ISRAEL JOURNAL OF MATHEMATICS **TBD** (2024), 1–42
DOI: 10.1007/s11856-024-2653-6

CONTINUITY PROPERTIES OF FOLDING ENTROPY

BY

GANG LIAO*

School of Mathematical Sciences
Center for Dynamical Systems and Differential Equations
Soochow University, Suzhou 215006, China
e-mail: lg@suda.edu.cn

AND

SHIROU WANG**

School of Mathematics, Jilin University, Changchun 130012, China
and
Department of Mathematical and Statistical Sciences, University of Alberta
Edmonton, AB T6G 2G1, Canada
e-mail: shirou@jlu.edu.cn

* G. Liao was partially supported by the National Key R&D Program of China (2022YFA1005802), NSFC(11701402, 11790274, 12071328, 12122113), BK20170327, BK20211595, Double Innovation Plan of Jiangsu Province of China and Tang Scholar.

** S. Wang was supported by NSFC(12201244, 12271204), and a faculty development grant from Jilin University.

Received June 13, 2021 and in revised form December 11, 2022

ABSTRACT

The folding entropy is a quantity originally proposed by Ruelle in 1996 during the study of entropy production in the non-equilibrium statistical mechanics [53]. As derived through a limiting process to the non-equilibrium steady state, the continuity of entropy production plays a key role in its physical interpretations. In this paper, the continuity of folding entropy is studied for a general (non-invertible) differentiable dynamical system with degeneracy. By introducing a notion called degenerate rate, it is proved that on any subset of measures with uniform degenerate rate, the folding entropy, and hence the entropy production, is upper semi-continuous. This extends the upper semi-continuity result in [53] from endomorphisms to all $C^r (r > 1)$ maps.

We further apply our result in the one-dimensional setting. In achieving this, an equality between the folding entropy and (Kolmogorov–Sinai) metric entropy, as well as a general dimension formula are established. The upper semi-continuity of metric entropy and dimension are then valid when measures with uniform degenerate rate are considered. Moreover, the sharpness of the uniform degenerate rate condition is shown by examples of C^r interval maps with positive metric (and folding) entropy.

1. Introduction

In the study of non-equilibrium statistical mechanics, the statistical mechanical entropy is exhibited to be persistently pumped out of systems during time evolutions due to the energy or heat exchange with the environment. The numerical experiments in practice indicate that the entropy production is non-negative and usually positive in accordance with the second law of thermodynamics [16, 17], and such phenomenon in the mathematical representations as well has been effectively discussed and justified to a certain extent in the stochastic process theory [25, 27, 50, 56, 62].

As a fundamental approach to interpret the thermodynamics, the entropy production theory was especially developed in the framework of dynamical systems [20, 21, 22, 23, 53, 54]. In [53], Ruelle investigated the entropy production for the standard dynamical system (M, f, μ) where M is a compact Riemannian manifold, the evolution $f : M \to M$ is either a diffeomorphism or a non-invertible differentiable map, and μ is the non-equilibrium steady state which is normally thought as an SRB (Sinai–Ruelle–Bowen) measure [23, 54]. The entropy production with respect to μ, denoted as $e_f(\mu)$, was illustrated through a limiting process in which a quantity called folding entropy naturally arises. More specifically, adopting the Shannon-type expression $S(\bar{\rho}) = -\int \rho(x) \log \rho(x) dx$ as the statistical mechanical entropy of a probability measure $\bar{\rho}$ with density ρ, the entropy pumped out of the system (to keep the energy fixed due to the non-conservative forces acting on the system) within one-time step is

$$e_f(\bar{\rho}) = -[S(f\bar{\rho}) - S(\bar{\rho})] := F_f(\bar{\rho}) - \int_M \log|\operatorname{Jac}(D_x f)|\rho(x)dx,$$

where $f\bar{\rho}$ denotes the image of $\bar{\rho}$ under f. The emerging term $F_f(\bar{\rho})$, called the folding entropy of $\bar{\rho}$, exactly expresses the complexities when f "folds" the different states according to their ρ-weights (or -masses). By assuming a general state μ as the limit of a sequence of absolutely continuous measures with probability densities $\{\rho_n\}_{n\geq 1}$, the entropy production with respect to μ is given by

$$e_f(\mu) = F_f(\mu) - \int_M \log|\operatorname{Jac}(D_x f)|d\mu.$$

Physically considering μ as an idealization of $\bar{\rho}_n$ when n is large enough, a natural question is what is the relationship between $e_f(\mu)$ and the limiting quantity (if it exists) $\lim_{n\to\infty} e_f(\bar{\rho}_n)$? Under some non-degenerate assumptions which essentially requires f to be of endomorphism-type, Ruelle [53] showed that the folding entropy is upper semi-continuous and hence[1]

$$e_f(\mu) \geq \limsup_{n\to\infty} e_f(\bar{\rho}_n). \tag{1.1}$$

We remark that when f is a diffeomorphism, the folding entropy is always zero since there is no space folding any more, and the entropy production is simply characterized by the phase volume contraction

$$-\int \log|\operatorname{Jac}(D_x f)|d\mu.$$

This relation was earlier pointed out in the discussion of more concrete non-equilibrium molecular dynamics models [15, 16, 24] and was theoretically revealed in [1, 21]. In this situation, the equality in the limiting process (1.1) naturally holds due to the continuity of the function $\log|\operatorname{Jac}(D_x f)|$.

Further progress depends on the study in the ergodic theory of differentiable dynamical systems. For a general (non-invertible) system beyond endomorphism, the difficulty in the analysis of entropy production is in the handling

[1] Under the endomorphism-type assumption in [53], $|\operatorname{Jac}(D_x f)|$ is uniformly away from 0. Hence, the term $\int \log|\operatorname{Jac}(D_x f)|d\mu$ is continuous with respect to μ.

of the possible accumulation of "foldings" due to the degeneracy (even with zero measure) in the approximation process. In this paper, f is assumed to be a $C^r(r > 1)$ map on a compact Riemannian manifold M. The main goal is to explore the mechanism of upper semi-continuity of folding entropy. By introducing a notion called degenerate rate which captures the complexity of f near the degenerate set $\Sigma_f := \{x \in M : \mathrm{Jac}(D_x f) = 0\}$, we shall establish the upper semi-continuity of folding entropy on any subset of measures with uniform degenerate rate. This then justifies (1.1) in the limiting process.

1.1. Folding entropy and degenerate rate. We begin with the precise definition of folding entropy. Let $\mathcal{P}(M)$ denote the set of all Borel probability measures on M. For $\mu \in \mathcal{P}(M)$, the folding entropy $F_f(\mu)$ is a conditional entropy measuring the complexities of preimages of f in terms of μ. More specifically, if we denote by $f^{-1}\epsilon = \{\{f^{-1}x\}\}_{x\in M}$ the preimage partition of f where $\epsilon := \{\{x\}\}_{x\in M}$ is the partition into single points, then the **folding entropy** of f with respect to μ is defined as the conditional entropy of ϵ relative to $f^{-1}\epsilon$ (see [53]),

$$F_f(\mu) = H_\mu(\epsilon | f^{-1}\,\epsilon) = \int_M H_{\tilde{\mu}_x}(\epsilon) d(f\mu),$$

where $\{\tilde{\mu}_x\}$ is a canonical family of conditional measures of μ disintegrated along the preimage sets $\{f^{-1}x\}$ for μ-a.e. $x \in M$; see [52] for more about the theory of conditional measure and conditional entropy.

Despite its seemingly abstract form, the folding entropy is in practice rather intuitive; see Figure 1(A). It is related to the number of preimages of f that if there are only finitely many preimage branches, say N, then the folding entropy is always bounded by $\log N$ and the equality is attained when the N preimages are of equal weight relative to μ. A simple but illuminating example is to consider $M = \mathbb{S}^1$ and $f : x \mapsto Nx \pmod 1$. Since f "evenly" expands the whole state space $\mathbb{S}^1$, the folding entropy with respect to the Lebesgue measure (which is an SRB measure) is $\log N$. Generalize a bit further for $N = 2$ by considering $f(x) = px$ for $x \in [0, 1/p)$ and $f(x) = \frac{p}{p-1}(1 - x)$ for $x \in [1/p, 1)$ where $1 < p \neq 2$. We see that the subintervals $[0, 1/p)$ and $[1/p, 1)$ are stretched by different scales. Still, the Lebesgue measure is an SRB measure, but the folding entropy now equals $-\frac{1}{p}\log\frac{1}{p} - \frac{p-1}{p}\log\frac{p-1}{p} < \log 2$. In [40], a relation between the folding entropy and the average logarithmic growth of the number of preimages was established for the equilibrium measure in the hyperbolic setting.

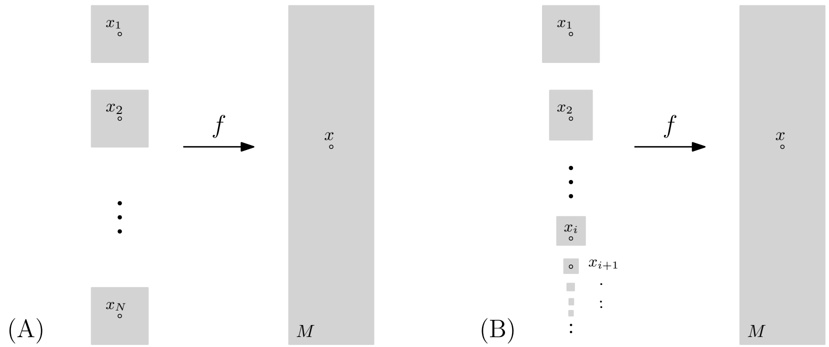

Figure 1. (A) There are N branches containing the N preimages, $x_1, \ldots, x_N$, of a typical $x \in M$, i.e., $f(x_1) = \cdots = f(x_N) = x$. (B) There are countably many branches containing the preimages, $x_1, \ldots, x_i, \ldots$, of a typical $x \in M$. The accumulation happens near the degenerate set.

As a consequence of the physically motivated and mathematically intuitive role played by the folding entropy, a folding-type entropy inequality on the backward process was proposed in [53] in contrast with the forward evolution. To be specific, for a $C^r(r>1)$ map f on a compact Riemannian manifold M and an f-invariant measure μ, the folding-type Ruelle inequality reads as

$$h_\mu(f) \leq F_f(\mu) - \int_M \sum_{\lambda_i(f,x)<0} \lambda_i(f,x) d\mu, \tag{1.2}$$

where the $\lambda_i(f,x)$'s denote the Lyapunov exponents which exist for μ-a.e. $x \in M$ by the Oseledets theorem [45]. The inequality (1.2) provides an approach to the non-negativity of entropy production [53] and was mathematically rigorously proved by Liu [33] for $C^r(r>1)$ maps under some polynomial-like degeneracy conditions and then by the authors [32] for all $C^r(r>1)$ maps. We would like to mention that in the classical Margulis–Ruelle inequality [52], the right-hand side is simply the sum of positive Lyapunov exponents. When it comes to the backward process however, since one trajectory may split into many branches, the term of folding entropy naturally arises and characterizes such "global" expansions of branch splittings, while the Lyapunov exponents term captures the "local" expansions inside each branch. Besides the entropy production in the forward process, Mihailescu and Urbański [39] studied the entropy production of the inverse SRB measures (Mihailescu [37]) which are the equilibrium states of the stable potential characterizing the backward equilibrium behaviors, and showed that the sign is nonpositive for a class of hyperbolic folded repellers.

In [53], it was shown that under the "endomorphism-type" assumptions, the folding entropy varies in an upper semi-continuous way.[2] In this situation, a single coarse-grained partition alone can exhaust all the complexities in the folding process. However, this does not work in general especially when arbitrarily many preimage branches are accumulated around the degenerate set; see Figure 1(B). A general mechanism on the influence of the degeneracy for systems beyond endomorphisms remains unknown yet. The main goal of this paper is to develop such a mechanism. To capture the emerging complexities in the refining coarse-graining process, a notion called degenerate rate is introduced as follows: take a decreasing sequence of open neighborhoods of Σ_f, denoted by $\mathcal{V} = \{V_m\}_{m\geq 1}$, such that

$$d_H(V_m, \Sigma_f) \to 0, \quad m \to \infty, \tag{1.3}$$

where d_H denotes the Hausdorff distance, and let $\bar{\eta} = \{\eta_m\}_{m\geq 1}$ be a sequence of positive real numbers approaching to zero. A measure $\mu \in \mathcal{P}(M)$ is said to admit **degenerate rate** $\bar{\eta}$ if

$$\left| \int_{V_m} \log |\operatorname{Jac}(D_x f)| d\mu \right| \leq \eta_m, \quad \forall m \geq 1. \tag{1.4}$$

Note that $\bar{\eta}$ characterizes how $|\operatorname{Jac}(D_x f)|$ diminishes on a sequence of shrinking domains around Σ_f.

