Ergod. Th. & Dynam. Sys. (2021), **41**, 1217–1249
doi:10.1017/etds.2019.114

On preimage entropy, folding entropy and stable entropy

WEISHENG WU† and YUJUN ZHU‡

† *Department of Applied Mathematics, Science College, China Agricultural University, Beijing 100083, P.R. China*
(e-mail: wuweisheng@cau.edu.cn)

‡ *School of Mathematical Sciences, Xiamen University, Xiamen 361005, P.R. China*
(e-mail: yjzhu@xmu.edu.cn)

(*Received* 5 *October* 2019 *and accepted in revised form* 2 *December* 2019)

Abstract. For non-invertible dynamical systems, we investigate how 'non-invertible' a system is and how the 'non-invertibility' contributes to the entropy from different viewpoints. For a continuous map on a compact metric space, we propose a notion of pointwise metric preimage entropy for invariant measures. For systems with uniform separation of preimages, we establish a variational principle between this version of pointwise metric preimage entropy and pointwise topological entropies introduced by Hurley [On topological entropy of maps. *Ergod. Th. & Dynam. Sys.* **15** (1995), 557–568], which answers a question considered by Cheng and Newhouse [Pre-image entropy. *Ergod. Th. & Dynam. Sys.* **25** (2005), 1091–1113]. Under the same condition, the notion coincides with folding entropy introduced by Ruelle [Positivity of entropy production in nonequilibrium statistical mechanics. *J. Stat. Phys.* **85**(1–2) (1996), 1–23]. For a C^1-partially hyperbolic (non-invertible and non-degenerate) endomorphism on a closed manifold, we introduce notions of stable topological and metric entropies, and establish a variational principle relating them. For C^2 systems, the stable metric entropy is expressed in terms of folding entropy (namely, pointwise metric preimage entropy) and negative Lyapunov exponents. Preimage entropy could be regarded as a special type of stable entropy when each stable manifold consists of a single point. Moreover, we also consider the upper semi-continuity for both of pointwise metric preimage entropy and stable entropy and give a version of the Shannon–McMillan–Breiman theorem for them.

Key words: preimage entropy, folding entropy, stable entropy, Shannon–Breiman–McMillan theorem, variational principle

2010 Mathematics Subject Classification: 37D35 (Primary); 37D30 (Secondary)

1. *Introduction*

Unlike the invertible dynamical systems, entropies, including topological entropy and measure-theoretic entropy, are more delicate for non-invertible dynamical systems. Let (X, d) be a compact metric space, f a continuous map on X and μ an f-invariant Borel probability measure. It is well known that when f is a homeomorphism we have $h_{\text{top}}(f) = h_{\text{top}}(f^{-1})$ and $h_\mu(f) = h_\mu(f^{-1})$. These equalities tell us that the structures of forward orbits and backward orbits have equal complexity from both topological and measure-theoretic points of view. However, when f is non-invertible, things become subtler and more complicated. Since the preimage of a given point is usually not single and even uncountable, the structure of the 'inverse orbits' of f is more complex. In recent years, to give quantitative estimates of how 'non-invertible' a system is, a number of different entropy-like invariants, including 'preimage entropy' and 'folding entropy', based on the preimage structure have been formulated and investigated.

Preimage entropies were introduced and studied in the 1990s by Langevin and Przytycki [**7**], Hurley [**6**], and Nitecki and Przytycki [**13**], in various forms from the topological point of view. Among these entropy-like invariants, there are two pointwise quantities,

$$\begin{aligned} h_m(f) &:= \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X} s(n, \varepsilon, f^{-n}x), \\ h_p(f) &:= \sup_{x\in X} \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log s(n, \varepsilon, f^{-n}x), \end{aligned} \tag{1}$$

where $s(n, \varepsilon, f^{-n}x)$ is the largest cardinality of any (n, ε)-separated subset of $f^{-n}x$. Clearly, $h_p(f) \le h_m(f) \le h_{\text{top}}(f)$. There are examples for which either of these inequalities is strict, and there are also many cases when the three invariants agree (see [**1**, **4**, **14**]). One can see that all these earlier works only considered the topological version of preimage entropies. A natural question is: can one introduce the counterpart of $h_m(f)$ or $h_p(f)$ from the measure-theoretic point of view, and obtain a variational principle relating them as has been done for the classical entropy? In [**2**], Cheng and Newhouse considered this topic and introduced another version of preimage entropies, $h_{\text{pre}}(f)$ and $h_{\text{pre},\mu}(f)$, in which $h_{\text{pre}}(f)$ is between $h_m(f)$ and $h_{\text{top}}(f)$, and obtained a variational principle. However, it is still unknown whether there is a variational principle for $h_m(f)$ or $h_p(f)$ (see [**2**, p. 1093]). In this paper, we introduce a new version of *pointwise metric preimage entropy*,

$$h_{m,\mu}(f) := \sup_{\alpha} \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\mathcal{B}),$$

where the supremum is taken over all finite measurable partitions and $\mathcal{B}$ is the Borel σ-algebra. Then we establish a variational principle relating $h_{m,\mu}(f)$ and $h_m(f)$ (respectively, $h_p(f)$) for f with uniform separation of preimages (Theorem B), which gives a positive answer to the above question. Moreover, we obtain the upper semi-continuity of $h_{m,\mu}(f)$ and give a version of the Shannon–McMillan–Breiman theorem for it.

Folding entropy was introduced by Ruelle [**17**] for C^1 maps to study the positivity of entropy production in non-equilibrium statistical mechanics. Now let $X = M$ be a closed Riemannian manifold and f a differentiable self-map on M. The folding entropy of an

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

f-invariant measure μ is defined as

$$\mathcal{F}_\mu(f) := H_\mu(\epsilon | f^{-1}\epsilon),$$

where ϵ is the partition of M into points. Liu [**11**] established a type of Ruelle inequality involving metric entropy, folding entropy and negative Lyapunov exponents for $C^{1+\alpha}$ maps with polynomial growth near degenerate points. Recently, Liao and Wang [**10**] established the inequality for any $C^{1+\alpha}$ maps. For non-degenerate C^2 maps, Liu [**12**] established the Pesin entropy formula; and then Shu [**18**] obtained a type of Ledrappier–Young formula [**8**, **9**] expressing metric entropy in terms of folding entropy, negative Lyapunov exponents and transversal Hausdorff dimensions along a hierarchy of stable manifolds. More precisely, assume that f is C^2 and non-degenerate, and μ is an ergodic measure with Lyapunov exponents $\lambda_1 < \lambda_2 < \cdots < \lambda_s < 0 \le \lambda_{s+1} < \cdots$. Then

$$h_\mu(f) = H_\mu(\xi | f^{-1}\xi) = \mathcal{F}_\mu(f) - \sum_{i=1}^{s} \lambda_i \gamma_i, \tag{2}$$

where ξ is a decreasing measurable partition subordinate to the stable lamination W^s of f, and γ_i $(1 \le i \le s)$ are the transversal Hausdorff dimensions of μ along a hierarchy of stable laminations W^i $(1 \le i \le s)$. Formula (2) tells us that the 'folding' contributes to the entropy.

Now there are two versions of measure-theoretic entropy-like invariants which describe the 'non-invertibility' of a system: folding entropy $\mathcal{F}_\mu(f)$ and pointwise metric preimage entropy $h_{m,\mu}(f)$. A natural question arises: is there a connection between them? In this paper, we show that they coincide with each other for continuous maps with uniform separation of preimages (Theorem A), particularly for non-degenerate C^1 endomorphisms.

The pointwise preimage entropies $h_m(f)$ and $h_p(f)$ are the quantities which measure the growth rate of the number of preimages of a single point, and they are automatically zero for homeomorphisms. In [**4**], Fiebig, Fiebig and Nitecki introduced another quantity which describes the dispersion of preimages of 'local stable sets', and this quantity may be non-zero and even equal to the topological entropy for some homeomorphisms. To be precise, let f be a continuous map on X, given $\varepsilon > 0$ and $x \in X$. Then the ε-stable set of x under f is

$$S(f, x, \varepsilon) = \{y \in X : d(f^n x, f^n y) \le \varepsilon \ \forall n = 0, 1, 2, \ldots\}.$$

The quantity in [**4**] is $\sup_{x \in M} h_s(f, x, \varepsilon)$ in which

$$h_s(f, x, \varepsilon) := \lim_{\delta \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s(n, \delta, S(f, x, \varepsilon)).$$

It is shown in [**4**, Theorem 4.1] that $\sup_{x \in M} h_s(f, x, \varepsilon) = h_{\mathrm{top}}(f)$ for all $\varepsilon > 0$ whenever X is a compact metric space of finite covering dimension. Via this quantity, Fiebig *et al* showed that when f is forward expansive, $h_p(f) = h_m(f) = h_{\mathrm{top}}(f)$.

Returning to differentiable dynamical systems, let $X = M$ be a closed Riemannian manifold and f a C^1 *partially hyperbolic endomorphism*. Since in this case stable manifolds exist for any $x \in M$ and they inherit the Riemannian structure from M, we can replace 'local stable sets' by 'local stable manifolds' in the above quantity in [**4**]. Note that stable manifolds are contained (usually properly) in stable sets. In this paper, we propose a notion of *stable entropy* for f, which is analogous to the notion of unstable entropy introduced by Hu, Hua and Wu [**5**] for C^1 partially hyperbolic diffeomorphisms. To be precise, we define the *stable metric entropy* of f with respect to an invariant measure μ as

$$h^s_\mu(f) := \sup_{\alpha,\eta} \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\eta),$$

where α and η range over all finite partitions of M with small enough diameter and certain measurable partitions subordinate to the W^s-foliation, respectively. Actually, we will show that $h^s_\mu(f)$ is independent of the choice of α and η and coincides with the partial entropy along stable foliation $h_\mu(f, \xi)$ studied by Shu [**18**] (Theorem C). Further, we obtain the upper semi-continuity of $h^s_\mu(f)$ and give a version of the Shannon–McMillan–Breiman theorem for it. We also define two types of *stable topological entropy* of f on M:

$$h^s_{p,\mathrm{top}}(f) := \lim_{\delta\to 0} \sup_{x\in M} \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log s(n, \varepsilon, f^{-n}\overline{W^s(x, \delta)})$$

and

$$h^s_{m,\mathrm{top}}(f) := \lim_{\delta\to 0} \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_{x\in M} \frac{1}{n} \log s(n, \varepsilon, f^{-n}\overline{W^s(x, \delta)}),$$

where $W^s(x, \delta)$ is the local stable manifold of x of radius δ. Then we establish a variational principle relating stable metric entropy and stable topological entropy (Theorem D). In fact, one can interpret preimage entropy as a special type of stable entropy in the scenario that each stable manifold consists of a single point, that is, f has trivial stable directions (but possibly large center directions). On the other hand, since topological entropy can be expressed as the entropy along preimages of stable sets as discussed above, we can say that the stable entropy lies between the preimage entropy and the topological entropy for C^1 partially hyperbolic endomorphisms.

Finally, combining the above properties about preimage entropy, folding entropy and stable entropy, we can get more information on how 'non-invertible' a system is and how the 'non-invertibility' contributes to the entropy, from different viewpoints. Moreover, we can observe the relation among topological entropy $h_{\mathrm{top}}(f)$ and pointwise topological preimage entropies $h_m(f)$ and $h_p(f)$ for a non-degenerate C^2 endomorphism f. Indeed, when all Lyapunov exponents of f are non-negative, $h_p(f) = h_m(f) = h_{\mathrm{top}}(f)$. By the upper semi-continuity of $h_{m,\mu}(f)$, there exists an ergodic measure of maximal preimage entropy. If such a measure has a negative Lyapunov exponent with positive transversal dimension, then we can conclude $h_p(f) = h_m(f) < h_{\mathrm{top}}(f)$ (Corollary D.1).

1.1. *Statement of main results.* We always assume that (X, d) is a compact metric space, f is a continuous self-map on X and $\mathcal{B}$ is the Borel σ-algebra. Let $\mathcal{M}_f(X)$ and $\mathcal{M}^e_f(X)$ respectively denote the set of all f-invariant and ergodic probability measures on X. Let $\mathcal{M}(X)$ denote the set of all probability measures on X.

In [**6**], Hurley introduced two pointwise preimage entropies $h_m(f)$ and $h_p(f)$ as in (1) to measure the non-invertibility of systems. It is clear that $h_p(f) \le h_m(f)$. In [**2**], Cheng and Newhouse defined the *metric preimage entropy* $h_{\mathrm{pre},\mu}(f)$ for an invariant measure

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

$\mu \in \mathcal{M}_f(X)$ by

$$h_{\text{pre},\mu}(f) = \sup_{\alpha} h_{\text{pre},\mu}(f, \alpha),$$

where α ranges over all finite partitions of X and

$$h_{\text{pre},\mu}(f, \alpha) = h_\mu(f, \alpha|\mathcal{B}^-) := \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|\mathcal{B}^-),$$

in which $\alpha_0^{n-1} = \bigvee_{i=0}^{n-1} f^{-i}\alpha$, $H_\mu(\cdot|\cdot)$ is the standard conditional entropy and $\mathcal{B}^-$ is the infinite past σ-algebra $\bigcap_{n\geq 0} f^{-n}\mathcal{B}$ related to $\mathcal{B}$. See §2.1 for more details on conditional entropy. Cheng and Newhouse also define the topological preimage entropy as

$$h_{\text{pre}}(f) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in X, k\geq n} s(n, \varepsilon, f^{-k}x).$$

Then a variational principle is obtained as follows [**2**, Theorem 2.5]:

$$h_{\text{pre}}(f) = \sup_{\mu\in\mathcal{M}_f(X)} h_{\text{pre},\mu}(f).$$

However, the variational principle for $h_m(f)$ or $h_p(f)$ is still unknown (see [**2**]). We introduce a new notion of pointwise metric preimage entropy of an invariant measure as follows.

Definition 1.1. The *pointwise metric preimage entropy of f with respect to $\mu \in \mathcal{M}_f(X)$ and a measurable partition α of X* is defined as

$$h_{m,\mu}(f, \alpha) := \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B}).$$

Then the *pointwise metric preimage entropy of f with respect to μ* is defined as

$$h_{m,\mu}(f) := \sup_{\alpha} h_{m,\mu}(f, \alpha),$$

where α ranges over all finite partitions of X.

It is easy to see that $h_{m,\mu}(f)$ is an invariant of measure-theoretic conjugacy. That is, if $(f, X, \mathcal{B}, \mu)$ and $(g, Y, \tilde{\mathcal{B}}, \nu)$ are measure-preserving transformations, and $\pi : X \to Y$ is a bimeasurable bijection (mod 0) such that $g\pi = \pi f$, then $h_{m,\mu}(f) = h_{m,\nu}(g)$.

Remark 1.2. As $\mathcal{B}^- \subset f^{-n}\mathcal{B}$ for any $n \in \mathbb{N}$, $h_{m,\mu}(f, \alpha) \leq h_{pre,\mu}(f, \alpha)$ and hence $h_{m,\mu}(f) \leq h_{pre,\mu}(f)$ for any $\mu \in \mathcal{M}_f(X)$. It is interesting to know when they coincide.

For any $\mu \in \mathcal{M}(X)$, we always consider the μ-completion of $\mathcal{B}$, which is also denoted by $\mathcal{B}$ for simplicity. For two measurable partitions α and β of X, we write $\alpha \leq \beta$ or $\beta \geq \alpha$ if β is a refinement of α. Let ϵ denote the partition of X into points. Then ϵ is a measurable partition which generates $\mathcal{B}$. In the following, we do not distinguish ϵ and $\mathcal{B}$ in the notation of conditional entropy. The following concept of folding entropy was introduced by Ruelle [**17**].

