UNIQUE EQUILIBRIUM STATES FOR FLOWS AND HOMEOMORPHISMS WITH NON-UNIFORM STRUCTURE

VAUGHN CLIMENHAGA AND DANIEL J. THOMPSON

ABSTRACT. Using an approach due to Bowen, Franco showed that continuous expansive flows with specification have unique equilibrium states for potentials with the Bowen property. We show that this conclusion remains true using weaker non-uniform versions of specification, expansivity, and the Bowen property. We also establish a corresponding result for homeomorphisms. In the homeomorphism case, we obtain the upper bound from the level-2 large deviations principle for the unique equilibrium state. The theory presented in this paper provides the basis for an ongoing program to develop the thermodynamic formalism in partially hyperbolic and non-uniformly hyperbolic settings.

1. INTRODUCTION

Let X be a compact metric space and $F = (f_t)_{t \in \mathbb{R}}$ a continuous flow on X. Given a potential function $\phi \colon X \to \mathbb{R}$, we study the question of existence and uniqueness of equilibrium states for (X, F, ϕ) – that is, invariant measures which maximize the quantity $h_\mu(f_1) + \int \phi \, d\mu$. We also study the same question for homeomorphisms $f \colon X \to X$. This problem has a long history [Bow75, Bow08, HK82, DKU90, Sar99, IT10, PSZ16, Pav16, CP16] and is connected with the study of global statistical properties for dynamical systems [Rue76, Kif90, PP90, BSS02, CRL11, Cli16].

For homeomorphisms, Bowen showed [Bow75] that (X, f, ϕ) has a unique equilibrium state whenever (X, f) is an expansive system with specification and ϕ satisfies a certain regularity condition (the Bowen property). Bowen's method was adapted to flows by Franco [Fra77]. Previous work by the authors established similar uniqueness results for shift spaces with a broad class of potentials [CT12, CT13], and for non-symbolic discrete-time systems in the case $\phi \equiv 0$ [CT14]. In this paper, we consider potential functions satisfying a non-uniform version of the Bowen property in both the discrete- and continuous-time case.

While we do not explore applications of this theory in this paper, we emphasize that the results are developed with a view to novel applications in the setting of smooth dynamical systems beyond uniform hyperbolicity. In

particular, the main theorems of this paper are applied to diffeomorphisms with weak forms of hyperbolicity in [CFT15] and to geodesic flows in non-positive curvature in [BCFT16].

We review the main points of our techniques for proving uniqueness of equilibrium states for maps, referring the reader to [CT13, CT14] for details. Our approach is based on weakening each of the three hypotheses of Bowen's theorem: expansivity, the specification property, and regularity of the potential. Instead of asking for specification and regularity to hold globally, we ask for these properties to hold on a suitable collection of orbit segments $\mathcal{G}$. Instead of asking for expansivity to hold globally, we ask that all measures with large enough free energy should observe expansive behavior.

These ideas lead naturally to a notion of orbit segments which are obstructions to specification and regularity, and measures which are obstructions to expansivity. The guiding principle of our approach is that if these obstructions have less topological pressure than the whole space, then a version of Bowen's strategy can still be developed. Some of the main points are as follows:

(1) For a discrete-time dynamical system, we work with $X \times \mathbb{N}$, which we think of as the space of orbit segments by identifying (x, n) with $(x, f(x), \ldots, f^{n-1}(x))$. At the heart of our approach is the concept of a *decomposition* $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ for $X \times \mathbb{N}$. We ask for specification and regularity to hold on a collection of 'good' orbit segments $\mathcal{G} \subset X \times \mathbb{N}$, while the collections of $\mathcal{P}, \mathcal{S} \subset X \times \mathbb{N}$ are thought of as 'bad' orbit segments which are obstructions to specification and regularity. We ask that any orbit segment can be decomposed as a 'good core' that is preceded and succeeded by elements of $\mathcal{P}$ and $\mathcal{S}$, respectively. More precisely, for any (x, n), there are numbers $p, g, s \in \mathbb{N} \cup \{0\}$ so that $p + g + s = n$, and

$$(x, p) \in \mathcal{P}, \quad (f^p x, g) \in \mathcal{G}, \quad (f^{p+g} x, s) \in \mathcal{S}.$$

The choice of the decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ depends on the setting of any given application, and the dynamics of the situation are encoded in this choice.

(2) We define a natural version of topological pressure for orbit segments, and we require that the topological pressure of $\mathcal{P} \cup \mathcal{S}$, which we think of as the pressure of the obstructions to specification and regularity, is less than that of the whole space.

(3) The positive expansivity property introduced in [CT14] is that for small ε, $\Gamma_\varepsilon(x) = \{y : d(f^n x, f^n y) \leq \varepsilon \text{ for all } n \geq 0\} = \{x\}$ for μ-almost every x, for any ergodic μ with $h_\mu(f) > h$, where h is a constant less than $h_{\text{top}}(f)$. We think of the smallest $h \geq 0$ so that this is true as the entropy of obstructions to expansivity.

Under these hypotheses, our strategy is then inspired by Bowen's: his main idea was to construct an equilibrium state with the Gibbs property,

and to show that this rules out the existence of a mutually singular equilibrium state. We obtain a certain Gibbs property which only applies to orbit segments in $\mathcal{G}$, and then we have to work to show that this is still sufficient to prove uniqueness of the equilibrium state.

The above strategy was carried out in [CT12, CT13, CT14] under the assumption that either (X, f) is a shift space or $\phi = 0$. In this paper, we work in the setting of a continuous flow or homeomorphism on a compact metric space, and a continuous potential function. This necessitates several new developments, which we now describe. For homeomorphisms and flows, we develop a theory for potential functions which are regular only on 'good' orbit segments. The lack of global regularity introduces fundamental technical difficulties not present in the classical theory or the symbolic setting. For flows, which are the main focus of this paper, we work with the space $X \times \mathbb{R}^+$, where the pair (x, t) is thought of as the orbit segment $\{f_s(x) \mid 0 \le s < t\}$. The main points addressed in this paper are:

(1) Our potentials are not regular on the whole space, and this forces us to introduce and control non-standard 'two-scale' partition sums throughout the proof (see §2.1).

(2) For flows, expansivity issues can be subtle and require new ideas beyond the discrete-time case. We introduce the notion of *almost expansivity* for a flow-invariant ergodic measure (§2.5), adapting a discrete-time version of this definition which was used in [CT14]. We also introduce the notion of *almost entropy expansivity* (§3.1) for a map-invariant ergodic measure. This is a natural analogue of entropy expansivity [Bow72], adapted to apply to almost every point in the space. Measures which are almost expansive for the flow are almost entropy expansive for the time-t map. Almost entropy expansivity plays a crucial role in our proof via Theorem 3.2, a general ergodic theoretic result that strengthens [Bow72, Theorem 3.5].

(3) Adapting the framework introduced in [CT14] to the case of flows requires careful control of small differences in transition times, particularly in Lemma 4.4.

(4) The unique equilibrium state we construct admits a weak upper Gibbs bound, which in many cases we use to obtain the upper bound from the level-2 large deviations principle, using results of Pfister and Sullivan (see §7).

We now state a version of our main result, which should be understood as a formalization of the strategy described previously. We introduce our notation, referring the reader to §2 for precise definitions: $P(\phi)$ is the standard topological pressure; the quantity $P_{\exp}^{\perp}(\phi)$ is the largest free energy of an ergodic measure which observes non-expansive behavior; the specification property and Bowen property are versions of the classic properties which apply only on $\mathcal{G}$ rather than globally; the expression $P([\mathcal{P}] \cup [\mathcal{S}], \phi)$ is the topological pressure of the obstructions to specification and regularity.

Theorem A. *Let (X, F) be a continuous flow on a compact metric space, and $\phi\colon X \to \mathbb{R}$ a continuous potential function. Suppose that $P^{\perp}_{\exp}(\phi) < P(\phi)$ and that $X \times \mathbb{R}^+$ admits a decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ with the following properties:*

  (I) $\mathcal{G}$ *has the weak specification property;*
  (II) ϕ *has the Bowen property on $\mathcal{G}$;*
 (III) $P([\mathcal{P}] \cup [\mathcal{S}], \phi) < P(\phi)$.

Then (X, F, ϕ) has a unique equilibrium state.

In fact, we will prove a slightly more general result, of which Theorem A is a corollary. The more general version, Theorem 2.9, applies under slightly weaker versions of our hypotheses, which we discuss and motivate in §2.6.

We also develop versions of our results that apply for homeomorphisms. These discrete-time arguments are analogous to, and easier than, the flow case, so we just outline the proof, highlighting any differences with the flow case. Our main results for homeomorphisms are Theorem 5.5, which is the analogue of Theorem A, and Theorem 5.6, which is the analogue for homeomorphisms of Theorem 2.9. Finally, in Theorem 5.7, we establish the upper level-2 large deviations principle for the unique equilibrium states provided by Theorem 5.5.

Structure of the paper. We collect our definitions, particularly for flows, in §2. Our main results for flows are proved in §§3–4. Our main results for maps are proved in §§5–6. In §7, we prove the large deviations results of Theorem 5.7. In §8, we prove Theorem 3.2, which is a self-contained result about measure-theoretic entropy for almost entropy expansive measures.

2. DEFINITIONS

In this section we give the relevant definitions for flows; the corresponding definitions for maps are given in §5.

2.1. **Partition sums and topological pressure.** Throughout, X will denote a compact metric space and $F = (f_t)_{t \in \mathbb{R}}$ will denote a continuous flow on X. We write $\mathcal{M}_F(X)$ for the set of Borel F-invariant probability measures on X. Given $t \geq 0$, $\delta > 0$, and $x, y \in X$ we define the *Bowen metric*

(2.1) $$d_t(x, y) := \sup\{d(f_s x, f_s y) \mid s \in [0, t]\},$$

and the *Bowen balls*

(2.2)
$$B_t(x, \delta) := \{y \in X \mid d_t(x, y) < \delta\},$$
$$\overline{B}_t(x, \delta) := \{y \in X \mid d_t(x, y) \leq \delta\}.$$

Given $\delta > 0$, $t \in \mathbb{R}^+$, and $E \subset X$, we say that E is (t, δ)-separated if for every distinct $x, y \in E$ we have $y \notin \overline{B}_t(x, \delta)$. Writing $\mathbb{R}^+ = [0, \infty)$, we view $X \times \mathbb{R}^+$ as the space of finite orbit segments for (X, F) by associating to each pair (x, t) the orbit segment $\{f_s(x) \mid 0 \leq s < t\}$. Our convention is

that $(x, 0)$ is identified with the empty set rather than the point x. Given $\mathcal{C} \subset X \times \mathbb{R}^+$ and $t \geq 0$ we write $\mathcal{C}_t = \{x \in X \mid (x, t) \in \mathcal{C}\}$.

Now we fix a continuous potential function $\phi \colon X \to \mathbb{R}$. Given a fixed scale $\varepsilon > 0$, we use ϕ to assign a weight to every finite orbit segment by putting

$$(2.3) \qquad \Phi_\varepsilon(x, t) = \sup_{y \in B_t(x, \varepsilon)} \int_0^t \phi(f_s y) \, ds.$$

In particular, $\Phi_0(x, t) = \int_0^t \phi(f_s x) \, ds$. The general relationship between $\Phi_\varepsilon(x, t)$ and $\Phi_0(x, t)$ is that

$$(2.4) \qquad |\Phi_\varepsilon(x, t) - \Phi_0(x, t)| \leq t \operatorname{Var}(\phi, \varepsilon),$$

where $\operatorname{Var}(\phi, \varepsilon) = \sup\{|\phi(x) - \phi(y)| \mid d(x, y) < \varepsilon\}$.

Given $\mathcal{C} \subset X \times \mathbb{R}^+$ and $t > 0$, we consider the *partition function*

$$(2.5) \qquad \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t) = \sup \left\{ \sum_{x \in E} e^{\Phi_\varepsilon(x, t)} \mid E \subset \mathcal{C}_t \text{ is } (t, \delta)\text{-separated} \right\}.$$

We will often suppress the function ϕ from the notation, since it is fixed throughout the paper, and simply write $\Lambda(\mathcal{C}, \delta, \varepsilon, t)$. When $\mathcal{C} = X \times \mathbb{R}^+$ is the entire system, we will simply write $\Lambda(X, \phi, \delta, \varepsilon, t)$ or $\Lambda(X, \delta, \varepsilon, t)$. We call a (t, δ)-separated set that attains the supremum in (2.5) *maximizing* for $\Lambda(\mathcal{C}, \delta, \varepsilon, t)$. We are only guaranteed the existence of such sets when $\mathcal{C} = X \times \mathbb{R}^+$, since otherwise $\mathcal{C}_t$ may not be compact.

The pressure of ϕ on $\mathcal{C}$ at scales δ, ε is given by

$$(2.6) \qquad P(\mathcal{C}, \phi, \delta, \varepsilon) = \varlimsup_{t \to \infty} \frac{1}{t} \log \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t).$$

Note that Λ is monotonic in both δ and ε, but in different directions; thus the same is true of P. Again, we write $P(\mathcal{C}, \phi, \delta)$ in place of $P(\mathcal{C}, \phi, \delta, 0)$ to agree with more standard notation, and we let

$$(2.7) \qquad P(\mathcal{C}, \phi) = \lim_{\delta \to 0} P(\mathcal{C}, \phi, \delta).$$

When $\mathcal{C} = X \times \mathbb{R}^+$ is the entire space of orbit segments, the topological pressure reduces to the usual notion of topological pressure on the entire system, and we write $P(\phi, \delta)$ in place of $P(\mathcal{C}, \phi, \delta)$, and $P(\phi)$ in place of $P(\mathcal{C}, \phi)$. The variational principle for flows [BR75] states that $P(\phi) = \sup\{h_\mu(f_1) + \int \phi \, d\mu \mid \mu \in \mathcal{M}_F(X)\}$, where $h_\mu(f_1)$ is the usual measure-theoretic entropy of the time-1 map of the flow. A measure achieving the supremum is called an *equilibrium state*.

Remark 2.1. The most obvious definition of partition function would be to take $\varepsilon = 0$ so that the weight given to each orbit segment is determined by the integral of the potential function along that exact orbit segment, rather than by nearby ones. To match more standard notation, we often write $\Lambda(\mathcal{C}, \phi, \delta, t)$ in place of $\Lambda(\mathcal{C}, \phi, \delta, 0, t)$. The partition sums $\Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t)$

arise throughout this paper, particularly in §4.1 and §4.6. The relationship between the two quantities can be summarised as follows.

(1) If (X, F) is expansive at scale ε, then $P(\mathcal{C}, \phi, \delta, \varepsilon) = P(\mathcal{C}, \phi, \delta, 0)$.
(2) If ϕ is Bowen at scale ε, then the two pressures above are equal, and moreover the ratio between $\Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t)$ and $\Lambda(\mathcal{C}, \phi, \delta, 0, t)$ is bounded away from 0 and ∞.
(3) In the absence of regularity or expansivity assumptions, we have the relationship

$$e^{-t \, \mathrm{Var}(\phi, \varepsilon)} \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t) \le \Lambda(\mathcal{C}, \phi, \delta, t) \le e^{t \, \mathrm{Var}(\phi, \varepsilon)} \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t),$$

and thus $|P(\mathcal{C}, \phi, \delta, \varepsilon) - P(\mathcal{C}, \phi, \delta)| \le \mathrm{Var}(\phi, \varepsilon)$. By continuity of ϕ, this establishes that $P(\mathcal{C}, \phi, \delta, \varepsilon) \to P(\mathcal{C}, \phi, \delta, 0)$ as $\varepsilon \to 0$, but does not give us the conclusions of (1) or (2).

Because our versions of expansivity and the Bowen property do not hold globally, we are in case (3) above, so a priori we cannot replace $\Lambda(\mathcal{C}, \phi, \delta, \varepsilon, t)$ with $\Lambda(\mathcal{C}, \phi, \delta, t)$ in the proofs.

Remark 2.2. We can restrict to (t, δ)-separated sets of maximal cardinality $E \subset \mathcal{C}_t$ in the definition of pressure: these always exist, even when $\mathcal{C}_t$ is non-compact, since the possible values for the cardinality are finite (by compactness of X). If E were not of maximal cardinality, we could just add in another point, which would increase the partition sum (2.5). Furthermore, a (t, δ)-separated set E of maximal cardinality is (t, δ)-spanning in the sense that $D_t \subset \bigcup_{x \in E} \overline{B}_t(x, \delta)$. If this were not so then we could add another point to E and increase the cardinality.

2.2. **Decompositions.** We introduce the notion of a decomposition for a sub-collection of the space of orbit segments.

Definition 2.3. *A* decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ *for* $\mathcal{D} \subseteq X \times \mathbb{R}^+$ *consists of three collections* $\mathcal{P}, \mathcal{G}, \mathcal{S} \subset X \times \mathbb{R}^+$ *and three functions* $p, g, s : \mathcal{D} \to \mathbb{R}^+$ *such that for every* $(x, t) \in \mathcal{D}$, *the values* $p = p(x, t)$, $g = g(x, t)$, *and* $s = s(x, t)$ *satisfy* $t = p + g + s$, *and*

(2.8) $(x, p) \in \mathcal{P}, \quad (f_p x, g) \in \mathcal{G}, \quad (f_{p+g} x, s) \in \mathcal{S}.$

If $\mathcal{D} = X \times \mathbb{R}^+$, *we say that* $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ *is a decomposition for* (X, F). *Given a decomposition* $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ *and* $M \in \mathbb{R}^+$, *we write* $\mathcal{G}^M$ *for the set of orbit segments* $(x, t) \in \mathcal{D}$ *for which* $p \le M$ *and* $s \le M$.

We make a standing assumption that $X \times \{0\} \subset \mathcal{P} \cap \mathcal{G} \cap \mathcal{S}$ to allow for orbit segments to be decomposed in 'trivial' ways; for example, (x, t) can belong 'purely' to one of the collections $\mathcal{P}, \mathcal{G}$, or $\mathcal{S}$ or can transition directly from $\mathcal{P}$ to $\mathcal{S}$ – note that formally the symbols $(x, 0)$ are identified with the empty set. This is implicit in our earlier work [CT12, CT13, CT14].