The degenerate rate is well-defined for any $\mu \in \mathcal{P}(M)$ if the integrable condition

$$\left| \int \log |\operatorname{Jac}(D_x f)| d\mu \right| < \infty \tag{1.5}$$

is satisfied. This is because (1.5) implies $\mu(\Sigma_f) = 0$ and if we denote $\eta_m = |\int_{V_m} \log |\operatorname{Jac}(D_x f)| d\mu|$, then the sequence $\eta_m \to 0$ as $m \to \infty$. We note that condition (1.5) holds for a bunch of examples. For instance, consider

[2] In [53], it was assumed that the state space (excluding the degenerate set) $M \backslash \Sigma_f$ can be divided into several pieces which have either identical or disjoint images and the Jacobian are uniformly away from zero. This rules out the possibility of accumulations of infinite "foldings" occurring near the degenerate set during the approximation process.

the family of maps $\{f_a := 1 - ax^2\}$ on $[-1,1]$. If it satisfies that for any $n \geq 1$, (i) $|(f_a^n)'(f_a(0))| \geq e^{c_1 n}$, and (ii) $|f_a^n(0)| \geq e^{-c_2 n}$ with certain $c_1, c_2 > 0$, then (1.5) holds for any invariant measure μ of f_a. We remark that (i) and (ii) hold for a set of parameter a near 2 with positive Lebesgue measure (Collet–Eckmann [11], Benedicks–Carleson [4]). To see (1.5), consider the partition $\{A_k\}_{k\geq 1}$ of $[-1,1]$ with

$$A_k = \Big\{x \in [-1,1] : \frac{1}{k+1} < |x| \leq \frac{1}{k}\Big\}.$$

For small $\delta \in (0, \frac{c_1}{10})$ and sufficiently large k, let

$$N = N(k) = [\frac{\log(k\delta)}{c_2 + \log(2a)}].$$

Then for any $x \in A_k$,

$$|f_a^n(0) - f_a^n(x)| \leq \frac{1}{k}(2a)^n \leq \delta e^{-c_2 n}, \quad 1 \leq n \leq N,$$

which implies

$$\begin{aligned}
\Big|\frac{(f_a^N)'(f_a(x))}{(f_a^N)'(f_a(0))}\Big| &= \Big|\frac{\prod_{n=1}^N f_a^n(x)}{\prod_{n=1}^N f_a^n(0)}\Big| = \prod_{n=1}^N \Big|\frac{f_a^n(x) - f_a^n(0)}{f_a^n(0)} + 1\Big| \\
&\geq e^{-\frac{c_1}{10}N} e^{-\sum_{n=1}^N |\frac{f_a^n(x) - f_a^n(0)}{f_a^n(0)}|} \\
&\geq e^{-\frac{c_1}{10}N} e^{-\delta N} \geq e^{-\frac{c_1}{5}N}.
\end{aligned}$$

Thus, there exists $k_0 > 0$ such that for any $k \geq k_0$, $x \in A_k$,

$$\frac{\log|(f_a^{N+1})'(x)|}{N+1} \geq \frac{\frac{4}{5}c_1 N - \log(k+1)}{N+1} \geq \frac{4}{5}c_1 - \frac{11}{10}c_2 - \ln(2a).$$

For any x, let a sequence $0 = N_0 < N_1 < N_2 < \cdots$ be such that

$$N_{i+1} = \begin{cases} N_i + N(k) + 1, & \text{if } f^{N_i}(x) \in A_k, \\ N_i + 1, & \text{otherwise.} \end{cases}$$

Denote $\hat{c} = \min\{\log \frac{2a}{k_0},\ \frac{4}{5}c_1 - \frac{11}{10}c_2 - \log(2a)\}$. If μ is ergodic, then for μ-a.e. x,

$$\begin{aligned}
\int \log|f'(x)| d\mu &= \lim_{n\to\infty} \frac{1}{n}\sum_{i=0}^{n-1} \log|f_a'(f^i(x))| \\
&= \lim_{n\to\infty} \frac{1}{N_n}\sum_{i=0}^{n-1} \log|(f_a^{N_{i+1}-N_i})'(f^{N_i}(x))| \geq \hat{c}.
\end{aligned}$$

If μ is an invariant measure, by the ergodic decomposition, we still have

$$\int \log|f'(x)|d\mu \geq \hat{c}.$$

Moreover, the upper bound estimation $\int \log|f'(x)|d\mu \leq \log(2a)$ always holds.

We are concerned with the probability measures with a uniform degenerate rate. Define

$$\mathcal{P}_{\bar{\eta}}(f) := \left\{\nu \in \mathcal{P}(M) : \left|\int_{V_m} \log|\operatorname{Jac}(D_x f)|d\mu\right| \leq \eta_m, \forall m \geq 1\right\},$$

i.e., $\mathcal{P}_{\bar{\eta}}(f)$ collects all the probability measures with the degenerate rate $\bar{\eta}$. Note that for any $\mu \in \mathcal{P}_{\bar{\eta}}(f)$, $\mu(\Sigma_f) = 0$ holds automatically. Before proceeding to the main theorems, two remarks on $\mathcal{P}_{\bar{\eta}}(f)$ are made as follows:

(i) $\mathcal{P}_{\bar{\eta}}(f)$ is a closed subset of $\mathcal{P}(M)$ in the weak*-topology. To see this, let $\{\mu_i\}_{i\geq 1}$ be a sequence of measures in $\mathcal{P}_{\bar{\eta}}(f)$ satisfying $\lim_{i\to\infty} \mu_i = \mu$. Without loss of generality, assume $|\operatorname{Jac}(D_x f)||_{V_m} < 1$ for all $m \geq 1$. Then for any $s > m$, $|\int_{V_m\setminus V_s} \log|\operatorname{Jac}(D_x f)|d\mu_i| \leq \eta_m$. Hence,

$$\left|\int_{V_m\setminus V_s} \log|\operatorname{Jac}(D_x f)|d\mu\right| = \lim_{i\to\infty}\left|\int_{V_m\setminus V_s} \log|\operatorname{Jac}(D_x f)|d\mu_i\right| \leq \eta_m$$

which, together with $\mu(\Sigma_f) = 0$, yields

$$\left|\int_{V_m} \log|\operatorname{Jac}(D_x f)|d\mu\right| \leq \eta_m, \quad \forall m \geq 1.$$

That is, $\mu \in \mathcal{P}_{\bar{\eta}}(f)$.

(ii) The uniform degenerate rate is independent of the choice of $\mathcal{V} = \{V_m\}_{m\geq 1}$. In fact, let $\mathcal{V}^{(i)} = \{V_m^{(i)}\}_{m\geq 1}$ $(i = 1, 2)$ be any two sequences of neighborhoods of Σ_f satisfying (1.3). Then the corresponding sequences

$$\left\{\left|\int_{V_m^{(i)}} \log|\operatorname{Jac}(D_x f)|d\mu\right|\right\}_{m\geq 1}, \quad i = 1, 2$$

are uniformly equivalent, i.e., for any $\gamma > 0, m_1 \geq 1$, there exists $m_2 > 0$ such that

$$\left|\int_{V_m^{(2)}} \log|\operatorname{Jac}(D_x f)|d\mu\right| < \left|\int_{V_{m_1}^{(1)}} \log|\operatorname{Jac}(D_x f)|d\mu\right| + \gamma, \quad \forall m \geq m_2,$$

and vice versa. Henceforth, we fix a certain decreasing sequence $\mathcal{V} = \{V_m\}_{m\geq 1}$ satisfying (1.3), and by a variation of $\bar{\eta} = \{\eta_m\}_{m\geq 1}$ with limit zero, all settings of uniform degenerate rate will be exhausted.

For a general $C^r(r>1)$ system, the upper semi-continuity of folding entropy is established when measures with uniform degenerate rate are considered.

THEOREM 1.1: *Let f be a $C^r(r>1)$ map on a compact Riemannian manifold M. Then the folding entropy*

$$\mu \mapsto F_f(\mu)$$

is upper semi-continuous on $\mathcal{P}_{\bar{\eta}}(f)$.

Note that the map $\mu \mapsto \int \log|\operatorname{Jac}(D_x f)| d\mu$ is continuous on $\mathcal{P}_{\bar{\eta}}(f)$ since the integral on the neighborhoods of Σ_f is uniformly controlled. Theorem 1.1 then directly yields the upper semi-continuity of the entropy production.

COROLLARY 1.2: *Let f be a $C^r(r>1)$ map on a compact Riemannian manifold M. Let $\{\mu_n\}_{n\geq 1}\subseteq\mathcal{P}_{\bar{\eta}}(f)$ be a sequence of measures such that $\lim_{n\to\infty}\mu_n=\mu$. Then it holds that*

$$e_f(\mu) \geq \limsup_{n\to\infty} e_f(\mu_n).$$

1.2. APPLICATIONS TO INTERVAL MAPS. It is well-known that for a measurable dynamical system, the (Kolmogorov–Sinai) metric entropy measures the complexities (or uncertainties) produced during the dynamical evolutions. Among the various properties of metric entropy that have been well studied, the upper semi-continuity has drawn much attention as it plays a key role for the existence of equilibrium states [6, 57]. In the context of interval maps, an equality between folding entropy and metric entropy is established; see Theorem 4.1. As an application of Theorem 1.1, the upper semi-continuity property of the latter one is then obtained when invariant measures with uniform degenerate rates are considered. Henceforth, for an interval map f, let $\mathcal{M}_{\mathrm{inv}}(f)$ denote the set of all f-invariant Borel probability measures, and still, $\bar{\eta}=\{\eta_m\}_{m\geq 1}$ denotes any sequence of positive real numbers approaching to zero. Denote

$$\mathcal{M}_{\bar{\eta},\mathrm{inv}}(f)=\mathcal{P}_{\bar{\eta}}(f)\cap\mathcal{M}_{\mathrm{inv}}(f)$$

the set of all f-invariant Borel probability measures with uniform degenerate rate $\bar{\eta}$.

THEOREM 1.3: *Let f be a $C^r(r>1)$ interval map. Then the metric entropy*

$$\mu \mapsto h_\mu(f)$$

is upper semi-continuous on $\mathcal{M}_{\bar{\eta},\mathrm{inv}}(f)$.

It has been known that for any C^∞ system, the upper semi-continuity of metric entropy always holds [44, 60] on the set of invariant measures. In fact, the C^∞ smoothness brings about the uniform control on the degeneracy of the map which, as has been shown by various examples [9, 42, 43, 55], can be destroyed if only a finite smoothness is satisfied instead. However, the uniform degenerate rate concerns not only the degeneracy of map, but also how measures are distributed around the degenerate set which can be very singular leading to a non-uniform degenerate rate among the measures. Thus, the uniform degenerate rate provides a rather different mechanism of upper semi-continuity of metric entropy for systems with finite smoothness from the local geometrical control in the C^∞ setting.

Besides the smoothness, hyperbolicity provides another mechanism of upper semi-continuity of metric entropy. For systems with some hyperbolicity, e.g., uniformly hyperbolic systems [5], partially hyperbolic systems with one-dimensional center [12, 13], and diffeomorphisms away from tangencies [31], the upper semi-continuity of metric entropy has been established. For non-uniformly hyperbolic systems, Newhouse proposed the notion of hyperbolic rate to characterize the level of hyperbolicity and proved that the upper semi-continuity of metric entropy holds on any subset of invariant measures with uniform hyperbolic rate [44]. The uniform degenerate rate proposed in this paper is in the similar spirit in terms of the level of degeneracy with respect to measures.

Not only related to the metric entropy, the folding entropy was recently proved by Wu and Zhu [59] to be equal to another two types of entropy: the measure-theoretic preimage entropy introduced by Cheng and Newhouse [10] and the pointwise metric preimage entropy introduced by Wu and Zhu [58], and by analyzing separation property of preimages, the upper semi-continuity was also studied.

Besides the entropy, dimension is another intrinsic characteristic of complexity in the dynamical theory. For a differentiable dynamical system, the interrelation among the various dynamical quantities has been intensively investigated [28, 30, 47, 52, 61]. In particular, for an interval map f and an ergodic f-invariant Borel probability measure μ, the following formula

$$h_\mu(f) = \lambda^+(\mu)\dim_H(\mu) \tag{1.6}$$

was established in [28] under certain assumptions on the degeneracy, where $\lambda^+(\mu) = \max\{\lambda(\mu), 0\}$ is the positive part of the Lyapunov exponent $\lambda(\mu)$, and $\dim_H(\mu)$ is the Hausdorff dimension of μ (the definition is given in Section 2). In Section 4, by developing an integrable version of the Brin–Katok formula, it is shown that for a $C^r(r > 1)$ interval map f, any hyperbolic (i.e., $\lambda(\mu) \neq 0$) ergodic f-invariant measure μ is exact dimensional, i.e., the pointwise dimension is constant μ-a.e. and hence coincides with the Hausdorff dimension $\dim_H(\mu)$, and the validity of (1.6) is established; see Theorem 4.2. We emphasize that the hyperbolicity of measure is necessary for the exact dimensional property. The Hausdorff dimension always exists, but for non-hyperbolic measures, the pointwise dimension may not exist [26, 29].

As a consequence of Theorem 1.3, the upper semi-continuity of dimension at all hyperbolic measures on the set of ergodic measures with uniform degenerate rate is obtained. For simplicity of notation, write

$$\mathcal{M}_{\bar{\eta},\mathrm{erg}}(f) = \mathcal{P}_{\bar{\eta}}(f) \cap \mathcal{M}_{\mathrm{erg}}(f),$$

where $\mathcal{M}_{\mathrm{erg}}(f)$ denotes the set of all ergodic f-invariant Borel probability measures.

Theorem 1.4: *Let f be a $C^r(r > 1)$ interval map. Then any hyperbolic ergodic measure of f is exact dimensional, and the Hausdorff dimension*

$$\mu \mapsto \dim_H(\mu)$$

is upper semi-continuous at all hyperbolic measures in $\mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$.

1.3. Sharpness of the uniform degenerate rate condition.

For the previous known examples concerning the loss of upper semi-continuity of metric entropy, the ergodic measures at which the upper semi-continuity fails are all atomic and hence admit zero metric entropy (see, for instance, [9, 43, 55]). Note that the positivity of metric entropy for an ergodic measure implies that the hyperbolicity holds on a nontrivial full-measure set (due to the Margulis–Ruelle inequality [52] and ergodicity). Taking in consideration that the hyperbolicity is a possible approach to the upper semi-continuity of metric entropy [5], there arises the following question (Burguet [8]):

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

We give an answer to the above question in Section 5 by constructing a type of modified examples for which the loss of upper semi-continuity of metric entropy occurs at a non-atomic ergodic measure, say μ, with positive entropy. The defect of upper semi-continuity of metric entropy at μ follows from the infinite "foldings" around a homoclinic tangency whose images in average converge to μ; see Figure 5. The approximation is realized by taking some generic point of μ, say x_0, and then elaborately choosing a sequence of ergodic measures whose generic points follow the orbit of x_0 with high frequency but do not admit a uniform degenerate rate. This indicates that in the setting of differentiable systems (with degeneracy), the uniform degenerate rate is an essential condition for the upper semi-continuity of folding entropy.