Definition 1.3. The *folding entropy of f with respect to μ* is defined as

$$\mathcal{F}_\mu(f) := H_\mu(\epsilon|f^{-1}\epsilon) = H_\mu(\mathcal{B}|f^{-1}\mathcal{B}).$$

Definition 1.4. f is said to have *uniform separation of preimages* if for some $\varepsilon_0 > 0$, $d(x, y) \leq \varepsilon_0$ and $fx = fy$ implies $x = y$. Then ε_0 is said to be an *exponent* of separation for f.

THEOREM A. *For any continuous map $f : X \to X$, we have*

$$h_{m,\mu}(f) \leq \mathcal{F}_\mu(f).$$

If we assume further that f has uniform separation of preimages, then

$$h_{m,\mu}(f) = \mathcal{F}_\mu(f).$$

COROLLARY A.1. *If $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$, then for any finite Borel partition α of X with* $\operatorname{diam}(\alpha) < \varepsilon_0$, *we have*

$$h_{m,\mu}(f, \alpha) = h_{m,\mu}(f) = \mathcal{F}_\mu(f).$$

THEOREM B. (Variational principle) *Let $f : X \to X$ be a continuous map. Then*

$$h_m(f) \geq \sup_{\mu \in \mathcal{M}_f(X)} h_{m,\mu}(f).$$

In particular, if f has uniform separation of preimages, then

$$h_p(f) = h_m(f) = \sup_{\mu \in \mathcal{M}_f(X)} h_{m,\mu}(f).$$

Moreover, $\mathcal{M}_f(X)$ can be replaced by $\mathcal{M}_f^e(X)$ in the above inequality/equality.

Remark 1.5. In [**4**], it is shown that $h_p(f) = h_m(f)$ if f is forward expansive. Theorem B generalizes this result for f with uniform separation of preimages.

Remark 1.6. For f with uniform separation of preimages, we can show that the function $\mu \mapsto h_{m,\mu}(f)$ is upper semi-continuous (Proposition 2.13). As a consequence, there exists a measure ν of maximal preimage entropy, that is, $h_{m,\nu}(f) = h_m(f)$. A version of the Shannon–McMillan–Breiman theorem can also be established (Theorem 2.14).

We now consider a C^1 endomorphism $f : M \to M$, where M is a C^∞ closed Riemannian manifold. In this paper, we focus on the case where f is non-degenerate, that is, $\det D_x f \neq 0$ for any $x \in M$. Then f has uniform separation of preimages (see Lemma 2.6). To give the definition of partially hyperbolic endomorphisms, we recall the inverse limit space as follows.

Let $M^{\mathbb{Z}}$ be the infinite product space of M endowed with the product topology and the metric $\tilde{d}(\tilde{x}, \tilde{y}) = \sum_{n=-\infty}^{\infty} 2^{-|n|} d(x_n, y_n)$ for $\tilde{x} = \{x_n\}_{n=-\infty}^{\infty}$ and $\tilde{y} = \{y_n\}_{n=-\infty}^{\infty}$. The *inverse limit space*, denoted by M^f, is a subspace of the product space $M^{\mathbb{Z}}$ such that $fx_n = x_{n+1}$, $n \in \mathbb{Z}$, for any $\tilde{x} = \{x_n\}_{n=-\infty}^{+\infty} \in M^f$. It is clear that M^f is a closed subset of $M^{\mathbb{Z}}$. Let $\Pi : M^f \to M$ be the projection such that for $\tilde{x} = \{x_n\}_{n=-\infty}^{+\infty}$, $\Pi(\tilde{x}) = x_0$. Let τ be the left shift map on M^f.

Consider the pullback bundle $E = \Pi^* TM$. The tangent map Df on TM induces a fiber-preserving map on E with respect to the left shift map τ, defined by $\Pi^* \circ Df \circ \Pi_*$ and also denoted by Df.

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Definition 1.7. A non-degenerate C^1 endomorphism $f : M \to M$ is said to be *partially hyperbolic*, if there exists a non-trivial continuous splitting $E = E^s \oplus E^c \oplus E^u$ of the pullback bundle E into stable, center, and unstable distributions on M^f such that:

(i) the splitting is Df-invariant, that is, $D_{\tilde{x}} f(E^\sigma(\tilde{x})) = E^\sigma(\tau(\tilde{x}))$ $(\sigma = c, s, u)$ for any $\tilde{x} \in M^f$;

(ii) all unit vectors $v^\sigma \in E^\sigma(\tilde{x})$ $(\sigma = c, s, u)$ with $\tilde{x} \in M^f$ satisfy

$$\|D_{\tilde{x}} f v^s\| < \|D_{\tilde{x}} f v^c\| < \|D_{\tilde{x}} f v^u\|;$$

(iii) $\|D_{\tilde{x}} f|_{E^s(\tilde{x})}\| < 1$ and $\|D_{\tilde{x}} f|_{E^u(\tilde{x})}\| > 1$ for any $\tilde{x} \in M^f$.

From (iii) in the definition, we know that the unstable distribution E^u may depend on the past. However, this cannot happen for the stable distribution, that is, $E^s(\tilde{x})$ depends only on x_0. Both stable and unstable distributions E^s and E^u are uniquely integrable to the stable and unstable foliations $\tilde{W}^s$ and $\tilde{W}^u$, with $T\tilde{W}^s = E^s$ and $T\tilde{W}^u = E^u$, respectively (see [**15**, p. 30]). $\tilde{W}^s$ only depends on the future, and there exists a stable foliation W^s on M with $W^s = \Pi \tilde{W}^s$. In contrast, generally there is no unstable foliation on M. We mainly work on the stable foliation W^s. Let d^s be the metric induced by the Riemannian structure on any stable manifold, and $W^s(x, \delta)$ denote the open ball in $W^s(x)$ of radius $\delta > 0$ with respect to d^s.

We study partial entropy caused by the stable direction for a C^1 partially hyperbolic endomorphism, which is a natural generalization of results by Ledrappier and Young [**8**, **9**] for the diffeomorphism case. We first recall the partial entropy following Liu [**12**] and Shu [**18**]. The definition involves a type of deceasing partitions subordinate to stable manifolds, which we describe as follows.

PROPOSITION 1.8. [**18**, Proposition 2.6] *Given $\mu \in \mathcal{M}_f(M)$, there exists a measurable partition ξ of M which has the following properties:*

(i) *ξ is decreasing, that is, $f^{-1}\xi \leq \xi$;*

(ii) *$\bigvee_{n=0}^{\infty} \tau^n(\Pi^{-1}\xi) = \epsilon$, where ϵ is the partition of M^f into single points;*

(iii) *ξ is subordinate to the W^s-foliation of f with respect to a measure μ, that is, for μ-almost every $x \in M$, $\xi(x)$ contains an open neighborhood of x in $W^s(x)$.*

We denote by $\mathcal{Q}^s = \mathcal{Q}^s(\mu)$ the set of all decreasing measurable partitions subordinate to the W^s-foliation as in Proposition 1.8. Define $h_\mu(f, \xi) := H_\mu(\xi | f^{-1}\xi)$, which is independent of the choice of ξ as long as $\xi \in \mathcal{Q}^s$.

Assume that an ergodic measure μ has negative Lyapunov exponents $\lambda_1 < \lambda_2 < \cdots < \lambda_s < 0$ in the stable direction. By Shu [**18**], if f is C^2, then

$$h_\mu(f) = h_\mu(f, \xi) = \mathcal{F}_\mu(f) - \sum_{i=1}^{s} \lambda_i \gamma_i,$$

where γ_i is the transversal Hausdorff dimension of μ on ith Pesin stable manifold $W^i(x)$ $(1 \leq i \leq s)$.

In the following, we construct a new type of measurable partition subordinate to the W^s-foliation. Fix $\varepsilon_0 > 0$ small enough and let α be a finite partition of M with $\mathrm{diam}(\alpha) \ll \varepsilon_0$. Denote

$$\eta(x) = \alpha(x) \cap W^s(x, \varepsilon_0), \quad \text{for all } x \in M,$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

where $\alpha(x)$ is the element of α containing x. Then $\eta = \{\eta(x) : x \in M\}$ is a measurable partition of M. By continuity of the W^s-foliation, if $\mu(\partial\alpha) = 0$, η is a measurable partition subordinate to the W^s-foliation, where $\partial\alpha = \bigcup_{A\in\alpha} \partial A$ and, for $B \subset M$, ∂B means the boundary of B. Let $\mathcal{P}$ denote the set of all finite partitions with small enough diameter and $\mathcal{P}^s$ denote the set of measurable partitions of M subordinate to the W^s-foliation which are induced by finite partitions in $\mathcal{P}$.

Definition 1.9. The *conditional entropy* of f for a finite measurable partition α of M with respect to $\eta \in \mathcal{P}^s$ is defined as

$$h_\mu(f, \alpha|\eta) = \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|f^{-n}\eta).$$

The conditional entropy of f with respect to η is defined as

$$h_\mu(f|\eta) = \sup_{\alpha\in\mathcal{P}} h_\mu(f, \alpha|\eta),$$

and the *stable metric entropy* of f is defined as

$$h_\mu^s(f) = \sup_{\eta\in\mathcal{P}^s} h_\mu(f|\eta).$$

THEOREM C. *Let f be a C^1 non-degenerate partially hyperbolic endomorphism and μ an ergodic measure of f. Then for any $\xi \in \mathcal{Q}^s$, $\eta \in \mathcal{P}^s$ and $\alpha \in \mathcal{P}$,*

$$h_\mu(f, \xi) = h_\mu(f, \alpha|\eta).$$

We may interpret Theorem C as a generalization of Theorem A in the following sense. If each stable manifold consists of a single point, $h_\mu(f, \xi)$ can be regarded as the folding entropy, while $h_\mu(f, \alpha|\eta)$ corresponds to the pointwise preimage entropy.

Two corollaries can be obtained directly as follows.

COROLLARY C.1. *We have $h_\mu^s(f) \le h_\mu(f)$. Moreover, if f is C^2, then $h_\mu(f) = h_\mu^s(f) - \sum_{\lambda_i^c<0} \lambda_i^c \gamma_i^c$. In particular, if there is no negative Lyapunov exponent in the center direction at μ-a.e. $x \in M$, then $h_\mu^s(f) = h_\mu(f)$.*

COROLLARY C.2. *Suppose that $\mu \in \mathcal{M}_f^e(M)$ is ergodic. Then for any $\alpha \in \mathcal{P}$ and $\eta \in \mathcal{P}^s$, we have*

$$h_\mu^s(f) = h_\mu(f, \alpha|\eta) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|f^{-n}\eta).$$

We now give the definition of stable topological entropy for endomorphisms. Let $W^s(x, \delta)$ be the open ball inside $W^s(x)$ centered at x of radius $\delta > 0$ with respect to the metric d^s.

Definition 1.10. We define two types of *stable topological entropy* of f on M. The first is defined as

$$h_{p,\mathrm{top}}^s(f) := \lim_{\delta\to 0} \sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x, \delta)}),$$

where

$$h_{\mathrm{top}}^s(f, \overline{W^s(x, \delta)}) := \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log s(n, \varepsilon, f^{-n}\overline{W^s(x, \delta)}).$$

The second is defined as

$$h_{m,\mathrm{top}}^s(f) := \lim_{\delta\to 0} \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \sup_{x\in M} \frac{1}{n} \log s(n, \varepsilon, f^{-n}\overline{W^s(x, \delta)}).$$

THEOREM D. *Let $f : M \to M$ be a C^1 non-degenerate partially hyperbolic endomorphism. Then*

$$h^s_{p,\mathrm{top}}(f) = h^s_{m,\mathrm{top}}(f) = \sup\{h^s_\mu(f) : \mu \in \mathcal{M}_f(M)\}.$$

Moreover,

$$h^s_{p,\mathrm{top}}(f) = h^s_{m,\mathrm{top}}(f) = \sup\{h^s_\nu(f) : \nu \in \mathcal{M}^e_f(M)\}.$$

COROLLARY D.1. *Assume that $f : M \to M$ is a C^2 non-degenerate endomorphism.*

(1) *If all Lyapunov exponents of f are non-negative, then $h_p(f) = h_m(f) = h_{\mathrm{top}}(f)$.*

(2) *Suppose that for an ergodic measure of maximal preimage entropy there exists a negative Lyapunov exponent with positive transversal dimension. Then $h_p(f) = h_m(f) < h_{\mathrm{top}}(f)$.*

2. *Pointwise preimage entropy*

2.1. *Preliminaries on conditional entropy.* Recall that for a measurable partition $\mathcal{P}$ of a measure space $(X, \mathcal{A})$ and a finite Borel measure μ on X, the *canonical system of conditional measures for μ and $\mathcal{P}$* is a family of probability measures $\{\mu^{\mathcal{P}}_x : x \in X\}$ with $\mu^{\mathcal{P}}_x(\mathcal{P}(x)) = 1$ for μ-a.e. x, such that for every measurable set $B \subset X$, $x \mapsto \mu^{\mathcal{P}}_x(B)$ is $\mathcal{B}(\mathcal{P})$-measurable and

$$\mu(B) = \int_X \mu^{\mathcal{P}}_x(B)\, d\mu(x),$$

where $\mathcal{B}(\mathcal{P})$ is the sub-σ-algebra of elements of $\mathcal{A}$ which are unions of elements of $\mathcal{P}$. See, for example, [**16**] for reference.

THEOREM 2.1. (Rokhlin, cf. [**16**]) *If $\mathcal{P}$ is a measurable partition, then there exists a system of conditional measures relative to $\mathcal{P}$. It is essentially unique in the sense that two such systems coincide in a set of full μ-measure.*

Assume that (X, d) is a compact metric space and $f : X \to X$ is a continuous map. The *information function* of a measurable partition α of X is defined as

$$I_\mu(\alpha)(x) = -\log \mu(\alpha(x)).$$

The *conditional information function* of α with respect to a measurable partition η of X is defined as

$$I_\mu(\alpha|\eta)(x) = -\log \mu^\eta_x(\alpha(x)).$$

Then the *entropy* $H_\mu(\alpha)$ and the *conditional entropy* $H_\mu(\alpha|\eta)$ are defined as

$$H_\mu(\alpha) = \int_X I_\mu(\alpha)(x)\, d\mu(x) \quad \text{and} \quad H_\mu(\alpha|\eta) = \int_X I_\mu(\alpha|\eta)(x)\, d\mu(x),$$

respectively.

The following facts about conditional entropy are important in our computation in the sequel. For the proof, see [**16**] or [**5**]. Assume that $f : X \to X$ is a homeomorphism in the remainder of this subsection.

LEMMA 2.2. *Let $\mu \in \mathcal{M}(X)$ and α, β, γ be measurable partitions of X with $H_\mu(\alpha|\gamma)$, $H_\mu(\beta|\gamma) < \infty$.*

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

(i) *If* $\alpha \leq \beta$, *then* $I_\mu(\alpha|\gamma)(x) \leq I_\mu(\beta|\gamma)(x)$ *and* $H_\mu(\alpha|\gamma) \leq H_\mu(\beta|\gamma)$.
(ii) $I_\mu(\alpha \vee \beta|\gamma)(x) = I_\mu(\alpha|\gamma)(x) + I_\mu(\beta|\alpha \vee \gamma)(x)$ *and* $H_\mu(\alpha \vee \beta|\gamma) = H_\mu(\alpha|\gamma) + H_\mu(\beta|\alpha \vee \gamma)$.
(iii) $H_\mu(\alpha \vee \beta|\gamma) \leq H_\mu(\alpha|\gamma) + H_\mu(\beta|\gamma)$.
(iv) *If* $\beta \leq \gamma$, *then* $H_\mu(\alpha|\beta) \geq H_\mu(\alpha|\gamma)$.