We will be interested in decompositions where $\mathcal{G}$ has specification, ϕ has the Bowen property on $\mathcal{G}$, and $\mathcal{P} \cup \mathcal{S}$ carries smaller pressure than the entire system. In the case of flows, a priori we must replace the collections $\mathcal{P}$

and $\mathcal{S}$ that appear in the decomposition with a related and slightly larger collection $[\mathcal{P}] \cup [\mathcal{S}]$, where given $\mathcal{C} \subset X \times \mathbb{R}^+$ we write

(2.9) $[\mathcal{C}] := \{(x, n) \in X \times \mathbb{N} \mid (f_{-s}x, n + s + t) \in \mathcal{C} \text{ for some } s, t \in [0, 1]\}.$

Passing from $\mathcal{P} \cup \mathcal{S}$ to $[\mathcal{P}] \cup [\mathcal{S}]$ ensures that the decomposition is well behaved with respect to replacing continuous time with discrete time. This issue occurs in Lemma 4.8.
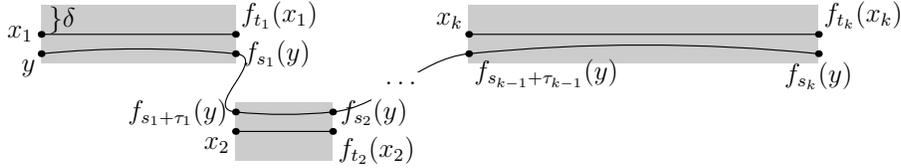
FIGURE 1. The specification property.

2.3. Specification.

We say that $\mathcal{G} \subset X \times \mathbb{R}^+$ has *weak specification at scale* δ if there exists $\tau > 0$ such that for every $\{(x_i, t_i)\}_{i=1}^k \subset \mathcal{G}$ there exists a point y and a sequence of "gluing times" $\tau_1, \ldots, \tau_{k-1} \in \mathbb{R}^+$ with $\tau_i \leq \tau$ such that for $s_j = \sum_{i=1}^j t_i + \sum_{i=1}^{j-1} \tau_i$ and $s_0 = \tau_0 = 0$, we have (see Figure 1)

(2.10) $\qquad d_{t_j}(f_{s_{j-1}+\tau_{j-1}}y, x_j) < \delta \text{ for every } 1 \leq j \leq k.$

We say that $\mathcal{G} \subset X \times \mathbb{R}^+$ has *weak specification at scale δ with maximum gap size τ* if we want to declare a value of τ that plays the role described above. We say that $\mathcal{G} \subset X \times \mathbb{R}^+$ has *weak specification* if it has weak specification at every scale $\delta > 0$.

Remark 2.4. We often write *(W)-specification* as an abbreviation for weak specification. Furthermore, since *(W)-specification* is the only version of the specification property considered in this paper, we henceforth use the term *specification* as shorthand for this property.

Intuitively, (2.10) means that there is some point y whose orbit shadows the orbit of x_1 for time t_1, then after a "gap" of length at most τ, shadows the orbit of x_2 for time t_2, and so on. Note that s_j is the time spent for the orbit y to shadow the orbit segments (x_1, t_1) up to (x_j, t_j). Note that we differ from Franco [Fra77] in allowing τ_j to take any value in $[0, \tau]$, not just one that is close to τ. This difference is analogous in the discrete time case to the difference between (S)-specification where we take the transition times exactly τ, or (W)-specification where the transition times are bounded above by τ. Franco also asks that the shadowing orbit y can be taken to be periodic, and that the gluing time τ_j does not depend on any of the orbit segments (x_i, t_i) with $i > j + 1$.

We can weaken the definition of specification so that it only applies to elements of $\mathcal{G}$ that are sufficiently long. This gives us some useful additional flexibility which we exploit in Lemma 2.10.

Definition 2.5. *We say that $\mathcal{G} \subset X \times \mathbb{R}^+$ has* tail (W)-specification at scale δ *if there exists $T_0 > 0$ so that $\mathcal{G} \cap (X \times [T_0, \infty))$ has weak specification at scale δ; i.e. the specification property holds for the collection of orbit segments*

$$\{(x_i, t_i) \in \mathcal{G} \mid t_i \geq T_0\}.$$

We also sometimes write "$\mathcal{G}$ has (W)-specification at scale δ for $t \geq T_0$" to describe this property.

2.4. **The Bowen property.** The Bowen property was first defined for maps in [Bow75], and extended to flows by Franco [Fra77]. We give a version of this definition for a collection of orbit segments $\mathcal{C}$.

Definition 2.6. *Given $\mathcal{C} \subset X \times \mathbb{R}^+$, a potential ϕ has the* Bowen property *on $\mathcal{C}$ at scale ε if there exists $K > 0$ so that*

$$(2.11) \qquad \sup\{|\Phi_0(x,t) - \Phi_0(y,t)| : (x,t) \in \mathcal{C}, y \in B_t(x,\varepsilon)\} \leq K.$$

We say ϕ has the Bowen property *on $\mathcal{C}$ if there exists $\varepsilon > 0$ so that ϕ has the Bowen property on $\mathcal{C}$ at scale ε.*

In particular, we say that $\phi \colon X \to \mathbb{R}$ has the *Bowen property* if ϕ has the Bowen property on $\mathcal{C} = X \times \mathbb{R}^+$; this agrees with the original definition of Bowen and Franco. This dynamically-defined regularity property is central to Bowen's proof of uniqueness of equilibrium states. For a uniformly hyperbolic system, every Hölder potential ϕ has the Bowen property. This is no longer true in non-uniform hyperbolicity; for example, the geometric potential $-\log f'$ for the Manneville–Pomeau map $f(x) = x + x^{1+\alpha}$ is a natural potential which is Hölder but not Bowen. Asking for the Bowen property to hold on a collection $\mathcal{G}$ rather than globally allows us to deal with non-uniformly hyperbolic systems where one only expects this kind of regularity to hold for those orbit segments which experience a definite amount of hyperbolicity, and where it may not be known whether natural potentials such as the geometric potential are Hölder [CFT15, BCFT16].

We sometimes call K the *distortion constant* for the Bowen property. Note that if ϕ has the Bowen property at scale ε on $\mathcal{G}$ with distortion constant K, then for any $M > 0$, ϕ has the Bowen property at scale ε on $\mathcal{G}^M$ with distortion constant given by $K(M) = K + 2M \operatorname{Var}(\phi, \varepsilon)$.

2.5. **Almost expansivity.** Given $x \in X$ and $\varepsilon > 0$, consider the set

$$(2.12) \qquad \Gamma_\varepsilon(x) := \{y \in X \mid d(f_t x, f_t y) \leq \varepsilon \text{ for all } t \in \mathbb{R}\},$$

which can be thought of as a two-sided Bowen ball of infinite order for the flow. Note that $\Gamma_\varepsilon(x) = \bigcap_{t \in \mathbb{R}} f_t \overline{B}_{2t}(x, \varepsilon)$ is compact for every x, ε.

Expansivity for flows was defined by Bowen and Walters; their definition, details of which can be found in [BW72], implies that for every $s > 0$, there exists $\varepsilon > 0$ such that

$$(2.13) \qquad \Gamma_\varepsilon(x) \subset f_{[-s,s]}(x) := \{f_t(x) \mid t \in [-s, s]\}$$

for every $x \in X$. Since points on a small segment of orbit always stay close for all time, (2.13) essentially says that the set $\Gamma_\varepsilon(x)$ is the smallest possible. Thus, we declare the set of non-expansive points to be those where (2.13) fails. We want to consider measures that witness expansive behaviour, so we declare an almost expansive measure to be one that gives zero measure to the non-expansive points. This is the content of the next definition.

Definition 2.7. *Given $\varepsilon > 0$, the set of* non-expansive points at scale ε *for a flow (X, F) is the set*

$$\mathrm{NE}(\varepsilon) := \{x \in X \mid \Gamma_\varepsilon(x) \not\subset f_{[-s,s]}(x) \text{ for any } s > 0\}.$$

We say that an F-invariant measure μ is almost expansive at scale ε *if $\mu(\mathrm{NE}(\varepsilon)) = 0$.*

A measure μ which is almost expansive at scale ε gives full measure to the set of points x for which there exists $s = s(x)$ for which (2.13) holds. We remark that in contrast to the Bowen-Walters definition, we allow $s(x)$ to be large or even unbounded. Furthermore, our hypotheses do not preclude the existence of fixed points for the flow; for expansive flows, fixed points can only be isolated [BW72, Lemma 1] and can hence be disregarded.

The following definition gives a quantity which captures the largest possible free energy of a non-expansive ergodic measure.

Definition 2.8. *Given a potential ϕ, the* pressure of obstructions to expansivity at scale ε *is*

$$P^\perp_{\exp}(\phi, \varepsilon) = \sup_{\mu \in \mathcal{M}^e_F(X)} \left\{ h_\mu(f_1) + \int \phi \, d\mu \mid \mu(\mathrm{NE}(\varepsilon)) > 0 \right\}$$

$$= \sup_{\mu \in \mathcal{M}^e_F(X)} \left\{ h_\mu(f_1) + \int \phi \, d\mu \mid \mu(\mathrm{NE}(\varepsilon)) = 1 \right\}.$$

We define a scale-free quantity by

$$P^\perp_{\exp}(\phi) = \lim_{\varepsilon \to 0} P^\perp_{\exp}(\phi, \varepsilon).$$

Note that $P^\perp_{\exp}(\phi, \varepsilon)$ is non-increasing as $\varepsilon \to 0$, which is why the limit in the above definition exists. It is essential that the measures in the first supremum are ergodic. If we took this supremum over invariant measures, and a non-expansive measure existed, we would include measures that are a convex combination of a non-expansive measure and a measure with large free energy, so the supremum would equal the topological pressure.

2.6. **Main results for flows.** Theorem A will be deduced from the following more general result, which is proved in §§3–4.

Theorem 2.9. *Let (X, F) be a continuous flow on a compact metric space, and $\phi \colon X \to \mathbb{R}$ a continuous potential function. Suppose there are $\delta, \varepsilon > 0$ with $\varepsilon > 40\delta$ such that $P^\perp_{\exp}(\phi, \varepsilon) < P(\phi)$ and there exists $\mathcal{D} \subset X \times \mathbb{R}^+$ which admits a decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ with the following properties:*

(**I$'$**) *For every $M \in \mathbb{R}^+$, $\mathcal{G}^M$ has tail (W)-specification at scale δ;*

(**II$'$**) *ϕ has the Bowen property at scale ε on $\mathcal{G}$;*

(**III$'$**) *$P(\mathcal{D}^c \cup [\mathcal{P}] \cup [\mathcal{S}], \phi, \delta, \varepsilon) < P(\phi)$.*

Then (X, F, ϕ) has a unique equilibrium state.

These hypotheses weaken those of Theorem A in two main directions.

(1) Theorem A requires that every orbit segment has a decomposition, while Theorem 2.9 permits a set of orbit segments $\mathcal{D}^c \subset X \times \mathbb{R}^+$ to have no decomposition, provided they carry less pressure than the whole system.

(2) The hypotheses of Theorem A require knowledge of the system at all scales: in particular, the specification condition (**I**) in Theorem A requires specification to hold at *every* scale $\delta > 0$. Here, we require a specification property to be verified only at a fixed scale δ, and all other hypotheses to be verified at a larger fixed scale ε. An example where this is useful is the Bonatti–Viana family of diffeomorphisms, where in [CFT15] we are able to verify the discrete-time version of these hypotheses at suitably chosen scales, but establishing them for arbitrarily small scales is difficult, and perhaps impossible.

We make a few more remarks on these hypotheses. By Remark 2.1, we can guarantee (**III$'$**) by checking the bound

$$(2.14) \qquad P(\mathcal{D}^c \cup [\mathcal{P}] \cup [\mathcal{S}], \phi, \delta) + \mathrm{Var}(\phi, \varepsilon) < P(\phi).$$

We do not claim that the relationship $\varepsilon > 40\delta$ is sharp, but we do not expect that it can be significantly improved using these methods. The number 40 does not have any special significance but it is unavoidable that we control the Bowen property and expansivity at a larger scale than where specification is assumed.

If we assume the hypotheses of Theorem A, we can verify the hypotheses of Theorem 2.9 by taking $\mathcal{D} = X \times \mathbb{R}^+$, and any suitably small $\delta, \varepsilon > 0$ with $\varepsilon > 40\delta$. The only hypothesis which is not immediate to verify from the hypotheses of Theorem A is (**I$'$**), and this is verified by the following lemma.

Lemma 2.10. *Suppose that $\mathcal{G} \subset X \times \mathbb{R}^+$ has tail specification at all scales $\delta > 0$, then so does $\mathcal{G}^M$ for every $M > 0$. In particular, (**I**) implies (**I$'$**).*

Proof. Given $M > 0$, let $\delta' = \delta'(M) > 0$ be such that $d(x, y) < \delta'$ implies that $d(f_t x, f_t y) < \delta$ for every $t \in [0, M]$. (Positivity of δ' follows from continuity of the flow and compactness of X.) Now let $T_0 > 0$ be such that $\mathcal{G} \cap (X \times [T_0, \infty))$ has specification at scale δ'. Given any $(x, t) \in \mathcal{G}^M$ with $t \geq T_0 + 2M$, we must have $g(x, t) \geq T_0$. Thus if $(x_1, t_1), \ldots, (x_k, t_k)$ is any collection of orbit segments in $\mathcal{G}^M$ with $t_i \geq T_0 + 2M$, then there are $s_i \in [0, M]$ and $t_i' \in [t_i - 2M, t_i]$ such that $\{(f^{s_i} x_i, t_i') \mid 1 \leq i \leq k\} \subset \mathcal{G}$. Since $t_i' \geq T_0$ we can use the specification property on $\mathcal{G}$ to get an orbit that shadows each $(f^{s_i} x_i, t_i')$ to within δ' (with transition times at most

$\tau = \tau(\delta')$). By our choice of δ', this orbit shadows each (x_i, t_i) to within δ (with transition times at most τ). We conclude that $\mathcal{G}^M$ has tail specification at scale δ.   □

We conclude that Theorem A is a corollary of Theorem 2.9, and we now turn our attention to proving this more general statement.

3. Weak expansivity and generating for adapted partitions

In this section, we develop some general preparatory results on generating properties of partitions in the presence of weak expansivity properties.

3.1. **Almost entropy expansivity.** It is well known that the time-t map of an expansive flow is entropy expansive. We develop an analogue of entropy expansivity for measures called *almost entropy expansivity*, which has the property that if μ is almost expansive for a flow, then it is almost entropy expansive for the time-t map of the flow. This property plays an important role in our proof, as entropy expansivity does for Franco, and is key to obtaining a number of results on generating for partitions.

Let X be a compact metric space and $f \colon X \to X$ a homeomorphism. Let μ be an ergodic f-invariant Borel probability measure. For a set $K \subset X$, let $h(K)$ denote the (upper capacity) entropy of K. That is, $h(K)$ corresponds to $P(\mathcal{C}, 0)$ for $\mathcal{C} = K \times \mathbb{N}$ as defined in §5, which is the natural analogue for maps of (2.5)–(2.7).

Given $x \in X$, consider the set

(3.1) $\Gamma_\varepsilon(x; f, d) := \{y \in X \mid d(f^n x, f^n y) \leq \varepsilon \text{ for all } n \in \mathbb{Z}\}.$

Recall from [Bow72] that the map f is said to be *entropy expansive* if $h(\Gamma_\varepsilon(x; f, d)) = 0$ for every $x \in X$. We will need the following weaker notion.

Definition 3.1. *We say that μ is* almost entropy expansive *at scale ε (in the metric d) with respect to f if $h(\Gamma_\varepsilon(x; f, d)) = 0$ for μ-a.e. $x \in X$.*

Our notation emphasizes the role of the metric d because later in the paper we will need to use this notion relative to various metrics d_t. Bowen proved that if f is entropy expansive at scale ε, then every partition $\mathcal{A}$ with diameter smaller than ε has $h_\mu(f, \mathcal{A}) = h_\mu(f)$. This result was obtained as an immediate consequence of the main part of [Bow72, Theorem 3.5], which shows that for any $\varepsilon > 0$ and any partition with $\text{diam}\,\mathcal{A} \leq \varepsilon$, we have

(3.2) $h_\mu(f) \leq h_\mu(f, \mathcal{A}) + \sup_{x \in X} h(\Gamma_\varepsilon(x; f, d)).$

Clearly, f is entropy expansive if the supremum is 0. Similarly, one sees immediately that μ is almost entropy expansive at scale ε if and only if the *essential* supremum

(3.3) $h^*(\mu, \varepsilon; f, d) = \sup\{\bar{h} \in \mathbb{R} \mid \mu\{x \mid h(\Gamma_\varepsilon(x; f, d)) > \bar{h}\} > 0\}$

vanishes, and we strengthen Bowen's result by showing that one can use the μ-essential supremum in (3.2). The following theorem is proved in §8.

Theorem 3.2. *Let X be a compact metric space and $f\colon X \to X$ a homeo-morphism. Let μ be an ergodic f-invariant Borel probability measure. If $\mathcal{A}$ is any partition with $\operatorname{diam}\mathcal{A} \leq \varepsilon$ in the metric d, then*

$$(3.4) \qquad h_\mu(f) \leq h_\mu(f,\mathcal{A}) + h^*(\mu,\varepsilon;f,d).$$

In particular, if μ is almost entropy expansive at scale ε, then every partition with diameter smaller than ε has $h_\mu(f) = h_\mu(f,\mathcal{A})$.

To apply Theorem 3.2 in the setting of our main results, we first relate almost expansivity for the flow with almost entropy expansivity for the time-t map of the flow.

Proposition 3.3. *If $\mu \in \mathcal{M}_F(X)$ is almost expansive at scale ε, then μ is almost entropy expansive (at scale ε in the metric d_t) with respect to the time-t map f_t.*

Proof. It is immediate from the definitions that $\Gamma_\varepsilon(x) = \Gamma_\varepsilon(x; f_t, d_t)$. Thus, if μ is almost expansive for F, then for μ-a.e. x, the set $\Gamma_\varepsilon(x; f_t, d_t)$ is contained in $f_{[-s,s]}(x)$ for some $s = s(x) \in \mathbb{R}^+$. Fix such an x and let $s = s(x)$. In what follows, we will show that $h(f_{[-s,s]}(x), f_t) = 0$. This shows that $h(\Gamma_\varepsilon(x; f_t, d_t), f_t) = 0$, and since this argument applies to μ-almost every x, it follows that μ is almost entropy expansive for f_t.