Theorem 1.5: *For any $1 < r < \infty$, there exist C^r interval maps admitting ergodic measures with positive entropy as the non-upper semi-continuity points of the metric (or folding) entropy.*

This paper is organized as follows. Section 2 reviews the basic notions in the ergodic and entropy theory that will be used throughout the paper. Section 3 proves Theorem 1.1. Section 4 studies the applications of Theorem 1.1 in the one-dimensional setting. In particular, an entropy formula (Theorem 4.1) and a dimension formula (Theorem 4.2) are established for all $C^r(r > 1)$ interval maps. Theorem 1.3 and Theorem 1.4 then follow almost immediately. Section 5 constructs a type of modified interval maps by developing a typical failing mechanism of the upper semi-continuity of metric entropy at an ergodic measure with positive entropy. The sharpness of the condition of uniform degenerate rate in Theorem 1.1 and Theorem 1.3 is then illustrated.

2. Preliminaries

In this section, basic notions and facts in the entropy and dimension theory for later discussions are reviewed. Readers may refer to [48, 51, 57] for more details.

2.1. Invariant measures, partitions and entropy. Let X be a compact metric space with the Borel σ-algebra $\mathcal{B}(X)$ and $f : X \to X$ be a measurable transformation. Recall in the introduction that $\mathcal{P}(X)$ denotes the set of all Borel probability measures on X endowed with the weak*-topology, i.e., $\mu_n \to \mu$

in $\mathcal{P}(X)$ if and only if for any continuous function φ, $\int \varphi d\mu_n \to \int \varphi d\mu$ as $n \to \infty$. For $\mu \in \mathcal{P}(X)$, the **image** of μ under f is given by

$$f\mu(B) = \mu(f^{-1}B), \quad \forall B \in \mathcal{B}(X),$$

and μ is said to be f**-invariant** if $f\mu = \mu$. Denote by $\mathcal{M}_{\mathrm{inv}}(f)$ the set of all f-invariant Borel probability measures.

A **partition** $\xi = \{A_\alpha\}_{\alpha \in \mathscr{A}}$ of X is a collection of disjoint elements of $\mathcal{B}(X)$ such that $\bigcup_{\alpha \in \mathscr{A}} A_\alpha = X$. In particular, ξ is called **finite** if the cardinality $\#\mathscr{A} < \infty$. Given two partitions $\xi = \{A_\alpha\}_{\alpha \in \mathscr{A}}$ and $\zeta = \{C_\gamma\}_{\gamma \in \mathscr{C}}$, the join of ξ and ζ is a partition of X, denoted as $\xi \vee \zeta$, such that

$$\xi \vee \zeta = \{A_\alpha \cap C_\gamma\}_{\alpha \in \mathscr{A}, \gamma \in \mathscr{C}}.$$

If $\xi \vee \zeta = \zeta$, i.e., each element of ξ is a union of elements of ζ, then we call ζ a **refinement** of ξ, and write $\xi \preceq \zeta$. A sequence of partitions $\{\xi_n\}_{n \geq 1}$ is said to be **refining** (or **increasing**) if $\xi_1 \preceq \xi_2 \preceq \cdots \preceq \xi_n \preceq \cdots$.

Given $\mu \in \mathcal{P}(X)$ and a finite partition ξ of X, the **entropy** of ξ with respect to μ is

$$H_\mu(\xi) = \int_X -\log \mu(\xi(x)) d\mu(x),$$

where $\xi(x)$ denotes the element of ξ containing x. For any two finite partitions ξ and ζ, the **conditional entropy** of ξ given ζ is

$$H_\mu(\xi|\zeta) = \sum_{A \in \xi, C \in \zeta} -\mu(A \cap C) \log\left(\frac{\mu(A \cap C)}{\mu(C)}\right) = \sum_{C \in \zeta} \mu(C) H_{\mu_C}(\xi|_C)$$

where $\mu_C(A) := \mu(A \cap C)/\mu(C)$ is the conditional measure of A given C, and $\xi|_C$ denotes the partition ξ restricted on C. For any $\mu \in \mathcal{M}_{\mathrm{inv}}(f)$, the **metric entropy** of f with respect to μ and ξ is

$$h_\mu(f, \xi) = \lim_{n \to \infty} H_\mu\left(\xi | \bigvee_{i=1}^{n} f^{-i}\xi\right) = \inf_{n \geq 1} H_\mu\left(\xi | \bigvee_{i=1}^{n} f^{-i}\xi\right), \tag{2.1}$$

and the metric entropy of f with respect to μ is

$$h_\mu(f) = \sup_\xi h_\mu(f, \xi),$$

where the supremum is taken over all finite partitions of X.

We remark that the entropy and conditional entropy can be defined in the more general setting where $(X, \mathcal{B}(X), \mu)$ is a Lebesgue space and ξ, ζ are measurable partitions [51]. In this setting, the folding entropy is a particular type of conditional entropy with ξ and ζ chosen as ϵ and $f^{-1}\epsilon$, respectively.

The definition of metric entropy is based on the following fundamental function:

$$\phi(x) = \begin{cases} -x\log x, & x \in (0,1], \\ 0, & x = 0, \end{cases} \tag{2.2}$$

from the information theory. The following simple fact on the concavity of ϕ will be used several times in the estimation of entropy in Section 3.

Proposition 2.1: *Let $\{p_i\}_{i=1}^n$ be such that $\sum_{i=1}^n p_i = 1$ and $p_i \geq 0, 1 \leq i \leq n$, and $\{x_i\}_{i=1}^n \subset [0,1]$. Then*

$$\sum_{i=1}^n p_i\phi(x_i) \leq \phi\Big(\sum_{i=1}^n p_i x_i\Big).$$

In particular, $\sum_{i=1}^n \phi(p_i) \leq \log n$.

2.2. Dimension of measures. For the compact metric space X with metric d, given any finite Borel measure μ, for μ-a.e. $x \in X$, the lower and upper local dimension are respectively defined as

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

where $B(x,\delta) = \{y : d(y,x) < \delta\}$. The measure μ is said to have **local dimension** at x if $\underline{\dim}_\mu(x) = \overline{\dim}_\mu(x)$. For a subset $Y \subset X$ and a number $t > 0$, the t-Hausdorff measure of Y is defined as

$$m_H(Y,t) = \lim_{\varepsilon\to 0} \inf_{\mathcal{U}} \sum\nolimits_{U\in\mathcal{U}} \operatorname{diam}^t(U),$$

where the infimum is taken over all finite or countable coverings $\mathcal{U}$ of Y of open sets with $\operatorname{diam}(\mathcal{U}) \leq \varepsilon$. The Hausdorff dimension of Y and then of μ are respectively defined as

$$\begin{aligned} \dim_H(Y) &= \inf\{t : m_H(Y,t) = 0\} = \sup\{t : m_H(Y,t) = \infty\}, \\ \dim_H(\mu) &= \inf\{\dim_H(Y) : \mu(Y) = 1\}. \end{aligned}$$

The following classical result states that under standard conditions, the local dimension exists and coincides with the Hausdorff dimension.

Proposition 2.2 (Young [61]): *Let X be a compact separable metric space of finite topological dimension, and μ be a finite Borel measure on X such that*

$$\underline{\dim}_\mu(x) = \overline{\dim}_\mu(x) = D, \quad \mu\text{-a.e. } x \in X. \tag{2.3}$$

Then $\dim_H(\mu) = D$.

Any measure μ satisfying (2.3) is said to be *exact dimensional* with the common value denoted by $\dim(\mu)$. Proposition 2.2 states that if a measure is exact dimensional, then $\dim(\mu)$ is well-defined and coincides with the Hausdorff dimension $\dim_H(\mu)$. The exact dimensionality and dimension formulas have been extensively investigated in the different contexts (see, for instance, [2, 3, 6, 14, 18, 19, 30, 35, 36, 38, 40, 41, 46, 48, 49]). In Section 4.2, any hyperbolic ergodic measure of a $C^r(r > 1)$ interval map is shown to be exact dimensional and a dimension formula relating dimension, Lyapunov exponents and metric entropy is established.

3. Upper semi-continuity of folding entropy

In this section, we prove Theorem 1.1. Throughout this section, let f be a $C^r(r > 1)$ map on a compact Riemannian manifold M. Before proceeding to the proof, we would like to note that the basic idea in the handling of the degeneracy of f is based on our previous work in [32] with some further refining estimations.

Note that the folding entropy is upper semi-continuous if there is no degeneracy [53]. For proving the upper semi-continuity when the degeneracy is present, our method is to take a sequence of appropriate partitions according to the degenerate rate so that on every element (excepting the degenerate one) of each partition, f displays like a local diffeomorphism with no more complexities arising from the foldings. By making the degeneracy approach to zero through the uniform degenerate rate, the complexities from the degenerate part of f is controlled. The upper semi-continuity is hence obtained by estimating the folding entropy via such a sequence of partitions.

3.1. Construction of refining partitions. For convenience of analysis, embed M into a Euclidean space $\mathbb{R}^N$. One can find tubular neighborhoods $T_1 \subset T_2$ of M in $\mathbb{R}^N$ and a C^r extension $g : T_1 \to T_2$ of f such that $g^{-1}(\overline{T}_1) \subset T_1$ and $M = \bigcap_{n\in\mathbb{N}} g^{-n}(T_1)$. For simplicity of notation, we still write $g|_{T_1}$ as f

without confusion. Let $\alpha = \min\{r-1, 1\}$. Then the derivative of f is α-Hölder continuous, i.e., there exists $K > 1$ such that for any $x, y \in T_1$,

$$\|D_x f - D_y f\| \leq Kd(x,y)^\alpha. \tag{3.1}$$

To approximate the folding entropy, we shall use a sequence of finite refining partitions to approximate the measurable partitions $\{\epsilon|_{f^{-1}\epsilon(x)}\}_{x\in T_1}$. To begin with, for each $k \in \mathbb{N}$, define a partition of $\mathbb{R}^N$ as

$$\Gamma_k = \Big\{\Big(\frac{q_1}{2^k}, \frac{q_1+1}{2^k}\Big)\overbrace{\times\cdots\times}^{N}\Big(\frac{q_N}{2^k}, \frac{q_N+1}{2^k}\Big) : q_1, \ldots, q_N \in \mathbb{Z}\Big\}.$$

Plainly, $\Gamma_1 \preceq \cdots \preceq \Gamma_k \preceq \cdots$. Since f only acts on T_1, henceforth, by Γ_k we mean $\Gamma_k|_{T_1}$. Then there exists a constant $C_1 > 0$ independent of k and N such that

$$\#\{A \in \Gamma_k : A \cap T_1 \neq \emptyset\} \leq C_1 2^{kN}, \quad \forall\, k \in \mathbb{N}. \tag{3.2}$$

To approximate the inverse partition $f^{-1}\epsilon$, we need to consider the pull-back of Γ_k under f. First, we have to deal with the degeneracy of f. Given any $\varepsilon > 0$, define

$$U_\varepsilon = \{x \in T_1 : m(D_x f) < \varepsilon\}, \quad G_\varepsilon = \{x \in T_1 : m(D_x f) \geq \varepsilon\},$$

where for a linear operator Φ, $m(\Phi) := \inf_{\|z\|=1}\|\Phi(z)\|$ denotes the small norm of Φ. By (3.1), if we take $r_\varepsilon = (\frac{\varepsilon^2}{4K})^{\frac{1}{\alpha}}$, we have for any $x \in G_\varepsilon$,

$$m(D_x f) > \frac{\varepsilon}{2}, \quad \forall y \in B(x, r_\varepsilon).$$

For each $k \in \mathbb{N}$, let $\varepsilon_k = \varepsilon_0\, 2^{-k\beta}$ with $\beta = \frac{\alpha}{2+\alpha}$ and $\varepsilon_0 = (2(4K)^{\frac{1}{\alpha}}\sqrt{N})^\beta$. Then $\{U_{\varepsilon_k}\}$ is a decreasing sequence of neighborhoods of the degenerate set Σ_f. Also, it is not hard to see that for each k,

$$f(B(x, r_{\varepsilon_k})) \supseteq \Gamma_k(fx), \quad \forall x \in G_{\varepsilon_k}, \tag{3.3}$$

where $\Gamma_k(y)$ denotes the element of Γ_k containing y.

For any $P \in \Gamma_k$, denote $P^{-1,c}$ the set of all the connected components of the preimage $f^{-1}P$. To collect all the components "near" Σ_f, let

$$B_k = \bigcup_{P\in\Gamma_k}\ \bigcup_{\substack{Q\in P^{-1,c},\\ Q\cap U_{\varepsilon_k}\neq\emptyset}} Q.$$

Plainly, $B_k \supset U_{\varepsilon_k}$; see Figure 2(A). In fact, as the following Lemma 3.1 shows, B_k is in the neighborhood of Σ_f with size the same order as U_{ε_k}.

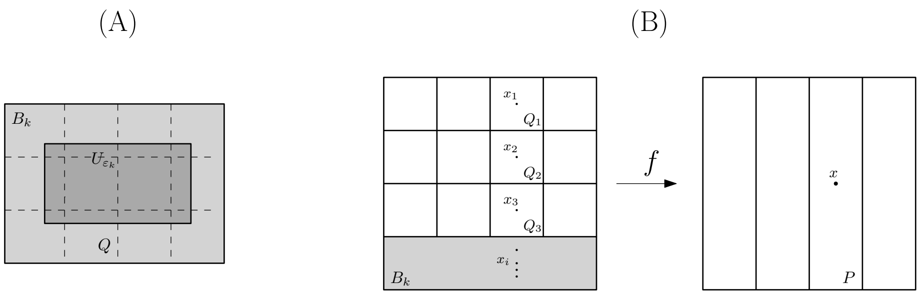

Figure 2. (A) The degenerate component B_k (in light grey) consists of all the components $Q \in \Gamma_k$ (in dashed line) which intersect with U_{ε_k} (in dark grey). (B) The left panel denotes the pullback partition $\Gamma_k^{-1,c}$; the right panel denotes the partition Γ_k. For a typical $P \in \Gamma_k$, $x \in P$, the preimages of x, $x_i (i = 1, 2, 3)$ lie in the different components $Q_i \in P^{-1,c} (i = 1, 2, 3)$. The other preimages $x_i (i = 4, \ldots)$ are all collected by the degenerate component $B_k \in \Gamma_k^{-1,c}$.

Lemma 3.1: *For any* $k \geq 1$,

$$B_k \subseteq U_{C_2 2^{-k\beta}}$$

for some constant $C_2 > 0$ *independent of* k, β.

