Nonlinearity **12** (1999) 79–101. Printed in the UK PII: S0951-7715(99)92923-7

Using Ulam's method to calculate entropy and other dynamical invariants

Gary Froyland†‡

Department of Mathematical Engineering and Information Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Received 31 March 1998, in final form 10 July 1998
Recommended by L Bunimovich

Abstract. Using a special form of Ulam's method, we estimate the measure-theoretic entropy of a triple (M, T, μ), where M is a smooth manifold, T is a $C^{1+\gamma}$ uniformly hyperbolic map, and μ is the unique physical measure of T. With a few additional calculations, we also obtain numerical estimates of (i) the physical measure μ, (ii) the Lyapunov exponents of T with respect to μ, (iii) the rate of decay of correlations for (T, μ) with respect to C^{γ} test functions, and (iv) the rate of escape (for repellors). Four main situations are considered: T is everywhere expanding, T is everywhere hyperbolic (Anosov), T is hyperbolic on an attracting invariant set (axiom A attractor), and T is hyperbolic on a non-attracting invariant set (axiom A non-attractor/repellor).

AMS classification scheme numbers: 28D20, 41A25, 58F11, 58F19

1. Introduction

We present a new rigorous result concerning the calculation of metric entropy of a smooth hyperbolic dynamical system $T : M \circlearrowleft$, on a compact Riemannian manifold M. The main difference between our method and current methods is the way in which the information to compute the estimates is gathered. In contrast to standard techniques which obtain all of their information from a single long orbit of the system, we use single-step dynamical information from all regions of phase space.

The matrix construction that we use in our approximation automatically provides us with a rigorous estimate of the Sinai–Bowen–Ruelle (SBR) or 'physical' measure μ of the system, and it is with respect to this measure that the metric entropy is calculated. The metric entropy of (T, μ) is intimately linked with the Lyapunov exponents of (T, μ), and we show that the sum of the positive Lyapunov exponents of (T, μ) may also be simply obtained from our construction.

In non-repelling systems, these are the Lyapunov exponents that are theoretically observed by following the orbits of almost all starting points in phase space. In repelling systems, we are also able to simply calculate the rate of escape (or pressure) of the system as an eigenvalue of our transition matrix. Bounds for the mixing rates (or rate of decay of correlations) of the systems are also easily computed as eigenvalues of our transition matrices. Convergence rates of our estimates to the true values are given.

† E-mail address: gary@sat.t.u-tokyo.ac.jp
‡ URL: www.sat.t.u-tokyo.ac.jp/∼gary/

0951-7715/99/010079+23$19.50 © 1999 IOP Publishing Ltd and LMS Publishing Ltd

2. Background definitions

Measure-theoretic (or metric) entropy is one of the most important indicators in dynamical systems, yet is also one of the most difficult to calculate. The difficulty arises from the nature of its definition. Let M be a compact Riemannian manifold, and $T : M \circlearrowleft$ be a C^1 map with Hölder continuous derivative. In order to speak of the metric entropy of T, we require a reference probability measure, invariant under the transformation T. Regarding our system (M, T) as a generator of information, it is the reference measure μ that distributes weight over the phase space M, and tells one how much information is generated from different regions of space. As we are after rigorous results, we shall be restricting ourselves to a situation where T has a unique reference measure μ that is generally considered to be the 'physical' or 'natural' measure of the system (M, T). This distinguished measure (the SBR measure), has the property that for the orbits of *Lebesgue* almost all starting points $x \in M$ distribute themselves over M according to μ. Formally,

$$\frac{1}{k}\sum_{i=0}^{k-1}\delta_{T^i x} \stackrel{\text{weakly}}{\to} \mu \qquad \text{as } k \to \infty \text{ for Lebesgue almost all } x \in M. \tag{1}$$

To assure the existence of an SBR measure, we suppose that T is either expanding, Anosov, or axiom A (see [1, 15]).

The entropy of T with respect to μ, denoted $h_\mu(T)$ is defined by

$$h_\mu(T) = \sup_{\mathfrak{P} \text{ a finite partition}} \lim_{N\to\infty} \frac{1}{N} \sum_{A \in \bigvee_{i=0}^{N-1} T^{-i}\mathfrak{P}} \mu(A) \log \mu(A), \tag{2}$$

where $\mathfrak{P}_1 \vee \mathfrak{P}_2$ denotes the join of two partitions $\mathfrak{P}_1$ and $\mathfrak{P}_2$.

3. Outline of our approximation method

Let $\mathfrak{P}_n = \{A_{n,1}, \ldots, A_{n,n}\}$ be a Markov partition of the system (M, T, μ). This statement will mean something slightly different for each of the four types of systems we are considering, namely, (i) expanding maps, (ii) Anosov maps, (iii) axiom A maps (attractors), and (iv) axiom A maps (non-attractors). In cases (i)–(iii), we construct a stochastic matrix of the form:

$$P_{n,ij} = \frac{m(A_{n,i} \cap T^{-1}A_{n,j})}{m(A_{n,i})}, \tag{3}$$

where m is the natural Riemannian volume measure on the manifold M. This matrix is a generalization to higher dimensions of a construction first put forward by Ulam† [17] as a computational method of approximating the absolutely continuous invariant measures of expanding interval maps. The idea is that one computes the (assumed unique) left eigenvector p_n of unit eigenvalue of P_n, and defines a measure on M by simply spreading weight $p_{n,i}$ evenly over each set $A_{n,i}$; that is,

$$\mu_n(A) = \sum_{i=1}^{n} \frac{m(A \cap A_{n,i})}{m(A_{n,i})} \cdot p_{n,i}. \tag{4}$$

The author has shown [7] that in the present setting $\mu_n \to \mu$ weakly, where $\mu = \mu_{\text{SBR}}$. In this paper, we go on to show that this convergence is in fact strong convergence in some situations and derive rates of convergence. It is also shown that with only a little extra work, one may use the matrix P_n to compute the entropy and the sum of the positive Lyapunov exponents of

† Ulam proposed an equipartition of $M = [0, 1]$ into subintervals, rather than a Markov partition.

the maps described above *with respect to the SBR measure*. Order of convergence results are stated in all cases.

Our approach is as follows. By using the relative volumes of intersection of the Markov partition sets (and in the Anosov and axiom A cases, the relative lengths of intersection of the unstable sides of the Markov partition sets are sometimes used), we obtain an estimate of the functions $\phi^{(u)}(x) = -\log|\det D_x T|$ (expanding case) and $\phi^{(u)}(x) = -\log|\det D_x T_{|E^u_x}|$ (Anosov and axiom A cases). We use the fact that the SBR measure is an *equilibrium state* [1, 15] for the weight function $\phi^{(u)}$, to compute μ_{SBR}. Our approximation ϕ_n of the function $\phi^{(u)}$ may itself be integrated with respect to the approximate SBR measure μ_n to give an estimate of $-\int_M \log|\det D_x T_{|E^u_x}|\,\mathrm{d}\mu_{\mathrm{SBR}}$, which is the sum of the positive Lyapunov exponents with respect to the SBR measure. In the expanding, Anosov, and axiom A attractor situations this is equal to the metric entropy of (T, μ_{SBR}) by the Pesin formula (see [11], for example). In the non-attracting situation, we use a result connecting the pressure, entropy, and Lyapunov exponents [1].

4. Types of maps considered

In this section, we detail the four types of maps that our results will apply to. The structure of the invariant sets and the SBR measures for these maps are slightly different in each case.

4.1. Expanding

The map. $T : M \circlearrowleft$ is $C^{1+\gamma}$ map of a compact d-dimensional Riemannian manifold M with the property that there is a $k > 0$ such that $\|D_x T^k(v)\| > \|v\|$ for all $x \in M$ and $v \in T_x M$.

The SBR measure. One is guaranteed a unique absolutely continuous invariant probability measure μ, the density of which is C^γ and strictly positive (see theorem III.1.1 in Mañé [11] for example). It is this absolutely continuous measure which is the analogue of the SBR measure, as μ satisfies equation (1).

4.2. Anosov

The map. $T : M \circlearrowleft$ is a transitive $C^{1+\gamma}$ $(0 < \gamma < 1)$ Anosov (uniformly hyperbolic on M) diffeomorphism of a compact two-dimensional Riemannian manifold M.

The SBR measure. One is guaranteed a unique invariant probability measure μ satisfying (1) (see [1, 11], for example). For our purposes, there are two types of measures satisfying (1): those that are equivalent to Lebesgue and those that are not. If

$$\det D_x T^p = 1 \text{ for all } x \in M \text{ such that } x = T^p x, \tag{5}$$

then T has an absolutely continuous invariant measure; if (5) does not hold, then μ is not absolutely continuous [1].

Remark 4.1. SBR measures satisfy three main properties:

(i) A set of points of full Lebesgue measure in M exhibit the SBR measure: property (1).

(ii) The expression $h_\nu(T) + \int_M -\log|\det D_x T_{|E^u_x}|\,\mathrm{d}\nu(x)$ (ν varying over the space of all T-invariant probability measures $\mathcal{M}(M, T)$) attains its maximum value (called the *pressure*

$P(T)$) at $\nu = \mu$, the SBR measure. Sometimes μ is called an *equilibrium state* for the function $\phi^{(u)} = -\log|\det D_x T_{|E_x^u}|$.

(iii) The conditional measures of the SBR measure along unstable directions are absolutely continuous with respect to Lebesgue measure.

(See theorems III.2.3 and the proof of IV.14.1 of [11] for the existence and uniqueness of a measure μ with these properties.)

4.3. Axiom A attractor

The map. Let M be a smooth compact two-dimensional Riemannian manifold, $U \subset M$ an open subset, $T : U \to M$ a transitive C^2 diffeomorphism onto $T(U)$, and $\Lambda \subset U$ be a compact T-invariant hyperbolic set. Λ is said to be a *locally maximal hyperbolic set* or a *basic set* if in addition, Λ arises as $\Lambda := \bigcap_{k\in\mathbb{Z}} T^k \bar{V}$ for some open neighbourhood V of Λ. In this section, we will assume that in fact $\Lambda := \bigcap_{k\in\mathbb{Z}^+} T^k \bar{V}$: we will then call Λ an *axiom A attractor*. Typically, Λ is locally the product of an interval with a Cantor set.

The SBR measure. There is a unique measure μ (the SBR measure) satisfying the properties (i)–(iii) of remark 4.1 (see [15]). This measure is singular with respect to (two-dimensional) Lebesgue measure, while its conditional measures in the unstable directions are absolutely continuous with respect to (one-dimensional) Lebesgue.

4.4. Axiom A non-attractor

The map. The setting is the same as in section 4.3, except that $\Lambda = \bigcap_{k\in\mathbb{Z}} T^k \bar{V}$ (inverse iterates are required). The invariant set Λ is not attracting, and is locally the product of two one-dimensional Cantor sets.