So, it just remains to show that the entropy of the finite orbit segment $f_{[-s,s]}(x)$ is 0 with respect to f_t. Let $r > 0$ be sufficiently small that $f_{[-r,r]}(y) \subset B(y,\varepsilon)$ for all $y \in X$ (this is possible by continuity of the flow and compactness of the space). Given $\delta > 0$, fix $N \in \mathbb{N}$ large enough such that $s/N < r$. Let $A = \{f_{kr}(x) \mid k = -N,\ldots,N\}$, and note that $f_t(f_{[(k-1)r,(k+1)r]}(x)) \subset B(f_{t+kr}x,\varepsilon)$ for all $t \in \mathbb{R}$ and all k. Thus, for every n, the set A is (nt,δ)-spanning under F for $f_{[-s,s]}(x)$. It follows that, in the metric d_t, A is (n,δ)-spanning under f_t, which gives $h(\Gamma_\varepsilon(x; f_t, d_t), f_t) = 0$. $\qquad\qquad\square$

The following proposition, which plays a similar role as [CT14, Proposition 2.6], is a consequence of Theorem 3.2 and Proposition 3.3.

Proposition 3.4. *If $\mu \in \mathcal{M}_F(X)$ is almost expansive at scale ε and $\mathcal{A}$ is a finite measurable partition of X with diameter less than ε in the d_t metric for some $t > 0$, then the time-t map f_t satisfies $h_\mu(f_t,\mathcal{A}) = h_\mu(f_t)$.*

3.2. **Adapted partitions and results on generating.** We extend Proposition 3.4 to some useful results on generating using the notion of an adapted partition. This terminology was introduced in [CT14], although the concept goes back to Bowen [Bow74].

Definition 3.5. *Let E_t be a (t,γ)-separated set of maximal cardinality. A partition $\mathcal{A}_t$ of X is* adapted *to E_t if for every $w \in \mathcal{A}_t$ there is $x \in E_t$ such that $B_t(x,\gamma/2) \subset w \subset \overline{B}_t(x,\gamma)$.*

Adapted partitions exist for any (t,γ)-separated set of maximal cardinality since the sets $B_t(x,\gamma/2)$ are disjoint and the sets $\overline{B}_t(x,\gamma)$ cover X.

Lemma 3.6. *If $\mu \in \mathcal{M}_F(X)$ is almost expansive at scale ε, and $\mathcal{A}_t$ is an adapted partition for a $(t, \varepsilon/2)$-separated set E_t of maximal cardinality, then $h_\mu(f_t, \mathcal{A}_t) = h_\mu(f_t)$.*

Proof. For any $w \in \mathcal{A}_t$, there exists x so that $w \subset \overline{B}_t(x, \varepsilon/2)$; this shows that $\operatorname{diam} \mathcal{A}_t \leq \varepsilon$ in the metric d_t. By Proposition 3.4, we have $h_\mu(f_t, \mathcal{A}_t) = h_\mu(f_t)$. □

The proof of the following proposition requires both Lemma 3.6 and a careful use of the almost expansivity property to take a crucial step of replacing a term of the form $\Phi_\varepsilon(x, t)$ with $\Phi_0(x, t)$.

Proposition 3.7. *If $P^\perp_{\exp}(\phi, \varepsilon) < P(\phi)$, then $P(\phi, \gamma) = P(\phi)$ for every $\gamma \in (0, \varepsilon/2]$.*

Proof. Given an ergodic μ, write $P_\mu(\phi) := h_\mu(f_1) + \int \phi \, d\mu$ for convenience. We prove the proposition by showing that $P(\phi, \varepsilon) \geq P_\mu(\phi)$ for every ergodic μ with $P_\mu(\phi) > P^\perp_{\exp}(\phi, \varepsilon)$. We do this by relating both $P(\phi, \varepsilon)$ and $P_\mu(\phi)$ to an adapted partition. In order to carry this out we first introduce a technical lemma that will be used both here and in the proof of Lemma 4.18.

Given a finite partition $\mathcal{A}$ and an F-invariant measure μ, for each $w \in \mathcal{A}$ with $\mu(w) > 0$ we define a function $\Phi = \Phi_\mu \colon \mathcal{A} \to \mathbb{R}$ by

$$(3.5) \qquad \Phi(w) := \frac{1}{\mu(w)} \int_w \Phi_0(x, t) \, d\mu.$$

Given $\alpha \in (0, 1)$, write $H(\alpha) := -\alpha \log \alpha - (1 - \alpha) \log(1 - \alpha)$.

Lemma 3.8. *Suppose $\mu \in \mathcal{M}_F(X)$ is almost expansive at scale ε, and let $\gamma \in (0, \varepsilon/2]$. Let $\mathcal{A}_t$ be an adapted partition for a maximizing (t, γ)-separated set for $\Lambda(X, \gamma, t)$. Let $D \subset X$ be a union of elements of $\mathcal{A}_t$. Then for every $t \in \mathbb{R}^+$ we have*

$$t\left(h_\mu(f_1) + \int \phi \, d\mu\right) \leq \mu(D) \log \sum_{\substack{w \in \mathcal{A}_t \\ w \subset D}} e^{\Phi(w)} + \mu(D^c) \log \sum_{\substack{w \in \mathcal{A}_t \\ w \subset D^c}} e^{\Phi(w)} + H(\mu(D))$$

where $D^c = X \setminus D$, and Φ is as in (3.5).

Proof. Abramov's formula [Abr59] gives $h_\mu(f_t) = t h_\mu(f_1)$ for all $t \in \mathbb{R}^+$, and Lemma 3.6 gives $h_\mu(f_t, \mathcal{A}_t) = h_\mu(f_t)$, so

$$t P_\mu(\phi) = h_\mu(f_t, \mathcal{A}_t) + \int \Phi_0(x, t) \, d\mu \leq \sum_{w \in \mathcal{A}_t} \mu(w)\left(-\log \mu(w) + \Phi(w)\right).$$

Let $\mathcal{W} = \{w \in \mathcal{A}_t \mid w \subset D\}$, and write $\mathcal{W}^c = \mathcal{A}_t \setminus \mathcal{W}$. Breaking up the above sum and normalizing, we have

$$
\begin{aligned}
tP_\mu(\phi) \leq{}& \sum_{w \in \mathcal{W}} \mu(w)\big(\Phi(w) - \log\mu(w)\big) + \sum_{w \in \mathcal{W}^c} \mu(w)\big(\Phi(w) - \log\mu(w)\big) \\
={}& \mu(D) \sum_{w \in \mathcal{W}} \frac{\mu(w)}{\mu(D)}\left(\Phi(w) - \log\frac{\mu(w)}{\mu(D)}\right) \\
& + \mu(D^c) \sum_{w \in \mathcal{W}^c} \frac{\mu(w)}{\mu(D^c)}\left(\Phi(w) - \log\frac{\mu(w)}{\mu(D^c)}\right) \\
& + (-\mu(D)\log\mu(D) - \mu(D^c)\log\mu(D^c)).
\end{aligned}
$$

Recall that for non-negative p_i with $\sum p_i = 1$ and arbitrary $a_i \in \mathbb{R}$ we have $\sum_i p_i(a_i - \log p_i) \leq \log \sum_i e^{a_i}$; the conclusion of Lemma 3.8 follows by applying this to the first sum with $p_w = \mu(w)/\mu(D)$, $a_w = \Phi(w)$, and the second sum with $p_w = \mu(w)/\mu(D^c)$, $a_w = \Phi(w)$. $\qquad\square$

Now we return to the proof of Proposition 3.7. Let $\varepsilon > 0$ be as in the hypothesis, and let μ be ergodic with $P_\mu(\phi) > P_{\exp}^\perp(\phi, \varepsilon)$, so that μ is almost expansive at scale ε. Fix $\alpha > 0$. Given $s \in \mathbb{R}^+$, consider the set

$$
X_s := \{x \in X \mid \Gamma_\varepsilon(x) \subset f_{[-s,s]}(x)\}.
$$

We have $\bigcup_s X_s = X \setminus \mathrm{NE}(\varepsilon)$, so there is s such that $\mu(X_s) > 1 - \alpha$.

Now, we fix $t = s/\alpha$, and for an arbitrary $r > 0$, we write

$$
B_{[-r,t+r]}(x, \varepsilon) := \{y : d(f_\tau x, f_\tau y) \leq \varepsilon \text{ for } \tau \in [-r, t+r]\}
$$

For any $x \in X \setminus \mathrm{NE}(\varepsilon)$, we have

$$
\Gamma_\varepsilon(x) = \bigcap_{r>0} B_{[-r,t+r]}(x, \varepsilon).
$$

In particular, given s as above and $x \in X_s$, we see that $\bigcup_{y \in f_{[-s,s]}(x)} B_t(y, \alpha)$ is an open set which contains $\Gamma_\varepsilon(x)$, so there is $r = r(x)$ so that $B_{[-r,t+r]}(x, \varepsilon) \subset \bigcup_{y \in f_{[-s,s]}(x)} B_t(y, \alpha)$. Now, for $r > 0$, let

$$
Y_r := \left\{ x \in X_s : B_{[-r,t+r]}(x, \varepsilon) \subset \bigcup_{y \in f_{[-s,s]}(x)} B_t(y, \alpha) \right\}.
$$

We have $\bigcup_{r>0} Y_r = X_s$, so we can fix r sufficiently large so that $\mu(Y_r) > 1 - \alpha$. We now pass to the set of points whose orbits spend a large proportion of time in Y_r. Given $n \in \mathbb{N}$, consider the set

$$
Z_n = \{x \in X \mid \mathrm{Leb}\{\tau \in [0, nt] \mid f_\tau(x) \in Y_r\} > (1 - \alpha)nt\},
$$

and note that $\lim_{n\to\infty} \mu(Z_n) = 1$ by the Birkhoff ergodic theorem. Take N large enough that $\mu(Z_n) > 1 - \alpha$ for all $n \geq N$. The following lemma gives us a regularity property for the potential ϕ for points in Z_n.

Lemma 3.9. *Given $x \in Z_n$ and $y \in B_{nt}(x, \varepsilon)$, we have*

$$(3.6) \qquad |\Phi_0(x, nt) - \Phi_0(y, nt)| \leq (8\alpha nt + 4r)\|\phi\| + nt \operatorname{Var}(\phi, \alpha).$$

Proof. Let $\mathbf{T} = \{\tau \in [0, nt] \mid f_\tau(x) \in Y_r\}$ and choose $\xi > 0$ such that $\operatorname{Leb}([0, nt] \setminus \mathbf{T}) + \xi n < \alpha nt$ (here we use that $x \in Z_n$). Define $\tau_1, \dots, \tau_k \in [r, nt - r]$ iteratively as follows: let $\tau_1' = \inf(\mathbf{T} \cap [r, nt])$, and then given τ_i', choose any $\tau_i \in \mathbf{T} \cap [\tau_i', \tau_i' + \xi]$, and put $\tau_{i+1}' = \inf(\mathbf{T} \cap [\tau_i + t + 2s, nt])$.
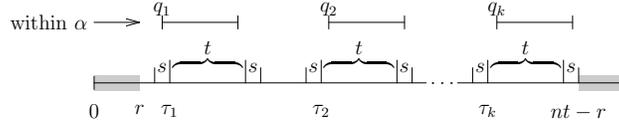

FIGURE 2. Proving Lemma 3.9.

It follows from the definition of τ_i and properties of $\mathbf{T}$ that

- $\tau_{i+1} > \tau_i + t + 2s$ for every i;
- $\sum_{i=1}^{k-1}(\tau_{i+1} - (\tau_i + t + 2s)) \leq \alpha nt$; and
- $f_{\tau_i}(x) \in Y_r$ for every i.

Since $f_{\tau_i} y \in B_{[-r, t+r]}(f_{\tau_i} x, \varepsilon)$, the third property gives $f_{\tau_i} y \in B_t(f_{q_i} x, \alpha)$ for some $q_i \in [\tau_i - s, \tau_i + s]$; see Figure 2. Thus

$$|\Phi_0(f_{\tau_i} y, t) - \Phi_0(f_{q_i} x, t)| \leq 2s\|\phi\| + t \operatorname{Var}(\phi, \alpha).$$

The first two properties give

$$\left| \Phi_0(y, nt) - \sum_{i=1}^{k} \Phi_0(f_{\tau_i} y, t) \right| \leq (\alpha nt + 2r + 2sn)\|\phi\|,$$

$$\left| \Phi_0(x, nt) - \sum_{i=1}^{k} \Phi_0(f_{q_i} x, t) \right| \leq (\alpha nt + 2r + 2sn)\|\phi\|,$$

and putting it all together we have

$$|\Phi_0(x, nt) - \Phi_0(y, nt)| \leq 2(\alpha nt + 2r + 2sn)\|\phi\| + 2sn\|\phi\| + nt \operatorname{Var}(\phi, \alpha)$$
$$\leq (8\alpha nt + 4r)\|\phi\| + nt \operatorname{Var}(\phi, \alpha),$$

which proves the lemma. $\qquad \square$

To finish the proof of Proposition 3.7, for $n \geq N$, let E_n be any maximizing $(nt, \varepsilon/2)$-separated set for $\Lambda(X, \varepsilon, nt)$, and let $\mathcal{A}_n$ be an adapted partition for E_n. Let $D_n = \bigcup\{w \mid w \in \mathcal{A}_n, w \cap Z_n \neq \emptyset\}$ and note that $\mu(D_n) > 1 - \alpha$. With $\alpha < 1/2$, Lemma 3.8 gives

$$nt P_\mu(\phi) \leq \log \sum_{\substack{w \in \mathcal{A}_t \\ w \subset D_n}} e^{\Phi(w)} + \alpha \log \sum_{\substack{w \in \mathcal{A}_t \\ w \subset D_n^c}} e^{\Phi(w)} + H(\alpha).$$

We need to get estimates on the sums. Note that there is a 1-1 correspondence between elements of E_n and elements of $\mathcal{A}_n$, and that given $x \in E_n$ and $w \in \mathcal{A}_n$ with $x \in w$, we have $\Phi(w) \leq \Phi_\varepsilon(x, nt)$. Thus

$$(3.7) \quad ntP_\mu(\phi) \leq \log \sum_{x \in E_n \cap D_n} e^{\Phi_\varepsilon(x,nt)} + \alpha \log \sum_{x \in E_n \cap D_n^c} e^{\Phi_\varepsilon(x,nt)} + H(\alpha).$$

To control the first sum, we use Lemma 3.9 and get

$$\log \sum_{x \in E_n \cap D_n} e^{\Phi_\varepsilon(x,nt)} \leq (8\alpha\|\phi\| + \mathrm{Var}(\phi,\alpha))nt + 4T\|\phi\| + \log \Lambda(X, \varepsilon, nt),$$

while for the second we can put $Q > P(\phi, \varepsilon, 2\varepsilon)$ and obtain a constant C such that for every n we have

$$\sum_{x \in E_n \cap D_n^c} e^{\Phi_\varepsilon(x,nt)} \leq \Lambda(X, \varepsilon, 2\varepsilon, nt) \leq Ce^{Qnt}.$$

Dividing both sides of (3.7) by nt and sending $n \to \infty$, these bounds give

$$P_\mu(\phi) \leq 8\alpha\|\phi\| + \mathrm{Var}(\phi,\alpha) + P(\phi, \varepsilon) + \alpha Q.$$

Since $\alpha > 0$ was arbitrary, we conclude that $P(\phi, \varepsilon) \geq P_\mu(\phi)$. Taking a supremum over all ergodic μ with $P_\mu(\phi) > P_{\exp}^\perp(\phi, \varepsilon)$ gives $P(\phi, \varepsilon) \geq P(\phi)$, which completes the proof of Proposition 3.7. □

3.3. **Approximating invariant sets with adapted partitions.** In addition to the generating properties from the previous section, adapted partitions are useful for approximating invariant sets with Bowen balls; this will be used in the proof of uniqueness in §§4.7–4.8. The following proposition is inspired by an approximation lemma of Bowen [Bow74, Lemma 2], which plays a key role in Franco's proof.

Proposition 3.10. *Let F be a continuous flow on a compact metric space X, and suppose $\mu \in \mathcal{M}_F(X)$ is almost expansive at scale ε. Let $\gamma \in (0, \varepsilon/2]$, and for each $t > 0$, let $\mathcal{A}_t$ be an adapted partition for a (t, γ)-separated set of maximal cardinality. Let $Q \subset X$ be a measurable F-invariant set. Then for every $\alpha > 0$, there exists t_0 so that if $t \geq t_0$, we can find $U \subset \mathcal{A}_t$ such that $\mu(U \triangle Q) < \alpha$.*

Proof. If $\mu(Q) = 0$ we take $U = \emptyset$, so from now on we assume $\mu(Q) > 0$. For a set $w \subset X$ and $s \in \mathbb{R}^+$, let

$$(3.8) \qquad \mathrm{diam}_{[-s,s]} w = \sup_{x_1, x_2 \in w} \inf_{t_1, t_2 \in [-s,s]} d(f_{t_1} x_1, f_{t_2} x_2),$$

and for a partition $\mathcal{A}$, let $\mathrm{diam}_{[-s,s]} \mathcal{A} = \max\{\mathrm{diam}_{[-s,s]} w : w \in \mathcal{A}\}$. Bowen proved that the conclusion of Proposition 3.10 holds if the partitions $\mathcal{A}_t$ satisfy $\mathrm{diam}_{[-s,s]} \mathcal{A}_t \to 0$ for some $s \in \mathbb{R}^+$. In our setting we do not have uniform convergence and so we first need to pass to a set of large measure. For $s \in \mathbb{R}^+$, let $X_s = \{x \mid \Gamma_\varepsilon(x) \subset f_{[-s,s]}(x)\}$. Fix $\beta > 0$. Since μ is almost expansive at scale ε, the arguments which precede Lemma 3.9

show that there exists s such that $\mu(X_s) > 1 - \beta$ and for every $x \in X_s$, $\text{diam}_{[-s,s]} B_{[-t,t]}(x,\varepsilon) \to 0$ as $t \to \infty$.