LEMMA 2.3. *Let* $\mu \in \mathcal{M}_f(X)$, *and* α, β, γ *be measurable partitions of* X *with* $H_\mu(\alpha|\gamma)$, $H_\mu(\beta|\gamma) < \infty$. *Then we have*
(i) $I_\mu(\beta_0^{n-1}|\gamma)(x) = I_\mu(\beta|\gamma)(x) + \sum_{i=1}^{n-1} I_\mu(\beta|f^i(\beta_0^{i-1} \vee \gamma))(f^i(x))$, *hence*

$$H_\mu(\beta_0^{n-1}|\gamma) = H_\mu(\beta|\gamma) + \sum_{i=1}^{n-1} H_\mu(\beta|f^i(\beta_0^{i-1} \vee \gamma));$$

(ii) $I_\mu(\alpha_0^{n-1}|\gamma)(x) = I_\mu(\alpha|f^{n-1}\gamma)(f^{n-1}(x)) + \sum_{i=0}^{n-2} I_\mu(\alpha|\alpha_1^{n-1-i} \vee f^i\gamma)(f^i(x))$, *hence*

$$H_\mu(\alpha_0^{n-1}|\gamma) = H_\mu(\alpha|f^{n-1}\gamma) + \sum_{i=0}^{n-2} H_\mu(\alpha|\alpha_1^{n-1-i} \vee f^i\gamma).$$

LEMMA 2.4. *Let* α *be a measurable partition of* X *and* $\{\zeta_n\}$ *be a sequence of increasing measurable partitions of* X *with* $\zeta_n \nearrow \zeta$. *If* $H_\mu(\alpha|\zeta_1) < \infty$, *then for* $\varphi_n(x) = I_\mu(\alpha|\zeta_n)(x)$, $\varphi^* := \sup_n \varphi_n \in L^1(\mu)$.

LEMMA 2.5. *Let* α *be a finite Borel partition of* X *and* $\{\zeta_n\}$ *be a sequence of increasing measurable partitions of* X *with* $\zeta_n \nearrow \zeta$. *Then:*
(i) $\lim_{n\to\infty} I_\mu(\alpha|\zeta_n)(x) = I_\mu(\alpha|\zeta)(x)$ *for* μ*-a.e.* $x \in X$; *and*
(ii) $\lim_{n\to\infty} H_\mu(\alpha|\zeta_n) = H_\mu(\alpha|\zeta)$.

2.2. *Connection to folding entropy.* Recall that $f : X \to X$ is said to have uniform separation of preimages if for some $\varepsilon_0 > 0$, $d(x, y) \leq \varepsilon_0$, and $fx = fy$ implies $x = y$. Denote the cardinality of a set A by $\#A$. The following lemma gives some examples and basic properties of such systems, whose proof is immediate from the very definition and omitted here.

LEMMA 2.6.
(1) *Assume that* f *is forward expansive, namely, for some* $\delta_0 > 0$, $d(f^n x, f^n y) \leq \delta_0$ *for any* $n \in \mathbb{N}$ *implies* $x = y$. *Then* f *has uniform separation of preimages.*
(2) *If* $f : M \to M$ *is a* C^1 *non-degenerate endomorphism, or, more generally, if* $f : M \to M$ *is a local homeomorphism, then* f *has uniform separation of preimages.*
(3) *If* f *has uniform separation of preimages with exponent* $\varepsilon_0 > 0$, *then* $f^{-n}x$ *is* (n, ε)*-separated for any* $x \in X$, $n \in \mathbb{N}$ *and* $0 < \varepsilon < \varepsilon_0$.
(4) *If* f *has uniform separation of preimages, then* $x \mapsto \#f^{-n}x$ *is upper semi-continuous, that is, if* $x_i \to x$ *as* $i \to \infty$, *then*

$$\limsup_{i\to\infty} \#f^{-n}x_i \leq \#f^{-n}x.$$

The following lemma is crucial. Recall that ϵ is the partition of X into points.

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

LEMMA 2.7. *Assume that $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$. Then for any finite Borel partition α of X with* $\mathrm{diam}(\alpha) < \varepsilon_0$, *we have*

$$\alpha_0^{n-1} \vee f^{-n}\epsilon = \epsilon$$

for any $n \in \mathbb{N}$.

Proof. It is clear that $\alpha_0^{n-1} \vee f^{-n}\epsilon \leq \epsilon$ for any $n \in \mathbb{N}$. It is sufficient to prove the other direction.

Let $y \in (\alpha_0^{n-1} \vee f^{-n}\epsilon)(x)$. Then $f^i y \in \alpha(f^i x)$ for any $0 \leq i \leq n-1$. As $\mathrm{diam}(\alpha) < \varepsilon_0$, we know that $d(f^i y, f^i x) < \varepsilon_0$ for any $0 \leq i \leq n-1$. On the other hand, $y \in (f^{-n}\epsilon)(x)$, that is, $f^n y = f^n x$. This, together with $d(f^{n-1}y, f^{n-1}x) < \varepsilon_0$, implies $f^{n-1}y = f^{n-1}x$, since f has the uniform separation of preimages. By induction, we have $y = x$. Thus $\epsilon \leq \alpha_0^{n-1} \vee f^{-n}\epsilon$, and the lemma follows. □

Proof of Theorem A. It is clear that $\epsilon > f^{-1}\epsilon > f^{-2}\epsilon > \cdots$. By invariance of μ, we have

$$H_\mu(\epsilon | f^{-1}\epsilon) = H_\mu(f^{-1}\epsilon | f^{-2}\epsilon) = H_\mu(f^{-2}\epsilon | f^{-3}\epsilon) = \cdots .$$

It follows that

$$\mathcal{F}_\mu(f) = H_\mu(\epsilon | f^{-1}\epsilon) = \frac{1}{n} H_\mu(\epsilon | f^{-n}\epsilon).$$

Since $\alpha_0^{n-1} \leq \epsilon$, we have

$$h_{m,\mu}(f, \alpha) = \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\epsilon) \leq \limsup_{n\to\infty} \frac{1}{n} H_\mu(\epsilon | f^{-n}\epsilon) = \mathcal{F}_\mu(f).$$

So

$$h_{m,\mu}(f) = \sup_\alpha h_{m,\mu}(f, \alpha) \leq \mathcal{F}_\mu(f).$$

Now assume that f has uniform separation of preimages with exponent ε_0. Then for any finite partition α with $\mathrm{diam}(\alpha) < \varepsilon_0$, $\alpha_0^{n-1} \vee f^{-n}\epsilon = \epsilon$ by Lemma 2.7. Hence

$$H_\mu(\alpha_0^{n-1} | f^{-n}\epsilon) = H_\mu(\alpha_0^{n-1} \vee f^{-n}\epsilon | f^{-n}\epsilon) = H_\mu(\epsilon | f^{-n}\epsilon) = n\mathcal{F}_\mu(f).$$

Therefore

$$h_{m,\mu}(f, \alpha) = \limsup_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\epsilon) = \mathcal{F}_\mu(f). \tag{3}$$

This completes the proof of Theorem A. □

Proof of Corollary A.1. The corollary follows from (3) in the proof of Theorem A. □

2.3. *Basic properties.*

In this subsection we collect some properties of pointwise metric preimage entropy. Some of these will be used in the proof of Theorem B.

2.3.1. *Power rule and product rule.*

LEMMA 2.8. *Assume that $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$, α is a finite partition of X with* $\mathrm{diam}(\alpha) < \varepsilon_0$, *and $\mu \in \mathcal{M}_f(X)$. Then $a_n := H_\mu(\alpha_0^{n-1} | f^{-n}\mathcal{B})$ is a subadditive sequence, that is, $a_{m+n} \leq a_m + a_n$ for any $m, n \geq 1$.*

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Proof. By Lemma 2.7, $f^{-m}\alpha_0^{n-1} \vee f^{-(m+n)}\mathcal{B} = f^{-m}\mathcal{B}$ for any $m, n \geq 1$. Then we have

$$\begin{aligned} a_{m+n} &= H_\mu(\alpha_0^{m+n-1} | f^{-(m+n)}\mathcal{B}) \\ &= H_\mu(f^{-m}\alpha_0^{n-1} | f^{-(m+n)}\mathcal{B}) + H_\mu(\alpha_0^{m-1} | f^{-m}\alpha_0^{n-1} \vee f^{-(m+n)}\mathcal{B}) \\ &= H_\mu(\alpha_0^{n-1} | f^{-n}\mathcal{B}) + H_\mu(\alpha_0^{m-1} | f^{-m}\mathcal{B}) \\ &= a_n + a_m, \end{aligned}$$

which proves the lemma. □

The following is a standard consequence of the subadditivity of a_n.

COROLLARY 2.9. *Assume that $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$, α is a finite partition of X with* diam$(\alpha) < \varepsilon_0$, *and $\mu \in \mathcal{M}_f(X)$. Then*

$$h_{m,\mu}(f, \alpha) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\mathcal{B}) = \inf_{n\geq 1} \frac{1}{n} H_\mu(\alpha_0^{n-1} | f^{-n}\mathcal{B}).$$

PROPOSITION 2.10. (Power rule) *Assume that $f : X \to X$ has uniform separation of preimages and $\mu \in \mathcal{M}_f(X)$. Then*

$$h_{m,\mu}(f^l) = l h_{m,\mu}(f)$$

for any $l \in \mathbb{N}$.

Proof. Let α be a finite partition of X with diam$(\alpha) < \varepsilon_0$, where ε_0 is an exponent of separation for f. Then

$$H_\mu\bigg(\bigvee_{i=0}^{n-1} f^{-il}\alpha_0^{l-1} | f^{-nl}\mathcal{B}\bigg) = H_\mu(\alpha_0^{nl-1} | f^{-nl}\mathcal{B}).$$

Dividing by n and taking the limit $n \to \infty$, by Corollary 2.9 we get

$$h_{m,\mu}(f^l, \alpha_0^{l-1}) = l h_{m,\mu}(f, \alpha).$$

By Corollary A.1, we have $h_{m,\mu}(f^l) = l h_{m,\mu}(f)$. □

PROPOSITION 2.11. (Product rule) *Let $(f, X, \mathcal{B}, \mu)$ and $(g, Y, \tilde{\mathcal{B}}, \nu)$ be two measure-preserving and continuous maps with uniform separation of preimages. Then*

$$h_{m,\mu\times\nu}(f \times g) = h_{m,\mu}(f) + h_{m,\nu}(g).$$

Proof. Note that for any finite partitions α, γ of X and $\tilde{\alpha}, \tilde{\gamma}$ of Y, we have

$$\begin{aligned} H_{\mu\times\nu}(\alpha \times \tilde{\alpha} | \gamma \times \tilde{\gamma}) &= H_{\mu\times\nu}((\alpha \times \tilde{\alpha}) \vee (\gamma \times \tilde{\gamma})) - H_{\mu\times\nu}(\gamma \times \tilde{\gamma}) \\ &= H_\mu(\alpha|\gamma) + H_\nu(\tilde{\alpha}|\tilde{\gamma}). \end{aligned}$$

Now choose two increasing sequences of finite Borel partitions $\gamma_1 \leq \gamma_2 \leq \cdots$ of X and $\tilde{\gamma}_1 \leq \tilde{\gamma}_2 \leq \cdots$ of Y with diameters tending to zero for which $\mathcal{B} = \mathcal{B}(\bigvee_{j=1}^\infty \gamma_j)$ and $\tilde{\mathcal{B}} = \mathcal{B}(\bigvee_{j=1}^\infty \tilde{\gamma}_j)$. Then we have, for any $n \in \mathbb{N}$,

$$H_{\mu\times\nu}((\alpha \times \tilde{\alpha})_0^{n-1} | (f \times g)^{-n}(\gamma_j \times \tilde{\gamma}_j)) = H_\mu(\alpha_0^{n-1} | f^{-n}\gamma_j) + H_\nu(\tilde{\alpha}_0^{n-1} | g^{-n}\tilde{\gamma}_j).$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

By Lemma 2.5(ii), taking the limit as $j \to \infty$ gives

$$H_{\mu\times\nu}((\alpha\times\tilde{\alpha})_0^{n-1}|(f\times g)^{-n}(\mathcal{B}\times\tilde{\mathcal{B}})) = H_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B}) + H_\nu(\tilde{\alpha}_0^{n-1}|g^{-n}\tilde{\mathcal{B}}).$$

Dividing by n and taking the limit as $n \to \infty$, we have

$$h_{m,\mu\times\nu}(f\times g, \alpha\times\tilde{\alpha}) = h_{m,\mu}(f,\alpha) + h_{m,\nu}(g,\tilde{\alpha}).$$

Since $h_{m,\mu\times\nu}(f\times g)$ can be computed as the supremum over product partitions $\alpha\times\tilde{\alpha}$, the product rule then follows. □

2.3.2. *Affinity and upper semi-continuity.* Recall that $\mathcal{M}_f(X)$ and $\mathcal{M}_f^e(X)$ denote the set of all f-invariant and ergodic probability measures on X, respectively. Let $\mathcal{M}(X)$ denote the set of all probability measures on X.

Any measurable partition γ generates a sub-σ-algebra $\mathcal{B}(\gamma)$, that is, $\mathcal{B}(\gamma)$ is the smallest sub-σ-algebra that contains the elements in the partition γ. Clearly, if $\{\gamma_j\}$ is a sequence of increasing measurable partitions, then $\{\mathcal{B}(\gamma_j)\}$ is a sequence of increasing sub-σ-algebras.

PROPOSITION 2.12. (Affinity) *For any finite partition α of X and $n \in \mathbb{N}$, the map $\mu \mapsto H_\mu(\alpha|f^{-n}\mathcal{B})$ from $\mathcal{M}(X)$ to $\mathbb{R}^+\cup\{0\}$ is concave. Furthermore, for any continuous map $f: X\to X$, the map $\mu\mapsto h_{m,\mu}(f)$ from $\mathcal{M}_f(X)$ to $\mathbb{R}^+\cup\{0\}$ is affine.*

Proof. Consider $\mu_1,\mu_2\in\mathcal{M}(X)$ and a convex combination $\mu = a_1\mu_1 + a_2\mu_2$, where $0\le a_1, a_2\le 1$ and $a_1+a_2=1$. Let α,γ be finite partitions of X. From the concavity of $s\mapsto -s\log s$, we have

$$a_1H_{\mu_1}(\alpha|\gamma) + a_2H_{\mu_2}(\alpha|\gamma) \le H_\mu(\alpha|\gamma) \le a_1H_{\mu_1}(\alpha|\gamma) + a_2H_{\mu_2}(\alpha|\gamma) + \log 2. \quad (4)$$

Choosing an increasing sequence of finite partitions $\gamma_1\le\gamma_2\le\cdots$ with diameters tending to zero for which $\mathcal{B} = \mathcal{B}(\bigvee_{j=1}^\infty \gamma_j)$, we have by Lemma 2.5(ii), for $i = 1, 2$,

$$H_\mu(\alpha|f^{-n}\mathcal{B}) = \lim_{j\to\infty} H_\mu(\alpha|f^{-n}\gamma_j) \quad\text{and}\quad H_{\mu_i}(\alpha|f^{-n}\mathcal{B}) = \lim_{j\to\infty} H_{\mu_i}(\alpha|f^{-n}\gamma_j). \quad (5)$$

Combining (4) and (5), we have

$$\begin{aligned} a_1H_{\mu_1}(\alpha|f^{-n}\mathcal{B}) + a_2H_{\mu_2}(\alpha|f^{-n}\mathcal{B}) &\le H_\mu(\alpha|f^{-n}\mathcal{B}) \\ &\le a_1H_{\mu_1}(\alpha|f^{-n}\mathcal{B}) + a_2H_{\mu_2}(\alpha|f^{-n}\mathcal{B}) + \log 2. \end{aligned}$$

This shows that the map $\mu\mapsto H_\mu(\alpha|f^{-n}\mathcal{B})$ from $\mathcal{M}(X)$ to $\mathbb{R}^+\cup\{0\}$ is concave.