Proof. We claim that for any $Q \in P^{-1,c}$, where $P \in \Gamma_k$, such that $Q \cap U_{\varepsilon_k} \neq \emptyset$, it holds that

$$B(x, r_{\varepsilon_k}) \cap U_{\varepsilon_k} \neq \emptyset, \quad \forall x \in Q. \tag{3.4}$$

In fact, suppose there exists some $x_0 \in Q$ with $B(x_0, r_{\varepsilon_k}) \cap U_{\varepsilon_k} = \emptyset$, by (3.3), $B(x_0, r_{\varepsilon_k})$ contains a certain preimage component of $f^{-1}(\Gamma_k(fx_0))$, say Q_0, such that $x_0 \in Q_0$. Then it must be that $Q = Q_0$. This contradicts $Q \cap U_{\varepsilon_k} \neq \emptyset$.

By the Hölder continuity of Df in (3.1), (3.4) yields

$$m(D_x f) \leq \varepsilon_k + K r_{\varepsilon_k}^{\alpha} \leq C_2 2^{-k\beta}, \quad \forall x \in B_k$$

for some constant $C_2 > 0$. Hence, $B_k \subset U_{C_2 2^{-k\beta}}$. ∎

Let

$$\Gamma_k^{-1,c} = \{B_k\} \bigcup_{P \in \Gamma_k} \{Q \in P^{-1,c} : Q \cap U_{\varepsilon_k} = \emptyset\}.$$

Note that $\Gamma_k^{-1,c}$ is an approximation of the pull-back of Γ_k in such a way that for any $P \in \Gamma_k$, all the components of $f^{-1}P$ "away from" Σ_f are separated by the different elements $Q \in \Gamma_k^{-1,c}$, while the components "near" Σ_f are integrated as a whole by the single set B_k; see Figure 2(B). We call B_k the **degenerate component** of $\Gamma_k^{-1,c}$.

Now, a sequence of refining partitions $\{\Gamma_k\}$ together with its pull-backs $\{\Gamma_k^{-1,c}\}$ have been constructed. In the following, we shall respectively use the finite partitions $f^{-1}\Gamma_k$ and $\Gamma_k^{-1,c}$ to approximate the measurable partitions $f^{-1}\epsilon$ and $\epsilon|_{f^{-1}\epsilon}$, and the conditional entropy $H_\mu(\Gamma_k^{-1,c}|f^{-1}\Gamma_k)$ to approximate the folding entropy $H_\mu(\epsilon|f^{-1}\epsilon)$ in the partition refining process.

3.2. Approximation of folding entropy. Throughout this section, we are concerned with any probability measure ν with a certain degenerate rate $\bar{\eta}$. It is easy to see that

$$\left|\int \log|\operatorname{Jac}(D_x f)\right| d\nu(x)| < \infty,$$

and hence $\nu(B_k) \to 0$ as $k \to \infty$. By construction, for ν-a.e. x, $\Gamma_k^{-1,c}$ is increasing to $\epsilon|_{(f^{-1}\epsilon)(x)}$ as $k \to \infty$. Therefore,

$$\begin{aligned} F_f(\nu) &= \lim_{k\to+\infty} H_\mu(\Gamma_k^{-1,c} \mid f^{-1}\epsilon) \\ &\le \limsup_{k\to+\infty} H_\mu(\Gamma_k^{-1,c} \mid f^{-1}\Gamma_k) \\ &= \limsup_{k\to\infty} \sum_{P\in\Gamma_k} \sum_{Q\in\Gamma_k^{-1,c}} -\nu(Q\cap f^{-1}P)\log\Big(\frac{\nu(Q\cap f^{-1}P)}{\nu(f^{-1}P)}\Big). \end{aligned} \tag{3.5}$$

Note that for any $Q \in \Gamma_k^{-1,c}$, either $Q = B_k$ or $Q \cap B_k = \emptyset$, and for the latter case, $f|_Q$ is a diffeomorphism with $fQ = P$ for certain $P \in \Gamma_k$. Then according to whether Q equals B_k or not, the right-hand side of (3.5) is split into two terms:

$$\Delta_k^{(1)}(\nu) = \sum_{P\in\Gamma_k} -\nu(B_k\cap f^{-1}P)\log\Big(\frac{\nu(B_k\cap f^{-1}P)}{\nu(f^{-1}P)}\Big),$$

$$\Delta_k^{(2)}(\nu) = \sum_{\substack{Q\in\Gamma_k^{-1,c}\setminus\{B_k\},\\ P=fQ}} -\nu(Q)\log\Big(\frac{\nu(Q)}{\nu(f^{-1}P)}\Big).$$

Thus, we have

$$F_f(\nu) \leq \limsup_{k\to\infty}(\Delta_k^{(1)}(\nu) + \Delta_k^{(2)}(\nu)). \tag{3.6}$$

Note that $\Delta_k^{(1)}$ and $\Delta_k^{(2)}$ characterize the complexities of folding entropy near and away from the degenerate set Σ_f, respectively. In the following, $\Delta_k^{(1)}$ and $\Delta_k^{(2)}$ are analyzed respectively.

3.2.1. *Estimation of* $\Delta_k^{(1)}$. In this subsection, we show that the complexity collected by the degenerate component B_k, i.e., the term $\Delta_k^{(1)}$, is bounded by the degenerate rate, and hence is uniformly small on any subset of measures with uniform degenerate rate.

The following lemma establishes the relation between the degeneracy of maps and the degenerate measures.

Lemma 3.2: *For any $k \geq 1$, it holds that*

$$k\nu(B_k) \leq C_3\nu(B_k) - C_4 \int_{B_k} \log|\operatorname{Jac}(D_xf)| d\nu(x),$$

where the constants $C_3, C_4 > 0$ are independent of k and ν.

Proof. By Lemma 3.1, if we denote $L = \max_{x\in T_1}\|D_xf\|$, then

$$|\operatorname{Jac}(D_xf)| \leq C_2L^{N-1}2^{-\beta k}, \quad \forall x \in B_k,$$

or equivalently,

$$-\log C_2L^{N-1} + k\beta\log 2 \leq -\log|\operatorname{Jac}(D_xf)|, \quad \forall x \in B_k. \tag{3.7}$$

Integrating both sides of (3.7) on B_k, we have

$$k\nu(B_k) \leq \frac{\nu(B_k)\log(C_2L^{N-1})}{\beta\log 2} - \frac{1}{\beta\log 2}\int_{B_k}\log|\operatorname{Jac}(D_xf)| d\nu(x).$$

The lemma is proved by taking $C_3 = \frac{\log(C_2L^{N-1})}{\beta\log 2}, C_4 = \frac{1}{\beta\log 2}$. ∎

As a direct corollary of Lemma 3.2, the following fact will be used several times throughout this section.

Lemma 3.3: *For any $\nu \in \mathcal{P}_{\bar\eta}(f)$, $\nu(B_k) \to 0$ uniformly as $k \to \infty$.*

Now, we are prepared to estimate $\Delta_k^{(1)}$.

Lemma 3.4: *For any $\varepsilon > 0$, there exists $k_1 \in \mathbb{N}$ such that for any $\nu \in \mathcal{P}_{\bar{\eta}}(f)$ and $k \geq k_1$, it holds that*

$$\Delta_k^{(1)}(\nu) < \varepsilon\,.$$

Proof. By noting that $\log \nu(f^{-1}P) \leq 0$, we have

$$\Delta_k^{(1)}(\nu) \leq \sum_{P\in\Gamma_k} -\nu(B_k \cap f^{-1}P) \log \nu(B_k \cap f^{-1}P),$$

which, by renormalizing on B_k, yields

$$\Delta_k^{(1)}(\nu) \leq \nu(B_k) \sum_{P\in\Gamma_k} \phi\Big(\frac{\nu(B_k \cap f^{-1}P)}{\nu(B_k)}\Big) + \phi(\nu(B_k)), \tag{3.8}$$

where, we recall, ϕ is defined in (2.2).

Applying Proposition 2.1 with $\{p_i\} = \{\frac{\nu(B_k\cap f^{-1}P)}{\nu(B_k)}\}_{P\in\Gamma_k}$, we have

$$\sum_{P\in\Gamma_k} \phi\Big(\frac{\nu(B_k \cap f^{-1}P)}{\nu(B_k)}\Big) \leq \log(\#\,\Gamma_k) \leq \log(C_1 2^{kN}),$$

where the last inequality is by (3.2). Hence,

$$\Delta_k^{(1)}(\nu) \leq \log(C_1^{1/k} 2^N) k\nu(B_k) + \phi(\nu(B_k)),$$

which, by Lemma 3.2, yields

$$\Delta_k^{(1)}(\nu) \leq \tilde{C}_3 \nu(B_k) - \tilde{C}_4 \int_{B_k} \log|\operatorname{Jac}(D_x f)| d\nu(x) + \phi(\nu(B_k)) \tag{3.9}$$

where the constants $\tilde{C}_3, \tilde{C}_4 > 0$ are independent of k, ν.

By Lemma 3.3 and the definition of uniform degenerate rate, all the three terms in the right-hand side of (3.9) converge to zero uniformly on $\mathcal{P}_{\bar{\eta}}(f)$. Hence, the convergence $\Delta_k^{(1)}(\nu) \to 0$ is uniform on $\mathcal{P}_{\bar{\eta}}(f)$. The proof of the lemma is concluded. ∎

3.2.2. *Analysis of $\Delta_k^{(2)}$.* To analyze $\Delta_k^{(2)}$, we need to investigate the partition refining process more carefully. Note that as k increases, the refining of the partition $\Gamma_{k-1}^{-1,c}$ by the more refined partition $\Gamma_k^{-1,c}$ consists of two different processes: (i) each element of $\Gamma_{k-1}^{-1,c} \backslash \{B_{k-1}\}$ is refined by the elements of partition $\Gamma_k^{-1,c}$; (ii) new separated components of $\Gamma_k^{-1,c}$ emerge from $(B_{k-1} \backslash B_k)$ which have been integrated by the degenerate component $B_{k-1} \in \Gamma_{k-1}^{-1,c}$; see Figure 3(A).

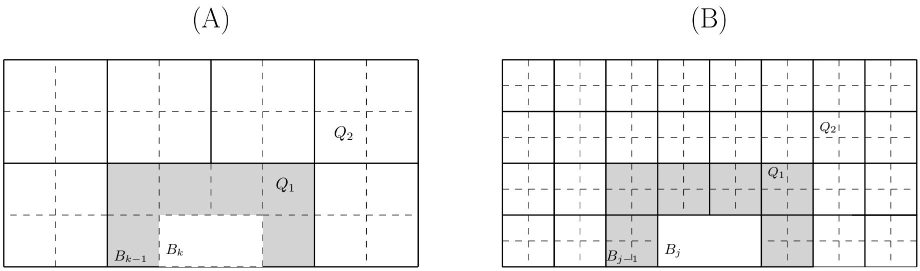

Figure 3. (A) The solid line denotes the partition $\Gamma_{k-1}^{-1,c}$; the dashed line denotes the refined partition $\Gamma_k^{-1,c}$. The set difference $B_{k-1}\backslash B_k$ (the region in grey) consists of the new-emerging components $Q_1 \in \Omega_k$. Outside B_{k-1}, the partition $\Gamma_{k-1}^{-1,c}$ is refined by the components $Q_2 \in \Gamma_{k,1}^{-1,c}$. (B) The solid line denotes the partition $\Gamma_j^{-1,c}$; the dashed line denotes the refined partition $\Gamma_k^{-1,c}$. The set difference $B_{j-1}\backslash B_j$ (the region in grey) is refined by components $Q_1 \in \Omega'_{k,j}$. Outside B_{j-1}, the partition $\Gamma_j^{-1,c}$ is refined by components $Q_2 \in \mathcal{W}_{k,j}$.

For each $k \geq 1$, define

$$\begin{aligned}\Gamma_{k,1}^{-1,c} &= \{Q \in \Gamma_k^{-1,c} : Q \text{ is contained in some } Q' \in \Gamma_{k-1}^{-1,c} \setminus \{B_{k-1}\}\},\\ \Omega_k &= \{Q \in \Gamma_k^{-1,c} \setminus \{B_k\} : Q \text{ is contained in } B_{k-1}\}.\end{aligned}$$

Then corresponding to the above refining processes (i) and (ii), $\Delta_k^{(2)}$ is further split into two terms:

$$\begin{aligned}\Delta_k^{(2)}(\nu) &= \sum_{\substack{Q\in\Gamma_{k,1}^{-1,c},\\ P=fQ}} -\nu(Q)\log\left(\frac{\nu(Q)}{\nu(f^{-1}P)}\right) + \sum_{\substack{Q\in\Omega_k,\\ P=fQ}} -\nu(Q)\log\left(\frac{\nu(Q)}{\nu(f^{-1}P)}\right)\\ &=: \Delta_k^{(2,1)}(\nu) + \Delta_k^{(2,2)}(\nu),\end{aligned} \tag{3.10}$$

where $\Delta_k^{(2,1)}$ captures the refining complexities outside B_{k-1}, and $\Delta_k^{(2,2)}$ collects the new complexities arising from the difference $(B_{k-1}\backslash B_k)$.

In the rest of this subsection, we show that the new emerging complexities $\Delta_k^{(2,2)}$ from the degenerate component during the refining process are uniformly small on any $\mathcal{P}_{\bar{\eta}}(f)$. In the next subsection, the (non-increasing) monotonicity related to the complexity $\Delta_k^{(2,1)}$ will be established.