Let $\mathcal{A}'_t := f_{t/2}\mathcal{A}_t$, and write $w_t(x)$ for the element of the partition $\mathcal{A}'_t$ which contains x, and observe that by the construction of $\mathcal{A}'_t$, for every $x \in X$ there exists a point x' so that $w_t(x) \subset B_{[-t/2,t/2]}(x',\gamma)$. It follows that $w_t(x) \subset B_{[-t/2,t/2]}(x,2\gamma)$, and thus $\text{diam}_{[-s,s]} w_t(x) \to 0$ for μ-a.e $x \in X_s$, and by Egorov's theorem there is $X'_s \subset X_s$ with $\mu(X_s \setminus X'_s) < \beta$ such that the convergence is uniform on X'_s.

By restricting our attention to $Q' := X'_s \cap Q$ we can follow Bowen's argument from [Bow74, Lemma 2].

Let $K_1 \subset Q'$ and $K_2 \subset X \setminus Q$ be compact sets with $\mu(Q' \setminus K_1) < \beta$ and $\mu(X \setminus (Q \cup K_2)) < \beta$. For $i = 1, 2$ let $K_i^s = \{f_t(x) \mid x \in K_i, t \in [-s,s]\}$ and note that K_1^s, K_2^s are compact; moreover, they are disjoint, since $K_1^s \subset Q$ and $K_2^s \subset X \setminus Q$ by F-invariance of Q. Thus K_1^s and K_2^s are uniformly separated by some distance $\gamma > 0$; that is,

$$(3.9) \qquad d(f_{t_1}(x_1), f_{t_2}(x_2)) \geq \gamma \text{ for all } x_i \in K_i \text{ and } t_i \in [-s, s].$$

This gives $d_{[-s,s]}(K_1, K_2) > \gamma$. By the uniform convergence obtained above on Q', there is $t_0 \in \mathbb{R}^+$ such that $\text{diam}_{[-s,s]} w_t(x) < \gamma$ for every $t \geq t_0$ and $x \in Q'$. It follows that if $t \geq t_0$, then every $w \in \mathcal{A}'_t$ with $w \cap K_1 \neq \emptyset$ has $w \cap K_2 = \emptyset$. Let $U' = \bigcup\{w \in \mathcal{A}'_t \mid w \cap K_1 \neq \emptyset\}$. Then $K_1 \subset U'$ and $K_2 \cap U' = \emptyset$. Hence

$$\mu(U' \triangle Q) = \mu(U' \setminus Q) + \mu(Q \setminus U') \leq \mu(X \setminus (Q \cup K_2)) + \mu(Q \setminus K_1)$$
$$\leq \beta + \mu(Q \setminus Q') + \mu(Q' \setminus K_1) \leq \beta + 2\beta + \beta = 4\beta,$$

and we were free to choose $\beta < \alpha/4$. To complete the proof of Proposition 3.10, let $U \subset \mathcal{A}_t$ be defined to be $U = f_{-t/2}U'$. By F-invariance of the measure μ and the set Q, we have

$$\mu(U \triangle Q) = \mu(f_{-t/2}(U \triangle Q)) = \mu(U' \triangle Q) < \alpha. \qquad \square$$

4. PROOF OF THEOREM 2.9

We adapt the strategy of the proof in [CT14], which in turn was inspired by the proof of Bowen [Bow75]. We use δ to represent the scale at which specification occurs, and ε for the scale at which expansivity and the Bowen property occur. We use γ, θ, etc. to represent other scales, which usually lie in between δ and ε. Many of the intermediate lemmas in this section hold without any specification or expansivity assumption, so we take care to state exactly which assumptions are required.

4.1. Lower bounds on X.

Lemma 4.1. *For every $\gamma > 0$ and $t_1, \ldots, t_k > 0$ we have*

$$(4.1) \qquad \Lambda(X, 2\gamma, t_1 + \cdots + t_k) \leq \prod_{j=1}^{k} \Lambda(X, \gamma, \gamma, t_j).$$

Proof. Write $T = t_1 + \cdots + t_k$. We will also need to consider the partial sums $r_j = t_1 + \cdots + t_{j-1}$ for each $2 \leq j \leq k$, where we put $r_1 = 0$. Let E be a maximizing $(T, 2\gamma)$-separated set, and similarly for each j let E_j be a maximizing (t_j, γ)-separated set. Each E_j is (t_j, γ)-spanning, and so we can define a map $\pi \colon E \to E_1 \times \cdots \times E_k$ by $\pi(x) = (x_1, \ldots, x_k)$, where $x_j \in E_j$ is such that $f_{r_j}(x) \in \overline{B}_{t_j}(x_j, \gamma)$.

We claim that π is injective. To see this, observe that for all distinct $x, y \in E$ there exists j such that $d_{t_j}(f_{r_j}x, f_{r_j}y) > 2\gamma$, and thus

$$d_{t_j}(x_j, y_j) \geq d_{t_j}(f_{r_j}x, f_{r_j}y) - d_{t_j}(f_{r_j}x, x_j) - d_{t_j}(f_{r_j}y, y_j)$$
$$> 2\gamma - \gamma - \gamma = 0,$$

so $x_j \neq y_j$ and thus $\pi(x) \neq \pi(y)$. It follows that

$$(4.2) \qquad \Lambda(X, 2\gamma, T) = \sum_{x \in E} e^{\Phi_0(x, T)} = \sum_{(x_1, \ldots, x_k) \in \pi(E)} e^{\Phi_0(\pi^{-1}(x_1, \ldots, x_k), T)}.$$

Given $x \in E$ with $\pi(x) = (x_1, \ldots, x_k)$, we observe that $f_{r_j}(x) \in \overline{B}_{t_j}(x_j, \gamma)$. Thus $\Phi_0(f_{r_j}x, t_j) \leq \Phi_\gamma(x_j, t_j)$, and so

$$\Phi_0(x, T) \leq \sum_{j=1}^{k} \Phi_0(f_{r_j}x, t_j) \leq \sum_{j=1}^{k} \Phi_\gamma(x_j, t_j).$$

Together with (4.2), this gives

$$\Lambda(X, 2\gamma, T) \leq \sum_{(x_1, \ldots, x_k) \in \pi(E)} e^{\sum_{j=1}^{k} \Phi_\gamma(x_j, t_j)}$$

$$\leq \sum_{x_1 \in E_1} \cdots \sum_{x_k \in E_k} \prod_{j=1}^{k} e^{\Phi_\gamma(x_j, t_j)} = \prod_{j=1}^{k} \Lambda(X, \gamma, \gamma, t_j),$$

completing the proof of Lemma 4.1. $\qquad\qquad\qquad\qquad\qquad\square$

Lemma 4.2. *Let ε satisfy $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$. For every $t \in \mathbb{R}^+$ and $0 < \gamma \leq \varepsilon/4$, we have the inequality $\Lambda(X, \gamma, \gamma, t) \geq e^{tP(\phi)}$.*

Proof. For every $k \in \mathbb{N}$, Lemma 4.1 gives $\Lambda(X, 2\gamma, kt) \leq (\Lambda(X, \gamma, \gamma, t))^k$. It follows that $\frac{1}{kt} \log \Lambda(X, 2\gamma, kt) \leq \frac{1}{t} \log \Lambda(X, \gamma, \gamma, t)$. The left hand side converges to $P(\phi, 2\gamma)$ as $k \to \infty$. The result follows by Proposition 3.7. $\quad\square$

This lemma demonstrates why we need to consider 'two-scale' pressure. If ϕ had the Bowen property globally (at scale γ), then by Remark 2.1, we would have

$$\Lambda(X, \gamma, t) \geq e^{-K} \Lambda(X, \gamma, \gamma, t) \geq e^{-K} e^{tP(\phi)},$$

so we can reduce to the standard partition sum. However, if φ is not Bowen globally, then we only obtain

$$\Lambda(X, \gamma, t) \geq e^{-t \operatorname{Var}(\phi, \gamma)} \Lambda(X, \gamma, \gamma, t) \geq e^{-t \operatorname{Var}(\phi, \gamma)} e^{tP(\phi)}.$$

which is why we must work with the 'two-scale' partition sums. In our setting, to obtain lower bounds on the standard partition sums $\Lambda(X, \gamma, t)$,

we require a more in-depth analysis, and all the hypotheses of our theorem. This is carried out in Proposition 4.10.

4.2. **Upper bounds on $\mathcal{G}$.** Now suppose that $\mathcal{G} \subset X \times \mathbb{R}^+$ has specification at scale $\delta > 0$ for $t \geq T_0$.

Proposition 4.3. *Suppose that $\mathcal{G}$ has specification at scale $\delta > 0$ for $t \geq T_0$ with maximum gap size τ, and $\zeta > \delta$ is such that ϕ has the Bowen property on $\mathcal{G}$ at scale ζ. Then, for every $\gamma > 2\zeta$, there is a constant $C_1 > 0$ so that for every $k \in \mathbb{N}$ and $t_1, \ldots, t_k \geq T_0$, writing $T := \sum_{i=1}^{k} t_i + (k-1)\tau$, and $\theta := \zeta - \delta$, we have*

$$(4.3) \qquad \prod_{j=1}^{k} \Lambda(\mathcal{G}, \gamma, t_j) \leq C_1^k \Lambda(X, \theta, T).$$

Proof. Let $\lambda_j < \Lambda(\mathcal{G}, \gamma, t_j)$ be arbitrary, and let $E_1, \ldots, E_k \subset \mathcal{G}$ be (t_j, γ)-separated sets such that $\sum_{x \in E_j} e^{\Phi_0(x,t)} > \lambda_j$.

By specification, for every $\mathbf{x} = (x_1, \ldots, x_k) \in \prod_{j=1}^{k} E_j$ there are $y = y(\mathbf{x}) \in X$ and $\boldsymbol{\tau} = \boldsymbol{\tau}(\mathbf{x}) = (\tau_1, \ldots, \tau_{k-1}) \in [0, \tau]^{k-1}$ such that $y = y(\mathbf{x}) \in B_{t_1}(x_1, \delta)$, $f_{t_1 + \tau_1}(y) \in B_{t_2}(x_2, \delta)$, and so on. Moreover, τ_j depends only on $x_1, \ldots, x_{j+1}$.

Let $E \subset X$ be maximizing (θ, T)-separated, and hence (θ, T)-spanning. Let $p \colon X \to E$ be the map that takes $x \in X$ to the point in E that is closest to it in the d_T-metric. Let $\pi = p \circ y \colon \prod E_j \to E$, as shown in the following commutative diagram.

$$(4.4) \qquad \begin{array}{ccc} E_1 \times \cdots \times E_k & & \\ {\scriptstyle (y,\boldsymbol{\tau})} \downarrow & \searrow {\scriptstyle \pi} & \\ X \times [0,\tau]^{k-1} & \xrightarrow{\ p\ } & E \end{array}$$

Writing $s_j(\mathbf{x}) = \sum_{i=1}^{j} t_i + \sum_{i=1}^{j-1} \tau_i$, and $s_0(\mathbf{x}) = \tau_0(\mathbf{x}) = 0$, we observe that for every $\mathbf{x} \in \prod E_j$ and every $1 \leq j \leq k$, we have

$$(4.5) \qquad d_{t_j}(x_j, f_{s_{j-1} + \tau_{j-1}}(\pi\mathbf{x})) \leq \delta + \theta = \zeta.$$

We will use the map π to compare the two sides of (4.3). For this we will need the following two lemmas controlling the multiplicity of π, and the difference between the weights of the orbit segments.

Lemma 4.4. *There is a constant $C_2 \in \mathbb{R}$ such that $\#\pi^{-1}(z) \leq C_2^k$ for every $z \in E$.*

Lemma 4.5. *There is a constant $C_3 \in \mathbb{R}$ such that every $\mathbf{x} \in \prod E_j$ has $\Phi_0(\pi\mathbf{x}, T) \geq -kC_3 + \sum_j \Phi_0(x_j, t_j)$.*

We remark that in addition to the dynamics (X, F), the constant C_2 depends on τ and γ, while C_3 depends on τ, $\|\phi\|$, and the constant K from

the Bowen property. Before proving Lemmas 4.4 and 4.5 we show how they complete the proof of Proposition 4.3. We see that

$$
\Lambda(X, \theta, 0, T) = \sum_{z \in E} e^{\Phi_0(z,T)} \geq C_2^{-k} \sum_{\mathbf{x} \in \prod_j E_j} e^{\Phi_0(\pi \mathbf{x}, T)}
$$

(4.6)
$$
\geq C_2^{-k} e^{-kC_3} \sum_{\mathbf{x} \in \prod_j E_j} \prod_j e^{\Phi_0(x_j,t_j)} \geq (C_2 e^{C_3})^{-k} \prod_j \lambda_j,
$$

where the first inequality uses Lemma 4.4 and the second uses Lemma 4.5. Since λ_j can be taken arbitrarily close to $\Lambda(\mathcal{G}, \gamma, t_j)$, this establishes (4.3), and so it only remains to prove the lemmas.

Proof of Lemma 4.4. Let $\gamma' = \gamma - 2\zeta > 0$, and let $m \in \mathbb{N}$ be large enough so that writing $\zeta' := \tau/m$, we have $d(x, f_s x) < \gamma'$ for every $x \in X$ and $s \in (-\zeta', \zeta')$. We partition the interval $[0, k\tau]$ into km sub-intervals $I_1, \ldots, I_{km}$ of length ζ', denoting this partition P. Given $\mathbf{x} \in \prod E_j$, take the sequence $n_1, \ldots, n_k$ so that

$$
\tau_1(\mathbf{x}) + \cdots + \tau_i(\mathbf{x}) \in I_{n_i} \text{ for every } 1 \leq i \leq k - 1.
$$

Now let $\ell_1 = n_1$ and $\ell_{i+1} = n_{i+1} - n_i$ for $1 \leq i \leq k - 2$, and let $\ell(\mathbf{x}) := (\ell_1, \ldots, \ell_{k-1})$. Since $\tau_{i+1}(\mathbf{x}) \in [0, \tau]$, we have $n_i \leq n_{i+1} \leq n_i + m$ for each i, and thus $\ell(\mathbf{x}) \in \{0, \ldots, m-1\}^{k-1}$.

Now given $\bar{\ell} \in \{0, \ldots, m-1\}^{k-1}$, let $E^{\bar{\ell}} \subset \prod E_j$ be the set of all $\mathbf{x}$ such that $\ell(\mathbf{x}) = \bar{\ell}$. Note that if $\mathbf{x}, \mathbf{x}' \in E^{\bar{\ell}}$ and $i \in \{1, \ldots, k-1\}$, then by construction, $\tau_1(\mathbf{x}) + \cdots + \tau_i(\mathbf{x})$ and $\tau_1'(\mathbf{x}) + \cdots + \tau_i'(\mathbf{x})$ belong to the same element of the partition P.

We show that π is 1-1 on each $E^{\bar{\ell}}$. Fix $\bar{\ell}$ and let $\mathbf{x}, \mathbf{x}' \in E^{\bar{\ell}}$ be distinct. Let j be the smallest index such that $x_j \neq x_j'$. Write $\tau_i = \tau_i(\mathbf{x})$ and $\tau_i' = \tau_i(\mathbf{x}')$. Let $r = \sum_{i=1}^j (t_i + \tau_i)$ and $r' = \sum_{i=1}^j (t_i + \tau_i')$. Since $\sum_{i=1}^j \tau_i$ and $\sum_{i=1}^j \tau_i$ belong to the same element of P, then $|r - r'| = |\sum_{i=1}^j \tau_i - \sum_{i=1}^j \tau_i'| < \zeta'$.

Because $x_j \neq x_j' \in E_j$ and E_j is (t_j, γ)-separated, we have $d_{t_j}(x_j, x_j') > \gamma$. Now we have

$$
d_T(\pi \mathbf{x}, \pi \mathbf{x}') \geq d_{t_j}(f_r \pi \mathbf{x}, f_r \pi \mathbf{x}') \geq d_{t_j}(f_r \pi \mathbf{x}, f_{r'} \pi \mathbf{x}') - \gamma',
$$

where the γ' term comes from the fact that $d_{t_j}(f_r \pi \mathbf{x}', f_{r'} \pi \mathbf{x}') \leq \gamma'$ by our choice of ζ'. For the first term, observe that

$$
d_{t_j}(f_r \pi \mathbf{x}, f_{r'} \pi \mathbf{x}') \geq d_{t_j}(x_j, x_j') - d_{t_j}(x_j, f_r \pi \mathbf{x}) - d_{t_j}(f_{r'} \pi \mathbf{x}', x_j').
$$

Using (4.5) gives

$$
d_T(\pi \mathbf{x}, \pi \mathbf{x}') > \gamma - 2\zeta - \gamma' = 0,
$$

and so $\pi \mathbf{x} \neq \pi \mathbf{x}'$, and thus π is 1-1 on each $E^{\bar{\ell}}$. Since there are m^{k-1} choices of $\bar{\ell}$, this completes the proof of Lemma 4.4. $\square$

Proof of Lemma 4.5. Let $\mathbf{x} \in \prod_j E_j$. Noting that $T - \sum_{j=1}^k t_j = (k-1)\tau$, we have

$$\Phi_0(\pi\mathbf{x}, T) = \int_0^T \phi(f_s(\pi\mathbf{x}))\,ds \geq -(k-1)\tau\|\phi\| + \sum_{j=1}^k \Phi_0(f_{s_{j-1}+\tau_{j-1}}(\pi\mathbf{x}), t_j).$$

Moreover, since from (4.5) we have $f_{s_{j-1}+\tau_{j-1}}(\pi\mathbf{x}) \in \overline{B}_\zeta(x_j, t_j)$ for every j, we can use the Bowen property at scale ζ to get

$$\Phi_0(f_{s_{j-1}+\tau_{j-1}}(\pi\mathbf{x}), t_j) \geq \Phi_0(x_j, t_j) - K.$$

We conclude that

$$\Phi_0(\pi\mathbf{x}, T) \geq -k(\tau\|\phi\| + K) + \sum_{j=1}^k \Phi_0(x_j, t_j),$$

which completes the proof of Lemma 4.5. $\qquad\square$

As explained above, this completes the proof of Proposition 4.3, by the computation in (4.6). $\qquad\square$

The following corollary extends the statement of Proposition 4.3 to the more general partition sums with $\eta > 0$.