Replacing α by α_0^{n-1}, we get

$$\begin{aligned} a_1H_{\mu_1}(\alpha_0^{n-1}|f^{-n}\mathcal{B}) + a_2H_{\mu_2}(\alpha_0^{n-1}|f^{-n}\mathcal{B}) \le H_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B}) \\ \le a_1H_{\mu_1}(\alpha_0^{n-1}|f^{-n}\mathcal{B}) + a_2H_{\mu_2}(\alpha_0^{n-1}|f^{-n}\mathcal{B}) + \log 2. \end{aligned}$$

Dividing by n and taking the limit, we have

$$h_{m,\mu}(f,\alpha) = a_1h_{m,\mu_1}(f,\alpha) + a_2h_{m,\mu_2}(f,\alpha).$$

Since the finite partition α is arbitrary, the second part of the proposition follows immediately. □

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

PROPOSITION 2.13. (Upper semi-continuity)

(a) *Let $\mu \in \mathcal{M}(X)$. For any $\alpha \in \mathcal{P}$ with $\mu(\partial\alpha) = 0$ and any $n \in \mathbb{N}$, the map $\mu \mapsto H_\mu(\alpha|f^{-n}\mathcal{B})$ from $\mathcal{M}(X)$ to $\mathbb{R}^+ \cup \{0\}$ is upper semi-continuous at μ, that is,*

$$\limsup_{\nu\to\mu} H_\nu(\alpha|f^{-n}\mathcal{B}) \le H_\mu(\alpha|f^{-n}\mathcal{B}).$$

(b) *Assume that $f : X \to X$ has uniform separation of preimages and $\mu \in \mathcal{M}_f(X)$. The pointwise preimage entropy map $\mu \mapsto h_{m,\mu}(f)$ from $\mathcal{M}_f(X)$ to $\mathbb{R}^+ \cup \{0\}$ is upper semi-continuous at μ, that is,*

$$\limsup_{\nu\to\mu} h_{m,\nu}(f) \le h_{m,\mu}(f).$$

Proof. (a) Let $n \in \mathbb{N}$. Choose an increasing sequence of finite partitions $\gamma_1 \le \gamma_2 \le \cdots$ with diameters tending to zero for which $\mathcal{B} = \mathcal{B}(\bigvee_{j=1}^{\infty} \gamma_j)$, and moreover, $\mu(\partial\gamma_j) = 0$ for each j. Since $\mu(\partial\alpha) = 0$, for any $\nu \in \mathcal{M}(X)$, we have

$$\lim_{\nu\to\mu} H_\nu(\alpha|f^{-n}\gamma_j) = H_\mu(\alpha|f^{-n}\gamma_j)$$

for each j. By Lemma 2.5(ii), $\lim_{j\to\infty} H_\nu(\alpha|f^{-n}\gamma_j) = H_\nu(\alpha|f^{-n}\mathcal{B})$. For any $\varepsilon > 0$, there exists $J \in \mathbb{N}$ such that

$$H_\mu(\alpha|f^{-n}\gamma_J) \le H_\mu(\alpha|f^{-n}\mathcal{B}) + \varepsilon.$$

We have

$$\limsup_{\nu\to\mu} H_\nu(\alpha|f^{-n}\mathcal{B}) \le \limsup_{\nu\to\mu} H_\nu(\alpha|f^{-n}\gamma_J) = H_\mu(\alpha|f^{-n}\gamma_J) \le H_\mu(\alpha|f^{-n}\mathcal{B}) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we get the inequality.

(b) Let f have uniform separation of preimages with exponent $\varepsilon_0 > 0$. By Corollary A.1, for any $\nu \in \mathcal{M}_f(X)$ and any finite Borel partition α of X with $\mathrm{diam}(\alpha) < \varepsilon_0$ and $\mu(\partial\alpha) = 0$, we have $h_{m,\nu}(f, \alpha) = h_{m,\nu}(f)$. Then by Corollary 2.9,

$$h_{m,\nu}(f) = \lim_{n\to\infty} \frac{1}{n} H_\nu(\alpha_0^{n-1}|f^{-n}\mathcal{B}) = \inf_{n\ge1} \frac{1}{n} H_\nu(\alpha_0^{n-1}|f^{-n}\mathcal{B}).$$

Let $\delta > 0$ be arbitrary. We can take $N \in \mathbb{N}$ large enough such that

$$\frac{1}{N} H_\mu(\alpha_0^{N-1}|f^{-N}\mathcal{B}) \le h_{m,\mu}(f) + \delta.$$

Therefore, apply the conclusion in part (a) with $n = N$ and α replaced by α_0^{N-1} to get

$$\begin{aligned}\limsup_{\nu\to\mu} h_{m,\nu}(f) &= \limsup_{\nu\to\mu} h_{m,\nu}(f, \alpha)\\ &= \limsup_{\nu\to\mu} \inf_{n\ge1} \frac{1}{n} H_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B})\\ &\le \limsup_{\nu\to\mu} \frac{1}{N} H_\nu(\alpha_0^{N-1}|f^{-N}\mathcal{B})\\ &\le \frac{1}{N} H_\mu(\alpha_0^{N-1}|f^{-N}\mathcal{B}) \le h_{m,\mu}(f) + \delta.\end{aligned}$$

Since $\delta > 0$ is arbitrary, we get the result. □

2.4. *Shannon–McMillan–Breiman theorem.*

THEOREM 2.14. (Shannon–McMillan–Breiman theorem) *Assume that $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$, $\mu \in \mathcal{M}_f(X)$. Then*

$$h_{m,\mu}(f) = \int_X \lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B})(x)d\mu(x),$$

for any finite partition α with $\operatorname{diam}(\alpha) < \varepsilon_0$. *Furthermore, for $\mu \in \mathcal{M}_f^e(X)$,*

$$h_{m,\mu}(f) = \lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B})(x), \quad \mu\text{-a.e. } x,$$

for any finite partition α with $\operatorname{diam}(\alpha) < \varepsilon_0$.

Proof. Let α be a finite partition of X with $\operatorname{diam}(\alpha) < \varepsilon_0$. By Corollary A.1, $h_{m,\mu}(f) = h_{m,\mu}(f, \alpha)$. Let $\varphi_{n-1} = I_\mu(\alpha|\alpha_1^{n-1} \vee f^{-n}\mathcal{B})$ for $n > 0$, and $\varphi_0 = I_\mu(\alpha|f^{-1}\mathcal{B})$. By Lemma 2.7,

$$\alpha_1^{n-1} \vee f^{-n}\mathcal{B} = f^{-1}(\alpha_0^{n-2} \vee f^{-(n-1)}\mathcal{B}) = f^{-1}\epsilon,$$

and hence $\varphi_k = \varphi_0 = I_\mu(\alpha|f^{-1}\mathcal{B})$ for any $0 < k \le n-1$. Note also that

$$H_\mu(\alpha|f^{-1}\mathcal{B}) = H_\mu(\alpha \vee f^{-1}\mathcal{B}|f^{-1}\mathcal{B}) = H_\mu(\mathcal{B}|f^{-1}\mathcal{B}) = \mathcal{F}_\mu(f) = h_{m,\mu}(f)$$

by Theorem A. Thus

$$\int \varphi_0\, d\mu = \int I_\mu(\alpha|f^{-1}\mathcal{B})\, d\mu = H_\mu(\alpha|f^{-1}\mathcal{B}) = h_{m,\mu}(f).$$

Observe that

$$\begin{aligned} I_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B}) &= I_\mu(\alpha_1^{n-1}|f^{-n}\mathcal{B}) + I_\mu(\alpha|\alpha_1^{n-1} \vee f^{-n}\mathcal{B}) \\ &= I_\mu(\alpha_0^{n-2}|f^{-n-1}\mathcal{B}) \circ f + I_\mu(\alpha|\alpha_1^{n-1} \vee f^{-n}\mathcal{B}) \\ &= \sum_{k=0}^{n-1} \varphi_{n-k-1} \circ f^k = \sum_{k=0}^{n-1} \varphi_0 \circ f^k. \end{aligned}$$

Therefore, the Birkhoff ergodic theorem yields the results. □

Remark 2.15. We do not know whether there is a Shannon–McMillan–Breiman theorem without the ‘uniform separation of preimages’ condition.

2.5. *Proof of Theorem B: the variational principle.* In this section we will give a variational principle for pointwise preimage entropies $h_{m,\mu}(f)$ and $h_m(f)$. In the proof, we use the standard and elegant strategy of [**19**], borrow some ideas from [**2**], and essentially use the assumption of the uniform separation of preimages of the map. We always assume that (X, d) is a compact metric space and f is a continuous map on X. First, we restate and prove the first part of Theorem B.

PROPOSITION 2.16. *Let $f : X \to X$ be a continuous map. Then*

$$h_{m,\mu}(f) \le h_m(f)$$

for all $\mu \in \mathcal{M}_f(X)$.

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Proof. Let $\mu \in \mathcal{M}_f(X)$. Let $\alpha = \{A_1, A_2, \ldots, A_s\}$ be any finite partition of X and choose $\rho > 0$ such that $\rho < (1/(s \log s))$. Since μ is regular, there exist compact sets $B_j \subset A_j$, $1 \le j \le s$, with $\mu(A_j \setminus B_j) < \rho$. Put $B_0 = X \setminus \bigcup_{j=1}^{s} B_j$. Note that $\mu(B_0) < s\rho$. So, for the compact partition $\beta = \{B_0, B_1, \ldots, B_s\}$, we have

$$H_\mu(\alpha|\beta) \le \mu(B_0) \log s < 1.$$

Therefore,

$$\begin{aligned} H_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B}) &\le H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) + H_\mu(\alpha_0^{n-1}|\beta_0^{n-1} \vee f^{-n}\mathcal{B}) \\ &= H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) + H_\mu(\alpha_0^{n-1}|\beta_0^{n-1}) \\ &\le H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) + \sum_{i=0}^{n-1} H_{\mu\circ f^{-i}}(\alpha|\beta) \\ &< H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) + n. \end{aligned}$$

Dividing by n and taking $n \to \infty$ gives

$$h_{m,\mu}(f, \alpha) \le h_{m,\mu}(f, \beta) + 1.$$

Thus, using a standard technique, it suffices to prove, for any compact partition β, that

$$h_{m,\mu}(f, \beta) \le h_m(f) + \log 2. \tag{6}$$

Indeed, once this is done, it follows that

$$h_{m,\mu}(f) \le h_m(f) + \log 2 + 1$$

for any f and $\mu \in \mathcal{M}_f(X)$. It therefore holds for f^q ($q \in \mathbb{N}$). By [**13**, Theorem 5.1], there is a power rule for $h_m(f)$, that is, $h_m(f^q) = q h_m(f)$, and hence by Proposition 2.10, we have

$$q h_{m,\mu}(f) = h_{m,\mu}(f^q) \le h_m(f^q) + \log 2 + 1 \le q h_m(f) + \log 2 + 1.$$

Dividing by q and taking $q \to \infty$ yields

$$h_{m,\mu}(f) \le h_m(f).$$

For (6), it suffices to show that there is an $\varepsilon > 0$ such that for any $n \in \mathbb{N}$,

$$H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) \le n \log 2 + \log \sup_{x \in X} s(n, \varepsilon, f^{-n}x). \tag{7}$$

In fact, following the arguments in step 1 of the proof of [**2**, Theorem 2.5] (see equation (25) there), one can get (7) immediately. For completeness, we give an outline of the proof as follows.

Let $\varepsilon > 0$ be such that any 4ε-ball meets at most two elements of β. Choose an increasing sequence of finite partitions $\gamma_1 \le \gamma_2 \le \cdots$ with diameters tending to zero for which $\mathcal{B} = \mathcal{B}\left(\bigvee_{j=1}^{\infty} \gamma_j\right)$. Then by Lemma 2.5(ii),

$$H_\mu(\beta_0^{n-1}|f^{-n}\mathcal{B}) = \lim_{j \to \infty} H_\mu(\beta_0^{n-1}|f^{-n}\gamma_j).$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

So it suffices to show that for sufficiently large j,

$$H_\mu(\beta_0^{n-1}|f^{-n}\gamma_j) \le n\log 2 + \log \sup_{x\in X} s(n, \varepsilon, f^{-n}x). \tag{8}$$

Using the fact that the decomposition $\{f^{-n}x : x \in f^n X\}$ is upper semi-continuous and arguing as in [**2**], we can choose j big enough such that for any $C \in f^{-n}\gamma_j$ the inequality

$$\#(\beta_0^{n-1}|_C) \le \sup_{x\in X} s(n, \varepsilon, f^{-n}x)\cdot 2^n$$

holds, where $\beta_0^{n-1}|_C := \{B \cap C : B \in \beta_0^{n-1}\}$. Therefore (8) is satisfied. □

The following proposition is based on Theorem 2.14, which holds for maps with uniform separation of preimages.

PROPOSITION 2.17. *Let $f : X \to X$ be a continuous map with uniform separation of preimages. Then*

$$h_{m,\mu}(f) \le h_p(f)$$

for all $\mu \in \mathcal{M}_f(X)$.

Proof. We only need to prove the proposition for ergodic measure $\nu \in \mathcal{M}_f^e(X)$. Indeed, let $\mu = \int_{\mathcal{M}_f^e(X)} \nu\, d\theta(\nu)$ be the unique ergodic decomposition where θ is a probability measure on the Borel subsets of $\mathcal{M}_f(X)$ and $\theta(\mathcal{M}_f^e(X)) = 1$. Since $\mu \mapsto h_{m,\mu}(f)$ is affine and upper semi-continuous by Propositions 2.12 and 2.13, we have

$$h_{m,\mu}(f) = \int_{\mathcal{M}_f^e(X)} h_{m,\nu}(f)\, d\theta(\nu) \tag{9}$$

by a classical result in convex analysis (cf. [**3**, Fact A.2.10 on p. 356]).