The SBR measure. In the non-attracting situation, there is strictly no SBR measure. That is, there is no measure that satisfies property (1) (as Λ is a *non-attracting* hyperbolic set) and there is no invariant measure that has absolutely continuous conditional measures along unstable directions (as any such measure must be supported on a set that is locally the product of two Cantor sets). Thus, no form of the Pesin formula holds for such measures, and we cannot simply equate Lyapunov exponents and metric entropy as before. There is, however, a notion which links the two. Consider an open neighbourhood V of Λ, and the rate of decrease in volume of the set $E_k := \{x : x, Tx, \dots, T^k x \in V\}$. It is known (proposition 4.8 [1]) that

$$\lim_{k\to\infty} \frac{1}{k} \log m(E_k) = P(T), \tag{6}$$

where

$$P(T) = \sup_{\nu\in\mathcal{M}(\Lambda,T)} \left\{ h_\nu(T) + \int_\Lambda \phi^{(u)}\, \mathrm{d}\nu \right\} \tag{7}$$

is the pressure of T as mentioned in remark 4.1 (ii). Sometimes $\mathrm{e}^{P(T)}$ is called the *escape rate* for the system. In the non-attracting situation, one still obtains a unique *equilibrium state* μ' for the weight function $\phi^{(u)}$; that is, an invariant probability measure μ' that maximizes the r.h.s. of (7). We shall describe a method of estimating $P(T)$ and $\int_\Lambda \phi^{(u)}\, \mathrm{d}\mu'$ (the sum of the positive Lyapunov exponents with respect to μ') and then calculate $h_{\mu'}(T)$ from (7). While the equilibrium state μ' does not satisfy properties (i) and (iii) of remark 4.1, it is the natural analogue of an SBR measure for these sorts of systems.

5. Statement of results

We begin by giving a suitable definition of a Markov partition.

Definition 5.1. A collection of sets $\mathfrak{P} = \{A_1, \ldots, A_n\}$ is called a *Markov partition* of (M, T) if:

(i) For Anosov and axiom A maps, each A_i is diffeomorphic to a closed unit square, with one pair of opposing sides denoted $\partial_u A_i$ and the other pair of opposing sides denoted $\partial_s A_i$. For expanding maps, each A_i is diffeomorphic to a closed d-dimensional unit cube.

(ii) $\mathrm{Int}(A_i) \cap \mathrm{Int}(A_j) = \emptyset$ if $i \neq j$.

(iii) For expanding and Anosov maps, $M = \bigcup_{i=1}^n A_i$, while for axiom A maps, $\Lambda \subset \bigcup_{i=1}^n A_i$ and $\Lambda \cap \mathrm{Int}(A_i) \neq \emptyset$, $i = 1, \ldots, n$.

(iv) In the Anosov and axiom A cases, the elements of $\mathfrak{P}$ transform as

(a) $T(\partial_s A_i) \subset \bigcup_{j=1}^n \partial_s A_j$,

(b) $T^{-1}(\partial_u A_i) \subset \bigcup_{j=1}^n \partial_u A_j$,

while in the expanding case, one has $T(A_i) = \bigcup_{j=1}^k A_{i_j}$, for some $i_1, \ldots, i_k$.

Remark 5.2. We have restricted ourselves to systems in two dimensions in the hyperbolic situations. The reason for this is that one is guaranteed to obtain Markov partition sets with smooth boundaries formed by segments of stable and unstable manifold. In dimensions three or higher, one may not even be able to define a length or area of a boundary of a Markov partition set. For example, let $T : \mathbb{T}^d \circlearrowleft$ be a hyperbolic toral automorphism defined by $x \mapsto Lx \pmod 1$, where L is a $d \times d$ integral hyperbolic matrix with $|\det L| = 1$. Cawley [3] shows that the only such T possessing a Markov partition with piecewise smooth boundary are those for which there is an integer $m \geqslant 0$ such that L^m is conjugate over the rationals to a block diagonal matrix, $(L_i)_{i=1}^{d/2}$ where each L_i is a 2×2 integral hyperbolic matrix with $|\det L_i| = 1$ (T is homomorphic to a direct product of toral automorphisms on $\mathbb{T}^2$).

However, for systems in higher dimensions for which it is known that suitable 'box-like' Markov partitions exist (see example 6.3), the techniques described here carry over easily.

Recall that one may refine an initial partition $\mathfrak{P}$ by taking joins with forward and inverse iterates of T to form

$$\mathfrak{P}^N = \bigvee_{i=-N}^{N} T^i \mathfrak{P}. \tag{8}$$

In the expanding case, it is sufficient to only take joins with inverse iterates.

Theorem 5.3 (main result). Let M be a compact two-dimensional (resp. d-dimensional) Riemannian manifold, and T be a transitive C^2 Anosov or axiom A diffeomorphism (resp. expanding map). Denote the SBR measure of T by μ, the Lyapunov exponents of (T, μ) by $\lambda^i(\mu)$, $i = 1, \ldots, r$, the metric entropy by $h_\mu(T)$, and the pressure by $P(T)$. Let $\{\mathfrak{P}_n\}_{n=n_0}^{\infty}$ be a sequence of Markov partitions with the property that $\max_{A_{n,i} \in \mathfrak{P}_n} \operatorname{diam} A_{n,i} \to 0$ as $n \to \infty$. Define a matrix

$$Q_{n,ij} = \frac{m(A_{n,i} \cap T^{-1} A_{n,j})}{m(A_{n,j})}, \tag{9}$$

and let ϱ_n denote the largest non-negative eigenvalue of Q_n, and v_n denote the corresponding *right* eigenvector. From Q_n, v_n, and ϱ_n, compute the stochastic matrix

$$P_{n,ij} = \frac{Q_{n,ij} v_{n,j}}{\varrho_n v_{n,i}}, \tag{10}$$

and its (unique) fixed left eigenvector p_n.

(i) As $n \to \infty$:

(a) The probability measures $\mu_n : \mathfrak{B}(M) \to [0,1]$

$$\mu_n(A) := \sum_{i=0}^{n} p_{n,i} \cdot \frac{m(A_{n,i} \cap A)}{m(A_{n,i})} \to \mu, \tag{11}$$

strongly (resp. weakly), if μ is absolutely continuous with respect to Lebesgue measure (resp. not absolutely continuous),

(b)

$$\lambda_n := -\sum_{i,j=1}^{n} p_{n,i} P_{n,ij} \log Q_{n,ij} \to \sum_{\lambda^i(\mu)>0} \lambda^i(\mu), \tag{12}$$

(c)

$$h_n := \log \varrho_n + \lambda_n \to h_\mu(T), \tag{13}$$

(d) $\log \varrho_n \to P(T)$,

(ii) Define a norm on the space of Borel probability measures by $\|\nu\|_w = \sup\{\int_M g \, d\nu : \|g\|_{L^1} = 1, \mathrm{Lip}(g) = 1\}$. If the Markov partitions are constructed as in (8), then:

(a)

$$\left.\begin{array}{ll} \mu \textit{ absolutely continuous:} & \|\mu_N - \mu\|_{L^1} \\ \mu \textit{ not absolutely continuous:} & \|\mu_N - \mu\|_w \end{array}\right\} = \mathrm{O}((s_\pm^{\gamma_1})^N), \tag{14}$$

(b) $|\lambda_N - \sum_{\lambda^i(\mu)>0} \lambda^i(\mu)| = \mathrm{O}((s_\pm^{\gamma_1})^N)$,

(c) $|h_N - h_\mu(T)| = \mathrm{O}((s_\pm^{\gamma_1})^N)$,

(d) $|\log \varrho_N - P(T)| = \mathrm{O}((s_\pm^{\gamma_1})^N)$,

where $0 < \gamma_1 < 1$,

$$s_\pm = \max\left\{1/\Big(\inf_{x\in M} \inf_{v\in E_x^u} \|D_x T(v)\|/\|v\|\Big), \sup_{x\in M} \sup_{v \in E_x^s} \|D_x T(v)\|/\|v\|\right\},$$

and E_x^u and E_x^s are respectively the unstable and stable subspaces at $x \in M$.

This theorem is proven in section 7. The weak convergence of $\mu_n \to \mu$ (without information on the rate of convergence) was shown in [7]. Most of the spectral approximation theory that we shall use in the proof of theorem 5.3 was developed in [8].

Remark 5.4.

(i) In all cases except the axiom A repellor, the matrix P_n is simply

$$P_{n,ij} = \frac{m(A_{n,i} \cap T^{-1}A_{n,j})}{m(A_{n,i})} = Q_{n,ij} \cdot \frac{m(A_{n,j})}{m(A_{n,i})}, \tag{15}$$

since $\varrho_n = 1$ and Q_n may be brought to the stochastic matrix P_n via a similarity transformation. This representation is now more transparently a generalization of the Ulam approximation.

(ii) Sometimes it may be more convenient to measure lengths in unstable directions rather than volumes. An alternative construction of Q_n is as follows. Let A be an element of some Markov partition, and define $\ell^u(A) = \text{length}\,(\partial_u A)/2$; this is simply the average of the length (as defined by the Riemannian metric) of the two connected curves that make up the unstable sides of the partition set A. Now put

$$Q_{n,ij} = \frac{\ell^u(A_{n,i} \cap T^{-1}A_{n,j})}{\ell^u(A_{n,j})}. \tag{16}$$

Theorem 5.3 holds for this alternative definition of Q_n, as does the analogous version of remark 5.4(i) above.

(iii) In the expanding or Anosov situations, T may by $C^{1+\gamma}$, where $0 < \gamma < 1$. The constant γ_1 appearing in theorem 5.3 must now satisfy $0 < \gamma_1 < \gamma$.

(iv) In the expanding case, $s_{\pm} = 1/(\inf_{x\in M} \inf_{v\in T_x M} \|D_x T(v)\|/\|v\|)$.

(v) The norm $\|\cdot\|_w$ generates the weak topology on the space of Borel probability measures on M, while $\|\cdot\|_{L^1}$ generates the standard strong topology.

(vi) We have introduced an abuse of notation regarding the cardinality and indexing of the partitions. If one has a sequence of Markov partitions which are *not* formed as a result of the process (8), we use the subscript n, for example, $\mathfrak{P}_n$, $A_{n,i}$, P_n, p_n, μ_n, h_n, ψ_n. If, however, we form a partition $\mathfrak{P}^N$ (with cardinality n_N) through the process (8), then for brevity, we denote quantities as $\mathfrak{P}^N$, $A_{N,i}$, P_N, p_N, μ_N, h_N, ψ_N.

(vii) Finally, we emphasize that the results of theorem 5.3 hold only for exact Markov partitions. In practice, Markov partitions will often need to be constructed numerically, and if there are small errors in the computed partition, the constructions of theorem 5.3 may produce inaccurate results. Sometimes the errors may be dealt with in a case-by-case basis, as in example 6.4, but we provide no general error analysis for estimates arising from approximate Markov partitions.