Corollary 4.6. *Let $\mathcal{G}, \delta, \zeta, \tau, \gamma, \phi, t_j, T, \theta, C_1$ be as in Proposition 4.3, and let $\eta_1, \eta_2 \geq 0$. Suppose further that the Bowen property for ϕ holds at scale η_1. Then*

$$(4.7) \qquad \prod_{j=1}^k \Lambda(\mathcal{G}, \gamma, \eta_1, t_j) \leq (e^K C_1)^k \Lambda(X, \theta, \eta_2, T),$$

where K is the distortion constant from the Bowen property.

Proof. Observe that replacing the term $\Lambda(X, \theta, 0, T)$ with $\Lambda(X, \theta, \eta_2, T)$ increases the right-hand side of (4.3), so it suffices to prove (4.7) when $\eta_2 = 0$. If ϕ has the Bowen property on $\mathcal{G}$ as scale η_1, then it is easy to show that

$$\Lambda(\mathcal{G}, \gamma, \eta_1, t) \leq e^K \Lambda(\mathcal{G}, \gamma, 0, t)$$

for every $\gamma, t > 0$. Thus, (4.3) yields the required inequality. $\qquad\square$

The key consequence of Proposition 4.3 is the following upper bound on partition sums over $\mathcal{G}$.

Proposition 4.7. *Suppose that $\mathcal{G} \subset X \times \mathbb{R}^+$ has tail specification at scale $\delta > 0$. Suppose that $\gamma > 2\delta$ and $\eta \geq 0$. Suppose that ϕ has the Bowen property on $\mathcal{G}$ at scale $\max\{\gamma/2, \eta\}$. Then there is C_4 such that for every $t \geq 0$ we have*

$$(4.8) \qquad\qquad \Lambda(\mathcal{G}, \gamma, \eta, t) \leq C_4 e^{tP(\phi)}.$$

Proof. Let $\zeta > 0$ satisfy $2\delta < 2\zeta < \gamma$, and thus by assumption, $\mathcal{G}$ has the Bowen property at scale ζ. Let $\theta := \zeta - \delta$. There is a $T_0 \in \mathbb{R}^+$ so that by Corollary 4.6, we have

$$(\Lambda(\mathcal{G}, \gamma, \eta, t))^k \leq (e^K C_1)^k \Lambda(X, \theta, k(t + \tau))$$

for every $k \in \mathbb{N}$ and $t \geq T_0$. Taking logarithms and dividing by k yields

$$\log \Lambda(\mathcal{G}, \gamma, \eta, t) \leq \log(e^K C_1) + (t + \tau) \frac{1}{k(t + \tau)} \log \Lambda(X, \theta, k(t + \tau)).$$

Sending $k \to \infty$ we get

$$\log \Lambda(\mathcal{G}, \gamma, \eta, t) \leq K + \log C_1 + (\tau + t) P(\phi, \theta);$$

since $P(\phi, \theta) \leq P(\phi)$ this gives

$$\Lambda(\mathcal{G}, \gamma, \eta, t) e^{-tP(\phi)} \leq C_1 e^{K + \tau P(\phi)}$$

for all $t \geq T_0$, which proves Proposition 4.7 with

$$C_4 = \max \left(C_1 e^{K + \tau P(\phi)}, \sup_{t \in [0, T_0]} \Lambda(\mathcal{G}, \gamma, \eta, t) e^{-tP(\phi)} \right). \qquad \square$$

4.3. Lower bounds on $\mathcal{G}$.

Lemma 4.8. *Fix a scale $\gamma > 2\delta > 0$, and let $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ be a decomposition for $\mathcal{D} \subseteq X \times \mathbb{R}^+$ such that*

(1) $\mathcal{G}$ has tail specification at scale δ;
(2) ϕ has the Bowen property on $\mathcal{G}$ at scale 3γ; and
(3) $P(\mathcal{D}^c, \phi, 2\gamma, 2\gamma) < P(\phi)$ and $P([\mathcal{P}] \cup [\mathcal{S}], \gamma, 3\gamma) < P(\phi)$,

Then for every $\alpha_1, \alpha_2 > 0$, there exists $M \in \mathbb{R}^+$ and $T_1 \in \mathbb{R}^+$ such that the following is true:

- *for any $t \geq T_1$ and $\mathcal{C} \subset X \times \mathbb{R}^+$ such that $\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) \geq \alpha_1 e^{tP(\phi)}$, we have $\Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t) \geq (1 - \alpha_2) \Lambda(\mathcal{C}, 2\gamma, 2\gamma, t)$.*

Proof. Fix $\beta_1 > 0$ such that

$$P(\mathcal{D}^c, \phi, 2\gamma, 2\gamma) < P(\phi) - 2\beta_1 \text{ and } P([\mathcal{P}] \cup [\mathcal{S}], \gamma, 3\gamma) < P(\phi) - 2\beta_1,$$

and so there is $C_5 \in \mathbb{R}^+$ such that

$$\begin{aligned}
(4.9) \qquad P([\mathcal{P}] \cup [\mathcal{S}], \gamma, 3\gamma) &\leq C_5 e^{t(P(\phi) - \beta_1)}, \\
\Lambda(\mathcal{D}^c, 2\gamma, 2\gamma, t) &\leq C_5 e^{t(P(\phi) - \beta_1)}
\end{aligned}$$

for all $t > 0$. We consider t sufficiently large so that

$$(4.10) \qquad C_5 e^{t(P(\phi) - \beta_1)} \leq \tfrac{1}{2} \alpha_1 \alpha_2 e^{tP(\phi)}.$$

For an arbitrary $\beta_2 > 0$, let $E_t \subset \mathcal{C}_t$ be a $(t, 2\gamma)$-separated set so that

$$(4.11) \qquad \sum_{x \in E_t} e^{\Phi_{2\gamma}(x, t)} > \Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) - \beta_2.$$

Writing $i \vee j$ for $\max\{i, j\}$, we split the sum in (4.11) into three terms

$$(4.12) \qquad \sum_{\substack{x \in E_t \\ (x,t) \in \mathcal{D}^c}} e^{\Phi_{2\gamma}(x,t)} + \sum_{\substack{x \in E_t \\ (p \vee s)(x,t) \leq M}} e^{\Phi_{2\gamma}(x,t)} + \sum_{\substack{x \in E_t \\ (p \vee s)(x,t) > M}} e^{\Phi_{2\gamma}(x,t)}$$

The first sum is at most $\Lambda(\mathcal{D}^c, 2\gamma, 2\gamma, t)$, and since $\{x \in E_t \mid (p \vee s)(x,t) \leq M\}$ is a $(t, 2\gamma)$-separated subset of $(\mathcal{C} \cap \mathcal{G}^M)_t$, the second sum is at most $\Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t)$.
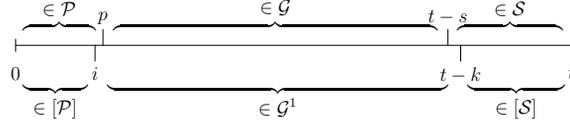
FIGURE 3. Decomposing an orbit segment.

We work on controlling the final sum in (4.12). Note that given $x \in E_t$ with $\lfloor p(x,t) \rfloor = i$ and $\lfloor s(x,t) \rfloor = k$, we have

$$x \in [\mathcal{P}]_i, \quad f_i x \in \mathcal{G}^1_{t-(i+k)}, \quad f_{t-k} x \in [\mathcal{S}]_k,$$

where $[\mathcal{P}], [\mathcal{S}]$ are as defined in (2.9); see Figure 3. Given $i, k \in \{1, \ldots, \lceil t \rceil\}$, let

$$C(i,k) = \{x \in E_t \mid (x,t) \in \mathcal{D} \text{ and } \lfloor p(x,t) \rfloor = i, \lfloor s(x,t) \rfloor = k\}.$$

Now for each $i \in \{1, \ldots, \lceil t \rceil\}$, let $E_i^P \subset [\mathcal{P}]_i$ be a (i, γ)-separated set of maximum cardinality; choose $E_j^G \subset \mathcal{G}_j^1$ and $E_k^S \subset [\mathcal{S}]_k$ similarly. As in Lemma 4.1, there is an injection $\pi \colon C(i,k) \to E_i^P \times E_{t-(i+k)}^G \times E_k^S$ given by $\pi(x) = (x_1, x_2, x_3)$, where

- $x_1 \in E_i^P$ is such that $x \in \overline{B}_i(x_1, \gamma)$;
- $x_2 \in E_{t-(i+k)}^G$ is such that $f_i x \in \overline{B}_{t-(i+k)}(x_2, \gamma)$;
- $x_3 \in E_k^S$ is such that $f_{t-k} x \in \overline{B}_k(x_3, \gamma)$.

Note that for $x \in C(i,k)$ we have

$$\Phi_{2\gamma}(x,t) \leq \Phi_{3\gamma}(\pi_1 x, i) + \Phi_{3\gamma}(\pi_2 x, t-(i+k)) + \Phi_{3\gamma}(\pi_3 x, k),$$

so we have the following bound on the final sum in (4.12):

$$\sum_{\substack{x \in E_t \\ (p \vee s)(x,t) > M}} e^{\Phi_{2\gamma}(x,t)} = \sum_{i \vee k > M} \sum_{x \in C(i,k)} e^{\Phi_{2\gamma}(x,t)}$$

$$\leq \sum_{i \vee k > M} \Lambda([\mathcal{P}], \gamma, 3\gamma, i) \Lambda(\mathcal{G}^1, \gamma, 3\gamma, t-i-k) \Lambda([\mathcal{S}], \gamma, 3\gamma, k).$$

Applying Proposition 4.7, we obtain

$$\sum_{\substack{x \in E_t \\ (p \vee s)(x,t) > M}} e^{\Phi_{2\gamma}(x,t)} \leq \sum_{i \vee k > M} \Lambda([\mathcal{P}], \gamma, 3\gamma, i) \Lambda([\mathcal{S}], \gamma, 3\gamma, k) C_4 e^{(t-i-k)P(\phi)}$$

$$\leq \sum_{i \vee k > M} C_4 C_5^2 e^{(i+k)(P(\phi)-\beta_1)} e^{(t-i-k)P(\phi)}$$

$$= C_4 C_5^2 e^{tP(\phi)} \sum_{i \vee k > M} e^{-(i+k)\beta_1},$$

where the second inequality uses (4.9). Let M be chosen large enough that the sum in the above expression is less than $\frac{1}{2}\alpha_1\alpha_2 C_4^{-1} C_5^{-2}$. Then, for t large enough so (4.10) holds, we have

$$\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) - \beta_2 < \Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t) + \frac{\alpha_1\alpha_2}{2} e^{tP(\phi)} + \Lambda(\mathcal{D}^c, 2\gamma, 2\gamma, t)$$

$$\leq \Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t) + \frac{\alpha_2}{2}\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) + C_5 e^{t(P(\phi)-\beta_1)}$$

$$\leq \Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t) + \frac{\alpha_2}{2}\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) + \frac{\alpha_1\alpha_2}{2} e^{tP(\phi)}$$

$$\leq \Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, t) + \alpha_2\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t).$$

Since $\beta_2 > 0$ was arbitrary and T_1, M were chosen independently of β_2, this completes the proof of Lemma 4.8. $\square$

4.4. **Consequences of lower bound on $\mathcal{G}$.** Throughout this section, we assume that $\mathcal{G} \subset X \times \mathbb{R}^+$ and $\delta, \varepsilon > 0$ are such that $\mathcal{G}$ has tail specification at scale δ and $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$. The following lemmas are consequences of Lemma 4.8.

Lemma 4.9. *Let δ, ε be as above. Fix $\gamma \in (2\delta, \varepsilon/8]$ such that ϕ has the Bowen property on $\mathcal{G}$ at scale 3γ. Then for every $\alpha > 0$ there are $M \in \mathbb{N}$ and $T_1 \in \mathbb{R}$ such that for every $t \geq T_1$ we have*

$$(4.13) \qquad \Lambda(\mathcal{G}^M, 2\gamma, 2\gamma, t) \geq (1-\alpha)\Lambda(X, 2\gamma, 2\gamma, t) \geq (1-\alpha)e^{tP(\phi)}.$$

Proof. The inequality $\Lambda(X, 2\gamma, 2\gamma, t) \geq e^{tP(\phi)}$ is true for any $0 < 2\gamma \leq \varepsilon/4$ by Lemma 4.2. Since $\gamma > 2\delta$ and ϕ is Bowen on $\mathcal{G}$ at scale 3γ, we can apply Lemma 4.8 to $\Lambda(X, 2\gamma, 2\gamma, t)$ with $\alpha_1 = 1$ and $\alpha_2 = \alpha$, obtaining M and T_1 such that (4.13) holds for $t \geq T_1$. $\square$

We now obtain the following lower bound for the standard partition sum $\Lambda(\mathcal{G}^M, 2\gamma, t)$.

Proposition 4.10. *Let $\varepsilon, \delta, \gamma$ be as in Lemma 4.9. Then there are $M, T_1, L \in \mathbb{R}^+$ such that for $t \geq T_1$,*

$$\Lambda(\mathcal{G}^M, 2\gamma, t) \geq e^{-L} e^{tP(\phi)}.$$

As a consequence, $\Lambda(X, 2\gamma, t) \geq e^{-L} e^{tP(\phi)}$ for $t \geq T_1$.

Proof. We apply Lemma 4.9 with $\alpha = 1/2$ to obtain M, T_1 so that

$$\Lambda(\mathcal{G}^M, 2\gamma, 2\gamma, t) \geq \tfrac{1}{2}e^{tP(\phi)}$$

for $t \geq T_1$. Now, since $\mathcal{G}$ has the Bowen property at scale 2γ, then $\mathcal{G}^M$ also has the Bowen property at scale 2γ. Letting $L := K(M) = K + 2M\operatorname{Var}(\phi, \varepsilon)$ be the distortion constant from the Bowen property for $\mathcal{G}^M$, we have $\Lambda(\mathcal{G}^M, 2\gamma, t) \geq e^{-L}\Lambda(\mathcal{G}^M, 2\gamma, 2\gamma, t)$, which gives us the required result. The elementary inequality $\Lambda(X, 2\gamma, t) \geq \Lambda(\mathcal{G}^M, 2\gamma, t)$ extends the result to partition sums $\Lambda(X, 2\gamma, t)$. $\square$

Lemma 4.11. *Let $\varepsilon, \delta, \gamma$ be as in Lemma 4.9. Then there is $C_6 \in \mathbb{R}^+$ such that $C_6^{-1}e^{tP(\phi)} \leq \Lambda(X, 2\gamma, t) \leq \Lambda(X, 2\gamma, 2\gamma, t) \leq C_6 e^{tP(\phi)}$ for every $t \in \mathbb{R}^+$.*

Proof. For the first inequality, take T_1 from Proposition 4.10. Let $\kappa_1 = \inf_{0 < t < T_1}\{\Lambda(X, 2\gamma, t)e^{-tP(\phi)}\}$. We have $\kappa_1 > 0$, since $\Lambda(X, 2\gamma, t) \geq e^{\inf \phi}$. Choose C_6 so that $C_6^{-1} \leq \min\{e^{-L}, \kappa\}$, and the first inequality follows from Proposition 4.10. The second inequality is immediate; the third comes by putting $\alpha = 1/2$ in Lemma 4.9 and taking M accordingly, then applying Proposition 4.7 to $\mathcal{G}^M$ to get for $t \geq T_1$,

$$\Lambda(X, 2\gamma, 2\gamma, t) \leq 2\Lambda(\mathcal{G}^M, 2\gamma, 2\gamma, t) \leq 2C_4 e^{tP(\phi)}.$$

Let $\kappa_2 = \sup_{0 < t < T_1}\{\Lambda(X, 2\gamma, t)e^{-tP(\phi)}\}$. Ensure that C_6 is chosen so that $C_6 \geq \max\{2C_4, \kappa_2\}$, and the result follows. $\square$

We note that the assumption that $t \geq T_1$ in Proposition 4.10 can be removed under a mild consistency condition on $\mathcal{G}^M$, which is automatically verified if $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ is a decomposition for the whole space (X, F). This is the content of the following lemma.

Lemma 4.12. *Let $\varepsilon, \delta, \gamma$ be as in Lemma 4.9, and M, T_1 as in Proposition 4.10. Suppose that $\mathcal{G}_t^M \neq \emptyset$ for all $0 < t < T_1$. Then there exists $C_6' \in \mathbb{R}^+$ so that for all $t > 0$,*

$$\Lambda(\mathcal{G}^M, 2\gamma, t) \geq C_6' e^{tP(\phi)}.$$

Proof. By Proposition 4.10, there exists L so that for $t \geq T_1$,

$$\Lambda(\mathcal{G}^M, 2\gamma, t) \geq e^{-L}e^{tP(\phi)}.$$

Let $\kappa_3 = \inf_{0 < t < T_1}\{\Lambda(\mathcal{G}^M, 2\gamma, t)e^{-tP(\phi)}\}$. Using the assumption that $\mathcal{G}_t^M \neq \emptyset$, we have $\kappa_3 > 0$, since $\Lambda(\mathcal{G}^M, 2\gamma, t) \geq e^{\inf \phi}$. Let $C_6' = \min\{e^{-L}, \kappa_3\}$, and the result follows. $\square$

4.5. **An equilibrium state with a Gibbs property.** From now on, we fix $\varepsilon > 40\delta > 0$ so the hypotheses of Theorem 2.9 are satisfied:

1. for every $M \in \mathbb{R}^+$ there are $T(M), \tau_M \in \mathbb{R}^+$ such that $\mathcal{G}^M$ has specification at scale δ for $t \geq T(M)$ with transition time τ_M;
2. ϕ is Bowen on $\mathcal{G}$ at scale ε (with distortion constant K);
3. $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$;
4. $P(\mathcal{D}^c \cup [\mathcal{P}] \cup [\mathcal{S}], \delta, \varepsilon) < P(\phi)$.