Let $\mu \in \mathcal{M}_f^e(X)$ and $\alpha = \{A_1, A_2, \ldots, A_s\}$ be a finite measurable partition of X such that $\operatorname{diam}(\alpha) < \varepsilon_0/100$, $\mu(\partial\alpha) = 0$ and $s \le C(\varepsilon_0)$ for some constant $C(\varepsilon_0) > 0$. Let $\rho > 0$ be arbitrarily small. Since $\mu(\partial\alpha) = 0$, we can take $\varepsilon > 0$ small enough such that $\mu(U_\varepsilon(\partial\alpha)) < \rho/(100s)$, where $U_\varepsilon(\partial\alpha) := \{y \in X : d(y, \partial\alpha) < \varepsilon\}$ is the ε-neighborhood of $\partial\alpha$. Then there exist compact sets $K_j \subset A_j$, $1 \le j \le s$, such that $\mu(A_j \setminus K_j) < \rho/s$ and $K_j \cap U_\varepsilon(\partial A_j) = \emptyset$. Put $K_0 = X \setminus \bigcup_{j=1}^s K_j$. Note that $\mu(K_0) < \rho$. Then we obtain a compact partition $\beta = \{K_0, K_1, \ldots, K_s\}$ of X. By shrinking ε if necessary, we can ensure that $d(K_i, K_j) > 4\varepsilon$ for any $1 \le i, j \le s$, $i \ne j$. Finally, define $\gamma = \alpha \vee \beta$.

Denote $h = h_{m,\mu}(f)$. Let B be the set of all $x \in X$ satisfying

(1) $h = \lim_{n\to\infty} (1/n) I_\mu(\alpha_0^{n-1}|f^{-n}\mathcal{B})(x)$;

(2) $\lim_{n\to\infty} (1/n) \sum_{i=0}^{n-1} \chi_{K_0}(f^i x) = \mu(K_0) < \rho$;

(3) $\mu_x^{f^{-n}\mathcal{B}}(B) = 1$.

By Theorem 2.14, the Birkhoff ergodic theorem and Rokhlin disintegration, $\mu(B) = 1$. Let $B_N \subset B$ be a set such that for any $x \in B_N$ and $n \ge N$,

$$\mu_x^{f^{-n}\mathcal{B}}(\alpha_0^{n-1}(x)) \le \exp(-n(h-\rho)) \tag{10}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

and

$$\sum_{i=0}^{n-1} \chi_{K_0}(f^i x) < n\rho. \tag{11}$$

Then $B = \bigcup_{N=1}^{\infty} B_N$. Choose N large enough such that $\mu(B_N) > 1 - \rho$.

Since $\mu(B_N) = \int_X \mu_x^{f^{-n}\mathcal{B}}(B_N)\, d\mu(x)$, one can find $x \in B_N$ such that $\mu_x^{f^{-n}\mathcal{B}}(B_N) > 1 - \rho$. Consider a maximal (n, ε)-separated set S of $B_N \cap f^{-n}z$ where $z = f^n x$; then $\#S \le s(n, \varepsilon, f^{-n}z)$. For every $z_k \in S \subset B_N$, the Bowen ball $B(z_k, n, \varepsilon)$ intersects at most $(1 + C(\varepsilon_0))^{n\rho} \cdot 2^n$ elements of γ_0^{n-1}, by (11) and the construction of γ. Thus

$$\begin{aligned} 1 - \rho < \mu_x^{f^{-n}\mathcal{B}}(B_N) &\le \mu_x^{f^{-n}\mathcal{B}}\bigg(\bigcup_{z_k \in S} B(z_k, n, \varepsilon)\bigg) \\ &\le (1 + C(\varepsilon_0))^{n\rho} \cdot 2^n \cdot s(n, \varepsilon, f^{-n}z) \cdot \exp(-n(h - \rho)) \end{aligned}$$

where in the last equality we used (10) and the fact $\gamma_0^{n-1}(y) \subset \alpha_0^{n-1}(y)$, $y \in B_N \cap f^{-n}z$. It follows that $h - \rho \le \rho \log(1 + C(\varepsilon_0)) + \log 2 + h_p(f)$. Letting $\rho \to 0$, we have

$$h \le \log 2 + h_p(f). \tag{12}$$

Since f can be replaced by f^q, $q \in \mathbb{N}$, in (12), the proposition follows immediately from standard arguments as before. □

Proof of Theorem B. We now prove the second part of the theorem. Assume that $f : X \to X$ has uniform separation of preimages with exponent $\varepsilon_0 > 0$. By Proposition 2.16, it is enough to show that

$$\sup_{\mu \in \mathcal{M}_f(X)} h_{m,\mu}(f) \ge h_m(f).$$

Given $0 < \varepsilon < \varepsilon_0$, we will find an f-invariant measure μ such that

$$h_{m,\mu}(f) \ge h_m(f, \varepsilon), \tag{13}$$

where $h_m(f, \varepsilon) := \limsup_{n\to\infty} (1/n) \log \sup_{x \in X} s(n, \varepsilon, f^{-n}x)$.

For each positive integer n, choose $x_n \in X$ and an (n, ε)-separated set E_n in $f^{-n}x_n$ such that $\#E_n = s(n, \varepsilon, f^{-n}x_n)$ and there exists a subsequence $n_k \to \infty$ satisfying

$$h_m(f, \varepsilon) = \lim_{k\to\infty} \frac{1}{n_k} \log \#E_{n_k}.$$

Now define probability measures ν_n on X as

$$\nu_n := \frac{1}{\#E_n} \sum_{x \in E_n} \delta_x,$$

where δ_x is the point mass at the point x. Let

$$\mu_n := \frac{1}{n} \sum_{l=0}^{n-1} f_*^l \nu_n.$$

Then there exists a subsequence of $\{n_k\}$, which is also denoted by $\{n_k\}$ for notational convenience, such that $\lim_{k\to\infty} \mu_{n_k} = \mu$. Clearly $\mu \in \mathcal{M}_f(X)$.

Next, choose a partition α of X with $\operatorname{diam}(\alpha) < \varepsilon$ and such that $\mu(\partial A) = 0$ for every $A \in \alpha$. In the following, we will show that, for every positive integer q,

$$h_{m,\mu}(f, \alpha_0^{q-1}) \geq q \lim_{k\to\infty} \frac{1}{n_k} \log \# E_{n_k}. \tag{14}$$

Indeed, dividing by q and taking $q \to \infty$ yields (13).

CLAIM. *For the above α and q, we have*

$$\frac{1}{n} \sum_{l=0}^{n-1} H_{\nu_n \circ f^{-l}}(\alpha_0^{q-1} | f^{-q}\mathcal{B}) \leq H_{\mu_n}(\alpha_0^{q-1} | f^{-q}\mathcal{B}) \tag{15}$$

and

$$\limsup_{k\to\infty} H_{\mu_{n_k}}(\alpha_0^{q-1} | f^{-q}\mathcal{B}) \leq H_{\mu}(\alpha_0^{q-1} | f^{-q}\mathcal{B}). \tag{16}$$

Proof of (14), assuming the claim. Note that ν_n is supported on $f^{-n}x_n$; the canonical system of conditional measures induced by ν_n with respect to $f^{-n}\mathcal{B}$ reduces to a single measure on the set $f^{-n}x_n$, which we may identify with ν_n. Hence for any finite partition γ, we have

$$H_{\nu_n}(\gamma | f^{-n}\mathcal{B}) = H_{\nu_n}(\gamma|_{f^{-n}x_n}).$$

Since each element of α_0^{n-1} has at most one point of E_n, we have

$$H_{\nu_n}(\alpha_0^{n-1}|_{f^{-n}x_n}) = \log \# E_n = \log s(n, \varepsilon, f^{-n}x_n).$$

Consider positive integers q, n with $n > q$ and let $a(j)$ denote the integer part of $(n-j)/q$ for $0 \leq j \leq q-1$. Then clearly,

$$\alpha_0^{n-1} = \bigvee_{r=0}^{a(j)-1} f^{-(rq+j)}\alpha_0^{q-1} \vee \bigvee_{t\in S_j} f^{-t}\alpha,$$

where $S_j = \{0, 1, \ldots, j-1\} \cup \{j + qa(j), \ldots, n-1\}$. Then $\# S_j \leq 2q$. We also denote

$$\rho(j) := n - j - qa(j).$$

For any $0 \leq r \leq a(j) - 2$,

$$f^{-n+j+rq}\mathcal{B} \vee \alpha_q^{n-j-rq-1} = f^{-q}(f^{-n+j+rq+q}\mathcal{B} \vee \alpha_0^{n-j-rq-q-1}) = f^{-q}\mathcal{B}$$

by Lemma 2.7. Then we have

$$\begin{aligned}
&H_{\nu_n}\left(f^{-(j+rq)}\alpha_0^{q-1} | f^{-n}\mathcal{B} \vee \alpha_{n-\rho(j)}^{n-1} \vee \bigvee_{s=r+1}^{a(j)-1} f^{-(j+sq)}\alpha_0^{q-1}\right)\\
&= H_{f_*^{j+rq}\nu_n}\left(\alpha_0^{q-1} | f^{-n+j+rq}\mathcal{B} \vee \alpha_{(a(j)-r)q}^{n-j-rq-1} \vee \bigvee_{s'=1}^{a(j)-1-r} f^{-s'q}\alpha_0^{q-1}\right)\\
&= H_{f_*^{j+rq}\nu_n}(\alpha_0^{q-1} | f^{-n+j+rq}\mathcal{B} \vee \alpha_q^{n-j-rq-1})\\
&= H_{f_*^{j+rq}\nu_n}(\alpha_0^{q-1} | f^{-q}\mathcal{B}).
\end{aligned} \tag{17}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Similarly, as $f^{-(q+\rho(j))}\mathcal{B} \vee \alpha_q^{q+\rho(j)-1} = f^{-q}\mathcal{B}$ by Lemma 2.7, we have

$$\begin{aligned}
&H_{\nu_n}(f^{-(j+(a(j)-1)q)}\alpha_0^{q-1}|f^{-n}\mathcal{B} \vee \alpha_{n-\rho(j)}^{n-1}) \\
&\quad = H_{f_*^{j+(a(j)-1)q}\nu_n}(\alpha_0^{q-1}|f^{-(q+\rho(j))}\mathcal{B} \vee \alpha_q^{q+\rho(j)-1}) \\
&\quad = H_{f_*^{j+(a(j)-1)q}\nu_n}(\alpha_0^{q-1}|f^{-q}\mathcal{B}).
\end{aligned} \tag{18}$$

Hence, combining (17) and (18), we have

$$\begin{aligned}
&\log s(n, \varepsilon, f^{-n}x_n) \\
&\quad = H_{\nu_n}(\alpha_0^{n-1}|f^{-n}\mathcal{B}) \\
&\quad \le H_{\nu_n}\bigg(\bigvee_{t\in S_j} f^{-t}\alpha|f^{-n}\mathcal{B}\bigg) + H_{\nu_n}(f^{-(j+(a(j)-1)q)}\alpha_0^{q-1}|f^{-n}\mathcal{B} \vee \alpha_{n-\rho(j)}^{n-1}) \\
&\qquad + \sum_{r=0}^{a(j)-2} H_{\nu_n}\bigg(f^{-(j+rq)}\alpha_0^{q-1}|f^{-n}\mathcal{B} \vee \alpha_{n-\rho(j)}^{n-1} \vee \bigvee_{s=r+1}^{a(j)-1} f^{-(j+sq)}\alpha_0^{q-1}\bigg) \\
&\quad \le 2q \log \#\alpha + \sum_{r=0}^{a(j)-1} H_{f_*^{j+rq}\nu_n}(\alpha_0^{q-1}|f^{-q}\mathcal{B}).
\end{aligned} \tag{19}$$

Sum this inequality over j from 0 to $q-1$ to get

$$\begin{aligned}
q \log s(n, \varepsilon, f^{-n}x_n) &\le 2q^2 \log \#\alpha + \sum_{j=0}^{q-1}\sum_{r=0}^{a(j)-1} H_{f_*^{j+rq}\nu_n}(\alpha_0^{q-1}|f^{-q}\mathcal{B}) \\
&\le 2q^2 \log \#\alpha + nH_{\mu_n}(\alpha_0^{q-1}|f^{-q}\mathcal{B})) \quad \text{(by (15))}.
\end{aligned}$$

Dividing by nq, we get

$$\frac{1}{n}\log s(n, \varepsilon, f^{-n}x_n) \le \frac{2q^2}{n}\log \#\alpha + \frac{1}{q}H_{\mu_n}(\alpha_0^{q-1}|f^{-q}\mathcal{B}).$$

Hence

$$\begin{aligned}
h_m(f, \varepsilon) &= \limsup_{k\to\infty} \frac{1}{n_k}\log s(n_k, \varepsilon, f^{-n_k}x_{n_k}) \\
&\le \frac{1}{q}\limsup_{k\to\infty} H_{\mu_{n_k}}(\alpha_0^{q-1}|f^{-q}\mathcal{B}) \\
&\le \frac{1}{q}H_\mu(\alpha_0^{q-1}|f^{-q}\mathcal{B}) \quad \text{(by (16))}
\end{aligned}$$

which proves (14). Letting $q \to \infty$, we get

$$h_m(f, \varepsilon) \le h_{m,\mu}(f, \alpha) \le h_{m,\mu}(f).$$

This proves (13).

Proof of the claim. The claim essentially follows from the first parts of Propositions 2.12 and 2.13, respectively; we sketch the proof for completeness. Firstly, from concavity of $s \mapsto -s \log s$, we have

$$\frac{1}{n}\sum_{l=0}^{n-1} H_{\nu_n\circ f^{-l}}(\alpha_0^{q-1}|f^{-q}\gamma) \le H_{\mu_n}(\alpha_0^{q-1}|f^{-q}\gamma) \tag{20}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

for any finite partition γ of X. Then choose an increasing sequence of finite partitions $\gamma_1 \leq \gamma_2 \leq \cdots$ with diameters tending to zero for which $\mathcal{B} = \mathcal{B}(\bigvee_{j=1}^{\infty} \gamma_j)$. Then by Lemma 2.5,

$$H_{\mu_n}(\beta_0^{n-1} | f^{-n}\mathcal{B}) = \lim_{j\to\infty} H_{\mu_n}(\beta_0^{n-1} | f^{-n}\gamma_j) \tag{21}$$

for each n, and

$$\frac{1}{n}\sum_{l=0}^{n-1} H_{\nu_n \circ f^{-l}}(\alpha_0^{q-1} | f^{-n}\mathcal{B}) = \lim_{j\to\infty} \frac{1}{n}\sum_{l=0}^{n-1} H_{\nu_n \circ f^{-l}}(\alpha_0^{q-1} | f^{-n}\gamma_j). \tag{22}$$

Combining (20), (21) and (22) and letting $j \to \infty$ yields (15).

Moreover, we can take the partitions $\{\gamma_j\}$ satisfying Lemma 2.5 and $\mu(\partial\gamma_j) = 0$ for each j, and hence

$$\lim_{n\to\infty} H_{\mu_n}(\alpha_0^{q-1} | f^{-q}\gamma_j) = H_{\mu}(\alpha_0^{q-1} | f^{-q}\gamma_j)$$

for each j. Similarly to the proof of (a) in Proposition 2.13, (replace ν, α, γ_j there by μ_n, α_0^{q-1}, $f^{-q}\gamma_j$, respectively), we get inequality (16).