Lemma 3.5: *For any $\varepsilon > 0$, there exists $k_2 \in \mathbb{N}$ such that for any $k \geq k_2$ and $\nu \in \mathcal{P}_{\bar{\eta}}(f)$,*

$$\sum_{k>k_2} \Delta_k^{(2,2)}(\nu) < \varepsilon\,.$$

Proof. Since for any $k \geq 1$, $\bigcup_{Q\in\Omega_k} Q = B_{k-1} \setminus B_k$,

$$\Delta_k^{(2,2)}(\nu) \leq \sum_{Q\in\Omega_k} -\nu(Q)\log\nu(Q). \tag{3.11}$$

Similar to (3.8), the right-hand side of (3.11) can be written as

$$\nu(B_{k-1} \setminus B_k) \sum_{Q\in\Omega_k} \phi\Big(\frac{\nu(Q)}{\nu(B_{k-1} \setminus B_k)}\Big) - \nu(B_{k-1}\setminus B_k)\log\nu(B_{k-1} \setminus B_k)$$
$$=: \Delta_k^{(2,2,1)}(\nu) + \Delta_k^{(2,2,2)}(\nu).$$

Thus, the estimation of $\Delta_k^{(2,2)}$ is reduced to that of $\Delta_k^{(2,2,1)}$ and $\Delta_k^{(2,2,2)}$, respectively.

We use a similar approach as in Lemma 3.4 to estimate $\Delta_k^{(2,2,1)}$. Since each $Q \in \Omega_k$ contains a ball with radius at least $(2^{k+1}L)^{-1}$, we have

$$\#\Omega_k \leq \tilde{C}_1^{N(k+1)} L^N$$

for some constant $\tilde{C}_1 > 0$ independent of k, L and N. Then applying Proposition 2.1 with

$$\{p_i\} = \Big\{\frac{\nu(Q)}{\nu(B_{k-1} \setminus B_k)}\Big\}_{Q\in\Omega_k},$$

we have

$$\Delta_k^{(2,2,1)}(\nu) \leq \log(\#\tilde{C}_1^{N(k+1)} L^N). \tag{3.12}$$

Applying Lemma 3.2 with B_k replaced by $(B_k\setminus B_{k-1})$,

$$\Delta_k^{(2,2,1)}(\nu) \leq \tilde{C}_3\nu(B_{k-1} \setminus B_k) - \tilde{C}_4 \int_{B_{k-1}\setminus B_k} \log|\operatorname{Jac}(D_x f)| d\nu(x),$$

where the constants $\tilde{C}_3, \tilde{C}_4 > 0$ are independent of k and ν. Thus, for any $k_2 \in \mathbb{N}$,

$$\sum_{k>k_2} \Delta_k^{(2,2,1)}(\nu) \leq \tilde{C}_3\nu(B_{k_2}) - \tilde{C}_4 \int_{B_{k_2}} \log|\operatorname{Jac}(D_x f)| d\nu(x). \tag{3.13}$$

For the estimation of $\Delta_k^{(2,2,2)}$, denote

$$I_1 = \{k \in \mathbb{N} : k \geq -\log\nu(B_k \setminus B_{k-1})\}, \quad I_2 = \{k \in \mathbb{N} : k < -\log\nu(B_k \setminus B_{k-1})\}.$$

Then for any $k_2 \in \mathbb{N}$,

$$\begin{aligned}\sum_{k>k_2} \Delta_k^{(2,2,2)} &= \sum_{k\in I_1, k>k_2} -\nu(B_k \setminus B_{k-1}) \log \nu(B_k \setminus B_{k-1}) \\ &\quad + \sum_{k\in I_2, k>k_2} -\nu(B_k \setminus B_{k-1}) \log \nu(B_k \setminus B_{k-1}) \\ &\le \sum_{k\in I_1, k>k_2} k\nu(B_k \setminus B_{k-1}) + \sum_{k\in I_2, k>k_2} ie^{-i}.\end{aligned}$$

Applying Lemma 3.2 for each $k \in I_1$ with $k > k_2$, we have

$$\sum_{k>k_2} \Delta_k^{(2,2,2)} \le C_3\nu(B_{k_2}) - C_4 \int_{B_{k_2}} \log|\operatorname{Jac}(D_x f)| d\nu(x) + \sum_{k\in I_2, k>k_2} ke^{-k}. \tag{3.14}$$

Note that $\sum_{k>k_2} ke^{-k} \to 0$ as $k_2 \to \infty$. Also, by Lemma 3.3, the convergence

$$\int_{B_{k_2}} \log|\operatorname{Jac}(D_x f)| d\nu(x) \to 0, \quad \nu(B_{k_2}) \to 0 \quad \text{as } k_2 \to \infty$$

is uniform on $\mathcal{P}_{\bar{\eta}}(f)$.

Combining (3.13) and (3.14), the lemma is proved. ∎

3.2.3. *Monotonicity by refining partitions.* In this subsection, we show that the complexities of folding entropy outside the degenerate component is (non-increasing) monotonic with respect to the refining partitions. Unlike the previous two subsections in which measures are required to have a uniform degenerate rate, in this subsection, the analysis applies to all measures in $\mathcal{P}(M)$.

In the same spirit of $\Gamma_{k,1}^{-1,c}$ and Ω_k defined in section 3.2.2, for any $1\le j\le k$, let

$$\begin{aligned}\mathcal{W}_{k,j} &= \{Q \in \Gamma_k^{-1,c} : Q \subseteq Q' \text{ for some } \; Q' \in \Gamma_j^{-1,c} \setminus \{B_j\}\}, \\ \Omega'_{k,j} &= \{Q \in \Gamma_k^{-1,c} \setminus \{B_k\} : Q \subseteq Q' \text{ for some } Q' \in \Omega_j\}.\end{aligned}$$

Note that $\mathcal{W}_{k,j}$ and $\Gamma_{k,1}^{-1,c}$ respectively correspond to the refinement of $\Gamma_j^{-1,c}\setminus\{B_j\}$ and Ω_j by elements from the more refined partition $\Gamma_k^{-1,c}$; see Figure 3(B). In particular,

$$W_{k,k-1} = \Gamma_{k,1}^{-1,c}, \quad \Omega'_{k,k} = \Omega_k \,.$$

Denote the complexities collected by $\mathcal{W}_{k,j}$ and $\Omega'_{k,j}$ as

$$I_{k,j}(\nu) = \sum_{\substack{Q\in\mathcal{W}_{k,j},\\ P=fQ}} -\nu(Q) \log\left(\frac{\nu(Q)}{\nu(f^{-1}P)}\right), \quad I'_{k,j}(\nu) = \sum_{\substack{Q\in\Omega'_{k,j},\\ P=fQ}} -\nu(Q) \log\left(\frac{\nu(Q)}{\nu(f^{-1}P)}\right).$$

As a natural link to $\Delta_k^{(2)}$, it is not hard to have the following.

Proposition 3.6: *Fix any $k_0 \in \mathbb{N}$. For any $k > k_0$,*

$$\Delta_k^{(2)}(\nu) = I_{k,k_0}(\nu) + \sum_{j=k_0+1}^{k} I'_{k,j}(\nu).$$

In particular, (3.10) *follows by taking $k_0 = k-1$, i.e.,*

$$\Delta_k^{(2)}(\nu) = I_{k,k-1}(\nu) + I'_{k,k}(\nu) = \Delta_k^{(2,1)} + \Delta_k^{(2,2)}.$$

The following lemma is the main result of this subsection on the monotonic properties of $I_{k,j}$ and $I'_{k,j}$.

Lemma 3.7: *Fix any $j \in \mathbb{N}$. For any $k \geq j$, $I_{k,j}(\nu)$ (resp. $I'_{k,j}(\nu)$) is non-increasing with respect to k. In particular, $I'_{k,j}(\nu) \leq I'_{j,j}(\nu) = \Delta_j^{(2,2)}(\nu)$.*

Proof. We only prove the non-increasing monotonicity of $I_{k,j}$, and the same argument applies to $I'_{k,j}$.

A key observation is that for any $k' > k \geq j$, the partition $\Gamma_{k'}$ and its pull-back $\Gamma_{k'}^{-1,c}$ respectively refine the partitions Γ_k and $\Gamma_k^{-1,c}$ in a corresponding way. More specifically, for any $P \in \Gamma_k$ and $P' \in \Gamma_{k'}$ with $P' \subset P$, there is a (unique) element $Q' \in \mathcal{W}_{k',j}$ such that $fQ' = P'$; see Figure 4.

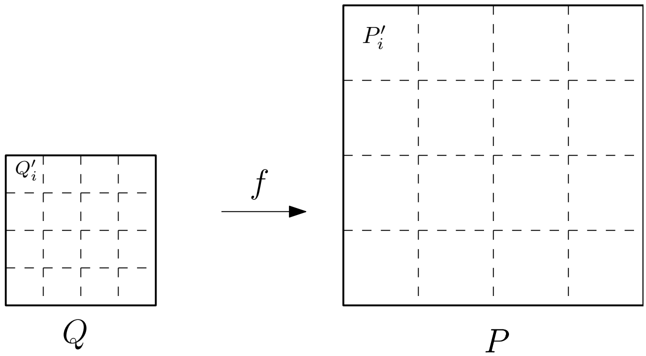

Figure 4. Any component $Q \in \mathcal{W}_{k,j}$ (resp. $P \in \Gamma_k$) is refined by the components $Q'_i \in \mathcal{W}_{k',j}$ (resp. $P'_i \in \Gamma_{k'}$) such that $P'_i = fQ'_i$ for each i.

Thus, $I_{k',j}$ can be written as

$$
\begin{aligned}
I_{k',j} &= \sum_{Q\in\mathcal{W}_{k,j}} \sum_{\substack{Q'\in\mathcal{W}_{k',j},\\ Q'\subseteq Q, P'=fQ'}} \nu(f^{-1}P')\phi\Big(\frac{\nu(Q')}{\nu(f^{-1}P')}\Big) \\
&= \sum_{\substack{Q\in\mathcal{W}_{k,j},\\ P=fQ}} \nu(f^{-1}P) \sum_{\substack{Q'\in\mathcal{W}_{k',j},\\ Q'\subseteq Q, P'=fQ'}} \frac{\nu(f^{-1}P')}{\nu(f^{-1}P)}\phi\Big(\frac{\nu(Q')}{\nu(f^{-1}P')}\Big).
\end{aligned}
\tag{3.15}
$$

Since $\{P'\}$ is a refinement of P, applying Proposition 2.1 with

$$
\{p_i\} = \Big\{\frac{\nu(f^{-1}P')}{\nu(f^{-1}P)}\Big\}_{P'}, \quad \{x_i\} = \Big\{\frac{\nu(Q')}{\nu(f^{-1}P')}\Big\}_{Q'},
$$

we have

$$
\sum_{\substack{Q'\in\mathcal{W}_{k',j},\\ Q'\subseteq Q, P'=fQ'}} \frac{\nu(f^{-1}P')}{\nu(f^{-1}P)}\phi\Big(\frac{\nu(Q')}{\nu(f^{-1}P')}\Big) \le \phi\Bigg(\sum_{\substack{Q'\in\mathcal{W}_{k',j},\\ Q'\subseteq Q,}} \frac{\nu(Q')}{\nu(f^{-1}P)}\Bigg) = \phi\Big(\frac{\nu(Q)}{\nu(f^{-1}P)}\Big).
$$

Therefore,

$$
\begin{aligned}
I_{k',j} &\le \sum_{\substack{Q\in\mathcal{W}_{k,j},\\ P=fQ}} \nu(f^{-1}P)\phi\Big(\frac{\nu(Q)}{\nu(f^{-1}P)}\Big) \\
&= \sum_{\substack{Q\in\mathcal{W}_{k,j},\\ P=fQ}} -\nu(f^{-1}P)\log\Big(\frac{\nu(Q)}{\nu(f^{-1}P)}\Big) = I_{k,j}. \qquad \blacksquare
\end{aligned}
$$

Now, fix any $k_0 \in \mathbb{N}$, and put

$$
I_{k_0}(\nu) = \lim_{k\to\infty} I_{k,k_0}(\nu) = \inf_{k\ge k_0} I_{k,k_0}(\nu).
$$

Since $\mathcal{W}_{k,k_0}$ only contains the components of $\Gamma_k^{-1,c}$ outside the degenerate component B_{k_0}, as k increases, I_{k,k_0} approximates the complexities of folding entropy away from the degenerate set Σ_f, i.e.,

$$
I_{k_0}(\nu) \le F_f(\nu). \tag{3.16}
$$

Proposition 3.8: *For each $k \ge 1$, $I_k(\cdot)$ is upper semi-continuous on $\mathcal{P}(M)$.*

Proof. Given any $\mu \in \mathcal{P}(M)$, without loss of generality (take a translation if necessary), we assume $\mu(\partial\Gamma_k) = \mu(\partial\,\Gamma_k^{-1,c}) = 0$ for all $k \in \mathbb{N}$.

Note that for any Borel set A with $\mu(\partial A) = 0$, we have $\nu_n(A) \to \mu(A)$ if $\nu_n \to \mu$. Hence, for any $k \geq k_0$,

$$I_{k,k_0}(\nu) = \sum_{\substack{Q\in\mathcal{W}_{k,k_0},\\ P=fQ}} -\nu(Q)\log\Big(\frac{\nu(Q)}{\nu(f^{-1}P)}\Big)$$

is continuous at μ. This implies the upper semi-continuity of $I_k = \inf_{k\geq k_0} I_{k,k_0}$ on $\mathcal{P}(M)$. ∎

3.3. Proof of Theorem 1.1. We are finally in a position to prove Theorem 1.1.

Given any $\varepsilon > 0$, by Lemma 3.4 and Lemma 3.5, for $k_0 = \min\{k_1, k_2\}$ we have

$$\Delta_k^{(1)}(\nu) < \varepsilon, \quad \sum_{k_0\leq j\leq k} \Delta_j^{(2,2)}(\nu) < \varepsilon, \quad \forall k \geq k_0.$$

By Proposition 3.6 and Lemma 3.7,

$$\Delta_k^{(2)}(\nu) \leq I_{k,k_0}(\nu) + \sum_{k_0<j\leq k} \Delta_j^{(2,2)}(\nu) < I_{k,k_0}(\nu) + \varepsilon\,.$$

It then follows from (3.6) that

$$F_f(\nu) < I_{k,k_0}(\nu) + 2\,\varepsilon, \quad \forall k \geq k_0, \tag{3.17}$$

where we enlarge k_0 if necessary.

