

Outside intervals $\{I_i\}_{i=1,2}$ and $\{J_k\}_{k \geq 1}$, we extend $f$ in a smooth way as depicted in Figure [1].

5.2. Properties of interval map $f$. In this subsection, we shall show that $\mu$ is approximated by any ergodic measures associated to a sequence of horseshoes with topological entropy having uniform upper gap from the metric entropy of $\mu$, which implies that the metric entropy is not upper semi-continuous at $\mu$, hence gives an answer to Burguet’s question in [2] since the ergodic $\mu$ admits positive metric entropy.

Given an interval map $g$ and integers $\ell \geq 2$, by an $\ell$-horseshoe of $g$ we mean a family of disjoint closed intervals $(K_1, \cdots, K_\ell)$ such that
$$g(K_i) \supseteq K_j, \quad \forall i, j = 1, \cdots, \ell.$$
If the union $\bigcup_{1 \leq i \leq \ell} K_i$ is contained in an interval $K$, we say $g$ admits an $\ell$-horseshoe on $K$. A basic fact about an $\ell$-horseshoe is that it conjugates to the one-sided shift of $\ell$ symbols.

Lemma 5.3. For each $k \geq 1$, $f^{n_k+N_i+1}(J_k) \supseteq J_k$, or $f^{n_k+N_i+1}(J_k) \supseteq J_k$. 

Proof. For each $k \geq 1$, since $x_0 \in f(J_k)$, we have $f^i(x_0) \in f^{i+1}(J_n)$ for any $i \geq 1$. Note that for $0 \leq i \leq n_k$,

$$|f^{i+1}(J_k)| = \lambda^i |f(J_k)| = \lambda^i \cdot 2A_k^r \leq \delta_0$$

and $|f^{n_k+1}(J_k)| = \delta_0$. Moreover, $f^{n_k}(x_0) \in [x_0 - \eta, x_0 + \eta]$, $\delta_1 < \delta_0$. By the statements below Lemma 5.2 we have

$$f^{N_1}(f^{n_k+1}(J_k)) \supset J_k \quad \text{or} \quad f^{N_2}(f^{n_k+1}(J_k)) \supset J_k.$$

□

By Lemma 5.3, we suppose $f^{n_k+N_1+1}(J_k) \supseteq J_k$ since the other case is similar. Then $f^{n_k+N_1+1}$ admits a $2M_k$-horseshoe on $J_k$. Denote

$$\Lambda_k = n_k+N_1 \bigcup_{j=0}^{j=t} f^j \left( \bigcap_{i \geq 0} f^{-(n_k+N_1+1)}(J_k) \right).$$

By Proposition 5.1, for each $k$, there exists an ergodic measures $\nu_k$ supported on $\Lambda_k$ such that

$$h_{\nu_k}(f) = h_{\text{top}}(f|\Lambda_k) = \frac{\log(2M_k)}{n_k+N_1+1}.$$

It is verified that

$$h_{\nu_k}(f) \to \frac{\log \lambda}{r}, \quad \text{as} \quad k \to +\infty. \quad (28)$$

We next show that $\nu_k \to \mu$, as $k \to +\infty$. Given any continuous functions $\varphi_1, \cdots, \varphi_s$ on $X$, for any $\varepsilon > 0$, there exists $\gamma > 0$ such that for any two points $x, y \in I$ such that $d(x, y) < \gamma$,

$$|\varphi_i(x) - \varphi_i(y)| < \varepsilon, \quad \forall 1 \leq i \leq s.$$ 

Let

$$t_k = \frac{\ln(\frac{\lambda}{2A_k^r})}{\ln \lambda}.$$ 

Then

$$|f^{j+1}(J_k)| \leq \gamma, \quad \forall 0 \leq j \leq t_k,$$

which implies

$$\lim_{k \to +\infty} \left| \int \varphi_i(x) d\nu_k(x) - \frac{\sum_{0 \leq j \leq n_k-1} \varphi_i(f^j(x_0))}{n_k} \right| \leq \varepsilon.$$ 

Since $(1/n_k) \sum_{0 \leq j \leq n_k-1} \delta f(x_0) \to \mu$ as $k \to +\infty$, the arbitrariness of $\varphi_1, \cdots, \varphi_s$ and $\varepsilon$ give rise the convergence of $\nu_k$ to $\mu$.

Combine (5.1) and the fact that $h_\mu(f) < \frac{\log \lambda}{r}$, we see that the metric entropy is not upper semi-continuous at $\mu$.

5.3. The folding and degenerate rate conditions is sharp. Now we show that the sequence of measures $\nu_k$ do not admit uniform folding and degenerate rate, which implies that the sufficient condition of Theorem 1.2 and Theorem 1.3 is sharp.
5.3.1. Calculation of folding rate. Recall that $\Sigma_f = \{ x \in I : f'(x) = 0 \}$. Note that in our example, $z_0 \in \Sigma_f$. Fix a sequence of neighborhoods $\mathcal{V} = \{ V_m \}_{m \geq 1}$ of $\Sigma_f$ with $d_H(V_m, \Sigma_f) \to 0$. For any given $m \geq 1$, we have $J_k \subseteq V_m$ for all $k$ sufficiently large.