6. Examples

6.1. Expanding

Example 6.1 (a bent baker's map). Our first example shall be a $C^{1+\mathrm{Lip}}$ expanding map of the circle, defined by

$$T(x) = \frac{4\sqrt{6}}{3}x^3 - 2\sqrt{6}x^2 + \left(2 + \frac{2\sqrt{6}}{3}\right)x \qquad (\mathrm{mod}\ 1).$$

It is easily verified that $T : S^1 \circlearrowleft$ is uniformly expanding with $s = (6+\sqrt{6})/10 \approx 0.8449$. The graph of T is shown in figure 1(a). Markov partitions $\mathfrak{P}^N$ were computed for $N = 0, 1, \ldots, 8$, and the associated transition matrices $P_0, P_1, \ldots, P_8$ constructed using (15). The partition $\mathfrak{P}^2$ is superimposed on the graph of T in figure 1(a). Approximations of the physical invariant measure μ were obtained from the fixed left eigenvectors $p_0, p_1, \ldots, p_8$, and estimates of the metric entropy $h_\mu(T)$ with respect to the physical measure μ were calculated using (12) and (13). The approximation of the physical invariant measure computed from $\mathfrak{P}_8$ is shown in figure 1(b), while the entropy estimates are displayed in table 1. Theorem 5.3 guarantees a geometric convergence rate of the estimates h_N to the true value $h_\mu(T)$. However, as we do not know the true value of $h_\mu(T)$, we cannot evaluate the error at the Nth step. To check this, we instead compute the Cauchy values $|h_{N+1} - h_N|$. These necessarily converge at a geometric rate no slower than the values $|h_N - h_\mu(T)|$. The Cauchy values are plotted on semilog axes in figure 1(c). A power-law fit of the form $|h_{N+1} - h_N| = \text{constant} \times (\text{rate})^N$ was performed, yielding a value rate ≈ 0.29. This value is less than $s = 0.8449$, as prescribed by theorem 5.3. Finally, as a quick check of the rate of decay of correlations of the system (T, μ) with respect to C^1 test functions, the spectrum of the matrix P_8 was plotted in figure 1(d). Clearly the spectral values are inside the easy bound of $|z| = s_{\pm} = 0.8449$ for the essential spectral radius (see Ruelle [16]). From this, we conclude that it is likely that the system (T, μ) possesses no isolated eigenvalues outside the disk $|z| = 0.845$, and a bound for the rate of decay (or mixing rate) is 0.845; see [8] for details.

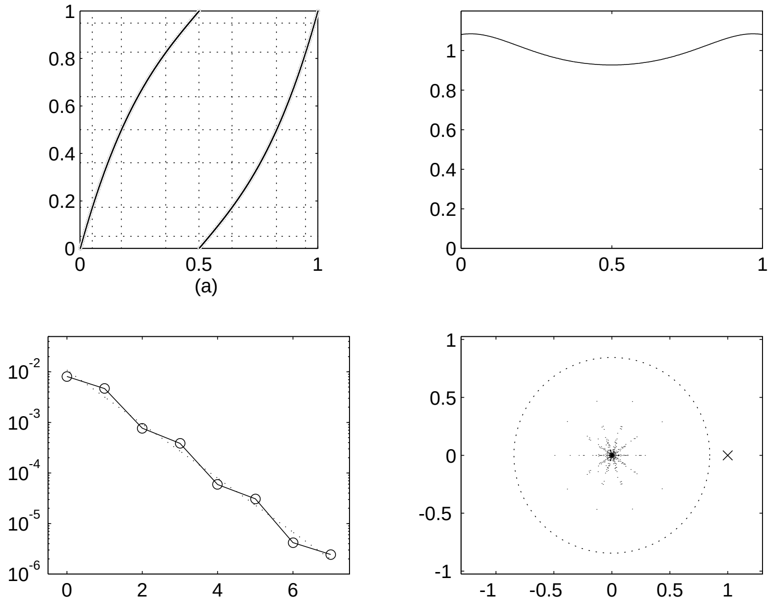

Figure 1. (*a*) Graph of T with partition $\mathfrak{P}^2$. (*b*) Graph of an approximation of the density of the physical measure μ (obtained from P_8). (*c*) Semilogarithmic graph of the Cauchy values $|h_{N+1} - h_N|$ versus N. The dotted line is a power-law fit of the form $|h_{N+1} - h_N| =$ constant $\times$ (rate)^N, where rate ≈ 0.29. (*d*) Plot of the spectrum of P_8 in the complex plane. All eigenvalues are clearly inside $|z| = s_{\pm} = 0.8449$ (shown by the dotted circle).

Table 1. Entropy estimates.

N	Number of partition sets n_N	Entropy estimate h_N
0	2	0.645 620
1	4	0.653 711
2	8	0.649 051
3	16	0.649 819
4	32	0.649 433
5	64	0.649 492
6	128	0.649 461
7	256	0.649 465
8	512	0.649 463

6.2. Anosov

Example 6.2 (the cat map). A simple example of an Anosov map is $T : S^1 \times S^1 \circlearrowleft$, $T(x, y) = (2x + y, x + y) \pmod 1$. The SBR measure of T is two-dimensional Lebesgue measure. Using a five-rectangle Markov partition of the torus (see [10], for example), we construct a matrix P_5 as in (15). From this we compute that the approximation of the SBR measure μ_5, is simply normalized two-dimensional Lebesgue measure—exactly the right answer because T is linear.

As for the entropy and positive Lyapunov exponent with respect to the SBR measure, we obtain $h_5 = \lambda_5 \approx 0.9624 = \log 1/s$ from (12) and (13), where $s = 1/|\det D_x T_{|E_x^u}| = 2/(3 + \sqrt{5})$ is the factor by which all of the rectangles are stretched in the unstable direction.

Both h_5 and λ_5 are the exact values, again because T is linear.

The eigenvalues of P_5 are $\{0, 0, 0, s^2, 1\}$, so an estimate of a bound for the rate of decay with respect to C^1 test functions is $\max\{s^{1/3}, s^{2/3}\} = s^{1/3} \approx 0.7256$ (see [8]). Crawford and Cary [4] show that in fact the rate is equal to s.

This example is rather trivial and we do not consider it as a serious test of our methods. Its inclusion it merely for illustrative purposes.

6.3. Axiom A attractor

Example 6.3 (the solenoid). An example of an axiom A attractor is the well known solenoid attractor (or Smale attractor) $T : S^1 \times D^2 \circlearrowleft$, defined by

$$T(\theta, x, y) = (2\theta, x/4 + (\cos 2\pi\theta)/2, y/4 + (\sin 2\pi\theta)/2).$$

The attractor $\Lambda := \bigcap_{k=0}^{\infty} T^k(S^1 \times D^2)$ is a connected set that is locally homeomorphic to the product of an interval with a (two-dimensional) Cantor set; see section 7.7 of [14] for details. A very simple Markov partition for Λ is $\mathfrak{P}^0 = \{[0, \frac{1}{2}) \times D^2, [\frac{1}{2}, 1) \times D^2\}$. As we refine $\mathfrak{P}^0$, we end up with $\mathfrak{P}^N := \bigvee_{i=-N}^{N} T^i \mathfrak{P}^0$ consisting of $2^{N+1} \times 2^N$ tubes of radius $(\frac{1}{4})^N$ and length approximately $2^{-(N+1)}\sqrt{1 + \frac{\pi^2}{4}}$.

Since T stretches uniformly, the transition matrix P_N from (15) will in fact be the same as the transition matrix $P_N^{(C)}$ for the circle doubling map $T^{(C)} : S^1 \circlearrowleft$, defined by $T^{(C)}(\theta) = 2\theta$ (mod 1); that is

$$P_{N,ij} = P_{N,ij}^{(C)} = \begin{cases} \frac{1}{2}, & j \equiv 2i - 1 \pmod{2^{2N+1}}, \\ \frac{1}{2} & j \equiv 2i \pmod{2^{2N+1}}, \\ 0 & \text{otherwise}. \end{cases} \tag{17}$$

The unique fixed vector of P_N is $p_N = [\frac{1}{2^{2N+1}}, \ldots, \frac{1}{2^{2N+1}}]$. Thus, equal weight of $1/2^{2N+1}$ will be given to each of the partition sets. As the lengths of the partition sets are all equal, an estimate of the entropy and positive Lyapunov exponent is easily computed from (12) and (13) as $h_N = \lambda_N = \log 2$. Because the stretching is uniform, we obtain the exact value for the entropy and the Lyapunov exponent immediately (at $N = 0$), as in example 6.2. While the entropy is estimated exactly, the approximation μ_N for the invariant measure μ is never exact for finite N. The (absolutely continuous) measures μ_N are blurred versions of μ, and have their support on the $2^{N+1} \times 2^N$ thin tubes that make up $\mathfrak{P}^N$; each tube $A_{N,i}$ receiving exactly the correct weight ($\mu_N(A_{N,i}) = \mu(A_{N,i})$ for every $i = 1, \ldots, 2^{2N+1}$ and $N \geqslant 0$). As the partition $\mathfrak{P}^N$ is refined, we obtain more, thinner tubes that look more like a Cantor set. However, the convergence of our absolutely continuous measures to the singular measure μ can only ever be weak convergence. The eigenvalues of P_N are $\{\underbrace{0, \ldots, 0}_{2^{2N+1}-1 \text{ times}}, 1\}$, so a bound for the rate of decay of C^1 functions is $s^{1/3} = 2^{-1/3}$ (see [8]).

6.4. Axiom A non-attractor

Example 6.4 (Hénon horseshoe). We analyse the map $T : \mathbb{R}^2 \circlearrowleft$, given by

$$T(x, y) = (5 - 0.3y - x^2, x). \tag{18}$$

In order to make T axiom A, we restrict the action of T to its set of non-wandering points $\Omega(T)$, given by $\Lambda = \bigcap_{i=-\infty}^{\infty} T^i S$, where S is the square $S = \{(x, y) \in \mathbb{R}^2 : |x| \leqslant R \text{ and } |y| \leqslant R\}$, $R = 0.65 + \sqrt{21.69}/2$. The fixed point of T in the third quadrant of $\mathbb{R}^2$ has coordinates

$(-R, -R)$ and is the lower left corner of the square S. It is shown in theorem 7.4.2 of [14] that (Λ, T) is a (non-attracting) axiom A system. We shall use theorem 5.3 to rigorously estimate the Lyapunov exponents of the map $T : \Lambda \circlearrowleft$ with respect to the equilibrium state μ' of the weight function $\phi^{(u)}(x) = -\log|\det D_x T_{|E^u_x}|$. In the process, we will also obtain a numerical estimate for the equilibrium state (or physical measure) itself, estimates of the pressure and metric entropy of (T, μ'), and bounds for the rate of decay of correlations with respect to μ'.

Remark 6.5. The definition of a Markov partition that we have been using so far is a little stronger than is necessary for our purposes. While the conditions (iv) (a)–(b) of definition 5.1 that the forward image of stable boundaries map onto stable boundaries and the reverse images of the unstable boundaries map onto unstable boundaries make for a tidy mental picture, this is not really what we are after. The main thing is that the inverse image of each partition set, lies across a collection of partition sets in a 'cross-like' formation. Formally, we may replace condition (iv) with:

(iv′) If $x \in \mathrm{Int}\, A_j$ and $T^{-1}x \in \mathrm{Int}\, A_i$, then

(a) $W^s(T^{-1}x) \cap A_i \subset T^{-1}(W^s(x) \cap A_j)$,

(b) $T^{-1}(W^u(x) \cap A_j) \subset W^u(T^{-1}x) \cap A_i$.

The above condition makes Markov partitions a little easier to find, and has a simpler interpretation for systems in more than two dimensions (cf example 6.3).