We fix a scale $\rho \in (5\delta, \varepsilon/8]$, and we also consider $\rho' := \rho - \delta$. For concreteness, to take scales which are all integer multiples of δ, we could take:

- $\varepsilon = 48\delta$
- $\rho = 6\delta$
- $\rho' = 5\delta$

Remark 4.13. The scale hypotheses in (2) and (4) above can be sharpened a little. For a scale $\rho > 5\delta$, where δ is the scale so that $\mathcal{G}^M$ has specification, the largest scale where we need to control expansivity occurs in §4.7 and §4.8, where we require $P_{\exp}^{\perp}(\phi, 8\rho) < P(\phi)$. The largest scale where we require the Bowen property on $\mathcal{G}$ is 3ρ (Lemma 4.9 applied with $\gamma = \rho$). The only place we require the estimate on pressure of $\mathcal{D}^c \cup [\mathcal{P}] \cup [\mathcal{S}]$ is Lemma 4.8, which is applied with $\gamma = \rho$, so it would suffice to assume that $P(\mathcal{D}^c, 2\rho, 2\rho) < P(\phi)$ and $P([\mathcal{P}] \cup [\mathcal{S}], \rho, 3\rho) < P(\phi)$. In particular, it would suffice to assume that $P(\mathcal{D}^c \cup [\mathcal{P}] \cup [\mathcal{S}], \delta) + \mathrm{Var}(\phi, 15\delta) < P(\phi)$.

The construction of an equilibrium state for ϕ is quite standard. For each $t \geq 0$, let $E_t \subset X$ be a maximizing (t, ρ')-separated set for $\Lambda(X, \rho', t)$. Then consider the measures

(4.14)
$$\nu_t := \frac{\sum_{x \in E_t} e^{\Phi_0(x,t)} \delta_x}{\sum_{x \in E_t} e^{\Phi_0(x,t)}},$$

$$\mu_t := \frac{1}{t} \int_0^t (f_s)_* \nu_t \, ds.$$

By compactness there is an integer valued sequence $n_k \to \infty$ such that μ_{n_k} converges in the weak* topology. Let $\mu = \lim_k \mu_{n_k}$.

Lemma 4.14. μ *is an equilibrium state for* (X, F, ϕ).

Proof. We compute $h_\mu(f_1) + \int \phi \, d\mu$. Recall that $\phi^{(1)}(x) = \int_0^1 \phi(f_t x) \, dt$, and for each n, note that

$$\nu_n = \frac{\sum_{x \in E_n} e^{S_n^{f_1} \phi^{(1)}(x)} \delta_x}{\sum_{x \in E_n} e^{S_n^{f_1} \phi^{(1)}(x)}}.$$

Let $\mu_n^{(1)} = \frac{1}{n} \sum_{k=0}^{n-1} (f_1)_*^k \nu_n$, and observe that

(4.15) $$\int_0^1 (f_s)_* \mu_n^{(1)} \, ds = \frac{1}{n} \sum_{k=0}^{n-1} \int_0^1 (f_{k+s})_* \nu_n \, ds = \frac{1}{n} \int_0^n (f_s)_* \nu_n \, ds = \mu_n.$$

Passing to a subsequence if necessary, let $\mu^{(1)} = \lim_k \mu_{n_k}^{(1)}$. The second part of the proof of [Wal82, Theorem 8.6] yields

$$h_{\mu^{(1)}}(f_1) + \int \phi^{(1)} \, d\mu^{(1)} = P(\phi^{(1)}, \rho'; f_1),$$

where we compute $P(\phi^{(1)}, \rho'; f_1)$ in the d_1 metric, and thus $P(\phi^{(1)}, \rho'; f_1) = P(\phi, \rho'; F) = P(\phi; F)$, by Proposition 3.7. Using (4.15) we get

$$h_\mu(f_1) + \int \phi \, d\mu = h_{\mu^{(1)}}(f_1) + \int \phi^{(1)} \, d\mu^{(1)} = P(\phi; F). \qquad \square$$

The previous lemma holds whenever $P(\phi, \rho') = P(\phi)$, and thus provides the existence of an equilibrium state under this hypothesis. Since this result has independent interest, we state it here as a self-contained proposition.

Proposition 4.15. *If $P_{\exp}^\perp(\phi) < P(\phi)$, then there exists an equilibrium state for ϕ.*

Proof. Since $P_{\exp}^\perp(\phi) < P(\phi)$, using Proposition 3.7, we can find a scale γ so that $P(\phi, \gamma) = P(\phi)$. We now construct μ as above but with γ in place of ρ', and carry out the argument of Lemma 4.14 to see that μ is an equilibrium state for ϕ. $\qquad \square$

The following lemma requires that $\rho \in (2\delta, \varepsilon/8]$ and ϕ has the Bowen property at scale 3ρ. This is true by our choice of ρ.

Lemma 4.16. *For sufficiently large M there is $Q_M > 0$ such that for every $(x, t) \in \mathcal{G}^M$ with $t \geq T(M)$ we have*

$$(4.16) \qquad \mu(B_t(x, \rho)) \geq Q_M e^{-tP(\phi) + \Phi_0(x,t)}.$$

Proof. We apply Proposition 4.10 with $\gamma = \rho$ to obtain M, T_1 and L so that

$$\Lambda(\mathcal{G}^M, 2\rho, t) > \tfrac{1}{2} e^{-L} e^{tP(\phi)}.$$

for all $t \geq T_1$. Without loss of generality assume that $T_1 \geq T(M)$. For every $u \geq T_1$ we can find a $(u, 2\rho)$-separated set $E_u' \subset \mathcal{G}_u^M$ such that

$$(4.17) \qquad \sum_{x \in E_u'} e^{\Phi_0(x,u)} \geq \tfrac{1}{2} e^{-L} e^{uP(\phi)}.$$

Given $(x, t) \in \mathcal{G}^M$ with $t \geq T(M)$, we estimate $\mu(B_t(x, \rho))$ by estimating $\nu_s(f_{-r} B_t(x, \rho))$ for $s \gg t$ and $r \in [\tau_M + T_1, s - 2\tau_M - 2T_1 - t]$. Given s and r, let $u_1 := r - \tau_M$ and $u_2 := s - r - t - \tau_M$; see Figure 4.
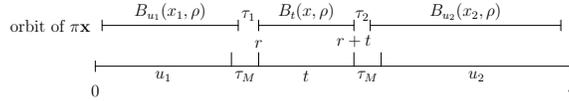
FIGURE 4. Proving the Gibbs property.

We use similar ideas to the proof of Proposition 4.3 to construct a map $\pi \colon E_{u_1}' \times E_{u_2}' \to E_t$ as follows. By the specification property, for each $\mathbf{x} \in E_{u_1}' \times E_{u_2}'$, we can find $y(\mathbf{x}), \tau_1(\mathbf{x}), \tau_2(\mathbf{x})$ so that

$$y(\mathbf{x}) \in B_{u_1}(x_1, \delta),$$
$$f_{u_1 + \tau_1(\mathbf{x})}(y(\mathbf{x})) \in B_t(x, \delta),$$
$$f_{u_1 + \tau_1(\mathbf{x}) + t + \tau_2(\mathbf{x})}(y(\mathbf{x})) \in B_{u_2}(x_2, \delta).$$

Recall that $\rho' = \rho - \delta$, so that $\rho' > 4\delta$, and that E_s denotes the maximizing (s, ρ')-separated set used in the construction of ν_s and μ_s. Let $\pi \colon E'_{u_1} \times E'_{u_2} \to E_s$ be given by choosing a point $\pi(\mathbf{x}) \in E_s$ such that

$$d_s(\pi(\mathbf{x}), f_{\tau_1(\mathbf{x}) - \tau_M} y(\mathbf{x})) \leq \rho'.$$

For any $\mathbf{x} \in E'_{u_1} \times E'_{u_2}$, we have

$$d_t(f_r(\pi\mathbf{x}), x) \leq d_t(f_r(\pi\mathbf{x}), f_{r + \tau_1 - \tau_M} y(\mathbf{x})) + d_t(f_{r + \tau_1 - \tau_M} y(\mathbf{x}), x) < \rho' + \delta = \rho.$$

This shows that

(4.18) $$\pi\mathbf{x} \in f_{-r} B_t(x, \rho).$$

The proof of Lemma 4.4 shows there is C_2 such that $\#\pi^{-1}(z) \leq C_2^3$ for every $z \in E_s$. Moreover, a mild adaptation of the proof of Lemma 4.5 gives the existence of C_7 such that

(4.19) $$\Phi_0(\pi\mathbf{x}, s) \geq -C_7 + \Phi_0(x_1, u_1) + \Phi_0(x_2, u_2) + \Phi_0(x, t).$$

Now we have the estimate

(4.20)
$$\nu_s(f_{-r} B_t(x, \rho)) = \frac{\sum_{z \in E_s} e^{\Phi_0(z, s)} \delta_z(f_{-r} B_t(x, \rho))}{\sum_{z \in E_s} e^{\Phi_0(z, s)}}$$
$$\geq C_2^{-3} C_6^{-1} e^{-sP(\phi)} \sum_{\mathbf{x} \in E'_{u_1} \times E'_{u_2}} e^{\Phi_0(\pi\mathbf{x}, s)},$$

where we use (4.18) and the multiplicity bound for the estimate on the numerator, and Lemma 4.11 for the estimate on the denominator. (We apply the lemma with $\gamma = \rho'/2$, so we need $\rho' > 4\delta$ and $\rho' \leq \varepsilon$.) By (4.19) and (4.17) we have

$$\sum_{\mathbf{x} \in E'_{u_1} \times E'_{u_2}} e^{\Phi_0(\pi\mathbf{x}, s)} \geq e^{-C_7} e^{\Phi_0(x, t)} \left(\sum_{x_1 \in E'_{u_1}} e^{\Phi_0(x_1, u_1)} \right) \left(\sum_{x_1 \in E'_{u_2}} e^{\Phi_0(x_2, u_2)} \right)$$
$$\geq \frac{1}{4} e^{-C_7} e^{\Phi_0(x, t)} e^{-2K} e^{u_1 P(\phi)} e^{u_2 P(\phi)}.$$

Together with (4.20) and the fact that $s = u_1 + u_2 + t + 2\tau_M$, this gives

$$\nu_s(f_{-r} B_t(x, \rho)) \geq C_8 e^{\Phi_0(x, t)} e^{(u_1 + u_2 - s)P(\phi)} = C_8 e^{-2\tau_M P(\phi)} e^{-tP(\phi) + \Phi_0(x, t)}$$

for every $r \in [\tau_M + T_1, s - 2\tau_M - 2T_1 - t]$. Integrating over r gives

$$\mu_s(B_t(x, \rho)) \geq \frac{1}{s} \int_{\tau_M + T_1}^{s - 2\tau_M - 2T_1 - t} \nu_s(f_{-r} B_t(x, \rho)) \, dr$$
$$\geq \left(1 - \frac{t + 3\tau_M + 3T_1}{s} \right) C_8 e^{-2\tau_M P(\phi)} e^{-tP(\phi) + \Phi_0(x, t)}.$$

Sending $s \to \infty$ completes the proof of Lemma 4.16. $\square$

Later we will prove that μ is ergodic. The proof of this will use the following lemma that generalizes Lemma 4.16. We also require here that $2\rho \leq \varepsilon$ and ϕ has the Bowen property at scale 3ρ, which is true by assumption.

Lemma 4.17. *For sufficiently large M there is $Q'_M > 0$ such that for every $(x_1, t_1), (x_2, t_2) \in \mathcal{G}^M$ with $t_1, t_2 \geq T(M)$ and every $q \geq 2\tau_M$ there is $q' \in [q - 2\tau_M, q]$ such that we have*

$$\mu(B_{t_1}(x_1, \rho) \cap f_{-(t_1 + q')} B_{t_2}(x_2, \rho)) \geq Q'_M e^{-(t_1 + t_2)P(\phi) + \Phi_0(x_1, t_1) + \Phi_0(x_2, t_2)}.$$

Furthermore, we can choose $N = N(F, \delta) \in \mathbb{N}$ such that q' can be taken to be of the form $q - \frac{2i\tau_M}{N}$ for some $i \in \{0, \ldots N\}$.

Proof. The proof closely resembles the proof of Lemma 4.16. We start in the same way, and, for an arbitrary $\beta > 0$, and every $u \geq T_1$ we take a $(u, 2\rho)$-separated set $E'_u \subset \mathcal{G}_u^M$ satisfying (4.17). Now given $q \geq 2\tau_M + T_1$, $s \gg t_1, t_2$, and $r \geq \tau_M + T_1$, choose u_1, u_2, u_3 such that

$$u_1 := r - \tau_M, \quad u_2 := q - 2\tau_M, \quad u_3 := s - r - t_1 - q - t_2 - \tau_M;$$

see Figure 5 for an illustration.
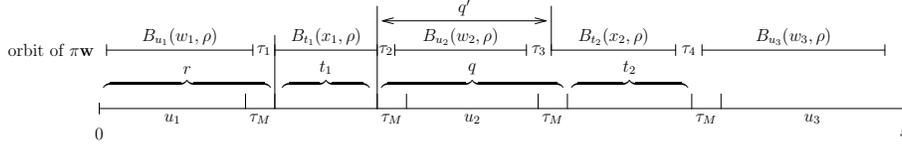
FIGURE 5. Proving Lemma 4.17.

We use similar ideas to the proof of Proposition 4.3 to construct a map $\pi \colon E'_{u_1} \times E'_{u_2} \times E'_{u_3} \to E_s$ as follows. By the specification property, for each $\mathbf{w} \in E'_{u_1} \times E'_{u_2} \times E'_{u_3}$, we can find $y = y(\mathbf{w})$, and $\tau_i = \tau_i(\mathbf{w})$ for $i = 1, 2, 3, 4$ so that

$$y \in B_{u_1}(w_1, \delta),$$
$$f_{u_1 + \tau_1}(y) \in B_{t_1}(x_1, \delta),$$
$$f_{u_1 + \tau_1 + t_1 + \tau_2}(y) \in B_{u_2}(w_2, \delta),$$
$$f_{u_1 + \tau_1 + t_1 + \tau_2 + u_2 + \tau_3}(y) \in B_{t_2}(x_2, \delta),$$
$$f_{u_1 + \tau_1 + t_1 + \tau_2 + u_2 + \tau_3 + t_2 + \tau_4}(y) \in B_{u_3}(w_3, \delta).$$

Recall that $\rho' = \rho - \delta$, so that $\rho' > 4\delta$, and that E_s denotes the maximizing (s, ρ')-separated set used in the construction of ν_s and μ_s. Let $\pi \colon E'_{u_1} \times E'_{u_2} \times E'_{u_3} \to E_s$ be given by choosing a point $\pi(\mathbf{w}) \in E_t$ such that

$$d_s(\pi(\mathbf{w}), f_{\tau_1(\mathbf{w}) - \tau_M} y(\mathbf{w})) \leq \rho'.$$

In particular,

(4.21)
$$d_{t_1}(f_r(\pi\mathbf{w}), x_1) < \rho,$$
$$d_{t_2}(f_{r + t_1 + \hat{q}(\mathbf{w})}(\pi\mathbf{w}), x_2) < \rho,$$

where we write $\hat{q}(\mathbf{w}) := \tau_2 + u_2 + \tau_3 \in [q - 2\tau_M, q]$. Note that similar to the proof of Lemma 4.4, by taking $A \subset [q - 2\tau_M, q]$ to be θ-dense and finite for a suitable small θ, we can replace $\hat{q}(\mathbf{w})$ with $q'(\mathbf{w}) \in A$ such that (4.21)

continues to hold. We can choose $\theta = \frac{2\tau_M}{N}$ for a suitably large N and set $A = \{q - i\theta : i \in \{0, \ldots, N\}\}$. Lemmas 4.4 and 4.5 adapt to give

$$\#\pi^{-1}(z) \leq C_9,$$

$$\Phi_0(\pi\mathbf{w}) \geq -C_{10} + \sum_{i=1}^{3} \Phi_0(w_i, u_i) + \Phi_0(x_1, t_1) + \Phi_0(x_2, t_2)$$

Thus we have the estimate

$$\sum_{q' \in A} \nu_s(f_{-r}B_{t_1}(x_1, \rho) \cap f_{-(r+t_1+q')}B_{t_2}(x_2, \rho))$$

$$= \frac{\sum_{z \in E_s} e^{\Phi_0(z,s)} \delta_z(f_{-r}B_{t_1}(x_1, \rho) \cap f_{-(r+t_1+q')}B_{t_2}(x_2, \rho))}{\sum_{z \in E_s} e^{\Phi_0(z,s)}}$$

$$\geq C_9^{-1} C_6^{-1} e^{-sP(\phi)} \sum_{\mathbf{w} \in \prod_i E'_{u_i}} e^{\Phi_0(\pi\mathbf{w}, s)}$$

$$\geq (C_9 C_6)^{-1} e^{-sP(\phi)} e^{-C_{10}} \sum_{\mathbf{w} \in \prod_i E'_{u_i}} e^{\sum_i \Phi_0(w_i, u_i)} e^{\Phi_0(x_1, t_1) + \Phi_0(x_2, t_2)}$$

$$\geq (C_9 C_6 e^{C_{10}})^{-1} e^{(u_1 + u_2 + u_3 - s)P(\phi)} e^{\Phi_0(x_1, t_1) + \Phi_0(x_2, t_2)}.$$

Observing that $u_1 + u_2 + u_3 \in [s - (t_1 + t_2) - 4\tau_M, s]$, integrating over r, and sending $s \to \infty$ gives

$$\sum_{q' \in A} \mu(B_{t_1}(x_1, \rho) \cap f_{-(t_1+q')}B_{t_2}(x_2, \rho)) \geq C_{11} e^{-(t_1+t_2)P(\phi)} e^{\Phi_0(x_1, t_1) + \Phi_0(x_2, t_2)}.$$

Taking $Q'_M = C_{11}/\#A$, this completes the proof of Lemma 4.17. $\square$

4.6. Adapted partitions and positive measure sets.

Recall that if E_t is a $(t, 2\gamma)$-separated set of maximal cardinality, then a partition $\mathcal{A}$ of X is *adapted* to E_t if for every $w \in \mathcal{A}$ there is $x \in E_t$ such that $B_t(x, \gamma) \subset w \subset \overline{B}_t(x, 2\gamma)$. We can write $x = x(w) \in E_t$ for the (unique) point corresponding to the partition element w. Conversely, given $x \in E_t$, let $w_x \in \mathcal{A}_t$ be the partition element such that

$$B_t(x, \gamma) \subset w_x \subset \overline{B}_t(x, 2\gamma).$$

We now state a crucial lemma, which gives us a growth rate on partition sums for sets with positive measure for an equilibrium measure. Note that this is another place in the proof where we are forced to work with the non-standard partition sums $\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t)$.