Let $\{\mu_n\} \subseteq \mathcal{M}_{\bar{\eta}}(f)$ such that $\mu_n \to \mu$ as $n \to \infty$. Applying (3.17) with ν being $\{\mu_n\}$, together with Proposition 3.8, we have for any $k \geq k_0$

$$\limsup_{n\to\infty} F_f(\mu_n) \leq \limsup_{n\to\infty} I_{k,k_0}(\mu_n) + 2\,\varepsilon \leq I_{k,k_0}(\mu) + 2\,\varepsilon \leq F_f(\mu) + 2\,\varepsilon,$$

where the last inequality is by (3.16).

By the arbitrariness of ε, the proof of Theorem 1.1 is concluded.

4. Applications to interval maps

This section focuses on the one-dimensional setting by considering a $C^r(r > 1)$ interval map f. As applications of Theorem 1.1, the upper semi-continuity of both the metric entropy $h_\mu(f)$ and Hausdorff dimension $\dim_H(\mu)$ are proved when measures with uniform degenerate rate are considered. In achieving this, a entropy formula and a dimension formula for general f are established.

4.1. Upper semi-continuity of metric entropy. For a $C^r(r>1)$ interval map f and $\mu \in \mathcal{M}_{\rm inv}(f)$, recall the following inequalities for the metric entropy $h_\mu(f)$:

$$h_\mu(f) \le \int \max\{\lambda(f,x),0\} d\mu(x), \tag{4.1}$$

$$h_\mu(f) \le F_f(\mu) - \int \min\{\lambda(f,x),0\} d\mu(x), \tag{4.2}$$

where $\lambda(f,x)$ is the Lyapunov exponent of f at x. The first inequality in (4.1) is the well-known Margulis–Ruelle inequality [52], and (4.2) is the folding-type Ruelle inequality (1.2) applied in the one-dimensional setting. The following theorem shows that in the one-dimensional setting, the metric entropy and folding entropy are equal.

Theorem 4.1: *Let f be a $C^r(r>1)$ map on an interval I. Then for any $\mu \in \mathcal{M}_{\rm inv}(f)$, it holds that*

$$h_\mu(f) = F_f(\mu). \tag{4.3}$$

Proof. First, we have that

$$h_\mu(f) \ge F_f(\mu), \quad \forall \mu \in \mathcal{M}_{\rm inv}(f). \tag{4.4}$$

To see this, let $\{\xi_n\}_{n\ge1}$ be a sequence of increasing finite partitions of I such that $\bigvee_{n\ge1}\xi_n = \epsilon \pmod \mu$. By (2.1), for each $n \ge 1$,

$$h_\mu(f,\xi_n) = H_\mu\left(\xi_n | \bigvee_{i=1}^{\infty} f^{-i}\xi_n\right) \ge H_\mu(\xi_n|f^{-1}\epsilon),$$

where the inequality is by

$$f^{-1}\epsilon \succeq \bigvee_{i=1}^{\infty} f^{-i}\xi_n = f^{-1}\left(\bigvee_{i=0}^{\infty} f^{-i}\xi_n\right).$$

Thus,

$$h_\mu(f) = \lim_{n\to\infty} h_\mu(f,\xi_n) \ge \lim_{n\to\infty} H_\mu(\xi_n|f^{-1}\epsilon) = H_\mu(\epsilon|f^{-1}\epsilon).$$

Now, it only remains to show $h_\mu(f) \le F_f(\mu)$. Denote

$$A_1 = \{x\in I : \lambda(f,x)\le 0\} \quad \text{and} \quad A_2 = \{x \in I : \lambda(f,x) > 0\}.$$

Without loss of generality, assume that A_1 and A_2 are both positive μ-measured, since otherwise, the situations would be easier. Let $\mu^{(i)} = \mu/\mu(A_i), i = 1,2$, and write $\mu = \sum_{i=1,2}\mu(A_i)\mu^{(i)}$.

By (4.1), $h_{\mu^{(1)}}(f) = 0$. Then by (4.4),

$$F_f(\mu^{(1)}) = h_{\mu^{(1)}}(f) = 0. \tag{4.5}$$

For $\mu^{(2)}$, (4.2) directly yields that

$$h_{\mu^{(2)}}(f) \le F_f(\mu^{(2)}). \tag{4.6}$$

Since both A_1 and A_2 are f-invariant (mod μ),

$$F_f(\mu) = \int_I H_{\tilde{\mu}_x}(\epsilon) d\mu = \sum_{i=1,2} \int_{A_i} H_{\tilde{\mu}_x}(\epsilon) d\mu = \sum_{i=1,2} \mu(A_i) F_f(\mu^{(i)}),$$

where we recall that $\tilde{\mu}_x$ is the disintegration of μ along the partition $\{f^{-1}x\}$. Now, combining (4.5) and (4.6),

$$h_\mu(f) = \sum_{i=1,2} \mu^{(i)}(A_i) h_{\mu^{(i)}}(f) \le \sum_{i=1,2} \mu^{(i)}(A_i) F_f(\mu^{(i)}) = F_f(\mu). \qquad \blacksquare$$

Proof of Theorem 1.3. By (4.3), Theorem 1.3 is a direct consequence of Theorem 1.1. ∎

4.2. Upper semi-continuity of dimension. For a general differentiable map f and any ergodic $\mu \in \mathcal{M}_{\mathrm{inv}}(f)$, the Lyapunov exponents of f are constants for μ-a.e. x, and particularly in the one-dimensional setting, f admits only one Lyapunov exponent which is denoted by $\lambda(\mu)$. The main goal of this subsection is to establish the following dimension formula for any $C^r (r > 1)$ interval map which relates the Hausdorff dimension with metric entropy through the Lyapunov exponent.

Theorem 4.2: *Let f be a $C^r (r > 1)$ map on an interval I and $\mu \in \mathcal{M}_{\mathrm{erg}}(f)$ be hyperbolic. Then then μ is exact dimensional and satisfies*

$$h_\mu(f) = \lambda^+(\mu) \dim_H(\mu), \tag{4.7}$$

where $\lambda^+(\mu) = \max\{\lambda(\mu), 0\}$.

Before proceeding to its proof, two remarks concerning Theorem 4.2 are given as follows: (i) The hyperbolicity assumption of μ is sharp. There have been both $C^r (r < \infty)$ and analytical examples showing that for a non-hyperbolic ergodic measure μ with zero exponent, the local dimension may not exist almost everywhere [29, 26]; (ii) The formula (4.7) was proved by Ledrappier [28] for an interval map f under the assumption that the entropy with respect to the partition by the degenerate sets of f and f' are finite. Theorem 4.2 is a general result in this respect.

A key step in the proof of Theorem 4.2 is a generalized version of the following classical Brin–Katok formula in the general setting beyond one-dimension. Let f be a continuous map on a compact metric space X. For any $\delta > 0$, $n \in \mathbb{N}$ and $x \in X$, define

$$B_n(x,\delta) = \{y \in X : d(f^i x, f^i y) < \delta, 0 \le i \le n-1\}.$$

The following local entropy formula was established by Brin and Katok in [7].

Proposition 4.3 (Brin–Katok [7]): *Let $\mu \in \mathcal{M}_{\mathrm{inv}}(f)$. Then for μ-a.e. $x \in X$,*

$$\lim_{\delta\to 0}\liminf_{n\to\infty} -\frac{\log \mu(B_n(x,\delta))}{n} = \lim_{\delta\to 0}\limsup_{n\to\infty} -\frac{\log \mu(B_n(x,\delta))}{n} := h_\mu(f,x), \tag{4.8}$$

where the local entropy $h_\mu(f,x)$ satisfies

$$\int_X h_\mu(f,x)d\mu = h_\mu(f).$$

In particular, if μ is ergodic, then

$$h_\mu(f,x) = h_\mu(f), \quad \mu\text{-a.e. } x \in X.$$

In this paper, an integrable version of Brin–Katok formula is established (see Lemma 4.5) in which δ is replaced by any log-integrable function. To be specific, let $\mathscr{S}$ be the set of all functions $\psi : X \to \mathbb{R}^+$ satisfying

$$\int -\log \psi(x) d\mu(x) < \infty.$$

For any $\psi \in \mathscr{S}$, $n \in \mathbb{N}$, and $x \in X$ define

$$B_n(x,\psi) = \{y \in X : d(f^i x, f^i y) < \psi(f^i x), 0 \le i \le n-1\}.$$

The following was shown by Mañé:

Proposition 4.4 (Mañé [34]): *For any $\psi \in \mathscr{S}$,*

$$\int \limsup_{n\to\infty} -\frac{\log \mu(B_n(x,\psi))}{n} d\mu(x) \le h_\mu(f).$$

Lemma 4.5: *Let f be a continuous map on a compact metric space X and $\mu \in \mathcal{M}_{\mathrm{inv}}(f)$. Then given any $\psi \in \mathscr{S}$, for μ-a.e. $x \in X$,*

$$\lim_{\delta\to 0}\liminf_{n\to\infty} -\frac{\log \mu(B_n(x,\psi^\delta))}{n} = \lim_{\delta\to 0}\limsup_{n\to\infty} -\frac{\log \mu(B_n(x,\psi^\delta))}{n} = h_\mu(f,x),$$

where $\psi^\delta(x) := \min\{\psi(x), \delta\}$ and $h_\mu(f,x)$ is as in Proposition 4.3.

Proof. For one thing, since $B_n(x,\psi^\delta)\subseteq B_n(x,\delta)$,

$$\lim_{\delta\to 0}\liminf_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}\geq h_\mu(f,x) \tag{4.9}$$

which, by (4.8), yields

$$\int\lim_{\delta\to 0}\liminf_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}d\mu\geq\int h_\mu(f,x)d\mu=h_\mu(f). \tag{4.10}$$

For another, by Proposition 4.4,

$$\begin{aligned}\int\lim_{\delta\to 0}\limsup_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}d\mu&=\lim_{\delta\to 0}\int\limsup_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}d\mu\\&\leq h_\mu(f).\end{aligned} \tag{4.11}$$

Combining (4.10) and (4.11), we have

$$\begin{aligned}\int\lim_{\delta\to 0}\limsup_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}d\mu&=\int\lim_{\delta\to 0}\liminf_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}d\mu\\&=\int h_\mu(f,x)d\mu=h_\mu(f).\end{aligned}$$

Then together with (4.9), for μ-a.e. $x\in X$, it holds that

$$\lim_{\delta\to 0}\limsup_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}=\lim_{\delta\to 0}\liminf_{n\to\infty}-\frac{\log\mu(B_n(x,\psi^\delta))}{n}=h_\mu(f,x).\quad\blacksquare$$

Proof of Theorem 4.2. Denote $\alpha=\min\{r-1,1\}$ and $L=\max_{x\in I}\{|f'(x)|,1\}$. Hence, f is $C^{1+\alpha}$, i.e., for some constant $K>0$,

$$|f'(x)-f'(y)|\leq K|x-y|^\alpha,\quad\forall x,y\in I.$$

Take small $\gamma\in(0,1)$, and for any $x\in I\backslash\Sigma_f$ let $\rho(x)=(\frac{\gamma|f'(x)|}{K})^{\frac{1}{\alpha}}$. Then it holds that

$$(1-\gamma)|f'(x)|\leq|f'(y)|\leq(1+\gamma)|f'(x)|,\quad\forall\, y\in B(x,\rho(x)),$$

which implies that $f|_{B(x,\rho(x))}$ is a local diffeomorphism.

By the definition of Lyapunov exponent, for μ-a.e. $x\in I$,

$$\lambda(\mu)=\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\log|f'(f^ix)|=\int\log|f'(x)|d\mu(x),$$

or equivalently,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\log|(f^{-1}\,|_{f(B(f^ix,\rho(f^ix)))})'(f^{i+1}x)|=-\lambda(\mu).$$

Thus, there exists $N(x) \in \mathbb{N}$ such that for any $n \geq N(x)$,

$$\prod_{i=0}^{n-1} |f'(f^i x)| \in (e^{n(\lambda(\mu)-\gamma)},\, e^{n(\lambda(\mu)+\gamma)}), \tag{4.12}$$

or equivalently

$$\prod_{i=0}^{n-1} |(f^{-1}\mid_{f(B(f^i x, \rho(f^i x)))})'(f^{i+1}x)| \in (e^{n(-\lambda(\mu)-\gamma)},\, e^{n(-\lambda(\mu)+\gamma)}). \tag{4.13}$$

Let $D(x) = \min_{0 \leq i \leq N(x)}\{\rho(f^i x), 1\}$. Note that the integrability of $\log|f'(x)|$ (and thus $\log \rho(x)$) implies $\mu(\Sigma_f) = 0$, which in turns yields $D(x) > 0$ for μ-a.e. $x \in I$.

Now, for any small $\delta \in (0,1)$, denote $A_\delta = \{x \in I : \rho(x) \leq \delta\}$, and for any $n \in \mathbb{N}$, define

$$\Phi_n(x) = \prod_{0 \leq i < n : f^i x \in A_\delta} \rho(f^i x).$$

Then $\delta\Phi_n(x) \leq \rho^\delta(f^i x), 0 \leq i < n$. By the Birkhoff ergodic theorem, for μ-a.e. $x \in I$,

$$\frac{\log \Phi_n(x)}{n} \to \int_{A_\delta} \log \rho(z) d\mu(z), \quad n \to \infty. \tag{4.14}$$

Now, for any $n \in \mathbb{N}$ let

$$\begin{aligned} \tilde{r}_n &= D(x) L^{-N(x)} \delta \Phi_n(x) (1+\gamma)^{-n} e^{-n(\lambda(\mu)+\gamma)}, \\ \hat{r}_n &= (1-\gamma)^{-n} e^{-n(\lambda(\mu)-\gamma)}. \end{aligned}$$

Obviously, $\tilde{r}_n, \hat{r}_n \to 0$ as $n \to \infty$.