Returning to our example, one could try to find a Markov partition of Λ by running out the stable and unstable manifolds of the fixed points (and periodic points, if necessary—see appendix 2 of [12]). However, for our purposes, we use a simple approximation that turns out to very close to a Markov partition, and is much easier to construct. We let our initial Markov partition be the trivial partition $\{S\}$ and use $\mathfrak{P}^1 = T^{-1}(S) \cap S \cap T(S)$ as a first refinement. $\mathfrak{P}^1$ consists of four connected sets, each diffeomorphic to a unit square in $\mathbb{R}^2$; see figure 2. The inverse images of each of these sets are shown as the long, thin, dotted-line 'rectangles'. The fact that the 'unstable' boundaries of these four sets are not exactly aligned with the corresponding unstable manifolds is evidenced by the curved 'unstable' ends of their inverse images. However, for our estimates, we are mainly interested in the width of intersection of $T^{-1}\mathfrak{P}^1$ with $\mathfrak{P}^1$ as defined by the 'stable' edges, and these provide us with good estimates†. To refine our partition, we define $\mathfrak{P}^2 = \bigcap_{i=-2}^{2} T^i(S)$. $\mathfrak{P}^2$ has 16 connected pieces, as shown in figure 3 (along with their inverse images). Once we have our partitions, it is a straightforward matter to calculate the matrices Q_1 and Q_2 using (9). For example,

$$Q_1 = \begin{pmatrix} 0.1785 & 0.2650 & 0 & 0 \\ 0 & 0 & 0.2604 & 0.1888 \\ 0 & 0 & 0.3286 & 0.2082 \\ 0.1982 & 0.3427 & 0 & 0 \end{pmatrix}.$$

From these matrices, we compute P_1 and P_2 using (10). For example, we find that

$$P_1 = \begin{pmatrix} 0.3551 & 0.6449 & 0 & 0 \\ 0 & 0 & 0.6228 & 0.3772 \\ 0 & 0 & 0.6538 & 0.3462 \\ 0.3211 & 0.6789 & 0 & 0 \end{pmatrix},$$

and $p_1 = [0.1159, 0.2327, 0.4187, 0.2327]$. Using (13) we obtain $\lambda^1(\mu')$, the largest Lyapunov exponent with respect to the equilibrium measure μ'. The estimates for the pressure are calculated from logarithms of the eigenvalues ϱ_N, and the entropy estimates are obtained from $h_N := \log \varrho_N + \lambda_N$. The numerical results are displayed in table 2. The escape rate for

† For an analysis of the errors incurred by using inexact Markov partitions in this example, see section 7.7.

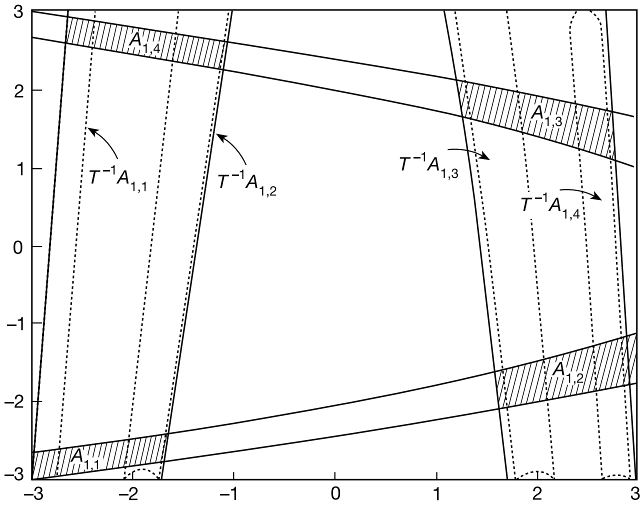

Figure 2. Hénon horseshoe. The solid lines are the portions of the boundaries of $T(S)$ and $T^{-1}(S)$ that are contained in S. Our Markov partition $\mathfrak{P}^1 = \{A_{1,1}, A_{1,2}, A_{1,3}, A_{1,4}\}$ consists of the four shaded 'rectangles' near each of the four corners of S. The basic set Λ is contained in $\bigcup_{i=1}^{4} A_{1,i}$. The dotted lines are the boundaries of $T^{-1}A_{1,i}$, $i = 1, 2, 3, 4$.

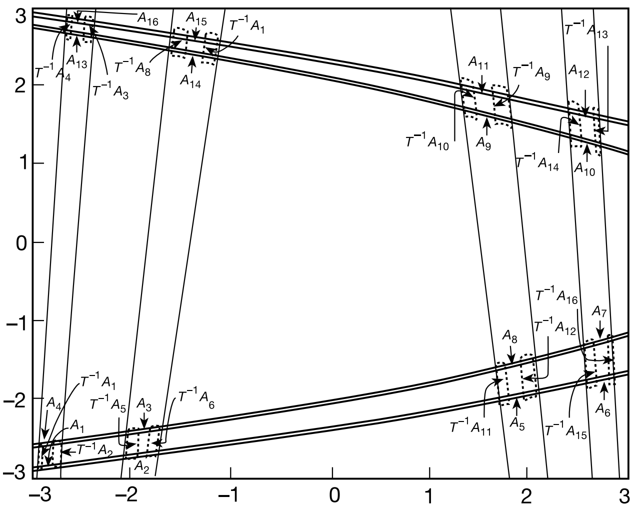

Figure 3. Hénon horseshoe. The solid lines are the portions of the boundaries of $T^2(S)$ and $T^{-2}(S)$ that are contained in S. Our Markov partition $\mathfrak{P}^2 = \{A_{2,1}, \ldots, A_{2,16}\}$ consists of the 16 'rectangles' with one lying in each of the four corners of the sets in $\mathfrak{P}^1$. Again Λ is contained in $\bigcup_{i=1}^{16} A_{2,i}$. The 16 long, thin dotted rectangles are the boundaries of $T^{-1}A_{2,i}$, $i = 1, \ldots, 16$.

the system is the rate at which the volume of an open neighbourhood of the invariant set Λ decreases. Estimates for the escape rate are the eigenvalues ϱ_N, and we obtain 0.503 for the 4-set partition, and 0.499 for the 16-set partition; this means that the volume of $\bigvee_{i=0}^{N-1} T^{-i}S$ (the long, thin rectangles extending from the 'bottom' of S to the 'top' of S) decreases by a factor of about a half, each time N increases by one.

Table 2. Estimates for the positive Lyapunov exponent, pressure, and entropy.

N	Number of partition sets n_N	Estimate of $\lambda^1(\mu')$ λ_N	Estimate of $P(T)$ ϱ_N	Estimate of $h_{\mu'}(T)$ h_N
1	4	1.334	−0.688	0.646
2	16	1.351	−0.696	0.655

Finally, an estimate for a bound for the rate of decay of correlations with respect to C^1 functions is $(1/1.7)^{1/3} \approx 0.8379$ (derived from the Ruelle bound for the essential spectral radius of the induced expanding map).

The Ruelle bound is used as the second largest eigenvalues of both P_1 and P_2 were not outside the disk $|z| \leqslant 1/1.7$ (see [8] for details).

We stress that the values for $\mu_N, h_N, \lambda_N, \log \varrho_N$, and the rate of decay are merely indications of what sort of numbers are produced by our approximation. It is not suggested that the values in table 2 are good approximations of the true values, as we have used a small number of partition sets for illustrative purposes.

6.5. Final remarks

While our technique is readily applicable in one-dimensional systems, we acknowledge that exact numerical construction of Markov partitions may be time-consuming, or simply impossible, particularly in higher dimensions, and we reiterate that the results of theorem 5.3 are proven only for exact Markov partitions. If one is in a situation where it is either too difficult to construct a Markov partition, or a Markov partition does not exist, a method for computing Lyapunov exponents that combines rigour and heuristics is described in [9]. Here, one uses any connected partition (usually a grid or triangulation) and constructs transition matrices in the same way. An estimate of the physical measure and all of the Lyapunov exponents with respect to that measure are then computed using a similar single-step method.

A method of computing the Lyapunov exponent of a piecewise C^2 expanding interval map that uses a construction similar to (9) may be found in Boyarsky [2].

In the case of piecewise C^2 expanding maps, (where the expansiveness depends to some extent on the (non-Markov) partition used), Ding and Zhou [5] show convergence of the invariant measure μ_n in (11) to the unique absolutely continuous measure μ.

7. Proofs

7.1. Coding and the Perron–Frobenius operator

We begin by outlining the machinery that is required for the proofs. We shall make use of the coding that is available to us through the Markov partitions. In this section, T will be an expanding mapping, and we think of the Markov partitions as those one has for expanding maps. Much of the following may be found in [8] in an expanded form. Define an $n \times n$ matrix B by

$$B_{ij} = \begin{cases} 1 & \text{if } m(A_{n,i} \cap T^{-1}A_{n,j}) > 0, \\ 0 & \text{if } m(A_{n,i} \cap T^{-1}A_{n,j}) = 0. \end{cases} \tag{19}$$

Further, define $\Sigma_n^+ = \{1, \ldots, n\}^{\mathbb{Z}^+}$, the space of all infinite sequences of symbols $1, \ldots, n$, and a subspace $\Sigma_B^+ = \{\xi \in \Sigma_n^+ : B_{\xi_i \xi_{i+1}} = 1, \text{ for all } i \geqslant 0\}$. The space Σ_B^+ is invariant under the (one-sided) left-shift $\sigma : \Sigma_n^+ \circlearrowleft$ defined by $[\sigma(\xi)]_i = \xi_{i+1}$. Denote by $[\xi_0, \ldots, \xi_{N-1}]$ the cylinder set $\{\eta \in \Sigma_B^+ : \eta_0 = \xi_0, \ldots, \eta_{N-1} = \xi_{N-1}\}$. Define the semiconjugacy $\pi : \Sigma_B^+ \to M$ by $\pi(\xi_0, \xi_1, \ldots) = x$, where $T^i x \in A_{n,\xi_i}$ for all $i \geqslant 0$. One has the following commutative diagram (see Bowen [1] for details):

$$\begin{array}{ccc} \Sigma_B^+ & \xrightarrow{\sigma} & \Sigma_B^+ \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ M & \xrightarrow{T} & M \end{array} \tag{20}$$

We define a metric on Σ_B^+ by $d_\theta(\xi, \eta) = \theta^N$, where N is the maximal integer for which $\xi_i = \eta_i$, $i = 0, \ldots, N-1$, and $0 < \theta < 1$ is fixed. A norm on the space of complex-valued functions on Σ_B^+ is given by

$$\|\rho\|_\theta = |\rho|_\infty + \sup_{N \geqslant 0} \sup_{\substack{\xi,\eta \in \Sigma_B^+ \\ \xi_0 = \eta_0, \ldots, \xi_{N-1} = \eta_{N-1}}} \frac{|\rho(\xi) - \rho(\eta)|}{\theta^N} := |\rho|_\infty + |\rho|_\theta.$$

Let $s = 1/(\inf_{x \in M} \inf_{v \in T_x M} \|D_x T(v)\|/\|v\|)$. The reader should imagine that $\theta = s^\gamma$ for the following reason. If $\rho : M \to \mathbb{C}$ is γ-Hölder on M $(0 < \gamma < 1)$, the function $\rho \circ \pi$ is Lipschitz on Σ_B^+ with respect to d_{s^γ} metric. Denote this space of Lipschitz functions on Σ_B^+ by

$$\mathfrak{F}_{s^\gamma}^+ = \{\rho : \Sigma_B^+ \to \mathbb{C} : |\rho|_{s^\gamma} < \infty\}.$$