Lemma 4.18. *Let ε, δ be as in the previous section, and let $\gamma \in (2\delta, \varepsilon/8]$. For every $\alpha \in (0, 1)$, there is $C_\alpha > 0$ such that the following is true. Consider any equilibrium state ν for ϕ and a family $\{E_t\}_{t>0}$ of maximizing $(t, 2\gamma)$-separated sets for $\Lambda(X, 2\gamma, t)$ with adapted partitions $\mathcal{A}_t$. If $t \in \mathbb{R}^+$ and*

$E'_t \subset E_t$ satisfy $\nu \left(\bigcup_{x \in E'_t} w_x \right) \geq \alpha$, then letting $\mathcal{C} = \{(x,t) : x \in E'_t\}$, we have

$$\Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) \geq C_\alpha e^{tP(\phi)}.$$

Proof. Fix $t > 0$, and let E_t, $\mathcal{A}_t$ and E'_t be as in the statement of the lemma. The proof is similar to [CT13, Proposition 5.4] and [CT14, Lemma 3.12]. Given $w \in \mathcal{A}_t$, write $\Phi(w) = \sup_{y \in w} \Phi_0(y,t)$, and note that if $x = x(w) \in E_t$ is such that $w \subset B_t(x, 2\gamma)$, then we have $\Phi(w) \leq \Phi_{2\gamma}(x,t)$. Using the fact that ν is an equilibrium state, and our assumption that $P^\perp_{\exp}(\phi, \varepsilon) < P(\phi)$ with $2\gamma \leq \varepsilon$, we can apply Lemma 3.8 and get

$$tP(\phi) \leq \nu(D) \log \sum_{x \in E'_t} e^{\Phi_{2\gamma}(x,t)} + \nu(D^c) \log \sum_{x \in E_t \setminus E'_t} e^{\Phi_{2\gamma}(x,t)} + \log 2,$$

where $D = \bigcup_{x \in E'_t} w_x$ and $D^c = X \setminus D$. Applying Lemma 4.11 gives

$$tP(\phi) \leq \nu(D) \log \Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) + (1 - \nu(D))(tP(\phi) + \log C_6) + \log 2,$$

and rearranging we get

$$-\log(2C_6) \leq \nu(D) \left(-\log C_6 - tP(\phi) + \log \Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) \right).$$

Since $\nu(D) \geq \alpha$, we get

$$-\log C_6 - tP(\phi) + \log \Lambda(\mathcal{C}, 2\gamma, 2\gamma, t) \geq -\frac{\log(2C_6)}{\nu(D)} \geq -\frac{\log(2C_6)}{\alpha},$$

which suffices to complete the proof of Lemma 4.18. $\qquad\square$

4.7. **No mutually singular equilibrium measures.** Let μ be the equilibrium state we have constructed, and suppose that there exists another equilibrium state $\nu \perp \mu$. Let $P \subset X$ be an F-invariant set such that $\mu(P) = 0$ and $\nu(P) = 1$, and let $\mathcal{A}_t$ be adapted partitions for maximizing $(t, 2\rho)$-separated sets E_t. To simplify notation, we use the same symbol (such as U) to denote both a set of partition elements ($U \subset \mathcal{A}$) and the union of those elements ($U \subset X$). Applying Proposition 3.10, there exists $U_t \subset \mathcal{A}_t$ such that $\frac{1}{2}(\mu + \nu)(U_t \triangle P) \to 0$. Note that to apply Proposition 3.10, we need $\frac{1}{2}(\mu + \nu)$ to be almost expansive at scale 4ρ. This is true, since we assumed that $4\rho \leq \varepsilon$ and $P^\perp_{\exp}(\phi, \varepsilon) < P(\phi)$. In particular, we have $\nu(U_t) \to 1$ and $\mu(U_t) \to 0$ and we can assume without loss of generality that $\inf_t \nu(U_t) > 0$, and so by Lemma 4.18, for $\mathcal{C} = \{(x,t) : x \in E_t \cap U_t\}$, we have

$$\Lambda(\mathcal{C}, 2\rho, 2\rho, t) \geq C e^{tP(\phi)},$$

and so by Lemma 4.8, there exists M so that

$$\Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\rho, 2\rho, t) \geq \frac{C}{2} e^{tP(\phi)}$$

for all sufficiently large t. In other words, letting $E_t^M = \{x \in E_t \mid (x, t) \in \mathcal{G}^M\}$, we have

$$\sum_{x \in E_t^M \cap U_t} e^{\Phi_{2\rho}(x,t)} \geq \frac{C}{2} e^{tP(\phi)},$$

and thus by the Bowen property for $\mathcal{G}$,

$$\sum_{x \in E_t^M \cap U_t} e^{\Phi_0(x,t)} \geq \frac{C}{2} e^{-K} e^{tP(\phi)}.$$

Finally, we use the Gibbs property in Lemma 4.16 to observe that since $U_t \supset B_t(x, \rho)$ for every $x \in E_t \cap U_t$, we have

$$\mu(U_t) \geq \sum_{x \in E_t^M \cap U_t} Q_M e^{-tP(\phi) + \Phi_0(x,t)} \geq Q_M e^{-K} C/2 > 0,$$

contradicting the fact that $\mu(U_t) \to 0$. This contradiction implies that any equilibrium state ν is absolutely continuous with respect to μ.

4.8. **Ergodicity.** The following result is the final ingredient needed to complete the proof of Theorem 2.9.

Proposition 4.19. *The equilibrium state μ constructed above is ergodic.*

Proof. Let $P, Q \subset X$ be measurable F-invariant sets. We show that if P and Q have positive μ-measure, then

$$\mu(P \cap Q) > 0.$$

We achieve this by showing there exists $s \in \mathbb{R}$ so that $\mu(P \cap f_{-s}Q) > 0$, noting the invariance of Q. The existence of an F-invariant set P with $0 < \mu(P) < 1$ would contradict this inequality (letting $Q = X \setminus P$), and thus we have ergodicity.

Let $\mathcal{A}_t$ be adapted partitions for maximizing $(t, 2\rho)$-separated sets E_t. Applying Proposition 3.10, there are sets $U_t \subset \mathcal{A}_t$ such that $\mu(U_t \triangle P) \to 0$, and $V_t \subset \mathcal{A}_t$ such that $\mu(V_t \triangle Q) \to 0$. Let $2\alpha_1 := \min\{\mu(P), \mu(Q)\}$. We can ensure that U_t and V_t are chosen so that for all t we have $\mu(U_t) \geq \alpha_1$, and $\mu(V_t) \geq \alpha_1$, so by Lemma 4.18, putting $C = C_{\alpha_1}$, for $\mathcal{C}^U = \{(x, t) : x \in E_t \cap U_t\}$ and $\mathcal{C}^V = \{(x, t) : x \in E_t \cap V_t\}$, we have

$$\Lambda(\mathcal{C}^U, 2\rho, 2\rho, t) \geq C e^{tP(\phi)}, \;\; \Lambda(\mathcal{C}^V, 2\rho, 2\rho, t) \geq C e^{tP(\phi)}$$

and so by Lemma 4.8, there exists M so that

$$\Lambda(\mathcal{C}^U \cap \mathcal{G}^M, 2\rho, 2\rho, t) \geq \frac{C}{2} e^{tP(\phi)} \text{ and } \Lambda(\mathcal{C}^V \cap \mathcal{G}^M, 2\rho, 2\rho, t) \geq \frac{C}{2} e^{tP(\phi)}$$

for all sufficiently large t. In other words, letting $E_t^M = \{x \in E_t \mid (x, t) \in \mathcal{G}^M\}$, we have

$$\sum_{x \in E_t^M \cap U_t} e^{\Phi_{2\rho}(x,t)} \geq \frac{C}{2} e^{tP(\phi)} \text{ and } \sum_{x \in E_t^M \cap V_t} e^{\Phi_{2\rho}(x,t)} \geq \frac{C}{2} e^{tP(\phi)}.$$

Fix $q > 2\tau_M$. Note that for every $x \in E_t^M \cap U_t$, $y \in E_t^M \cap V_t$, we have

$$(4.22) \qquad B_t(x, \rho) \cap f_{-(t+q)} B_t(y, \rho) \subset U_t \cap f_{-(t+q)} V_t.$$

By Lemma 4.17, there exists N so for each such x, y there exists $q' \in [q - 2\tau_M, q]$ with $q' = q - \frac{2i\tau_M}{N}$ for some $i \in \{0, \dots N\}$ so that

$$\mu(B_t(x, \rho) \cap f_{-(t+q')} B_t(y, \rho)) \geq Q'_M e^{-2tP(\phi) + \Phi_0(x,t) + \Phi_0(y,t)}.$$

Summing over $x \in E_t^M \cap U_t$ and $y \in E_t^M \cap V_t$, we see that for all $t > 0$, there is some value of $q' = q'(t) \in [q - 2\tau_M, q]$ for which

$$\mu(U_t \cap f_{-(t+q')} V_t) \geq \frac{Q'_M}{N} \sum_{x \in E_t^M \cap U_t} \sum_{y \in E_t^M \cap V_t} e^{-2tP(\phi) + \Phi_0(x,t) + \Phi_0(y,t)} \geq \frac{Q'_M}{N} \frac{C^2}{4}.$$

Denote this final quantity by α_2. Now choose t_0 large enough so that

$$\mu(U_{t_0} \triangle P) < \frac{\alpha_2}{2} \text{ and } \mu(V_{t_0} \triangle Q) < \frac{\alpha_2}{2},$$

and write $s = t_0 + q'(t_0)$. Note that

$$(U_{t_0} \cap f_{-s} V_{t_0}) \setminus (P \cap f_{-s} Q) \subset (U_{t_0} \setminus P) \cup f_{-s}(V_{t_0} \setminus Q),$$

which yields

$$\mu(P \cap Q) > \mu(U_{t_0} \cap f_{-s} V_{t_0}) - \alpha_2 \geq 0,$$

and the proposition follows. $\qquad \square$

In conclusion, §4.7 showed that any equilibrium state ν is absolutely continuous with respect to μ, and since μ is ergodic, this in turn implies that $\nu = \mu$, which completes the proof of uniqueness in Theorem 2.9.

4.9. **Upper Gibbs bound.** We establish a weak upper Gibbs bound on μ. We prove a version that applies globally which involves the term $\Phi_\gamma(x, t)$, rather than the term $\Phi_0(x, t)$ that appears in the usual version of the Gibbs property. This yields a uniform upper bound in terms of $\Phi_0(x, t)$ on each $\mathcal{G}^M$. First, we require a slightly different characterization of μ.

Lemma 4.20. *Suppose the hypothesis of Theorem 2.9 apply, and let $\eta \leq \varepsilon/2$. let $E'_t \subset X$ be a maximizing (t, η)-separated set for $\Lambda(X, \eta, t)$. Then consider the measures*

$$(4.23) \qquad \nu'_t := \frac{\sum_{x \in E'_t} e^{\Phi_0(x,t)} \delta_x}{\sum_{x \in E'_t} e^{\Phi_0(x,t)}}, \quad \mu'_t := \frac{1}{t} \int_0^t (f_s)_* \nu'_t \, ds.$$

Then μ'_t converges in the weak* topology to the unique equilibrium measure μ provided by the conclusion of Theorem 2.9.

Proof. Let $\mu'_{t_k} \to \nu$ for some $t_k \to \infty$. The proof of Lemma 4.14 applies with ρ' replaced by η, and shows that ν is an equilibrium measure for φ; the only place in that proof where the scale is used is in the application of Proposition 3.7, which requires $\eta \leq \varepsilon/2$. Since the equilibrium state for φ is unique by Theorem 2.9, then $\nu = \mu$, so μ is the only limit point of $\{\mu'_t\}_{t \in \mathbb{R}^+}$, whence $\mu'_t \to \mu$. $\qquad \square$

Now we establish the upper Gibbs property for μ.

Proposition 4.21. *Suppose the hypotheses of Theorem 2.9 apply and let $\gamma \in (4\delta, \varepsilon/4]$. Then there is $Q > 0$ such that for every $(x, t) \in X \times \mathbb{R}^+$, the unique equilibrium state μ satisfies*

$$\mu(B_t(x, \gamma)) \leq Qe^{-tP(\phi) + \Phi_\gamma(x,t)}.$$

Furthermore, for each $M \in \mathbb{R}^+$, there exists $Q(M) > 0$ so that for every $(x, t) \in \mathcal{G}^M$,

$$\mu(B_t(x, \gamma)) \leq Q(M)e^{-tP(\phi) + \Phi_0(x,t)}.$$

Proof. Let $\eta = 2\gamma$ and note that $\eta \leq \varepsilon/2$. For each $s \in \mathbb{R}^+$, let E'_s be a maximizing (t, η)-separated set for $\Lambda(X, \eta, t)$; then let ν'_s, μ'_s be as in (4.23). Then $\mu = \lim_k \mu'_k$ by Lemma 4.20. To obtain the upper Gibbs bound we want to estimate $\nu'_s(f_{-r}B_t(x, \gamma))$ with $s \gg t$ and $r \in [0, s - t]$. As in the proof of Lemma 4.1, let E''_r be a maximizing (r, γ)-separated set and consider the map $\pi\colon E'_s \cap f_{-r}B_t(x, \gamma) \to E''_r \times E''_{s-t-r}$ given by $d_r(\pi(x)_1, x) < \gamma$ and $d_{s-t-r}(\pi(x)_2, f_{r+t}x) < \gamma$. Injectivity of π follows here just as it did there. Observing that $\gamma/2 \in (2\delta, \varepsilon/8]$, we see from Lemma 4.11 that

$$C_6^{-1}e^{\tau P(\phi)} \leq \Lambda(X, \gamma, \tau) \leq \Lambda(X, \gamma, \gamma, \tau) \leq C_6 e^{\tau P(\phi)}$$

for all $\tau \in \mathbb{R}^+$, and thus

$$\begin{aligned}
\nu'_s(f_{-r}B_t(x, \gamma)) &= \frac{\sum_{x \in E'_s \cap f_{-r}B_t(x,\gamma)} e^{\Phi_0(x,t)}}{\sum_{x \in E'_s} e^{\Phi_0(x,t)}} \\
&\leq C_6 e^{-sP(\phi)} \sum_{z_1 \in E''_r} \sum_{z_2 \in E''_{s-r-t}} e^{\Phi_\gamma(z_1, r)} e^{\Phi_\gamma(x,t)} e^{\Phi_\gamma(z_2, s-r-t)} \\
&\leq C_6 e^{-sP(\phi)} \Lambda(X, \gamma, \gamma, r) \Lambda(X, \gamma, \gamma, s - r - t) e^{\Phi_\gamma(x,t)} \\
&\leq (C_6)^3 e^{-sP(\phi)} e^{rP(\phi)} e^{(s-r-t)P(\phi)} e^{\Phi_\gamma(x,t)} \\
&= (C_6)^3 e^{-tP(\phi) + \Phi_\gamma(x,t)}.
\end{aligned}$$

Averaging over r and sending $s \to \infty$ completes the proof of Proposition 4.21. $\qquad\square$

5. Unique equilibrium states for homeomorphisms

Let X be a compact metric space and $f\colon X \to X$ a homeomorphism. Everything below can also be formulated for non-invertible continuous maps; the only difference is that then we must consider almost positive expansivity in place of almost expansivity, as described in [CT14].

5.1. **Partition sums.** Given $n \in \mathbb{N}$ and $x, y \in X$ we write

$$(5.1) \qquad d_n(x, y) := \sup\{d(f^k x, f^k y) \mid 0 \leq k < n\}.$$

The *Bowen ball* of order n and radius δ centred at $x \in X$ is

(5.2)
$$B_n(x, \delta) := \{y \in X \mid d_n(x, y) < \delta\},$$
$$\overline{B}_n(x, \delta) := \{y \in X \mid d_n(x, y) \leq \delta\}.$$

Given $\delta > 0$, $n \in \mathbb{N}$, and $E \subset X$, we say that E is (n, δ)-separated if for every distinct $x, y \in E$ we have $y \notin \overline{B}_n(x, \delta)$.

We view $X \times \mathbb{N}$ as the space of finite orbit segments for (X, f) by associating to each pair (x, n) the orbit segment $\{f^k(x) \mid 0 \leq k < n\}$. Given $\mathcal{C} \subset X \times \mathbb{N}$ and $n \in \mathbb{N}$ we write $\mathcal{C}_n = \{x \in X \mid (x, n) \in \mathcal{C}\}$.

Now we fix a continuous potential function $\phi \colon X \to \mathbb{R}$. Given a fixed scale $\varepsilon > 0$, we use ϕ to assign a weight to every finite orbit segment by putting

(5.3)
$$\Phi_\varepsilon(x, n) = \sup_{y \in B_n(x, \varepsilon)} \sum_{k=0}^{n-1} \phi(f^k y).$$

In particular, $\Phi_0(x, n) = \sum_{k=0}^{n-1} \phi(f^k x)$.

Given $\mathcal{C} \subset X \times \mathbb{N}$ and $n \in \mathbb{N}$, we consider the *partition function*

(5.4) $\quad \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, n) = \sup \left\{ \sum_{x \in E} e^{\Phi_\varepsilon(x, n)} \mid E \subset \mathcal{C}_n \text{ is } (n, \delta)\text{-separated} \right\},$

and we write $\Lambda(\mathcal{C}, \phi, \delta, n)$ in place of $\Lambda(\mathcal{C}, \phi, \delta, 0, n)$. We can restrict the supremum to sets $E \subset \mathcal{C}_n$ of maximal cardinality, and thus which are (n, δ)-spanning. When $\mathcal{C} = X \times \mathbb{N}$ is the entire system, we will simply write $\Lambda(X, \phi, \delta, \varepsilon, n)$.