Claim 4.6: *For μ-a.e. $x \in I$ and $n \geq N(x)$,*

$$B(x, \tilde{r}_n) \subset B_n(x, \rho^\delta) \subset B(x, \hat{r}_n). \tag{4.15}$$

We postpone the proof of Claim 4.6 for the time being and proceed to finish the proof of Theorem 4.2. By (4.15) we have

$$\begin{aligned} \overline{\dim}_\mu(x) &= \limsup_{n\to\infty} \frac{\log \mu(B(x, \hat{r}_n))}{\log \hat{r}_n} \\ &\leq \limsup_{n\to\infty} \frac{\log \mu(B_n(x, \rho^\delta))}{-n} \frac{1}{\log(1-\gamma) + \lambda(\mu) - \gamma}. \end{aligned}$$

Then Lemma 4.5, together with the arbitrariness of δ and γ, yields

$$\overline{\dim}_\mu(x) \le \frac{h_\mu(f,x)}{\lambda(\mu)}.$$

Since μ is ergodic, it follows from Propositions 2.2 and 4.3 that

$$\overline{\dim}_\mu(x) \le \frac{h_\mu(f)}{\lambda(\mu)}. \tag{4.16}$$

On the other hand, (4.15) also yields

$$\begin{aligned}\underline{\dim}_\mu(x) &= \liminf_{n\to\infty} \frac{\log \mu(B(x,\tilde{r}_n))}{\tilde{r}_n} \\ &\ge \liminf_{n\to\infty} \frac{\log \mu(B_n(x,\rho^\delta))}{-n} \frac{1}{\frac{\log \Phi_n(x)}{n} + \log(1+\gamma) + \lambda(\mu) + \gamma} \\ &= \liminf_{n\to\infty} \frac{\log \mu(B_n(x,\rho^\delta))}{-n} \cdot \frac{1}{\int_{A_\delta} \log \rho(z) d\mu(z) + \lambda(\mu) + \gamma},\end{aligned}$$

where the last equality is from (4.14). Note that $\mu(A_\delta) \to 0$ as $\delta \to 0$. Thus, $\int_{A_\delta} \log \rho(z) d\mu(z)$ can be arbitrarily small by choosing small $\delta \in (0,1)$. Again, applying Lemma 4.5, Propositions 2.2 and 4.3, together with the arbitrariness of δ and γ, we have

$$\underline{\dim}_\mu(x) \ge \frac{h_\mu(f)}{\lambda(\mu)}. \tag{4.17}$$

Combining (4.16), (4.17) and Proposition 2.2, Theorem 4.2 is concluded. ∎

Proof of Claim 4.6. To prove $B(x,\tilde{r}_n) \subset B_n(x,\rho^\delta)$, we only need to show that

$$f^i(B(x,\tilde{r}_n)) \subset B(f^i x, \rho^\delta(f^i x)), \quad 0 \le i < n. \tag{4.18}$$

First, it is not hard to see that (4.18) holds for $0 \le i \le N(x)$. Now, assume that for $N(x) < i \le n$, (4.18) holds for $j = 0,1,\ldots,i-1$. Then for any $y \in B(x,\tilde{r}_n)$, by (4.12) we have

$$\prod_{j=0}^{i-1} |f'(f^j y)| \le \left(\prod_{j=0}^{i-1} |f'(f^j x)|\right)(1+\gamma)^i \le e^{i(\lambda(\mu)+\gamma)}(1+\gamma)^i \le e^{n(\lambda(\mu)+\gamma)}(1+\gamma)^n.$$

Thus,

$$f^i(B(x,\tilde{r}_n)) \subset B(f^i x, D(x)L^{-N(x)}\delta\Phi_n(x)) \subset B(f^i x, \delta\Phi_n(x)) \subset B(f^i x, \rho^\delta(f^i x)).$$

Then the induction establishes (4.18).

To establish $B_n(x, \rho^\delta) \subset B(x, \hat{r}_n)$, note that for any $y \in f^n(B_n(x, \rho^\delta))$,

$$f^i y \in B(f^i x, \rho(f^i x)), \quad 0 \leq i \leq n,$$

and hence for $0 \leq i \leq n$,

$$|f^{-1}|_{f(B(f^i(x),\rho(f^i x)))})'(f^{i+1}y)| \leq (1-\gamma)^{-1}|f^{-1}|_{f(B(f^i(x),\rho(f^i x)))})'(f^{i+1}x)|.$$

Then (4.13) yields

$$\prod_{0 \leq i < n} |(f^{-1}|_{f(B(f^i(x),\rho(f^i x)))})'(f^{i+1}y)| \leq (1-\gamma)^{-n} e^{-n(\lambda(\mu)-\gamma)}.$$

By noting that $f^n|_{B_n(x,\rho^\delta)}$ is a local diffeomorphism, we have

$$B_n(x, \rho^\delta) \subset B(x, \hat{r}_n).$$

This completes the proof of Claim 4.6. ∎

Proof of Theorem 1.4. First, note that the Lyapunov exponent

$$\lambda(\nu) = \int \log |f'(x)| d\nu(x)$$

is continuous on $\mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$. Then, if let $\{\mu_i\} \subseteq \mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$ be any sequence converging to some hyperbolic $\mu \in \mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$, we have that for any i sufficiently large, μ_i is hyperbolic such that $\lambda(\mu_i) > 0$ (resp. < 0) if $\lambda(\mu) > 0$ (resp. < 0). Note that if $\lambda(\mu) < 0$, then both μ and μ_i's ($i \gg 1$) are supported on some contracting periodic orbit, and hence the corresponding $\dim_H(\mu)$ and $\dim_H(\mu_i)$'s are all equal to zero, and the upper semi-continuity automatically holds. Now, assume $\lambda(\mu) > 0$. By Theorem 4.2, the dimension formula (4.7) applies to μ and all the μ_i's ($i \gg 1$). The upper semi-continuity of $\dim_H(\cdot)$ on $\mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$ is hence obtained from the upper semi-continuity of metric entropy by Theorem 1.3 and the continuity property of Lyapunov exponent on $\mathcal{M}_{\bar{\eta},\mathrm{erg}}(f)$. ∎

5. An example of interval maps

In this section, for each $r \in (1, \infty)$, we construct a C^r interval map f which admits ergodic measures with positive entropy as points of non-upper semi-continuity for the metric (or folding entropy). Before describing the example, the following quantitative properties on the metric entropy of a one-sided shift is recalled.

Proposition 5.1 (see Theorem 4.26 and its Remark in [57]): *Consider the k-full shift (Σ_k, σ). For any probability vector $(p_0, \ldots, p_{k-1})$, the Markov measure associated with the $(p_0, \ldots, p_{k-1})$-shift is ergodic with the entropy equal to $\sum_{i=0}^{k-1} -p_i \log p_i$. In particular, the metric entropy of the $(\frac{1}{k}, \ldots, \frac{1}{k})$-shift achieves the topological entropy of the full shift (Σ_k, σ) which equals $\log k$.*

By Proposition 5.1, for any full shift with topological entropy h, any real number in $(0, h]$ can be attained by an ergodic invariant measure with full support.

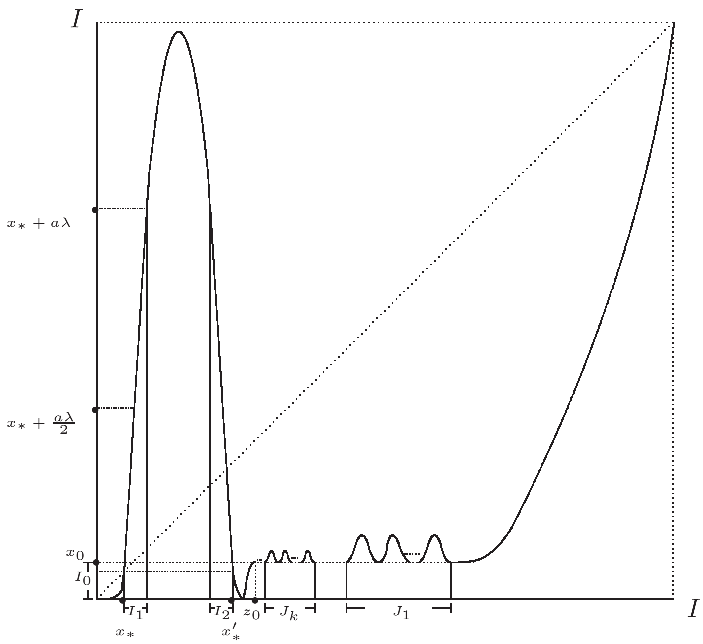

Figure 5. Accumulation of small horseshoes.

5.1. Example description. In this subsection, we describe in detail how the interval map f is constructed on the interval $I = [0, 1]$; see Figure 5 for the qualitative depiction. In Figure 5, I_1 and I_2 are two subintervals of I on which f acts linearly with slope λ 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, Λ) is uniformly expanding and conjugate to the (one-sided) full shift of two symbols. The topological entropy $h_{\text{top}}(f|_\Lambda) = \log 2$.

In the following, for any subinterval $J \subset I$, 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_* and x_*' be the left and right end point of I_1 and I_2, respectively. It is further required that x_* is a fixed point of f such that $f(x_*') = f(x_*) = x_*$. For simplicity, set $a = x_* = |I_1| = |I_2|$ and $d_H(I_1, I_2) = 3a$. Also, let λ be large enough so that

$$I_1 \cup I_2 \subset \left[x_*, x_* + \frac{a\lambda}{2}\right].$$

Since (f, Λ) is conjugate to the (one-sided) 2-full shift, by Proposition 5.1, for any $c \in (0, c_0)$ where $c_0 = \min\{\log 2, \frac{1}{r}\log\lambda\}$, there exsits $\mu \in \mathcal{M}_{\text{erg}}(f)$ supported on Λ such that

$$h_\mu(f) = c < \frac{1}{r}\log\lambda. \tag{5.1}$$

Take $\delta_0 = a/(2\lambda)$. Observing that $x_* \in \Lambda$, we can choose a generic point $x_0 \in \operatorname{supp}\mu \cap [x_*, \delta_0/2]$ such that

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

in the weak*-topology. As denoted in Figure 5, z_0 is a preimage point of x_0 lying on the right side of x_*' such that $|x_*' - z_0| = a$. Also, f can be required to remain linear with slope λ in the δ_0-neighborhood of $I_1 \cup I_2$ because $d_H(I_1, I_2) = 3a$.

Lemma 5.2: *Let $\delta_1 = x_0 - x_*$. 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. Since x_* is a fixed point, the expanding property of f on I_1 gives a certain $N_1 \in \mathbb{N}$ such that

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

By the continuity of f, there exists $\delta \in (0, \delta_1)$ such that

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

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

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

Now, consider the f-iterates of $[x_0, x_0 + \delta_1]$ which is successively cut by the interval $[x_* - \delta_0, x'_* + \delta_0]$. Then 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 δ_0 with $f^i(x_0)$ being one of its end points. In particular, let $t > t_0$. Now, we proceed according to the cases in which $f^t(x_0)$ is the left and right end point of R_t, respectively. If $f^t(x_0)$ is the left end point, we have $R_t \supset [x_0 - \delta_1 + \delta, x_0]$, and hence

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

Then $N_2 = t + N_1$ is as desired. If $f^t(x_0)$ is the right end point of R_t, we have

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

By the expanding property of f on I_1, there exists $t' \in \mathbb{N}$ such that

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

and $N_2 = t + t'$ is as desired. The proof is concluded. ∎

By Lemma 5.2 and the continuity of f, there exists $\eta > 0$ such that for any $x \in [x_0 - \eta, x_0 + \eta]$,

$$f^{N_1}([x - \delta_1, x]) \supset [z_0, z_0 + 2a], \quad f^{N_2}([x, x + \delta_1]) \supset [z_0, z_0 + 2a]. \tag{5.2}$$

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 $f^{n_k}(x_0) \in [x_0-\eta, x_0+\eta]$. Now, let $\{J_n\}_{n\geq 1}$ be a sequence of disjoint subintervals of I accumulating to z_0 from the right, satisfying the following properties:

(i) On each interval J_k, $f|_{J_k} = A_k^r \cos \omega_k(x-c_k)+(x_0+A_k^r)$, where $\{c_k\}_{k\geq 1}$ is a sequence of real numbers decreasing to z_0, and

$$A_k = \left(\frac{\delta_0}{2}\lambda^{-n_k}\right)^{\frac{1}{r}}, \quad \omega_k = \frac{L}{A_k}$$

in which $L \geq \lambda$ is chosen to satisfy $\max_{1\leq i\leq r} \sup_{x\in I} |f^{(i)}(x)| \leq L^r$.

(ii) On each interval J_k, f oscillates M_k times with

$$M_k = \frac{L\gamma_0}{2\pi k^2}\left(\frac{2}{\delta_0}\lambda^{n_k}\right)^{\frac{1}{r}}$$

in which γ_0 is a small real number.

Obviously, $x_0 \in f(J_k)$. We see that, by (i) and (ii),

$$|J_k| = \frac{2\pi M_k}{\omega_k} = \frac{\gamma_0}{k^2}.$$

Thus, by letting $\gamma_0 > 0$ sufficiently small, we have

$$\sum_{k=1}^{\infty} |J_k| = \gamma_0 \sum_{k=1}^{\infty} \frac{1}{k^2} < a.$$

Make

$$\bigcup\nolimits_{k\geq 1} J_k \subset [z_0, z_0 + 2a].$$

Outside the intervals $\{I_i\}_{i=1,2}$ and $\{J_k\}_{k\geq 1}$, f is extended smoothly as depicted in Figure 5.

5.2. Analysis of the example. In this subsection, μ is approximated by the ergodic measures supported on a sequence of horseshoes with the topological entropies having a uniform gap from the metric entropy of μ. This yields that the metric entropy is not upper semi-continuous at μ.

For an interval map g and integer $\ell \geq 2$, by an ℓ-horseshoe of g we mean a family of disjoint closed intervals $(K_1, \ldots, K_\ell)$ such that

$$g(K_i) \supseteq K_j, \quad \forall\, i, j \in \{1, \ldots, \ell\}.$$

For an interval K, g is said to admit an ℓ-horseshoe on K if $\bigcup_{1\leq i\leq \ell} K_i \subset K$. Recall that an ℓ-horseshoe is conjugate to the one-sided shift of ℓ symbols.