We are now in a position to define a Perron–Frobenius (or transfer) operator on our space $(\mathfrak{F}_{s^\gamma}^+, \|\cdot\|_{s^\gamma})$. Define $\mathcal{L} : (\mathfrak{F}_{s^\gamma}^+, \|\cdot\|_{s^\gamma}) \circlearrowleft$ by

$$(\mathcal{L}\rho)(\xi) = \sum_{\eta \in \sigma^{-1}\xi} \frac{\rho(\eta)}{|\det D_{\pi(\eta)} T|}. \tag{21}$$

This operator is the analogue of the standard Perron–Frobenius operator defined on the smooth space M. It is known that $\mathcal{L}$ has a unique fixed point $\Psi \in \mathfrak{F}_{s^\gamma}^+$. We also define an approximate Perron–Frobenius operator on $\mathfrak{F}_{s^\gamma}^+$ by

$$(\mathcal{L}_{\varphi_N}\rho)(\xi) = \sum_{\eta \in \sigma^{-1}\xi} \exp(\varphi_N(\eta))\rho(\eta), \tag{22}$$

where φ_N is a piecewise constant approximation (constant on N-cylinders $[\xi_0, \ldots, \xi_{N-1}]$) to $\varphi := -\log|\det D_{\pi(\eta)} T|$, given by

$$\varphi_N = -\log\left(\frac{m(\pi([\xi_0, \ldots, \xi_{N-2}]) \cap T^{-1}\pi([\xi_1, \ldots, \xi_{N-1}]))}{m(\pi([\xi_1, \ldots, \xi_{N-1}]))}\right). \tag{23}$$

To achieve operator norm convergence of $\mathcal{L}_{\varphi_N}$ to $\mathcal{L}$, we relax our norm $\|\cdot\|_\theta$ (and enlarge our function space) by increasing $\theta = s^\gamma$ to $\theta' = s^{\gamma'}$. In terms of our smooth space, we allow the Hölder exponent to decrease. Lemma 4.1 and proposition 4.2 of [8] show that

$$\|\varphi - \varphi_N\|_{\theta'} = \mathrm{O}((\theta/\theta')^N)) \qquad \text{for any } 0 < \theta < \theta' < 1. \tag{24}$$

and hence by proposition 4.4 [8],

$$\|\mathcal{L} - \mathcal{L}_{\varphi_N}\|_{\theta'} = \mathrm{O}((\theta/\theta')^N) \qquad \text{for any } 0 < \theta < \theta' < 1. \tag{25}$$

For our purposes, θ' may be taken as close to 1 as desired. As we have norm convergence of $\mathcal{L}_{\varphi_N}$ to $\mathcal{L}$, standard perturbation theory tells us that for sufficiently large N, $\mathcal{L}_{\varphi_N}$ has a unique fixed point Ψ_{φ_N}, and that $\Psi_{\varphi_N} \to \Psi$. In fact,

Lemma 7.1. $\|\Psi_{\varphi_N} - \Psi\|_{\theta'} = \mathrm{O}(\|\mathcal{L}_{\varphi_N} - \mathcal{L}\|_{\theta'})$.

Proof. See Chatelin [6] of theorem 6.7. □

The importance of this fact is that we may project these fixed points down from Σ_B^+ to M and obtain $\psi = \Psi \circ \pi^{-1}$ and $\psi_N = \Psi_{\varphi_N} \circ \pi^{-1}$.

7.2. The Rokhlin/Pesin formula and Lyapunov exponents

The second main ingredient in the proof of (13) is the Rokhlin/Pesin formula. We state the result for the case of Anosov maps (see [11] p 294), which also covers the expanding case.

Theorem 7.2. Let $T : M \circlearrowleft$ be a $C^{1+\gamma}$ transitive Anosov map of a compact Riemannian manifold M with SBR measure μ. Then

$$h_\mu(T) = \int_M \log|\det D_x T|_{E_x^u}| \, \mathrm{d}\mu\,(x). \tag{26}$$

7.3. Proof of the expanding case

7.3.1. Invariant measure estimates. The convergence in (11) was shown to occur weakly in [7]. This was based on the fact that $m(A_i \cap T^{-1}A_j)/m(A_j)$ approximates the value of $|\det D_x T|^{-1}$ on the set $A_i \cap T^{-1}A_j$ (lemma 4.1† in [7]).

We now strengthen this result using techniques developed in [8]. Lemma 7.1 gives us the stronger result that $\|\Psi_N - \Psi\|_{\theta'} = \mathrm{O}((\theta/\theta')^N)$ for any $0 < \theta < \theta' < 1$. That is, we have convergence at the sequence space level of the approximate invariant density $\Psi_N = \psi_N \circ \pi$ to $\Psi = \psi \circ \pi$. Recall that we set $\theta = s^\gamma$ and $\theta' = s^{\gamma'}$, where $\gamma' < \gamma$. Now,

$$\begin{aligned} &\|\psi_N - \psi\|_{L^1} \leqslant |\psi_N - \psi|_\infty = |\psi_N \circ \pi - \psi \circ \pi|_\infty \\ &\leqslant \|\psi_N \circ \pi - \psi \circ \pi\|_{\theta'} = \mathrm{O}((\theta/\theta')^N) \qquad \text{for every } 0 < \theta < \theta' < 1. \end{aligned}$$

By putting θ' close to one, we may as well say that

$$\|\psi_N - \psi\|_{L^1} = \mathrm{O}(\theta_1^N) = \mathrm{O}((s^{\gamma_1})^N), \tag{27}$$

where $\theta_1 = s^{\gamma_1}$ and we think of $\gamma_1 < \gamma$ as being very close γ.

A word on notation. For the purposes of this proof, we will assume that our original partition $\mathfrak{P}$ has n sets, and that the partition $\mathfrak{P}^N$ has n_N sets. When we write an element of $\mathfrak{P}^N$ as $A_{i_0 i_1 \cdots i_{N-1}}$, we mean $\bigcap_{j=0}^{N-1} T^{-j} A_{n,i}$. If we do not care how an element of $\mathfrak{P}^N$ was generated, we will sometimes write $A_{N,i} \in \mathfrak{P}^N$, $1 \leqslant i \leqslant n_N$.

7.3.2. Lyapunov exponent and entropy estimates. We define a piecewise constant function $\kappa_N : M \to \mathbb{R}^+$ that will approximate $|\det D_x T|^{-1}$ by

$$\kappa_N(x) = Q_{N,ij} = P_{N,ij} \cdot \frac{m(A_{N,i})}{m(A_{N,j})}, \qquad \text{where } x \in A_{N,i} \cap T^{-1}A_{N,j}.$$

Since $\det D_x T$ is C^γ and bounded away from zero, one may find a constant $C_1 < \infty$ such that

$$|-\log|\det D_x T| - \log \kappa_N| \leqslant C_1 \max_{i_0,\ldots,i_N \in \{1,\ldots,n\}} (\mathrm{diam}(A_{i_0 i_1 \cdots i_N}))^\gamma.$$

Moreover, one may find a universal constant $C_2 < \infty$ such that

$$\max_{i_0,\ldots,i_{N-1} \in \{1,\ldots,n\}} \mathrm{diam}(A_{i_0 i_1 \cdots i_{N-1}}) \leqslant C_2 s^N;$$

see Ruelle [15]. Thus, noting that $\log \varrho_N = 0$, for all $N \geqslant 0$,

$$|h_N - h_\mu(T)| = \left| -\sum_{i,j=1}^{n_N} p_{N,i} P_{N,ij} \log\left(P_{N,ij} \cdot \frac{m(A_{N,i})}{m(A_{N,j})}\right) - \int_M \log|\det D_x T| \, \mathrm{d}\mu\,(x)\right|$$

† The statement in lemma 4.1 contains two misprints: the correct version is that $\phi_n = \log(m(f^{-1}(\Omega_{x_1}^n \cap \Omega_{x_0}^n)) / m(\Omega_{x_0}^n)) \circ \pi^{-1} \to \phi^{(u)}$.

$$
\begin{aligned}
&= \left| \sum_{i,j=1}^{n_N} \left(\int_{A_{N,i} \cap T^{-1} A_{N,j}} -\log \kappa_N(x)\, \mathrm{d}\mu_N\,(x) \right.\right. \\
&\quad \left.\left. - \int_{A_{N,i} \cap T^{-1} A_{N,j}} \log |\det D_x T|\, \mathrm{d}\mu\,(x) \right) \right| \\
&\leqslant \sum_{i,j=1}^{n_N} \int_{A_{N,i} \cap T^{-1} A_{N,j}} |-\log \kappa_N(x) - \log|\det D_x T||\, \mathrm{d}\mu_N\,(x) \\
&\quad + \sum_{i,j=1}^{n_N} \left| \int_{A_{N,i} \cap T^{-1} A_{N,j}} \log |\det D_x T|\, \mathrm{d}\mu_N\,(x) \right. \\
&\quad \left. - \int_{A_{N,i} \cap T^{-1} A_{N,j}} \log |\det D_x T|\, \mathrm{d}\mu\,(x) \right| \\
&\leqslant C_1 (C_2 s^{N+1})^{\gamma} + |\log|\det D_x T||_\infty \cdot \|\mu_N - \mu\|_{L_1} \\
&\leqslant C_1 C_2^{\gamma} (s^{\gamma})^{N+1} + C_3 (s^{\gamma_1})^N \qquad \text{by (27)}, \\
&= \mathrm{O}((s^{\gamma_1}))^N,
\end{aligned}
$$

as required.

7.4. *Proof of the Anosov case*

There are two things that we must deal with. First we must prove convergence of the approximate invariant measures μ_N to the physical measure μ, and secondly we must prove the entropy estimates h_n converge to the true entropy $h_\mu(T)$. As in the expanding case, we have $\log \varrho_N = 0$ for all $N \geqslant 0$, so convergence of λ_N to the sum of the positive Lyapunov exponents follows from $h_N \to h_\mu(T)$ and the Pesin equality (26). Most of the work is in showing the invariant measure approximations converge. We point out that much of the material in sections 7.4.1 and 7.4.2 is well known. However, we must alter the usual constructions slightly for use in the proofs in sections 7.4.3 and 7.4.4, and it is for this reason that we review some standard theory in sections 7.4.1 and 7.4.2.

7.4.1. Induced expanding maps and their invariant measures (smooth space). In order to make use of the result that we have for expanding maps, it is necessary to somehow convert our hyperbolic map into an expanding map. The standard way to do this is to project the dynamics onto segments of unstable manifolds running across our Markov partition sets. To be precise, we form a subset of M that contains the points that stay in the interiors of the partition sets for all time:

$$M'_N = \bigcap_{j=0}^{\infty} T^j \left(M \Big\backslash \bigcup_{i=1}^{N_n} \partial A_{N,i} \right).$$

This set is T-invariant and has full μ-measure. For each $A_{N,i}$, we choose and fix an $x_i \in \mathrm{Int}\, A_{N,i}$, and define an 'expanding space'

$$X_N = \left(\bigcup_{i=1}^{N_n} (W^u(x_i) \cap A_{N,i}) \right) \cap M'_N,$$

which is the disjoint union of segments of unstable manifold. Each time we refine our partitions $\mathfrak{P}^N$, we will produce a new expanding space X_N.