This completes the proof of the claim and the second part of Theorem B. For the 'moreover' part of the theorem, we only need to show that if f has uniform separation of preimages, then

$$h_m(f) \leq \sup_{\nu \in \mathcal{M}_f^e(X)} h_{m,\nu}(f).$$

Let $\rho > 0$ be sufficiently small. Then there exists an invariant measure μ such that $h_{m,\mu}(f) > h_m(f) - \rho/2$. By (9), there exists an ergodic measure ν such that

$$h_{m,\nu}(f) > h_{m,\mu}(f) - \rho/2 > h_m(f) - \rho.$$

Since ρ is arbitrary, we have $h_m(f) \leq \sup\{h_{m,\nu}(f) : \nu \in \mathcal{M}_f^e(M)\}$. This completes the proof of Theorem B. □

3. *Stable entropy*

3.1. *Equivalence of two definitions of stable metric entropy.*

In this subsection we give the proof of Theorem C, namely, that the two definitions of stable metric entropy are equivalent when μ is ergodic. The proof involves the relationship between two types of measurable partitions, η and ξ, where η, constructed in §1, is a measurable partition subordinate to the W^s-foliation induced by a finite measurable partition, while ξ is a decreasing measurable partition subordinate to the W^s-foliation as in Proposition 1.8.

Recall that $\Pi : M^f \to M$ is the natural projection map. We denote by a tilde the objects in M^f pulled back from M by Π, namely, $\tilde{\mu} = \mu \circ \Pi$, $\tilde{\xi} = \Pi^{-1}\xi$, $\tilde{\eta} = \Pi^{-1}\eta$, $\tilde{\alpha} = \Pi^{-1}\alpha$, etc. Then

$$h_\mu(f, \alpha | \eta) = \limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1} | \tau^{-n}\tilde{\eta})$$

and

$$H_\mu(\xi | f^{-1}\xi) = H_{\tilde{\mu}}(\tilde{\xi} | \tau^{-1}\tilde{\xi}).$$

LEMMA 3.1. *For any* $\alpha, \beta \in \mathcal{P}$ *and* $\eta \in \mathcal{P}^s$,

$$\lim_{n\to\infty} \frac{1}{n} H_\mu(\tilde{\alpha}_0^{n-1} | \tilde{\beta}_0^{n-1} \vee \tau^{-n}\tilde{\eta}) = 0.$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Proof. Applying Lemma 2.3(ii) with $\gamma = \tilde{\beta}_0^{n-1} \vee \tau^{-n}\tilde{\eta}$, we have

$$H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tilde{\beta}_0^{n-1} \vee \tau^{-n}\tilde{\eta}) \leq H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\beta}^0_{-(n-1)} \vee \tau^{-1}\tilde{\eta}) + \sum_{i=0}^{n-2} H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\alpha}_1^{n-1-i} \vee \tilde{\beta}_{-i}^{n-i-1} \vee \tau^{-n+i}\tilde{\eta}).$$

Since $\tilde{\beta}_0^{n-i-1} \vee \tau^{-n+i}\tilde{\eta} \geq \widetilde{\beta^s}$, where $\beta^s \in \mathcal{P}^s$ is induced by $\beta \in \mathcal{P}$ and $\mathrm{diam}(\tilde{\beta}^0_{-i} \vee \widetilde{\beta^s}) \to 0$ as $i \to \infty$, we know that the term in the summation above tends to 0 as $i \to \infty$. Thus the lemma follows. □

PROPOSITION 3.2. *$h_\mu(f, \alpha|\eta)$ is independent of $\eta \in \mathcal{P}^s$ and $\alpha \in \mathcal{P}$.*

Proof. First, let us show that for $\eta_1, \eta_2 \in \mathcal{P}^s$, we have $h_\mu(f, \alpha|\eta_1) = h_\mu(f, \alpha|\eta_2)$.

By Lemma 2.2, we have

$$\begin{aligned} &H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_1) + H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_2|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta}_1) \\ &\quad = H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_2 \vee \tau^{-n}\tilde{\eta}_1) + H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_2|\tau^{-n}\tilde{\eta}_1). \end{aligned} \tag{23}$$

Similarly,

$$\begin{aligned} &H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_2) + H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_1|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta}_2) \\ &\quad = H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_1 \vee \tau^{-n}\tilde{\eta}_2) + H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_1|\tau^{-n}\tilde{\eta}_2). \end{aligned} \tag{24}$$

By the construction of η_1 and η_2, we know that there are two finite partitions $\alpha_1, \alpha_2 \in \mathcal{P}$ such that $\eta_j(x) = \alpha_j(x) \cap W^s(x, \varepsilon_0)$, $j = 1, 2$, for all $x \in M$. Let N_1 and N_2 be the cardinality of α_1 and α_2, respectively. For any $z \in M$, $\eta_1(z)$ intersects at most N_2 elements of α_2, hence intersects at most N_2 elements of η_2. Thus, we have

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_2|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta}_1) \leq \lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_2|\tau^{-n}\tilde{\eta}_1) \\ &\quad = \lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\eta}_2|\tilde{\eta}_1) = \lim_{n\to\infty} \frac{1}{n} H_{\mu}(\eta_2|\eta_1) \leq \lim_{n\to\infty} \frac{1}{n} \log N_2 = 0. \end{aligned}$$

Similarly, we have

$$H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_1|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta}_2) \leq \lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tau^{-n}\tilde{\eta}_1|\tau^{-n}\tilde{\eta}_2) \leq \lim_{n\to\infty} \frac{1}{n} \log N_1 = 0.$$

Hence by (23) and (24), we get

$$\limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_1) = \limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}_2).$$

Next we show that for any $\beta, \gamma \in \mathcal{P}$,

$$\limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tau^{-n}\tilde{\eta}) = \limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\gamma}_0^{n-1}|\tau^{-n}\tilde{\eta}).$$

Again, by Lemma 2.2, we have

$$H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tau^{-n}\tilde{\eta}) \leq H_{\tilde{\mu}}(\tilde{\gamma}_0^{n-1}|\tau^{-n}\tilde{\eta}) + H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tilde{\gamma}_0^{n-1} \vee \tau^{-n}\tilde{\eta}). \tag{25}$$

By Lemma 3.1, we know that

$$\lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tilde{\gamma}_0^{n-1} \vee \tau^{-n}\tilde{\eta}) = 0. \tag{26}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

By (25) and (26), we have

$$\limsup_{n\to\infty}\frac{1}{n}H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tau^{-n}\tilde{\eta})\le\limsup_{n\to\infty}\frac{1}{n}H_{\tilde{\mu}}(\tilde{\gamma}_0^{n-1}|\tau^{-n}\tilde{\eta}).$$

Interchanging β and γ, we in fact obtain

$$\limsup_{n\to\infty}\frac{1}{n}H_{\tilde{\mu}}(\tilde{\beta}_0^{n-1}|\tau^{-n}\tilde{\eta})=\limsup_{n\to\infty}\frac{1}{n}H_{\tilde{\mu}}(\tilde{\gamma}_0^{n-1}|\tau^{-n}\tilde{\eta}).$$

This proves the proposition. □

We present a construction of the decreasing partition ξ, which will be crucial in subsequent steps. The reader can refer to [**18**, §2.4] for more details. Given an ergodic $\mu\in\mathcal{M}_f^e(M)$, we can choose $x_*\in M$ and positive constants $\hat{\epsilon},\hat{r}$ such that

$$B:=B(x_*,\hat{\epsilon}\hat{r}/2)=\{x\in M: d(x,x_*)<\hat{\epsilon}\hat{r}/2\}$$

has positive μ measure and the following construction of a partition ξ satisfies Proposition 1.8.

For each $r\in[\hat{r}/2,\hat{r}]$, put

$$S_{s,r}=\bigcup_{x\in B}S_s(x,r),$$

where $S_s(x,r)=\{y\in W^s_{\mathrm{loc}}(x): y\in B(x_*,r)\}$. Then we can define a partition $\hat{\xi}_{x_*}$ of M such that

$$(\hat{\xi}_{x_*})(y)=\begin{cases}S_s(x,r), & y\in S_s(x,r)\text{ for some } x\in B,\\ M\setminus S_{s,r} & \text{otherwise.}\end{cases}$$

Next we can choose an appropriate $r\in[\hat{r}/2,\hat{r}]$ such that

$$\xi=\bigvee_{j=0}^{\infty}f^{-j}\hat{\xi}_{x_*}$$

is subordinate to the W^s-foliation. Thus $\xi\in\mathcal{Q}^s$. The notation $\hat{\xi}_0^k=\bigvee_{j=0}^k f^{-j}\hat{\xi}_{x_*}$ will be used in the following steps. For notational convenience, $\Pi^{-1}\hat{\xi}_0^k$ is denoted by $\hat{\xi}_0^k$ if there is no confusion.

The following lemma concerning $\hat{\xi}_0^k$ will be useful for the proof of our results.

LEMMA 3.3. *Let $\mu\in\mathcal{M}_f^e(M)$ be an ergodic measure. Suppose $\eta\in\mathcal{P}^s$ is subordinate to the W^s-foliation, and $\hat{\xi}_0^k$ is a partition described as above, where $k\in\mathbb{N}\cup\{\infty\}$. Then for $\tilde{\mu}$-almost every $\tilde{x}\in M^f$, there exists $N=N(\tilde{x})>0$ such that for any $j>N$, we have*

$$(\hat{\xi}_0^{k+j}\vee\tau^{-j}\tilde{\eta})(\tau^{-j}\tilde{x})=(\hat{\xi}_0^{k+j})(\tau^{-j}\tilde{x}).$$

Hence, for any partition β of M with $H_{\tilde{\mu}}(\tilde{\beta}|\hat{\xi}_0^k)<\infty$,

$$I_{\tilde{\mu}}(\tilde{\beta}|\hat{\xi}_0^{k+j}\vee\tau^{-j}\tilde{\eta})(\tau^{-j}\tilde{x})=I_{\tilde{\mu}}(\tilde{\beta}|\hat{\xi}_0^{k+j})(\tau^{-j}\tilde{x}),$$

which implies that

$$\lim_{j\to\infty}H_{\tilde{\mu}}(\tilde{\beta}|\hat{\xi}_0^{k+j}\vee\tau^{-j}\tilde{\eta})=H_{\tilde{\mu}}(\tilde{\beta}|\tilde{\xi}).$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

In particular, if we take $k = \infty$, then the above two equalities become

$$I_{\tilde{\mu}}(\tilde{\beta}|\tilde{\xi} \vee \tau^{-j}\tilde{\eta})(\tau^{-j}\tilde{x}) = I_{\tilde{\mu}}(\tilde{\beta}|\tilde{\xi})(\tau^{-j}\tilde{x})$$

and

$$\lim_{j\to\infty} H_{\tilde{\mu}}(\tilde{\beta}|\tilde{\xi} \vee \tau^{-j}\tilde{\eta}) = H_{\tilde{\mu}}(\tilde{\beta}|\tilde{\xi}).$$

Proof. Since η is subordinate to W^s, for $\tilde{\mu}$-a.e. $\tilde{x}$, there exists $\rho = \rho(\tilde{x}) > 0$ such that $B^s(\Pi\tilde{x}, \rho) \subset \eta(\Pi\tilde{x})$. Since $\tilde{\mu}$ is ergodic, for $\tilde{\mu}$-a.e. $\tilde{x} \in M^f$, there are infinitely many $n > 0$ such that $\tau^{-n}\tilde{x} \in \Pi^{-1}S_{s,r}$. Take $N = N(\tilde{x})$ large enough such that

$$\tau^{-N}\tilde{x} \in \Pi^{-1}S_{s,r}$$

and

$$\tau^{N}(\hat{\xi}(\tau^{-N}\tilde{x})) \subset \Pi^{-1}B^s(\Pi\tilde{x}, \rho) \subset \tilde{\eta}(\tilde{x}), \tag{27}$$

where we write $\hat{\xi} := \Pi^{-1}\hat{\xi}_{x_*}$ for short. Then for any $j \geq N$, we have by (27) that

$$\hat{\xi}^{k+j}(\tau^{-j}\tilde{x}) = \left(\bigvee_{l=0}^{k+j} \tau^{-l}\hat{\xi}\right)(\tau^{-j}\tilde{x}) \subset (\tau^{-j+N}\hat{\xi})(\tau^{-j}\tilde{x}) \subset \tau^{-j}(\tilde{\eta}(\tilde{x})) = (\tau^{-j}\tilde{\eta})(\tau^{-j}\tilde{x}).$$

Thus

$$(\hat{\xi}^{k+j} \vee \tau^{-j}\tilde{\eta})(\tau^{-j}\tilde{x}) = \hat{\xi}^{k+j}(\tau^{-j}\tilde{x}).$$

This proves the first statement in the lemma.

Following the proof of Lemma 2.11 in [**5**], we can prove the remaining results in the lemma, where Fatou's lemma and Lemma 2.5 are needed. We omit the details of the computation here. □

The proof of the following fact is analogous to that in [**5**], and the reader can refer to the proof of Lemma 2.10 in [**5**] for more details.

LEMMA 3.4. *Suppose that $\mu \in \mathcal{M}_f^e(M)$ is an ergodic measure and $\alpha \in \mathcal{P}$. Then for any $k \in \mathbb{N}$,*

$$\lim_{n\to\infty} H_{\tilde{\mu}}(\tilde{\alpha}|\hat{\xi}_{-n}^{k}) = 0.$$

Proof. By the Poincaré recurrence theorem, for $\tilde{\mu}$-a.e. $\tilde{x} \in M^f$, there exist $n_j \to \infty$ such that $\tau^{-n_j}\tilde{x} \in \Pi^{-1}S_{s,r}$. Thus $\mathrm{diam}(\Pi(\tau^{n_j}\hat{\xi})(\tilde{x})) \to 0$ as $j \to \infty$. Hence $\hat{\xi}_{-\infty}^{k}(\tilde{x}) = \{\Pi^{-1}\Pi\tilde{x}\}$, that is, $\hat{\xi}_{-\infty}^{k} = \tilde{\epsilon}$.