Lemma 5.3: *For each $k \geq 1$, either $f^{n_k+N_1+1}(J_k) \supseteq J_k$, or $f^{n_k+N_2+1}(J_k) \supseteq J_k$.*

Proof. For each $k \geq 1$, since $x_0 \in f(J_k)$, we have for any $i \geq 1$, $f^i(x_0) \in f^{i+1}(J_k)$. 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 (5.2), we have

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

In the following, only the case $f^{n_k+N_1+1}(J_k) \supseteq J_k$ is considered since the other one is similar. We see that $f^{n_k+N_1+1}|_{J_k}$ admits a $2M_k$-horseshoe. Let

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

By Proposition 5.1, for each k, there exists an ergodic measure ν_k supported on Λ_k such that

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

Thus,

$$h_{\nu_k}(f) \to \frac{1}{r}\log\lambda, \quad k \to \infty. \tag{5.3}$$

Now, we show that $\nu_k \to \mu$ as $k \to \infty$. Given any continuous functions $\varphi_1, \ldots, \varphi_s$ on X, for any $\varepsilon > 0$, there exists $\gamma > 0$ such that for any $x, y \in I$ satisfying $|x - y| < \gamma$,

$$|\varphi_i(x) - \varphi_i(y)| < \frac{\varepsilon}{2}, \quad i = 1, \ldots, s.$$

Let

$$t_k = \ln(\gamma/2A_k^r)/\ln\lambda = \frac{\ln\gamma - \ln\delta_0}{\ln\lambda} + n_k.$$

Then

$$\lim_{k\to\infty} \frac{t_k}{n_k} = 1 \quad \text{and} \quad |f^{j+1}(J_k)| \le \gamma, \quad \forall\, 0 \le j \le t_k,$$

which, by letting k large, yields

$$\lim_{k\to\infty} \left| \int \varphi_i(x) d\nu_k(x) - \frac{1}{n_k} \sum_{0 \le j \le n_k - 1} \varphi_i(f^j(x_0)) \right| \le \varepsilon, \quad i = 1, \ldots, s.$$

Since $(1/n_k)\sum_{0\le j\le n_k-1} \delta_{f^j x_0} \to \mu$ as $k \to \infty$, then together with the arbitrariness of $\varphi_1, \ldots, \varphi_s$ and ε, the convergence $\nu_k \to \mu$ is obtained. On the other hand, by (5.1) and (5.3),

$$\lim_{k\to\infty} h_{\nu_k}(f) = \frac{1}{r}\log\lambda > h_\mu(f).$$

Therefore, the metric entropy, and hence the folding entropy, is not upper semi-continuous at μ. In the following, we show that the ν_n's do not admit a uniform degenerate rate. This then demonstrates that the condition of uniform degenerate rate in Theorem 1.1 is sharp.

Recall $\Sigma_f = \{x \in I : f'(x) = 0\}$. In our example, $z_0 \in \Sigma_f$. Given a decreasing sequence of neighborhoods $\mathcal{V} = \{V_m\}_{m\ge1}$ of Σ_f, for any fixed $m \ge 1$, we have $J_k \subseteq V_m$ for all k sufficiently large. Let $y_k \in J_k$ be a generic point of ν_k, i.e.,

$$\frac{1}{t}\sum_{i=0}^{t-1} \delta_{f^t y_k} \to \nu_k, \quad t \to \infty.$$

Note that

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

Then for any fixed V_m and all large k, we have

$$\begin{aligned}
\int_{V_m} \log|f'(x)| d\nu_k &\le \frac{1}{n_k + N_1 + 1} \log(A_k^r \omega_k) \\
&= \frac{1}{n_k + N_1 + 1} \log(\frac{\delta_0}{2} \lambda^{-n_k})^{\frac{r-1}{r}} \\
&= \frac{\frac{r-1}{r} \log(\frac{\delta_0}{2})}{n_k + N_1 + 1} + \frac{\frac{r-1}{r} n_k \log \lambda^{-1}}{n_k + N_1 + 1},
\end{aligned}$$

which, as $k \to \infty$, converges to $\frac{r-1}{r} \log \lambda^{-1}$. Thus, for any $m \ge 1$,

$$\left| \int_{V_m} \log|f'(x)| d\nu_k \right| > \frac{r-1}{2r} \log \lambda, \quad \text{for all } k \text{ sufficiently large.}$$

That is, the sequence of measures $\{\nu_k\}_{k\ge 1}$ does not have a uniform degenerate rate.

Finally, we remark that in this example, the folding entropy is not upper semi-continuous on the whole space of invariant measures, but by Theorem 1.3, the upper semi-continuity holds on every subset of invariant measures with a uniform degenerate rate.

Acknowledgement. The authors are grateful to the anonymous referees whose thoughtful comments helped improve the exposition.

References

[1] L. Andrey, *The rate of entropy change in non-hamiltonian systems*, Physics Letters. A **111** (1985), 45–46.

[2] B. Bárány and A Käenmäki, *Ledrappier–Young formula and exact dimensionality of self-affine measures*, Advances in Mathematics **318** (2017), 88–129.

[3] L Barreira, Y. Pesin and J Schmeling, *Dimension and product structure of hyperbolic measures*, Annals of Mathematics **149** (1999), 755–783.

[4] M. Benedicks and L. Carleson, *On iterations of* $1-ax^2$ *on* $(-1,1)$, Annals of Mathematics **122** (1985), 1–25.

[5] R. Bowen, *Entropy expansive maps*, Transactions of the American Mathematical Society **164** (1972), 323–331.

[6] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer, Berlin, 2008.

[7] K. Brin and A. Katok, *On local entropy*, in *Geometric Dynamics (Rio de Janeiro, 1981)*, Lecture Notes in Mathematics, Vol. 1007, Springer, Berlin, 1983, pp. 30–38.

[8] D. Burguet, *Existence of measures of maximal entropy for C^r interval maps*, Proceedings of the American Mathematical Society **142** (2014), 957–968.

[9] J. Buzzi, *Intrinsic ergodicity of smooth interval maps*, Israel Journal of Mathematics **100** (1997), 125–161.

[10] W.-C. Cheng and S. Newhouse, *Pre-image entropy*, Ergodic Theory and Dynamical Systems **25** (2005), 1091–1113.

[11] P. Collet and J.-P. Eckmann, *On the abundance of aperiodic behaviour for maps on the interval*, Communications in Mathematical Physics **73** (1980), 115–160.

[12] W. Cowieson and L.-S. Young, *SRB measures as zero-noise limits*, Ergodic Theory and Dynamical Systems **25** (2005), 1115–1138.

[13] L. Díaz, T. Fisher, M. Pacifico and J. Vieitez, *Entropy-expansiveness for partially hyperbolic diffeomorphisms*, Discrete and Continuous Dynamical Systems **32** (2012), 4195–4207.

[14] J-P. Eckmann and D. Ruelle, *Ergodic theory of chaos and strange attractors*, Reviews of Modern Physics **57** (1985), 617–656.

[15] D. J. Evans, *Response theory as a free-energy extremum*, Physical Review. A **32** (1985), 2923–2925.

[16] D. J. Evans, E. G. D. Cohen and G. P. Morriss, *Viscosity of a simple fluid from its maximal Lyapunov exponents*, Physical Review. A **42** (1990), 5990–5997.

[17] D. J. Evans, E. G. D. Cohen and G. P. Morriss, *Probability of second law violations in shearing steady flows*, Physical Review Letters **71** (1993), 2401–2404.

[18] K. Falconer and X. Jin, *Exact dimensionality and projections of random self-similar measures and sets*, Journal of the London Mathematical Society **90** (2014), 388–412.

[19] D. J. Feng and H. Hu, *Dimension theory of iterated function systems*, Communications in Pure and Applied Mathematics **62** (2009), 1435–1500.

[20] G. Gallavotti, *Entropy production and thermodynamics of nonequilibrium stationary states: a point of view*, Chaos **14** (2004), 680–690.

[21] G. Gallavotti and E. G. D. Cohen, *Dynamical ensembles in nonequilibrium statistical mechanics*, Physical Review Letters **74** (1995), 2694–2697.

[22] G. Gallavotti and E. G. D. Cohen, *Dynamical ensembles in stationary states*, Journal of Statistical Physics **80** (1995), 931–970.

[23] G. Gallavotti and D. Ruelle, *SRB states and nonequilibrium statistical mechanics close to equilibrium*, Communications in Mathematical Physics **190** (1997), 279–285.

[24] W. G. Hoover and H. A. Posch, *Direct measurement of equilibrium and nonequilibrium Lyapunov spectra*, Physics Letters. A **123** (1987), 227–230.

[25] D.-Q. Jiang, M. Qian and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States*, Lecture Notes in Mathematics, Vol. 1833, Springer, Berlin, 2004.

[26] B. Kalinin and V. Sadovskaya, *On pointwise dimension of non-hyperbolic measures*, Ergodic Theory and Dynamical Systems **22** (2002), 1783–1801.

[27] A. I. Khinchin, *Mathematical Foundations of Statistical Mechanics*, Dover, New York, 1949.

[28] F Ledrappier, *Some relations between dimension and Lyapunov exponents*, Communications in Mathematical Physics **81** (1981), 229–238.

[29] F. Ledrappier and M. Misiurewicz, *Dimension of invariant measures for maps exponent zero*, Ergodic Theory and Dynamical Systems **5** (1985), 595–610.

[30] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula; II. Relations between entropy, exponents and dimension*, Annals of Mathematics **122** (1985), 509–539; 540–574.

[31] G. Liao, M. Viana and J. Yang, *The entropy conjecture for diffeomorphisms away from tangencies*, Journal of the European Mathematical Society **15** (2013), 2043–2060.

[32] G. Liao and S. Wang, *Ruelle inequality of folding type for $C^{1+\alpha}$ maps*, Mathematische Zeitshrift **290** (2018), 509–519.

[33] P.-D. Liu, *Ruelle inequality relating entropy, folding entropy and negative Lyapunov exponents*, Communications in Mathematical Physics **240** (2003), 531–538.

[34] R. Mañé, *A proof of Pesin's formula*, Ergodic Theory and Dynamical Systems **1** (1981), 95–102.

[35] R. Mañé, *The Hausdorff dimension of invariant probabilities of rational maps*, in *Dynamical Systems, Valparaiso 1986*, Lecture Notes in Mathematics, Vol. 1331, Springer, Berlin, 1986, pp. 86–117.

[36] A. Manning, *The dimension of the maximal measure for a polynomial map*, Annals of Mathematics **119** (1984), 425–430.

[37] E. Mihailescu, *Physical measures for multivalued inverse iterates near hyperbolic repellors*, Journal of Statistical Physics **139** (2010), 800–819.

[38] E. Mihailescu, *Thermodynamic formalism for invariant measures in iterated function systems with overlaps*, Communications in Contemporary Mathematics **24** (2022), 931–970.

[39] E. Mihailescu and M. Urbański, *Entropy production for a class of inverse SRB measures*, Journal of Statistical Physics **150** (2013), 881–888.

[40] E. Mihailescu and M. Urbański, *Measure-theoretic degrees and topological pressure for non-expanding transformations*, Journal of Functional Analysis **267** (2014), 2823–2845.

[41] E. Mihailescu and M. Urbański, *Random countable iterated function systems with overlaps and applications*, Advances in Mathematics **298** (2016), 726–758.

[42] M. Misiurewicz, *On non-continuity of topological entropy*, Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques **19** (1971), 319–320.

[43] M. Misiurewicz, *Diffeomorphism without any measure with maximal entropy*, Bulletin de l'Académie Polonaise des Sciences. Série des Sciences Mathématiques, Astronomiques et Physiques **21** (1973), 903–910.

[44] S. Newhouse, *Continuity properties of entropy*, Annals of Mathematics **129** (1989), 215–235.

[45] V. I. Oseledec, *A multiplicative ergodic theorem*, Trudy Moskovskogo Matematičeskogo Obščestva **19** (1968), 179–210.

[46] Y. Pesin and H. Weiss, *On the dimension of deterministic and random cantor-like sets, symbolic dynamics, and the Eckmann–Ruelle conjecture*, Communications in Mathematical Physics **182** (1996), 105–153.

[47] Y. B. Pesin, *Characteristic Lyapunov exponents and smooth ergodic theory*, Russian Mathematical Surveys **32** (1977), 55–114.

[48] Y. B. Pesin, *Dimension Theory in Dynamical Systems*, Chicago Lectures in Mathematics, University of Chicago Press, Chicago, IL, 1997.
[49] F. Przytycki and M. Urbański, *Conformal Fractals: Ergodic Theory Methods*, London Mathematical Society Lecture Note Series, Vol. 371, Cambridge University Press, Cambridge, 2010.
[50] M.-P. Qian, M. Qian and G.-L. Gong, *The reversibility and the entropy production of Markov processes*, in *Probability Theory and its Applications in China*, Contemporary Mathematics, Vol. 118, American Mathematical Society, Providence, RI, 1991, pp. 255–261.
[51] V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Russian Mathematical Surveys **22** (1967), 1–52.
[52] D. Ruelle, *An inequality for the entropy of differentiable maps*, Boletim da Sociedade Brasileira de Matemática **9** (1978), 83–87.
[53] D. Ruelle, *Positivity of entropy production in nonequilibrium statistical mechanics*, Journal of Statistical Physics **85** (1996), 1–23.
[54] D. Ruelle, *Smooth dynamics and new theoretical ideas in nonequilibrium statistical mechanics*, Journal of Statistical Physics **95** (1999), 393–468.
[55] S. Ruette, *Mixing C^r maps of the interval without maximal measure*, Israel Journal of Mathematical **127** (2002), 253–277.
[56] U. Seifert, *Stochastic thermodynamics, fluctuation theorems and molecular machines*, Reports on Progress in Physics **75** (2012), Article no. 126001.
[57] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, Vol. 79, Springer, New York–Berlin, 1982.
[58] W. Wu and Y. Zhu, *On preimage entropy, folding entropy and stable entropy*, Ergodic Theory and Dynamical Systems **41** (2021), 1217–1249.
[59] W. Wu and Y. Zhu, *Entropy via preimage structure*, Advances in Mathematics **406** (2022), Article no. 108483.
[60] Y. Yomdin, *Volume growth and entropy*, Israel Journal of Mathematics **57** (1987), 285–300.
[61] L.-S. Young, *Dimension, entropy and Lyapunov exponents*, Ergodic Theory and Dynamical Systems **2** (1982), 109–124.
[62] X.-J. Zhang, H. Qian and M. Qian, *Stochastic theory of nonequilibrium steady states and its applications. Part I*, Physics Reports **510** (2012), 1–86.