We project the dynamics onto these unstable segments. Define a projection $\Upsilon_N : M'_N \to X_N$ by $\Upsilon_N(x) = (W^s(x) \cap A_{N,i}) \cap X_N$, where $x \in \mathrm{Int}\, A_{N,i}$, and introduce the induced expanding map $T_{E,N} : X_N \circlearrowleft$ by $T_{E,N}x = \Upsilon_N(Tx)$. Note that the projection of a Markov partition for T on M'_N is a Markov partition on X_N for T_E. The map $T_{E,N}$ will have a unique absolutely continous invariant probability measure μ_E (where for brevity we leave out the subscript N).

As in the expanding case, we will approximate the weight function $-\log|\det D_xT_{E,N}|$ using functions that are constant on cylinder sets of increasing length.

7.4.2. Constructions on sequence space. At this stage, we introduce the two-sided sequence space $\Sigma_B = \{\xi \in \{1,\ldots,n\}^{\mathbb{Z}} : B_{\xi_i\xi_{i+1}} = 1,\ \text{for all } i\}$, and the metric $d_\theta(\xi,\eta) = \theta^N$, where N is the maximal integer for which $\xi_i = \eta_i$, $|i| < N$. Define $\mathrm{var}_N(\varphi) = \sup\{|\varphi(\xi) - \varphi(\eta)| : \xi_i = \eta_i, |i| < N\}$. A norm on the space of complex-valued functions on Σ_B is given by

$$\|\rho\|_\theta = |\rho|_\infty + \sup_{N\geqslant 0}\frac{\mathrm{var}_N(\rho)}{\theta^N} := |\rho|_\infty + |\rho|_\theta.$$

For the two-sided shift, we think of θ as equal to $s_\pm^\gamma$, where $s_\pm$ is defined in theorem 5.3.

There are two steps to go through in order to connect the equilibrium state for $\phi^{(u)}$ (a measure on M) with the fixed eigenfunction of some transfer operator $\mathcal{L}_{E,\vartheta_N}$ for $T_{E,N}$, (the density of a measure on X_N). Theorem 1.4 of [1] proves the existence of an equilibrium state μ_φ on Σ_B for a weight function $\varphi \in \mathfrak{F}_\theta$ (in our case, $\varphi = \phi^{(u)} \circ \pi = -\log|\det D_{\pi(\xi)}T_{|E^u_{\pi(\xi)}}|$ and $\mu_\varphi = \mu \circ \pi$, where μ is the SBR measure). However, we cannot attempt to find μ_φ as the natural extension of some equilibrium state on Σ_B^+, because our weight function φ strictly depends on past coordinates and cannot be defined on Σ_B^+. To get around this, lemma 1.6 [1] shows that one can always find another function $\vartheta : \Sigma_B^+ \to \mathbb{R}^+$ such that the equilibrium state for φ is equal to the equilibrium state for ϑ (still treating the two-sided shift on Σ_B): symbolically, $\mu_\varphi = \mu_\vartheta$ (†). In fact, we will construct a sequence $\{\vartheta_N\}$ of such functions, each of which satisfies (†). Unfortunately, the definition of the function ϑ_N is rather complicated, and we would prefer to approximate φ directly using our piecewise constant approximation φ_N, where φ_N in the Anosov case is defined by

$$\varphi_N = -\log\left(\frac{m(\pi([\xi_{-N}\ldots\xi_{N-1}]) \cap T^{-1}\pi([\xi_{-N+1}\ldots\xi_N]))}{m(\pi([\xi_{-N+1}\ldots\xi_N]))}\right). \tag{28}$$

Compare this definition with equation (23). In the Anosov case, φ_N is piecewise constant on $(2N+1)$-cylinders $[\xi_{-N}\ldots\xi_N]$ in Σ_B. What we need to do is show that both $\|\vartheta_N - \varphi\|_\theta$ and $\|\varphi - \varphi_N\|_\theta$ are small, so that the perturbation theory of the expanding case will tell us that the eigenfunctions Ψ_{E,ϑ_N} and Ψ_{E,φ_N} are close. The first eigenfunction is the density of the projection of the SBR measure μ_φ onto the one-sided space Σ_B^+, and the density Ψ_{E,φ_N} will be close to this. Both of these measures (μ_{E,ϑ_N} and μ_{E,φ_N}) may then be extended to the two-sided space Σ_B to give μ_φ and μ_{N,φ_N}, respectively. The following chain of implications outlines our proof procedure:

$$\begin{aligned}&(\vartheta_N \text{ close to } \varphi \text{ in } \Sigma_B) \text{ and } (\varphi_N \text{ close to } \varphi \text{ in } \Sigma_B) \Longrightarrow \varphi_N \text{ close to } \vartheta_N\\ &\Longrightarrow \mathcal{L}_{E,\vartheta_N} \text{ close to } \mathcal{L}_{E,\varphi_N} \Longrightarrow \Psi_{E,\vartheta_N} \text{ close to } \Psi_{E,\varphi_N} \Longrightarrow \mu_{E,\vartheta_N} \text{ close to } \mu_{E,\varphi_N}.\end{aligned} \tag{29}$$

In view of (†), μ_{E,ϑ_N} is simply a projection of μ_φ onto the induced expanding system. We then proceed as:

$$(\mu_\varphi \text{ and } \mu_{E,\vartheta_N} \text{ coincide on } N\text{-cylinders}) \text{ and } (29) \Longrightarrow \mu_{E,\varphi_N} \text{ close to } \mu_\varphi \text{ on } N\text{-cylinders}.$$

Finally, it follows straight from the definition that $\mu_N \circ \pi$ and μ_{E,φ_N} coincide on N-cylinders in Σ_B. Letting $N \to \infty$ will give us the required convergence of $\mu_N \circ \pi$ to μ_φ.

We shall now make this rigorous.

It will be useful to define B_N to be the $n_N \times n_N$ transition matrix for the partition $\mathfrak{P}^N$. We will denote the corresponding two-sided and one-sided sequence spaces by Σ_{B_N} and $\Sigma^+_{B_N}$ respectively. We may identify a partition set $A_{N,i} \in \mathfrak{P}^N$ with the one-cylinder $[i] \subset \Sigma^+_{B_N}, \Sigma_{B_N}$, via the projection $\pi_N : \Sigma^+_{B_N} \to M$, as in (20). We will now define a function ω_N on $\Sigma^+_{B_N}$ that turns a one-sided sequence into a two-sided sequence in $\Sigma^+_{B_N}$. For every $1 \leqslant i \leqslant n_N$, choose an allowable past $(\ldots \xi^i_{-2}\xi^i_{-1} i) \in \Sigma^-_{B_N}$ and define $\omega_N : \Sigma^+_{B_N} \to \Sigma_{B_N}$ by

$$\omega_N(\xi_0\xi_1\ldots) = (\ldots \xi^{\xi_0}_{-2}\xi^{\xi_0}_{-1}\xi_0\xi_1\xi_2\ldots).$$

We now show the existence of a family of functions $\{\vartheta_N\}$ with property (†) that converge to φ.

Lemma 7.3. There exists a family of functions $\vartheta_N : \Sigma^+_{B_N} \to \mathbb{R}$ that have the same equilibrium states on Σ_{B_N} as $\varphi : \Sigma_{B_N} \to \mathbb{R}^+$, and approach φ according to the geometric rate:

$$\|\vartheta_N - \varphi\|_{\theta'} \leqslant |\varphi|_\theta \left(\left(1 + \frac{1}{1-\theta}\right)\theta^N + \left(1 + \frac{1}{1-\theta/\theta'}\right)\left(\frac{\theta}{\theta'}\right)^N \right), \qquad (30)$$

for $0 < \theta < \theta' < 1$ and $N \geqslant 0$.

Proof. Let $\tau_N : \Sigma_{B_N} \to \Sigma_B$ be the map that takes a sequence in Σ_{B_N} to the corresponding one in Σ_B; that is, $\pi_N(\xi) = \pi(\tau_N(\xi))$ for $\xi \in \Sigma_{B_N}$. We begin by noting that since functions such as $\varphi \circ \omega_N$ depend only on future (and current) coordinates in Σ_{B_N} (positions $0 \to \infty$), when considered as functions on Σ_B, they depend on coordinates from positions $-N \to \infty$. This is because $\tau_N([\xi_0])$ is a $(2N+1)$-cylinder in Σ_B. For the purposes of this proof, we shall be sticking to our original coding on Σ_B, and all functions will be acting on Σ_B. It is shown in proposition 1.2 [13] that one may define

$$\vartheta_N(\xi) = \varphi(\omega_N(\xi)) + \sum_{i=0}^{\infty}(\varphi(\sigma^{i+1}\omega_N(\xi)) - \varphi(\sigma^i\omega_N(\sigma\xi)), \qquad (31)$$

($\varphi = f$ and $\vartheta_N = g$ in their notation), and that the respective equilibrium states μ_{ϑ_N} and μ_φ are identical (see also lemmas 1.5 and 1.6 of [1]). We first bound the sup norm of the difference $\vartheta_N - \varphi$.

$$|\varphi(\xi) - \vartheta_N(\xi)| \leqslant |\varphi(\xi) - \varphi(\omega_N(\xi))| + \sum_{i=0}^{\infty} |(\varphi(\sigma^{i+1}\omega_N(\xi)) - \varphi(\sigma^i\omega_N(\sigma\xi)))|.$$

$$\leqslant \operatorname{var}_N(\varphi) + \sum_{i=0}^{\infty} \operatorname{var}_{N+i}\varphi \leqslant |\varphi|_\theta\theta^N + \theta^N\frac{|\varphi|_\theta}{1-\theta}.$$

Now we move onto the rate of variation (in what follows, $0 < \theta < \theta' < 1$ as before).

$$\operatorname{var}_k(\varphi(\xi) - \vartheta_N(\xi)) \leqslant \operatorname{var}_k(\varphi(\xi) - \varphi(\omega_N(\xi))) + \sum_{i=0}^{\infty}\operatorname{var}_k(\varphi(\sigma^{i+1}\omega_N(\xi)) - \varphi(\sigma^i\omega_N(\sigma\xi)))$$

$$\leqslant \begin{cases} |\varphi|_\theta\theta^N, & (\text{if } k \leqslant N) \\ \operatorname{var}_k\varphi, & (\text{if } k > N) \end{cases} + \begin{cases} \sum_{i=0}^{\infty}|\varphi|_\theta\theta^{N+i}, & (\text{if } k \leqslant N+i) \\ \sum_{i=0}^{\infty}\operatorname{var}_k\varphi, & (\text{if } k > N+i) \end{cases}$$

$$\leqslant \begin{cases} |\varphi|_\theta\theta^N\theta'^{k-N}, & (\text{if } k \leqslant N) \\ |\varphi|_\theta\theta^N\theta'^{k-N}, & (\text{if } k > N) \end{cases}$$

$$+\begin{cases} \sum_{i=0}^{\infty} |\varphi|_\theta \theta^{N+i} \theta'^{k-N-i}, & (\text{if } k \leqslant N+i) \\ \sum_{i=0}^{\infty} |\varphi|_\theta \theta^{N+i} \theta'^{k-N-i} & (\text{if } k > N+i) \end{cases}$$

$$= |\varphi|_\theta \left(\frac{\theta}{\theta'}\right)^N \left(1 + \frac{1}{1-\theta/\theta'}\right) \cdot \theta'^k.$$

Thus

$$\|\vartheta_N - \varphi\|_{\theta'} \leqslant |\varphi|_\theta \left(\left(1 + \frac{1}{1-\theta}\right) \theta^N + \left(1 + \frac{1}{1-\theta/\theta'}\right) \left(\frac{\theta}{\theta'}\right)^N \right).$$

□

We now combine this with previous results.