Since $H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\epsilon}) = H_{\mu}(\alpha|\epsilon) = 0$, applying Lemma 2.5(ii) with $\zeta_n = \hat{\xi}_{-n}^{k}$ and $\zeta = \tilde{\epsilon}$, we prove the lemma. □

PROPOSITION 3.5. *$h_{\mu}(f, \alpha|\eta) \leq h_{\mu}(f, \xi)$ for any $\eta \in \mathcal{P}^s$ and $\xi \in \mathcal{Q}^s$.*

Proof. By Lemma 2.3(ii), with $\gamma = \tau^{-n}\tilde{\eta}$ and $\alpha = \hat{\xi}_0^k$, we have for any $\eta \in \mathcal{P}^s$, $n > 0$,

$$\begin{aligned}\frac{1}{n}H_{\tilde{\mu}}((\hat{\xi}_0^k)_0^{n-1}|\tau^{-n}\tilde{\eta}) &= \frac{1}{n}H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\eta}) + \frac{1}{n}\sum_{i=0}^{n-2} H_{\tilde{\mu}}(\hat{\xi}_0^k|\hat{\xi}_1^{n-1-i+k} \vee \tau^{i-n}\tilde{\eta})\\ &= \frac{1}{n}H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\eta}) + \frac{1}{n}\sum_{j=2}^{n} H_{\tilde{\mu}}(\hat{\xi}_0^k|\hat{\xi}_1^{j-1+k} \vee \tau^{-j}\tilde{\eta}).\end{aligned} \tag{28}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

By Lemma 3.3, the second term on the right-hand side of (28) converges to $H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\xi})$ as $j \to \infty$. It is clear that each element of $f^{-1}\eta$ intersects at most $l \cdot 2^{k+1}$ elements of $\hat{\xi}_0^k$, where $l = \# f^{-1}x$ for any $x \in M$. So we have

$$H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\eta}) \le \log(l \cdot 2^{k+1}),$$

which implies that

$$\lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\eta}) = 0.$$

Thus we get

$$\lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}((\hat{\xi}_0^k)_0^{n-1}|\tau^{-n}\tilde{\eta}) = H_{\tilde{\mu}}(\hat{\xi}_0^k|\tau^{-1}\tilde{\xi}) \le H_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi}). \tag{29}$$

By Lemma 2.3(ii) with $\gamma = (\hat{\xi}_0^k)_0^{n-1}$ and the fact that

$$\tau^j(\hat{\xi}_0^k)_0^{n-1} = \hat{\xi}_{-j}^{k+n-j-1},$$

we know that

$$\begin{aligned} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|(\hat{\xi}_0^k)_0^{n-1}) &= H_{\tilde{\mu}}(\tilde{\alpha}|\hat{\xi}_{-n+1}^k) + \sum_{i=0}^{n-2} H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\alpha}_1^{n-1-i} \vee \hat{\xi}_{-i}^{k+n-1-i}) \\ &\le H_{\tilde{\mu}}(\tilde{\alpha}) + \sum_{i=0}^{n-2} H_{\tilde{\mu}}(\tilde{\alpha}|\hat{\xi}_{-i}^k). \end{aligned}$$

By Lemma 3.4, we have

$$\lim_{i\to\infty} H_{\tilde{\mu}}(\tilde{\alpha}|\hat{\xi}_{-i}^k) = 0.$$

Then we get

$$\lim_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|(\hat{\xi}_0^k)_0^{n-1}) = 0. \tag{30}$$

By Lemma 2.2, we have

$$H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \le H_{\tilde{\mu}}((\hat{\xi}_0^k)_0^{n-1}|\tau^{-n}\tilde{\eta}) + H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|(\hat{\xi}_0^k)_0^{n-1}). \tag{31}$$

Thus by (29), (30) and (31) we have

$$\begin{aligned} h_\mu(f, \alpha|\eta) &= \limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \\ &\le H_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi}) \\ &= h_\mu(f, \xi) \end{aligned}$$

which finishes the proof of the proposition. □

PROPOSITION 3.6. *Suppose μ is an ergodic measure. Then for any $\eta \in \mathcal{P}^s$ and $\xi \in \mathcal{Q}^s$,*

$$h_\mu(f, \xi) \le \sup_{\alpha\in\mathcal{P}} h_\mu(f, \alpha|\eta).$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Proof. Choose a sequence of finite Borel partitions α_n of M such that

$$\mathcal{B}(\alpha_n) \nearrow \mathcal{B}(\xi) \quad \text{as } n \to \infty,$$

which implies

$$\lim_{n\to\infty} H_{\tilde{\mu}}(\tilde{\alpha}_n|\tau^{-1}\tilde{\xi}) = H_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi}).$$

Thus, we have

$$\sup_{\alpha\le\xi} H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi}) = H_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi}).$$

For any α with $\alpha \le \xi$, we have that for any $j > 0$, $\alpha_1^{j-1} \le \xi_1^{j-1} = f^{-1}\xi$. Then by Lemma 2.3(ii), we have

$$\begin{aligned} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) &= H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\eta}) + \sum_{i=0}^{n-2} H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\alpha}_1^{n-1-i} \vee \tau^{i-n}\tilde{\eta}) \\ &= H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\eta}) + \sum_{j=2}^{n} H_{\tilde{\mu}}(\tilde{\alpha}|\tilde{\alpha}_1^{j-1} \vee \tau^{-j}\tilde{\eta}) \\ &\ge H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\eta}) + \sum_{j=2}^{n} H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi} \vee \tau^{-j}\tilde{\eta}). \end{aligned}$$

Then by Lemma 3.3, we have

$$\lim_{j\to\infty} H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi} \vee \tau^{-j}\tilde{\eta}) = H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi}),$$

which implies that

$$\limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \ge \liminf_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \ge H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi}).$$

So we have

$$\begin{aligned} \sup_{\alpha\in\mathcal{P}} h_\mu(f, \alpha|\eta) &\ge \sup_{\alpha\le\xi} h_\mu(f, \alpha|\eta) \\ &= \sup_{\alpha\le\xi} \limsup_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \\ &\ge \sup_{\alpha\le\xi} \liminf_{n\to\infty} \frac{1}{n} H_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta}) \\ &\ge \sup_{\alpha\le\xi} H_{\tilde{\mu}}(\tilde{\alpha}|\tau^{-1}\tilde{\xi}) \\ &= H_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi}). \end{aligned} \tag{32}$$

Namely, $h_\mu(f, \xi) \le \sup_{\alpha\in\mathcal{P}} h_\mu(f, \alpha|\eta)$. □

Proof of Theorem C. We complete the proof of Theorem C. By Proposition 3.2, $h_\mu(f, \alpha|\eta)$ is independent of α, meaning that

$$h_\mu(f, \alpha|\eta) = \sup_{\beta\in\mathcal{P}} h_\mu(f, \beta|\eta)$$

for any $\alpha \in \mathcal{P}$. Combining Propositions 3.5 and 3.6, we are done. □

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Proof of Corollary C.1. $h_\mu^s(f) \leq h_\mu(f)$ follows from the definition of stable entropy. When f is C^2, the Ledrappier–Young formula (2) by Shu can be applied. Combining with Theorem C, we get

$$h_\mu(f) = h_\mu^s(f) - \sum_{\lambda_i^c < 0} \lambda_i^c \gamma_i^c.$$

If there is no negative Lyapunov exponent in the center direction at μ-a.e. $x \in M$, the sum above vanishes, and hence $h_\mu^s(f) = h_\mu(f)$. □

Proof of Corollary C.2. The equality $h_\mu^s(f) = h_\mu(f, \alpha|\eta)$ follows from Theorem C. Now all inequalities in (32) become equalities and 'sup' can be dropped, so $h_\mu(f, \alpha|\eta) = \lim_{n\to\infty} (1/n) H_\mu(\alpha_0^{n-1}|f^{-n}\eta)$. □

3.2. *Properties of stable entropy.* In this subsection we prepare some lemmas about the properties of stable metric and topological entropy, which are useful in the proof of Theorem D.

LEMMA 3.7. *Assume that $f : M \to M$ is a C^1 non-degenerate partially hyperbolic endomorphism. Then for any $\alpha \in \mathcal{P}$, $\eta \in \mathcal{P}^s$, we have for any $n \in \mathbb{N}$,*

$$\alpha_0^{n-1} \vee f^{-n}\eta \geq \alpha^s,$$

where α^s is the partition in $\mathcal{P}^s$ induced by $\alpha \in \mathcal{P}$.

Proof. Let $y \in (\alpha_0^{n-1} \vee f^{-n}\eta)(x)$. Then $f^i y \in \alpha(f^i x)$ for any $0 \leq i \leq n-1$. As $\operatorname{diam}(\alpha) \ll \varepsilon_0$, we know that $d(f^i y, f^i x) \ll \varepsilon_0$ for any $0 \leq i \leq n-1$. On the other hand, $y \in (f^{-n}\eta)(x)$, that is, $f^n y \in W^s(f^n x, \varepsilon_0)$. This, together with $d(f^{n-1}y, f^{n-1}x) < \varepsilon_0$, implies that $f^{n-1}y \in W^s(f^{n-1}x, \varepsilon_0) \cap \alpha(f^{n-1}x)$, since f is non-degenerate and the expansion of f^{-1} is bounded. With sufficient repetition of this argument, we have $y \in W^s(x, \varepsilon_0) \cap \alpha(x)$. Thus $y \in \alpha^s(x)$, and the lemma follows. □

PROPOSITION 3.8. (Power rule) *Assume that $f : M \to M$ is a C^1 non-degenerate partially hyperbolic endomorphism, $\mu \in \mathcal{M}_f(M)$, $\alpha \in \mathcal{P}$ and $\eta \in \mathcal{P}^s$. Then:*

(1) *$a_n := H_\mu(\alpha_0^{n-1}|f^{-n}\eta) + H_\mu(\eta|\alpha^s)$ is a subadditive sequence, that is, $a_{m+n} \leq a_m + a_n$ for any $m, n \geq 1$;*

(2)

$$\begin{aligned} h_\mu(f, \alpha|\eta) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|f^{-n}\eta) \\ &= \inf_{n\geq 1} \frac{1}{n}(H_\mu(\alpha_0^{n-1}|f^{-n}\eta) + H_\mu(\eta|\alpha^s)); \end{aligned}$$

(3) *$h_\mu^s(f^l) = l h_\mu^s(f)$ for any $l \in \mathbb{N}$.*

Proof. The proof is analogous to those of Lemma 2.8 and Proposition 2.10, with minor modifications. Note that $H_\mu(\eta|\alpha^s)$ is finite since $H_\mu(\eta|\alpha^s) \leq \log \#\eta^{\tilde{s}}$, where $\eta^{\tilde{s}} \in \mathcal{P}$ induces $\eta \in \mathcal{P}^s$. Instead of Lemma 2.7, we use Lemma 3.7 to obtain the subadditivity of a_n. The power rule uses the first equality in the second item. □

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

PROPOSITION 3.9. (Affinity) *Assume that* $f : M \to M$ *is a* C^1 *non-degenerate partially hyperbolic endomorphism. For any* $\alpha \in \mathcal{P}$ *and* $\eta \in \mathcal{P}^s$*, the map* $\mu \mapsto H_\mu(\alpha | f^{-n}\eta)$ *from* $\mathcal{M}(M)$ *to* $\mathbb{R}^+ \cup \{0\}$ *is concave. Furthermore, the map* $\mu \mapsto h^s_\mu(f)$ *from* $\mathcal{M}_f(M)$ *to* $\mathbb{R}^+ \cup \{0\}$ *is affine.*

Proof. The proof proceeds along the same lines as the proof of Proposition 2.12. We only need to use an increasing sequence of finite partitions $\gamma_1 \leq \gamma_2 \leq \cdots$ to approximate $\eta \in \mathcal{P}^s$ now. □

PROPOSITION 3.10. (Upper semi-continuity) *Assume that* $f : M \to M$ *is a* C^1 *non-degenerate partially hyperbolic endomorphism.*

(1) *Let* $\nu \in \mathcal{M}_f(M)$*. For any* $\alpha \in \mathcal{P}$ *and* $\eta \in \mathcal{P}^s$ *with* $\mu(\partial\alpha) = 0$ *and* $\mu(\partial\eta^{\tilde{s}}) = 0$ *(*$\eta^{\tilde{s}}$ *is the partition in* $\mathcal{P}$ *inducing* $\eta \in \mathcal{P}^s$*), the map* $\mu \mapsto H_\mu(\alpha | f^{-n}\eta)$ *from* $\mathcal{M}(M)$ *to* $\mathbb{R}^+ \cup \{0\}$ *is upper semi-continuous at* μ*, that is,*
$$\limsup_{\nu \to \mu} H_\nu(\alpha | f^{-n}\eta) \leq H_\mu(\alpha | f^{-n}\eta).$$
(2) *The stable entropy map* $\mu \mapsto h^s_\mu(f)$ *from* $\mathcal{M}_f(M)$ *to* $\mathbb{R}^+ \cup \{0\}$ *is upper semi-continuous at* μ*, that is,*
$$\limsup_{\nu \to \mu} h^s_\nu(f) \leq h^s_\mu(f).$$

Proof. The proof is analogous to the proof of Proposition 2.13. Indeed, to prove (1), we only need to use an increasing sequence of finite partitions $\gamma_1 \leq \gamma_2 \leq \cdots$ to approximate $\eta \in \mathcal{P}^s$ now. For the proof of (2), we need to use (1) and (2) of Proposition 3.8. □

The following result is the Shannon–McMillan–Breiman theorem for stable metric entropy.

THEOREM 3.11. *Let* f *be a* C^1 *partially hyperbolic endomorphism, and* μ *an ergodic measure of* f*. Then for any* $\alpha \in \mathcal{P}$*,* $\eta \in \mathcal{P}^s$ *and* μ*-a.e.* $x \in M$*, we have*
$$h^s_\mu(f) = \lim_{n\to\infty} \frac{1}{n} I_\mu(\alpha_0^{n-1} | f^{-n}\eta)(x).$$

Proof. Let $\mu \in \mathcal{M}^e_f(M)$ be ergodic. The following lemmas are counterparts of those in [**5**], but the proof need modifications. □

LEMMA 3.12. (See [**5**, Lemma 3.7]) *For any* $\eta \in \mathcal{P}^s$ *and* $\xi \in \mathcal{Q}^s$*, we have for* μ*-a.e.* $x \in M$*,*
$$h_\mu(f, \xi) = \lim_{n\to\infty} \frac{1}{n} I_\mu(\xi | f^{-n}\eta)(x).$$

Proof. With $\alpha = \tilde{\xi}$, $\gamma = \tau^{-n}\tilde{\eta}$, we use Lemma 2.3(ii) to get
$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1} | \tau^{-n}\tilde{\eta})(\tilde{x}) \\
&= \lim_{n\to\infty} \frac{1}{n}\left[I_{\tilde{\mu}}(\tilde{\xi} | \tau^{-1}\tilde{\eta})(\tau^{n-1}\tilde{x}) + \sum_{i=0}^{n-2} I_{\tilde{\mu}}(\tilde{\xi} | \tilde{\xi}_1^{n-1-i} \vee \tau^{i-n}\tilde{\eta})(\tau^i(\tilde{x}))\right] \\
&= \lim_{n\to\infty} \frac{1}{n}\left[I_{\tilde{\mu}}(\tilde{\xi} | \tau^{-1}\tilde{\eta})(\tau^{n-1}\tilde{x}) + \sum_{j=2}^{n} I_{\tilde{\mu}}(\tilde{\xi} | \tau^{-1}\tilde{\xi} \vee \tau^{-j}\tilde{\eta})(\tau^{n-j}(\tilde{x}))\right].
\end{aligned}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

By Lemma 3.3, for μ-a.e. x, there exist $N > 0$ such that for any $j > N$,

$$I_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi} \vee \tau^{-j}\tilde{\eta})(\tau^{n-j}(\tilde{x})) = I_{\tilde{\mu}}(\tilde{\xi}|\tau^{-1}\tilde{\xi})(\tau^{n-j}(\tilde{x})).$$

Therefore, the limit is equal to $h_\mu(f, \xi)$.

By Lemma 3.12, we need compare $I_\mu(\xi_0^{n-1}|f^{-n}\eta)$ and $I_\mu(\alpha_0^{n-1}|f^{-n}\eta)$. □

LEMMA 3.13. (See [**5**, Lemma 3.8]) *Let $\alpha \in \mathcal{P}$, $\eta \in \mathcal{P}^s$, $\xi \in \mathcal{Q}^s$. Then for $\tilde{\mu}$-a.e. $\tilde{x}$, we have*

$$\lim_{n\to\infty} \frac{1}{n} I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta})(\tilde{x}) = 0 = \lim_{n\to\infty} \frac{1}{n} I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tilde{\xi}_0^{n-1} \vee \tau^{-n}\tilde{\eta})(\tilde{x}).$$

Proof. We prove the first equality; the second is analogous. By Lemma 2.3(ii) with $\alpha = \tilde{\xi}$, $\gamma = \tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta}$,

$$\begin{aligned} &I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tilde{\alpha}_0^{n-1} \vee \tau^{-n}\tilde{\eta})(\tilde{x}) \\ &\quad = I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\alpha}_{-1}^{-(n-1)} \vee \tau^{-1}\tilde{\eta})(\tau^{n-1}(\tilde{x})) + \sum_{i=0}^{n-2} I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\xi}_1^{n-1-i} \vee \tilde{\alpha}_{-i}^{n-1-i} \vee \tau^{i-n}\tilde{\eta})(\tau^i(\tilde{x})) \\ &\quad \le I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\alpha}_{-1}^{-(n-1)} \vee \tau^{-1}\tilde{\eta})(\tau^{n-1}(\tilde{x})) + \sum_{j=2}^{n} I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\xi}_1^{j-1} \vee \tilde{\alpha}_1^{j-1} \vee \tau^{-j}\tilde{\eta})(\tau^{n-j}(\tilde{x})) \\ &\quad \le I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\alpha}_{-1}^{-(n-1)} \vee \tau^{-1}\tilde{\eta})(\tau^{n-1}(\tilde{x})) + \sum_{j=2}^{n} I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\xi}_1^{j-1} \vee \tilde{\alpha}_1^{j-1})(\tau^{n-j}(\tilde{x})), \end{aligned} \tag{33}$$

where in the last inequality we applied Lemma 3.3 with $\tilde{\xi}$ replaced by an even finer partition.