Choose $y_k \in J_k$ be a generic point of $\nu_k$, i.e., $\frac{1}{t} \sum_{i=0}^{t-1} \delta_{f^i(y_k)} \to \nu_k$, $t \to +\infty$. From the constructions in Section 5.2 for $m$ large enough, the set of times at which the orbit of $y_k$ lying in $V_m$ are $\{(n_k + N_1 + 1)\ell : \ell \in \mathbb{Z}^+\}$ or $\{(n_k + N_2 + 1)\ell : \ell \in \mathbb{Z}^+\}$. We suppose the first case. The number of preimages of $f((n_k + N_1 + 1)\ell(y_k))$ lying in $V_m$ is at least $2M_k$ since there are $M_k$ oscillations at each interval $J_k$. Therefore,

$$F_m(\nu_k) = \int_{V_m} \log^+ \sharp \{ f^{-1}(x) \cap (V_m \setminus \Sigma(f)) \} d\nu_k \geq \frac{1}{n_k + N_1 + 1} \log(2M_k) \to \frac{\log \lambda}{r}, \text{ as } k \to +\infty.$$ 

So for any fixed $m$, when $k$ large, $F_m(\nu_k)$ are uniformly away from zero, and thus $\{\nu_k\}_{k \geq 1}$ does not have uniform folding rate. This implies the condition of folding rate in Theorem 1.2 cannot be removed.

5.3.2. Calculation of degenerate rate. The non-uniformity of degenerate rate is similar. Note that

$$|f'(x)| \leq A_k^r \omega_k, \forall x \in J_k.$$ 

Then for fixed $V_m$ and all sufficiently large $k$,

$$\int_{V_m} \log |f'(x)| d\nu_k \leq \frac{1}{n_k + N_1 + 1} \log(A_k^r \omega_k) = \frac{1}{n_k + N_1 + 1} \log\left( \frac{\delta_0}{2} \lambda^{-n_k} r^{-1} \right) = \frac{r-1}{r} n_k \log \lambda^{-1}.$$ 

which goes to $\frac{r-1}{r} \log \lambda^{-1}$ when $k \to +\infty$.

Thus,

$$D_m(\nu_k) = \left| \int_{V_m} \log |f'(x)| d\nu_k \right| \geq \frac{r-1}{r} \log \lambda.$$ 

That is, $\{\nu_k\}_{k \geq 1}$ does not have uniform degenerate rate. This implies that the condition of uniform degenerate rate in Theorem 1.3 is sharp.

REFERENCES

[1] R. Bowen, Entropy expansive maps, Trans. Amer. Math. Soc., 164, 323-331, 1972
[2] D. Burguet, Existence of measures of maximal entropy for $C^r$ interval maps, Proc. Amer. Math. Soc., 142, 957-968, 2014.
[3] J. Buzzi, Intrinsic ergodicity of smooth interval maps, Israel J. Math., 100, 125-161, 1997.
[4] G. Liao, W. Sun and S. Wang, Upper semi-continuity of entropy map for nonuniformly hyperbolic systems, Nonlinearity, 28, 2977-2992, 2015.
[5] G. Liao and S. Wang, Ruelle inequality of folding type for $C^{1+\alpha}$ maps, Math. Z., 290, 509-519, 2018.
[6] G. Liao, M. Viana and J. Yang, The entropy conjecture for diffeomorphisms away from tangencies, J. Eur. Math. Soc., 15, 2043-2060, 2013.
[7] P. Liu, Ruelle inequality relating entropy, folding entropy and negative Lyapunov exponents, Commun. Math. Phys., 240, 531-538, 2003.
[8] M. Misiurewicz, Diffeomorphism without any measure with maximal entropy, Bull. Acad. Polon. Sci., 21, 903-910, 1973.
[9] M. Misiurewicz and W. Szlenk, Entropy of piecewise monotone mappings, Studia Math., 67, 45-63, 1980.
[10] S. Newhouse, Continuity properties of entropy, Ann. Math., 129, 215-235, 1989.
[11] V. I. Oseledets, A multiplicative ergodic theorem, Trans. Moscow Math. Soc., 19, 197-231, 1968.
[12] D. Ruelle, Lectures on the entropy theory of measure-preserving transformations, Russian Mathematical Surveys, 22, 1-54, 1967.
[13] D. Ruelle, An inequality for the entropy of differentiable maps, Bol. Soc. Brus. Mat., 9, 83-88, 1978.
[14] D. Ruelle, Positivity of entropy production in nonequilibrium statistical mechanics, J. Stat. Phys., 85, 1-23, 1996.
[15] S. Ruette, Mixing C’ maps of the interval without maximal measure, Israel J. Math., 127, 253-277, 2002.
[16] P. Walters, An introduction to ergodic theory, Springer Verlag, 1982.
[17] Y. Yomdin, Volume growth and entropy, Israel J. Math., 57, 285-300, 1987.

E-mail address: lg@suda.edu.cn
E-mail address: shirou@ualberta.ca