Lemma 7.4.

$$\|\Psi_{E,\vartheta_N} - \Psi_{E,\varphi_N}\|_{\theta'} = \mathrm{O}((\theta/\theta')^N), \tag{32}$$

where Ψ_{E,ϑ_N} and Ψ_{E,φ_N} are the functions left invariant by $\mathcal{L}_{E,\vartheta_N}$ and $\mathcal{L}_{E,\varphi_N}$ respectively.

Proof. Since $\varphi \in \mathfrak{F}_\theta$, one may show that $|\varphi - \varphi_N|_\infty \leqslant |\varphi|_\theta \cdot \theta^N$, $N \geqslant 0$, using the ideas contained in the proof of lemma 7.1 of [7] and following the proof of lemma 4.1 of [8] for the expanding case. By proposition 1.3 of [13] we obtain

$$|\varphi - \varphi_N|_{\theta'} \leqslant |\varphi|_\theta \cdot (\theta/\theta')^N. \tag{33}$$

Combining (33) with (30), we see that $\|\vartheta_N - \varphi_N\|_{\theta'} = \mathrm{O}(\theta/\theta')$, and hence $\|\mathcal{L}_{E,\vartheta_N} - \mathcal{L}_{E,\varphi_N}\|_{\theta'} = \mathrm{O}((\theta/\theta')^N)$. By lemma 7.1 we obtain the result. □

Natural extensions of measures on sequence space. On sequence space the natural extension of a measure $\nu_{E,N}$ on the one-sided space $\Sigma^+_{B_N}$ to a measure ν_N on the two-sided space Σ_{B_N} is carried out as follows (see the remarks following lemma 1.13 of [1]).

$$\int_{\Sigma_{B_N}} f \, \mathrm{d}\nu_N := \lim_{k\to\infty} \int_{\Sigma^+_{B_N}} \omega_N(f \circ \sigma^k) \, \mathrm{d}\nu_{E,N} \qquad \text{for every } f \in C(\Sigma_{B_N}, \mathbb{R}). \tag{34}$$

In particular, setting $f = \chi_{[\xi_0]}$, we see that

$$\nu_N([\xi_0]) = \nu_{E,N}([\xi_0]), \tag{35}$$

where the one-cylinder $[\xi_0]$ is considered as a subset of Σ_{B_N} on the l.h.s. of the equality and as a subset of $\Sigma^+_{B_N}$ on the r.h.s. In order to compare this with the definition given earlier on smooth space, note that the analogue of Υ_N on Σ_{B_N} is simply ω_N^{-1} (it is strictly not ω_N^{-1} as ω_N does not have Σ_{B_N} as its range; we really mean the function which just deletes the past), and in fact the following diagram commutes.

$$\begin{array}{ccc} \Sigma_{B_N} & \xrightarrow{\omega_N^{-1}} & \Sigma^+_{B_N} \\ \downarrow{\scriptstyle \pi_N} & & \downarrow{\scriptstyle \pi_N} \\ M'_N & \xrightarrow{\Upsilon_N} & X_N \end{array} \tag{36}$$

Lemma 1.16 [1] states that the natural extension of the measure μ_{E,ϑ_N} on $\Sigma^+_{B_N}$ (with density Ψ_{E,ϑ_N}) is equal to $\mu_{N,\varphi} = \mu \circ \pi_N$ for every $N \geqslant 0$. As for the natural extension of

μ_{E,φ_N} we are really only concerned with its value on one-cylinder sets in Σ_{B_N}. In fact, we will not perform a complete extension as described in the earlier paragraph, but simply define a measure on Σ_{B_N} by treating one-cylinders $[\xi_0] \in \Sigma^+_{B_N}$ as one-cylinders in Σ_{B_N} and spreading the weight $\mu_{E,\varphi_N}([\xi_0])$ evenly over each one-cylinder for $1 \leqslant \xi_0 \leqslant n_N$. The projections of our one-cylinders on M'_N will shrink in size exponentially, so we will only need to show that $|\mu_{\vartheta_N}([\xi_0]) - \mu_{\varphi_N}([\xi_0])| \to 0$ quickly enough as $N \to \infty$. To end this section, we have the following lemma.

Lemma 7.5. $\sum_{i=1}^{n_N} |\mu(A_{N,i}) - \mu_N(A_{N,i})| = \mathrm{O}((s_\pm^{\gamma_1})^N)$, for $\gamma_1 < \gamma$.

Proof. Recall that $\mu_N := \mu_{\phi_N} := \mu_{\varphi_N} \circ \pi_N^{-1}$ and $\mu := \mu_{\phi^{(u)}} := \mu_{N,\varphi} \circ \pi_N^{-1}$. We note that

(i) $\mu_N(A_{N,i}) := \mu_{\varphi_N}([i]) = \mu_{E,\varphi_N}([i]) = \int_{[i]} \Psi_{E,\varphi_N}(\xi)\, \mathrm{d}\,(m \circ \pi_N)(\xi)$, and

(ii) $\mu(A_{N,i}) := \mu_{N,\varphi}([i]) = \mu_{E,\vartheta_N}([i]) = \int_{[i]} \Psi_{E,\vartheta_N}(\xi)\, \mathrm{d}\,(m \circ \pi_N)(\xi)$.

So then

$$\sum_{i=1}^{n_N} |\mu(A_{N,i}) - \mu_N(A_{N,i})| \leqslant \sum_{i=1}^{n_N} |\Psi_{E,\vartheta_N} - \Psi_{E,\varphi_N}|_\infty \cdot m(A_{N,i}) = |\Psi_{E,\vartheta_N} - \Psi_{E,\varphi_N}|_\infty$$
$$= \mathrm{O}((\theta/\theta')^N) = \mathrm{O}((\theta_1)^N) = \mathrm{O}((s_\pm^{\gamma_1})^N),$$

where we put θ' close to 1, so that $s_\pm^\gamma = \theta < \theta_1 = s_\pm^{\gamma_1}$, with θ_1 close to θ (and $\gamma_1 < \gamma$ close to γ). □

7.4.3. Proof of convergence of invariant measure estimates. We define a measure $\bar{\mu}_N$ by

$$\bar{\mu}_N = \sum_{i=1}^{n_N} \frac{m(A \cap A_{N,i})}{m(A)} \cdot \mu(A_{N,i}).$$

The measure $\bar{\mu}_N$ simply spreads the weight given to each partition set evenly over each set.

We proceed first for systems whose SBR measure μ is absolutely continuous with respect to Lebesgue measure (and have density ψ). Note that we may find a universal constant $C_2 < \infty$ such that

$$\max_{A_{N,i} \in \mathfrak{P}^N} \mathrm{diam}(A_{N,i}) \leqslant C_2 s_\pm^N. \tag{37}$$

We split our estimate of $\|\mu - \mu_N\|_{L^1}$ into two parts. We first compare μ with $\bar{\mu}_N$, and then compare $\bar{\mu}_N$ with μ_N. For the former comparison, we simply note that as ψ (the density of μ) is C^γ, and μ and $\bar{\mu}_N$ have the same integral over each partition set, a simple Hölder continuity estimate will provide us with convergence. As for the latter comparison, we use lemma 7.5.

$$\begin{aligned}\|\mu - \mu_N\|_{L^1} &\leqslant \|\mu - \bar{\mu}_N\|_{L^1} + \|\bar{\mu}_N - \mu_N\|_{L^1} \\ &\leqslant |\psi|_\gamma (\max_{1\leqslant i \leqslant n_N} \mathrm{diam}(A_{N,i}))^\gamma + \sum_{i=1}^{n_N} |\mu(A_{N,i}) - \mu_N(A_{N,i})| \\ &\leqslant |\psi|_\gamma (C_2 s_\pm^N)^\gamma + \mathrm{O}((s_\pm^{\gamma_1})^N) = \mathrm{O}((s_\pm^{\gamma_1})^N).\end{aligned}$$

We now treat the case of Anosov systems with non-absolutely continuous SBR measure μ. In this situation μ does not have a density ψ, and we lose the easy Hölder continuity bound provided by ψ. We instead prove that μ_N converges to μ *weakly*, but at the same rate as before. We again split the difference $\mu - \mu_N$ into two parts:

$$\|\mu - \mu_N\|_w \leqslant \|\mu - \bar{\mu}_N\|_w + \|\bar{\mu}_N - \mu_N\|_w. \tag{38}$$

We may treat the second term as above, noting that $\|\bar{\mu}_N - \mu_N\|_w \leqslant \|\bar{\mu}_N - \mu_N\|_{L^1}$. For the first term, we proceed as follows:

$$\begin{aligned}\|\mu - \bar{\mu}_N\|_w &= \sup\left\{ \sum_{i=1}^{n_N} \left(\int_{A_{N,i}} g \,\mathrm{d}\mu - \int_{A_{N,i}} g \,\mathrm{d}\bar{\mu}_N \right) : \|g\|_{L^1} = 1, \mathrm{Lip}\, g = 1 \right\} \\ &\leqslant \sum_{i=1}^{n_N} \mathrm{diam}(A_{N,i})\mu(A_{N,i}) \leqslant \max_{1\leqslant i\leqslant n_N} \mathrm{diam}(A_{N,i}) \leqslant C_2(s_\pm^\gamma)^N. \qquad (39)\end{aligned}$$

This completes the proof for the order of convergence of the approximate measures μ_N to the SBR measure μ.

7.4.4. Lyapunov exponent and entropy estimates. We now move on to the order of convergence of the approximate metric entropy to the true value. The proof follows in a similar manner to that for the expanding case. We define a piecewise constant function $\kappa_N : M \to \mathbb{R}^+$ that will approximate† $|\det D_x T_{|E_x^u}|^{-1}$ by

$$\kappa_N(x) = Q_{N,ij} = P_{N,ij} \cdot \frac{\ell^u(A_{N,i})}{\ell^u(A_{N,j})}, \qquad \text{where } x \in A_{N,i} \cap T^{-1}A_{N,j}.$$

Since $\det D_x T_{|E_x^u}$ is C^γ and bounded away from zero, one may find a constant $C_1 < \infty$ such that

$$|-\log|\det D_x T_{|E_x^u}| - \log \kappa_N| \leqslant C_1 \max_{1\leqslant i\leqslant n_N} (\mathrm{diam}((A_{N,i}))^\gamma.$$

Moreover, as before, one may find a universal constant $C_2 < \infty$ such that

$$\max_{1\leqslant i\leqslant n_N} \mathrm{diam}(A_{N,i}) \leqslant C_2 s_\pm^N. \qquad (40)$$