Take $\phi_n(\tilde{x}) = I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\alpha}_1^{n-1} \vee \tilde{\xi}_1^{n-1})(\tilde{x})$ for $n \ge 2$. Since $\operatorname{diam}(\tilde{\alpha}_1^{n-1} \vee \tilde{\xi}_1^{n-1})(\tilde{x}) \to 0$ as $n \to \infty$, $\phi_n \to 0$ as $n \to \infty$ almost everywhere.

Also, by Lemma 2.4, $\phi^* = \sup_n \phi_n \in L^1(\mu)$. Hence we can apply [**5**, Proposition 3.5] to get that for $\tilde{\mu}$-a.e. $\tilde{x}$,

$$\lim_{n\to\infty} \frac{1}{n} \left[\sum_{i=0}^{n-2} I_{\tilde{\mu}}(\tilde{\xi}|\tilde{\xi}_1^{n-1-i} \vee \tilde{\alpha}_1^{n-1-i})(\tau^i(\tilde{x})) \right] = 0.$$

By (33), we get the result of the lemma. □

We are now ready to prove the theorem. By Lemma 2.2, we have

$$\begin{aligned} I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) &\le I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1} \vee \tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) \\ &= I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) + I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tilde{\xi}_0^{n-1} \vee \tau^{-n}\tilde{\eta})(\tilde{x}). \end{aligned}$$

Then by Lemmas 3.13 and 3.12 and Theorem C, we have

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) &\le \limsup_{n\to\infty} \frac{1}{n} I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) \\ &= H_\mu(\xi|f^{-1}\xi) = h_\mu^s(f). \end{aligned} \tag{34}$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

Similarly,

$$\begin{aligned}I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) &\le I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}\vee\tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x})\\ &= I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}) + I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tilde{\alpha}_0^{n-1}\vee\tau^{-n}\tilde{\eta})(\tilde{x}).\end{aligned}$$

Again by Lemmas 3.13 and 3.12 and Theorem C, we have

$$\begin{aligned}h_\mu^s(f) = H_\mu(\xi|f^{-1}\xi) &= \liminf_{n\to\infty}\frac{1}{n}I_{\tilde{\mu}}(\tilde{\xi}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x})\\ &\le \liminf_{n\to\infty}\frac{1}{n}I_{\tilde{\mu}}(\tilde{\alpha}_0^{n-1}|\tau^{-n}\tilde{\eta})(\tilde{x}).\end{aligned} \tag{35}$$

Combining (34) and (35), we compete the proof of the theorem. □

The following lemma on stable topological entropy is important for the proof of Theorem D.

LEMMA 3.14. *For any $\delta > 0$,*

$$h_{p,\mathrm{top}}^s(f) = \sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x,\delta)})$$

$$h_{m,\mathrm{top}}^s(f) = \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\sup_{x\in M}\frac{1}{n}\log s(n,\varepsilon, f^{-n}\overline{W^s(x,\delta)}).$$

Proof. We prove the first equality; the second is similar. It is easy to see that $h_{p,\mathrm{top}}^s(f) \le \sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x,\delta)})$ for any $\delta > 0$ since $\delta \mapsto \sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x,\delta)})$ is increasing.

Let us prove the other direction for some fixed $\delta > 0$. For any $\rho > 0$, let $y \in M$ be such that

$$\sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x,\delta)}) \le h_{\mathrm{top}}^s(f, \overline{W^s(y,\delta)}) + \frac{\rho}{3}. \tag{36}$$

We can choose $\varepsilon_0 > 0$ such that

$$\begin{aligned}h_{\mathrm{top}}^s(f, \overline{W^s(y,\delta)}) &= \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon, f^{-n}\overline{W^s(y,\delta)})\\ &\le \limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon_0, f^{-n}\overline{W^s(y,\delta)}) + \frac{\rho}{3}.\end{aligned} \tag{37}$$

Choose $\delta_1 > 0$ small enough such that $\delta_1 < \delta$ and

$$h_{p,\mathrm{top}}^s(f) \ge \sup_{x\in M} h_{\mathrm{top}}^s(f, \overline{W^s(x,\delta_1)}) - \frac{\rho}{3}. \tag{38}$$

Then there exist $y_j \in \overline{W^s(y,\delta)}$, $1 \le j \le N$, where N only depends on δ, δ_1 and the Riemannian structure on $\overline{W^s(y,\delta)}$, such that

$$\overline{W^s(y,\delta)} \subset \bigcup_{j=1}^N \overline{W^s(y_j,\delta_1)}.$$

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

It follows that

$$\begin{aligned}
\limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon_0,f^{-n}\overline{W^s(y,\delta)}) &\le \limsup_{n\to\infty}\frac{1}{n}\log\bigg(\sum_{j=1}^{N}s(n,\varepsilon_0,f^{-n}\overline{W^s(y_j,\delta_1)})\bigg)\\
&\le \limsup_{n\to\infty}\frac{1}{n}\log Ns(n,\varepsilon_0,f^{-n}\overline{W^s(y_i,\delta_1)})\\
&= \limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon_0,f^{-n}\overline{W^s(y_i,\delta_1)})\\
&\le \lim_{\varepsilon\to 0}\limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon,f^{-n}\overline{W^s(y_i,\delta_1)})\\
&= h^s_{\mathrm{top}}(f,\overline{W^s(y_i,\delta_1)}) \qquad (39)
\end{aligned}$$

for some $1\le i\le N$. Combining (36)–(39),

$$\begin{aligned}
\sup_{x\in M} h^s_{\mathrm{top}}(f,\overline{W^s(x,\delta)}) &\le h^s_{\mathrm{top}}(f,\overline{W^s(y,\delta)})+\frac{\rho}{3}\\
&\le \limsup_{n\to\infty}\frac{1}{n}\log s(n,\varepsilon_0,f^{-n}\overline{W^s(y,\delta)})+\frac{2\rho}{3}\\
&\le h^s_{\mathrm{top}}(f,\overline{W^s(y_i,\delta_1)})+\frac{2\rho}{3}\\
&\le \sup_{x\in M} h^s_{\mathrm{top}}(f,\overline{W^s(x,\delta_1)})+\frac{2\rho}{3}\\
&\le h^s_{p,\mathrm{top}}(f)+\rho.
\end{aligned}$$

Since $\rho>0$ is arbitrary, we have $\sup_{x\in M} h^s_{\mathrm{top}}(f,\overline{W^u(x,\delta)})\le h^s_{p,\mathrm{top}}(f)$. □

3.3. *Variational principle for stable entropy.*

Sketch of proof of Theorem D. The proof is analogous to that of Theorem B, with necessary modifications. We just point out the main difference here.

For the first part showing that $h^s_\mu(f)\le h^s_{p,\mathrm{top}}(f)$ for all $\mu\in\mathcal{M}_f(M)$, we follow the same lines as Proposition 2.17. The crucial point is that Theorem 3.11 as well as affinity and upper semi-continuity (Propositions 3.9 and 3.10) hold for stable metric entropy. We mention that the first part of Lemma 3.14 is also used.

The proof of the second part is more involved. We want to show that for any $\rho>0$, there exists $\mu\in\mathcal{M}_f(M)$ such that $h^s_\mu(f)\ge h^s_{m,\mathrm{top}}(f)-\rho$.

Firstly, we should carefully construct $\eta\in\mathcal{P}^s$. Fix some $\delta>0$ small enough. By the second part of Lemma 3.14, take $\varepsilon>0$ small enough, $x_n\in M$ and S_n an (n,ε)-separated set of $f^{-n}\overline{W^s(x_n,\delta)}$ with cardinality $s(n,\varepsilon,f^{-n}\overline{W^s(x_n,\delta)})$ such that there exists a subsequence $n_k\to\infty$ satisfying

$$h^s_{m,\mathrm{top}}(f,\varepsilon)=\lim_{k\to\infty}\frac{1}{n_k}\log\#S_{n_k}.$$

Define

$$\nu_n:=\frac{1}{\#S_n}\sum_{y\in S_n}\delta_y$$

and

$$\mu_n := \frac{1}{n}\sum_{i=0}^{n-1} f^i \nu_n.$$

There exists a subsequence of $\{n_k\}$, also denoted by n_k, such that $\lim_{k\to\infty}\mu_{n_k} = \mu$. Obviously $\mu \in \mathcal{M}_f(M)$.

Choose a subsequence of $\{x_{n_k}\}$ converging to some point $y_0 \in M$, also denoted by $\{x_{n_k}\}$ for simplicity. Without loss of generality, assume that $x_{n_k} \in B(y_0, \delta)$ for all x_{n_k}. As δ is very small, we can choose a partition $\beta \in \mathcal{P}$ such that $B(y_0, 100\delta) \subset \beta(y_0)$. Consequently, $\overline{W^s(x_{n_k}, \delta)} \subset B(y_0, 100\delta) \cap \overline{W^s(x_{n_k}, \delta)} \subset \eta(x_{n_k})$ for all x_{n_k}, where $\eta = \beta^s \in \mathcal{P}^s$. That is, $\overline{W^s(x_{n_k}, \delta)}$ is contained in a single element of η for any x_{n_k}. Then choose $\alpha \in \mathcal{P}$ such that $\mu(\partial\alpha) = 0$ and $\operatorname{diam}(\alpha) \ll \varepsilon$. In this way, we have $\log s(n_k, \varepsilon, f^{-n_k}\overline{W^s(x_{n_k}, \delta)}) = H_{\nu_{n_k}}(\alpha_0^{n_k-1} | f^{-n_k}\eta)$.

Secondly, the computation in (17)–(19) should now be modified accordingly for the conditional entropy $H_{\nu_n}(\alpha_0^{n-1}|f^{-n}\eta)$. The key is (17). Using Lemma 3.7 instead, we can obtain similarly

$$H_{\nu_n}\left(f^{-(j+rq)}\alpha_0^{q-1} | f^{-n}\eta \vee \alpha_{n-\rho(j)}^{n-1} \vee \bigvee_{s=r+1}^{a(j)-1} f^{-(j+sq)}\alpha_0^{q-1}\right) = H_{f_*^{j+rq}\nu_n}(\alpha_0^{q-1}|f^{-q}\alpha^s).$$

The remaining issue is the proof of (15) and (16). But this essentially follows from the affinity and upper semi-continuity of stable entropy, which are formulated in Propositions 3.9 and 3.10(1). We skip the details of the proof. □

Proof of Corollary D.1. Suppose that all Lyapunov exponents of f are non-negative. Then for any ergodic measure μ, $h_{m,\mu}(f) = h_\mu(f)$ by the Ledrappier–Young formula (2) by Shu and Theorem A. By Theorem B and the variational principle for classical entropy, we have $h_p(f) = h_m(f) = h_{\text{top}}(f)$. which proves the first item.

Now suppose that ν is an ergodic measure of maximal preimage entropy, for which there exists a negative Lyapunov exponent with positive transversal dimension. By (2) and Theorem A, we know that $h_{m,\nu}(f) < h_\nu(f)$. Then $h_p(f) = h_m(f) = h_{m,\nu}(f) < h_\nu(f) \le h_{\text{top}}(f)$. The proof of the corollary is complete. □

Acknowledgements. W. Wu is supported by NSFC (No. 11701559), Y. Zhu is supported by NSFC (No. 11771118). This work is also supported by NSFC (No. 11826104).

REFERENCES

[**1**] J. Bobok and Z. Nitecki. Topological entropy of m-fold maps. *Ergod. Th. & Dynam. Sys.* **25** (2005), 375–401.
[**2**] W.-C. Cheng and S. Newhouse. Pre-image entropy. *Ergod. Th. & Dynam. Sys.* **25** (2005), 1091–1113.
[**3**] T. Downarowicz. *Entropy in Dynamical Systems (New Mathematical Monographs, 18)*. Cambridge University Press, Cambridge, 2011.
[**4**] D. Fiebig, U. Fiebig and Z. Nitecki. Entropy and preimage sets. *Ergod. Th. & Dynam. Sys.* **23** (2003), 1785–1806.

https://doi.org/10.1017/etds.2019.114 Published online by Cambridge University Press

[**5**] H. Hu, Y. Hua and W. Wu. Unstable entropies and variational principle for partially hyperbolic diffeomorphsims. *Adv. Math.* **321** (2017), 31–68.
[**6**] M. Hurley. On topological entropy of maps. *Ergod. Th. & Dynam. Sys.* **15** (1995), 557–568.
[**7**] R. Langevin and F. Przytycki. Entropie de l'image inverse d'une application. *Bull. Soc. Math. France* **120** (1992), 237–250.
[**8**] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms: part I: Characterization of measures satisfying Pesin's entropy formula. *Ann. of Math. (2)* **122** (1985), 509–539.
[**9**] F. Ledrappier and L.-S. Young. The metric entropy of diffeomorphisms: part II: Relations between entropy, exponents and dimension. *Ann. of Math. (2)* **122** (1985), 540–574.
[**10**] G. Liao and S. Wang. Ruelle inequality of folding type for $C^{1+\alpha}$ maps. *Math. Z.* **290**(1–2) (2018), 509–519.
[**11**] P.-D. Liu. Ruelle inequality relating entropy, folding entropy and negative Lyapunov exponents. *Commun. Math. Phys.* **240**(3) (2003), 531–538.
[**12**] P.-D. Liu. Invariant measures satisfying an equality relating entropy, folding entropy and negative Lyapunov exponents. *Commun. Math. Phys.* **284**(2) (2008), 391–406.
[**13**] Z. Nitecki and F. Przytycki. Preimage entropy for mappings. *Internat. J. Bifur. Chaos* **9** (1999), 1815–1843.
[**14**] Z. Nitecki. Topological entropy and the preimage structure of maps. *Real Anal. Exchange* **29** (2003/2004), 7–39.
[**15**] M. Qian, J.-S. Xie and S. Zhu. *Smooth Ergodic Theory for Endomorphisms*. Springer, Berlin, 2009.
[**16**] V. A. Rokhlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Transl.* **1** (1962), 107–150.
[**17**] D. Ruelle. Positivity of entropy production in nonequilibrium statistical mechanics. *J. Stat. Phys.* **85**(1–2) (1996), 1–23.
[**18**] L. Shu. The metric entropy of endomorphisms. *Commun. Math. Phys.* **291**(2) (2009), 491–512.
[**19**] P. Walters. *An Introduction to Ergodic Theory*. Springer, New York, 1982.