Noting that $\log \varrho_N = 0$ for $N \geqslant 0$, we have

$$\begin{aligned}|h_N - h_\mu(T)| &= \left| -\sum_{i,j=1}^{n_N} p_{N,i} P_{N,ij} \log P_{N,ij} \cdot \frac{\ell^u(A_{N,i})}{\ell^u(A_{N,j})} - \int_M \log|\det D_x T_{|E_x^u}| \,\mathrm{d}\mu\,(x) \right| \\ &= \left| \sum_{i,j=1}^{n_N} \left(\int_{A_{N,i}\cap T^{-1}A_{N,j}} -\log \kappa_N(x) \,\mathrm{d}\mu_N\,(x) \right.\right. \\ &\quad \left.\left. - \int_{A_{N,i}\cap T^{-1}A_{N,j}} \log|\det D_x T_{|E_x^u}| \,\mathrm{d}\mu\,(x) \right) \right| \\ &\leqslant \sum_{i,j=1}^{n_N} \left| \int_{A_{N,i}\cap T^{-1}A_{N,j}} -\log \kappa_N(x) \,\mathrm{d}\mu_N\,(x) + \int_{A_{N,i}\cap T^{-1}A_{N,j}} \log \kappa_N(x) \,\mathrm{d}\mu\,(x) \right| \\ &\quad + \sum_{i,j=1}^{n_N} \int_{A_{N,i}\cap T^{-1}A_{N,j}} |-\log \kappa_N(x) - \log|\det D_x T_{|E_x^u}|| \,\mathrm{d}\mu\,(x) \\ &= \sum_{i,j=1}^{n_N} \left| \int_{\Upsilon_N(A_{N,i}\cap T^{-1}A_{N,j})} -\log \kappa_N(x) \,\mathrm{d}\mu_{E,N}\,(x) \right. \\ &\quad \left. + \int_{\Upsilon_N(A_{N,i}\cap T^{-1}A_{N,j})} \log \kappa_N(x) \,\mathrm{d}\mu_E\,(x) \right| \\ &\quad + \sum_{i,j=1}^{n_N} \int_{A_{N,i}\cap T^{-1}A_{N,j}} |-\log \kappa_N(x) - \log|\det D_x T_{|E_x^u}|| \,\mathrm{d}\mu\,(x)\end{aligned}$$

† For convenience, we use the definition of Q_N and P_N given in remark 5.4 (ii). The same result is true for Q_N and P_N constructed as in (9) and (10) (see [7]).

$$\leqslant |\log \kappa_N|_\infty \|\mu_{E,N} - \mu_E\|_{L^1} + C_1(C_2 s_\pm^{N+1})^\gamma$$
$$\leqslant |\det D_x T_{|E_x^u}|_\infty \|\mu_{E,N} - \mu_E\|_{L^1} + C_1(C_2 s_\pm^{N+1})^\gamma$$
$$= \mathrm{O}((s_\pm^{\gamma_1})^N) + C_1 C_2^\gamma (s_\pm^\gamma)^{N+1} \text{ as } \|\mu_{E,N} - \mu_E\|_{L^1} = \mathrm{O}((s_\pm^{\gamma_1})^N). = \mathrm{O}((s_\pm^{\gamma_1})^N).$$

7.5. Proof of the axiom A (attracting) case

We again project the dynamics onto the unstable sides of the Markov partition sets, and use the results for expanding maps to obtain estimates for the unique absolutely continuous measure on the unstable sides. The measure on each unstable side is then distributed back over the entire set to give an estimate of the SBR measure. Recall that our Markov partition sets are blurry images of pieces of the attractor, which is really a product of an interval with a Cantor set. The weight that is given to each partition set is an estimate of the SBR measure of the parts of the attractor contained in that partition set. As our partition is refined, it approximates the Cantor structure more finely. Inequalities (37) and (40) still hold, and the result concerning the rate of convergence for the SBR measure estimates follows as in the Anosov case for non-absolutely continuous μ. The proof of the rate of convergence of the entropy estimates also follows as in the Anosov case; we are able to equate the sum of the Lyapunov exponents with the entropy by following the final paragraph in the next section, noting that theorem 4.11 of Bowen [1] tells us that the pressure $\log \varrho = P(T) = 0$ for attractors.

7.6. Proof of the axiom A (non-attracting) case

The proof follows along the lines of the axiom A attractor case. As in the non-attracting case, our Markov partitions are a blurry picture of a Cantor set, and contain lots of points that should not really be there. And again, the weight given to one of our Markov partition sets will approximate the weight given to that part of the Cantor set contained in the partition set.

In the case of a non-attractor, we *must* construct φ_N directly, using the matrix Q_N in equation (9) rather than indirectly through P_N in (15) and a similarity transformation, as in the other three situations. The proof that $\mu_N \to \mu$ follows exactly as described in section 7.4.3 for Anosov systems with non-absolutely continuous μ, and equation (12) goes through as in the proof of convergence of entropy in the Anosov case, with the appropriate rate of convergence.

One difference in this case is that the maximal eigenvalue of $\mathcal{L}$ is not unity, but some number ϱ strictly less than 1. The facts that this maximal eigenvalue ϱ is simple, positive, and equals $\exp(P(\sigma))$, where $P(\sigma)$ is the pressure of the shift σ, may be found in theorem 2.2 of [13]. The maximal eigenvalues ϱ_N of the matrices Q_N converge to ϱ by standard spectral approximation theory, as we have norm convergence of $\mathcal{L}_{E,\varphi_N}$ to $\mathcal{L}_{E,\vartheta_N}$. The result that $\varrho_N \to \varrho$ follows along the lines of the proof of proposition 7.3 of [8]. The rate of convergence is $\mathrm{O}((\theta/\theta')^N)$, or equivalently, $\mathrm{O}((s_\pm^{\gamma_1})^N)$, $\gamma_1 < 1$.

We need now only connect the Lyapunov exponents and pressure with the metric entropy; this will follow from the variational principle. Theorem 3.5 of [13] states that $h_{\mu_\varphi}(\sigma) + \int_{\Sigma_B^+} \varphi \, \mathrm{d}\mu_\varphi = \log \varrho$. Projecting this equality down to our smooth space, we obtain $h_{\mu'}(T) = \log \varrho + \int_M \phi^{(u)} \, \mathrm{d}\mu' = \log \varrho + \sum_{\lambda^i(\mu')>0} \lambda^i(\mu')$. Combining this equality with the convergence orders for λ_N and ϱ_N, we see that $h_N \to h_{\mu'}(T)$ like $\mathrm{O}((s_\pm^{\gamma_1})^N)$.

7.7. Errors from using inexact Markov partitions in example 6.4

We give a brief error analysis for the use of inexact Markov partitions in example 6.4. The matrices Q_N provide us with an approximation of the Perron–Frobenius (or transfer) operator of T. In fact, we have seen that Q_N is a representation of a nearby Perron–Frobenius operator $\mathcal{L}_{\varphi_N}$. By using an approximate Markov partition, we now have two errors in the weight function φ. First, the error from using φ_N, the piecewise constant approximation to φ, and secondly because of the the extra error in this piecewise constant approximation due to the slightly inaccurate Markov partition sets. We have shown that the former error decays exponentially with N. The error due to the inexact Markov partitions may be dealt with as follows.

In order to produce a good estimate of $|\det D_x T|_{E_x^u}|^{-1}$ from (9) or (16), it is important that remark 6.5 (iv′)(b) is satisfied (the 'stable' edges of $T^{-1}A_j$ are inside the 'stable' edges of A_i if $A_i \cap T^{-1}A_j \neq \{\emptyset\}$). It is simple to show by induction that this property holds for each of the approximate Markov partitions defined by $\mathfrak{P}^N = \bigvee_{i=-N}^{N} T^i S$. In particular, it is clear from figures 2 and 3 that the 'stable' vertical edges of $T^{-1}\mathfrak{P}^1$ are inside those of $\mathfrak{P}^1$, and similarly for $T^{-1}\mathfrak{P}^2$ and $\mathfrak{P}^2$. It is also important that the 'unstable' edges quickly align themselves with the true unstable directions, so that we are measuring the compression of T^{-1} in the correct direction. We now outline the rate at which this happens. Define an *unstable cone* at x, $C^u(x) = \{(v_x, v_y) \in T_x\mathbb{R}^2 : |v_y| \leqslant t|v_x|\}$, $t = \frac{1}{1.7}$. It is shown in section 7.4.1 of [14] that the maximum angle between vectors in $\bigcap_{k=0}^{N} D_{T^{-k}x}T^k(C^u(T^{-k}x))$ is bounded by constant $\times$ $(0.3/1.7^2)^N$. Given a point x on the boundary of an element A of $\mathfrak{P}^N$, this result provides a bound for the angle between the tangent vector of ∂A at x, and E_x^u, the true unstable direction at x. Because $\max_{1\leqslant i\leqslant n_N} \operatorname{diam} A_{N,i} \leqslant \text{constant}/1.7^N$, the boundaries of elements of $\mathfrak{P}^N$ become increasingly linear (as they have a bounded curvature), so that the elements are almost-parallelograms. By combining (i) the exponential alignment of the boundaries of $\mathfrak{P}^N$ with the unstable directions (and similarly, the stable directions), and (ii) the exponential approach to linearity of the boundaries, we see that the extra error incurred in the entries of Q_N (and hence in the approximation φ_N) decays exponentially with N.

Acknowledgments

The author is grateful to two referees for their helpful comments. This work was completed at the Centre for Applied Dynamics and Optimization, Department of Mathematics, The University of Western Australia and was partially supported by a grant from the Australian Research Council.

References

[1] Bowen R 1975 Equilibrium states and the ergodic theory of Anosov diffeomorphisms *Lecture Notes in Mathematics 470* (Berlin: Springer)

[2] Boyarsky A 1988 A matrix method for estimating the Liapunov exponent of one-dimensional systems *J. Stat. Phys.* **50** 213–29

[3] Cawley E 1991 Smooth Markov partitions and toral automorphisms *Ergod. Theor. Dynam. Syst.* **11** 633–51

[4] Crawford J and Cary J 1983 Decay of correlations in a chaotic measure-preserving transformation *Physica D* **6** 223–32

[5] Ding J and Zhou A 1996 Finite approximations of Frobenius–Perron operators. a solution of Ulam's conjecture to multi-dimensional transformations *Physica D* **92** 61–8

[6] Chatelin F 1983 Spectral approximation of linear operators *Computer Science and Applied Mathematics* (New York: Academic)

[7] Froyland G 1995 Finite approximation of Sinai–Bowen–Ruelle measures of Anosov systems in two dimensions *Rand. Comput. Dynam.* **3** 251–64

[8] Froyland G 1997 Computer-assisted bounds for the rate of decay of correlations *Commun. Math. Phys.* **189** 237–57

[9] Froyland G, Judd K and Mees A I 1995 Estimation of Lyapunov exponents of dynamical systems using a spatial average *Phys. Rev.* E **51** 2844–55

[10] Katok A and Hasselblatt B 1995 Introduction to the modern theory of dynamical systems *Encyclopedia of Mathematics and its Applications* vol 54 (Cambridge: Cambridge University Press)
[11] Mañé R 1987 *Ergodic Theory and Differentiable Dynamics* (Berlin: Springer)
[12] Palis J and Takens F 1993 Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations *Cambridge Studies in Advanced Mathematics* vol 35 (Cambridge: Cambridge University Press)
[13] Parry W and Pollicott M 1990 *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics* vol 187–8 (Paris: Société Mathématique de France)
[14] Robinson C 1995 *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos* (Boca Raton, FL: Chemical Rubber Company)
[15] Ruelle D 1976 A measure associated with axiom-A attractors *Am. J. Math.* **98** 619–54
[16] Ruelle D 1989 The thermodynamic formalism for expanding maps *Commun. Math. Phys.* **125** 239–62
[17] Ulam S M 1964 *Problems in Modern Mathematics* (New York: Interscience)