The pressure of ϕ on $\mathcal{C}$ at scales δ, ε is given by

$$P(\mathcal{C}, \phi, \delta, \varepsilon) = \varlimsup_{t \to \infty} \frac{1}{n} \log \Lambda(\mathcal{C}, \phi, \delta, \varepsilon, n).$$

When $\varepsilon = 0$, we simplify our notation and write $P(\mathcal{C}, \phi, \delta)$ instead of $P(\mathcal{C}, \phi, \delta, 0)$. We let

$$P(\mathcal{C}, \phi) = \lim_{\delta \to 0} P(\mathcal{C}, \phi, \delta).$$

When $\mathcal{C} = X \times \mathbb{N}$ is the entire space of orbit segments we obtain the usual notion of topological pressure on the entire system, and simply write $P(\phi, \delta, \varepsilon)$ and $P(\phi)$.

5.2. Decompositions. We write $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Definition 5.1. *A decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ for $\mathcal{D} \subseteq X \times \mathbb{N}$ consists of three collections $\mathcal{P}, \mathcal{G}, \mathcal{S} \subset X \times \mathbb{N}_0$ and three functions $p, g, s \colon \mathcal{D} \to \mathbb{N}_0$ such that for every $(x, n) \in \mathcal{D}$, the values $p = p(x, n)$, $g = g(x, n)$, and $s = s(x, n)$ satisfy $n = p + g + s$, and*

(5.5) $\qquad (x, p) \in \mathcal{P}, \ (f^p(x), g) \in \mathcal{G}, \ (f^{p+g}(x), s) \in \mathcal{S}.$

If $\mathcal{D} = X \times \mathbb{N}$, we say that $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ is a decomposition for (X, f). Given a decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ and $M \in \mathbb{N}$, we write $\mathcal{G}^M$ for the set of orbit segments $(x, n) \in \mathcal{D}$ for which $p \leq M$ and $s \leq M$.

We make a standing assumption that $X \times \{0\} \subset \mathcal{P} \cap \mathcal{G} \cap \mathcal{S}$, where each symbol $(x, 0)$ is identified with the empty set. This allows for orbit segments to be decomposed in 'trivial' ways; for example, (x, n) can belong 'purely' to one of the collections $\mathcal{P}, \mathcal{G}$ or $\mathcal{S}$ or can transition directly from $\mathcal{P}$ to $\mathcal{S}$. This is implicit in our earlier work [CT12, CT13, CT14].

5.3. **Specification.** Say that $\mathcal{G} \subset X \times \mathbb{N}$ has *(W)-specification at scale δ* if there exists $\tau \in \mathbb{N}$ such that for every $\{(x_i, n_i)\}_{i=1}^k \subset \mathcal{G}$ there exists a point y and a sequence of "gluing times" $\tau_1, \ldots, \tau_{k-1} \in \mathbb{N}$ with $\tau_i \leq \tau$ such that writing $N_j = \sum_{i=1}^j n_i + \sum_{i=1}^{j-1} \tau_i$, and $N_0 = \tau_0 = 0$, we have

(5.6) $d_{n_j}(f^{N_{j-1}+\tau_{j-1}}y, x_j) < \delta$ for every $1 \leq j \leq k$.

Note that N_j is the time spent for the orbit y to shadow the orbit segments (x_1, n_1) up to (x_j, n_j).

If it is possible to let $\tau_i = \tau$ for all "gluing times", then we say that $\mathcal{G}$ has *(S)-specification at scale δ*. The only specification property used in this paper is (W)-specification. Henceforth, when we write *specification*, we mean the (W)-specification property. We also introduce tail specification in the discrete time setting.

Definition 5.2. *We say that $\mathcal{G} \subset X \times \mathbb{N}$ has* tail specification at scale δ *if there exists $N_0 \in \mathbb{N}$ so that $\mathcal{G} \cap (X \times [N_0, \infty))$ has specification at scale δ; i.e. the specification property holds for orbit segments $(x_i, n_i) \in \mathcal{G}$ with $n_i \geq N_0$. We also sometimes write "$\mathcal{G}$ has specification at scale δ for $n \geq N_0$" to describe this property.*

5.4. **The Bowen property.** The following property agrees with the usual Bowen property in the case when $\mathcal{C} = X \times \mathbb{N}$.

Definition 5.3. *We say that $\phi \colon X \to \mathbb{R}$ has the* Bowen property at scale ε *on $\mathcal{C} \subset X \times \mathbb{N}$ if there exists $K \in \mathbb{R}$ such that $|\Phi_0(x, n) - \Phi_0(y, n)| \leq K$ for all $(x, n) \in \mathcal{C}$ and $y \in B_n(x, \varepsilon)$.*

We sometimes call K the *distortion constant* for the Bowen property. Note that if ϕ has the Bowen property at scale ε on $\mathcal{G}$ with distortion constant K, then for any $M > 0$, ϕ has the Bowen property at scale ε on $\mathcal{G}^M$ with distortion constant given by $K(M) = K + 2M \operatorname{Var}(\phi, \varepsilon)$.

5.5. **Expansivity.** Given $x \in X$ and $\varepsilon > 0$, consider the set

(5.7) $\Gamma_\varepsilon(x) := \{y \in X \mid d(f^n x, f^n y) \leq \varepsilon \text{ for all } n \in \mathbb{Z}\}.$

The map is expansive if there exists $\varepsilon > 0$ such that

(5.8) $\Gamma_\varepsilon(x) = \{x\}$

for every $x \in X$. Following [BF13, CT14], we define the set of non-expansive points at scale ε as $\mathrm{NE}(\varepsilon) := \{x \in X \mid \Gamma_\varepsilon(x) \neq \{x\}\}$, and we say that an f-invariant measure μ is *almost expansive at scale ε* if $\mu(\mathrm{NE}(\varepsilon)) = 0$.

Definition 5.4. *Given a potential ϕ, the* pressure of obstructions to expansivity at scale ε *is*

$$P_{\exp}^{\perp}(\phi, \varepsilon) = \sup_{\mu \in \mathcal{M}_f^e(X)} \left\{ h_\mu(f) + \int \phi \, d\mu \mid \mu(\mathrm{NE}(\varepsilon)) > 0 \right\}$$

$$= \sup_{\mu \in \mathcal{M}_f^e(X)} \left\{ h_\mu(f) + \int \phi \, d\mu \mid \mu(\mathrm{NE}(\varepsilon)) = 1 \right\}.$$

We define a scale-free quantity by

$$P_{\exp}^{\perp}(\phi) = \lim_{\varepsilon \to 0} P_{\exp}^{\perp}(\phi, \varepsilon).$$

5.6. **Results for homeomorphisms.** The following result is the analogue of Theorem A for homeomorphisms.

Theorem 5.5. *Let X be a compact metric space, $f \colon X \to X$ a homeomorphism, and $\phi \colon X \to \mathbb{R}$ a continuous potential function. Suppose that $P_{\exp}^{\perp}(\phi) < P(\phi)$ and $X \times \mathbb{N}$ admits a decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ with the following properties:*

(I) *$\mathcal{G}$ has (W)-specification at any scale $\delta > 0$;*
(II) *ϕ has the Bowen property on $\mathcal{G}$;*
(III) *$P(\mathcal{P} \cup \mathcal{S}, \phi) < P(\phi)$.*

Then (X, f, ϕ) has a unique equilibrium state μ.

Again, we actually prove a version of this theorem with slightly weaker hypotheses, analogous to the more general version of our theorem for flows.

Theorem 5.6. *Let X be a compact metric space, $f \colon X \to X$ a homeomorphism, and $\phi \colon X \to \mathbb{R}$ a continuous potential function. Suppose there are $\delta, \varepsilon > 0$ with $\varepsilon > 40\delta$ such that $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$ and there exists $\mathcal{D} \subset X \times \mathbb{N}$ which admits a decomposition $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ with the following properties:*

(I$'$) *For every $M \in \mathbb{N}$, $\mathcal{G}^M$ has tail (W)-specification at scale δ;*
(II$'$) *ϕ has the Bowen property at scale ε on $\mathcal{G}$;*
(III$'$) *$P(\mathcal{P} \cup \mathcal{S} \cup \mathcal{D}^c, \phi, \delta, \varepsilon) < P(\phi)$.*

Then (X, f, ϕ) has a unique equilibrium state.

One can also formulate versions of these results for non-invertible continuous maps. The key difference is that in this case one should use a one-sided version of almost expansivity, as in [CT14].

The proof of Theorem 5.6 follows the same outline as that of Theorem 2.9. Certain arguments simplify because some technical points associated to the flow case do not arise in discrete time. In view of this, we only provide a sketch proof in §6, pointing out differences with the flow case, and referring the details to corresponding arguments in §4. The argument for showing

that Theorem 5.5 is a corollary of Theorem 5.6 is analogous to the flow case, using the discrete-time analogue of Lemma 2.10.

We obtain the upper level-2 large deviations principle for the equilibrium states provided by the Theorem 5.5.

Theorem 5.7. *Suppose μ is a unique equilibrium state for (X, f, ϕ) provided by Theorem 5.5. Then, μ satisfies the following upper large deviations bound: if A is any weak*-closed and convex subset of the space of Borel probability measures on X, then*

$$(5.9) \qquad \overline{\lim_{n \to \infty}} \frac{1}{n} \log \mu(\mathcal{E}_n^{-1}(A)) \leq \sup_{\nu \in A \cap \mathcal{M}_f(X)} \left(h_\nu(f) + \int \phi \, d\nu - P(\phi) \right),$$

where $\mathcal{E}_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k x}$ is the nth empirical measure associated to x.

The proof of Theorem 5.7 is given in §7. Although we cannot derive the large deviations estimate (5.9) under the weaker hypotheses of Theorem 5.6 in complete generality, we can do so as long as either: ϕ has the Bowen property globally (in particular, when $\phi = 0$, so μ is the measure of maximal entropy), or (X, f) is expansive (see §7).

5.7. **Adapted partitions and generating.** We need some results from [CT14] (which are significantly easier to prove than their counterparts in §3.2 and §3.3). We give the statements here formulated for homeomorphisms. The first result appears as Proposition 2.6 of [CT14], although we note that the corresponding statement in [CT14] is erroneously missing the factor of 2. This factor of 2 propagates through the proof, so the proof in that paper should yield a factor of 56 in the main theorem rather than 28. The proof in this paper, where scales are considered more carefully, brings the smallest integer valued factor back down to 41.

Proposition 5.8. *If μ is almost expansive at scale ε and $\mathcal{A}$ is a measurable partition of X such that every element of $\mathcal{A}$ is contained in $B(x, \varepsilon/2)$ for some $x \in X$, then $\mathcal{A}$ is (two-sided) generating for μ.*

In particular, for an adapted partition $\mathcal{A}_n$ for an (n, ε)-separated set of maximal cardinality, we have $h_\mu(f^n, \mathcal{A}_n) = h_\mu(f^n)$. (This also follows from Theorem 3.2.) The next result is an analogue of Proposition 3.7; for $\phi = 0$ it was proved [CT14, Proposition 2.7], with the same caveat as above regarding the scales.

Proposition 5.9. *If $P_{\exp}^\perp(\phi, \varepsilon) < P(\phi)$, then $P(\phi, \varepsilon/2) = P(\phi)$.*

The following approximation lemma is [CT14, Lemma 3.14].

Lemma 5.10. *Let μ be a finite Borel measure on a compact metric space X and let $P \subset X$ be measurable. Suppose $\mathcal{A}'_n$ is a sequence of partitions such that $\bigcap_n \mathcal{A}'_n(x) = \{x\}$ for every $x \in P$, and let $\delta > 0$. Then for all sufficiently large n there exists a collection $U \subset \mathcal{A}'_n$ such that $\mu(U \triangle P) < \delta$.*

In our setting, this proposition is applied when $\mathcal{A}'_n = f^{-\lfloor n/2 \rfloor} \mathcal{A}_n$, where $\mathcal{A}_n$ is an adapted partition, and P is a positive measure f-invariant set for an almost expansive measure. By invariance of the set P, we can find a collection $U \subset \mathcal{A}_n$ so that $\mu(U \bigtriangleup P) < \delta$, recovering the discrete-time analogue of Proposition 3.10.

6. Sketch Proof of Theorem 5.6

The proof is a simplified version of the proof for flows. We give a sketch proof which can be expanded into a full proof by using the details from §3 and [CT14], and which highlights the differences with the flow case.

Remark 6.1. An alternative approach, which we do not pursue, is to use a suspension flow argument to build a flow from the homeomorphism (see [BW72, §4]), and derive the result as a corollary of Theorem 2.9.

6.1. **Lower bounds on X.** The statements and proofs of the following lemmas are almost identical to those of Lemma 4.1 and Lemma 4.2.

Lemma 6.2. *For every $\gamma > 0$, and $n_1, \ldots, n_k \geq 0$ we have*

$$(6.1) \qquad \Lambda(X, 2\gamma, n_1 + \cdots + n_k) \leq \prod_{j=1}^{k} \Lambda(X, \gamma, \gamma, n_j).$$

Lemma 6.3. *Let ε satisfy $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$. For every $n \in \mathbb{N}$ and $0 < \gamma \leq \varepsilon/4$, we have the inequality $\Lambda(X, \gamma, \gamma, n) \geq e^{nP(\phi)}$.*

The remark after Lemma 4.2 about the relationship between 'two-scale' pressure and the Bowen property also applies in the discrete time case.

6.2. **Upper bounds on $\mathcal{G}$.** Now suppose that $\mathcal{G} \subset X \times \mathbb{N}_0$ has specification at scale $\delta > 0$ for $n \geq N_0$.

Proposition 6.4. *Suppose that $\mathcal{G}$ has specification at scale $\delta > 0$ for $n \geq N_0$ with maximum gap size τ, and $\zeta > \delta$ is such that ϕ has the Bowen property on $\mathcal{G}$ at scale ζ. Then, for every $\gamma > 2\zeta$, there is a constant $C_1 > 0$ so that for every $n_1, \ldots, n_k \geq N_0$, writing $N := \sum_{i=1}^{k} n_i + (k-1)\tau$, and $\theta := \zeta - \delta$, we have*

$$(6.2) \qquad \prod_{j=1}^{k} \Lambda(\mathcal{G}, \gamma, n_j) \leq C_1^k \Lambda(X, \theta, N).$$

The proof is a simplification of that of Proposition 6.4. In the discrete time case, we just show that the map π is 1-1 on each possible configuration of gluing times. With no significant changes to the proofs from the flow case, we also obtain the following results (see Corollary 4.6 and Proposition 4.7).

Corollary 6.5. *Let $\mathcal{G}, \delta, \zeta, \tau, \gamma, \phi, n_j, N, \theta, C_1$ be as in Proposition 6.4, and let $\eta_1, \eta_2 \geq 0$. Suppose further that the Bowen property for ϕ holds at scale η_1. Then*

$$(6.3) \qquad \prod_{j=1}^{k} \Lambda(\mathcal{G}, \theta, \eta_1, n_j) \leq (e^K C_1)^k \Lambda(X, \gamma, \eta_2, N),$$

where K is the distortion constant from the Bowen property.

Lemma 6.6. *Suppose that $\mathcal{G} \subset X \times \mathbb{N}$ has tail specification at scale $\delta > 0$. Suppose that $\gamma > 2\delta$, and $\eta \geq 0$. Suppose that ϕ has the Bowen property on $\mathcal{G}$ at scale $\max\{\gamma/2, \eta\}$. Then there is C_4 such that for every n we have*

$$(6.4) \qquad \qquad \Lambda(\mathcal{G}, \gamma, \eta, n) \leq C_4 e^{nP(\phi)}.$$

6.3. Lower bounds on $\mathcal{G}$.

Lemma 6.7. *Fix a scale $\gamma > 2\delta > 0$. Let $(\mathcal{P}, \mathcal{G}, \mathcal{S})$ be a decomposition for $\mathcal{D} \subseteq X \times \mathbb{N}$ such that*

   (1) $\mathcal{G}$ has tail specification at scale δ ;
   (2) ϕ has the Bowen property on $\mathcal{G}$ at scale 3γ; and
   (3) $P(\mathcal{D}^c, \phi, 2\gamma, 2\gamma) < P(\phi)$ and $P(\mathcal{P} \cup \mathcal{S}, \phi, \gamma, 3\gamma) < P(\phi)$,

Then for every $\alpha_1, \alpha_2 > 0$ there exists $M \in \mathbb{N}$ and $N_1 \in \mathbb{R}$ such that the following is true:

   • For any $n \geq N_1$ and $\mathcal{C} \subset X \times \mathbb{N}$ such that $\Lambda(\mathcal{C}, 2\gamma, 2\gamma, n) \geq \alpha_1 e^{nP(\phi)}$, we have $\Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\gamma, 2\gamma, n) \geq (1 - \alpha_2)\Lambda(\mathcal{C}, 2\gamma, 2\gamma, n)$.

   This proof is like that of Lemma 4.8, but is significantly simpler since we can partition points in $\mathcal{C}_n$ according to the integer values assigned by the decomposition.

6.4. Consequences of lower bound on $\mathcal{G}$.
Again, throughout this section we assume that $\mathcal{G} \subset X \times \mathbb{N}$ and $\delta, \varepsilon > 0$ are such that $\mathcal{G}$ has tail specification at scale δ and $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$. The following lemmas are consequences of Lemma 6.7. The proofs are identical to their counterparts in §4.4.

Lemma 6.8. *Let ε, δ be as above and fix $\gamma \in (2\delta, \varepsilon/8]$ such that ϕ has the Bowen property on $\mathcal{G}$ at scale 3γ. Then for every $\alpha > 0$ there is $M \in \mathbb{N}$ and $N_1 \in \mathbb{N}$ such that for every $n \geq N_1$ we have*

$$(6.5) \qquad \Lambda(\mathcal{G}^M, 2\gamma, 2\gamma, n) \geq (1 - \alpha)\Lambda(X, 2\gamma, 2\gamma, n) \geq (1 - \alpha)e^{nP(\phi)}.$$

Proposition 6.9. *Let $\varepsilon, \delta, \gamma$ be as in Lemma 6.8. Then there are M, N_1, L such that for $n \geq N_1$,*

$$\Lambda(\mathcal{G}^M, 2\gamma, n) \geq e^{-L}e^{nP(\phi)}.$$

As a consequence, $\Lambda(X, 2\gamma, t) \geq e^{-L}e^{nP(\phi)}$ for $n \geq N_1$.

Lemma 6.10. *Let $\varepsilon, \delta, \gamma$ be as in Lemma 6.8. Then there is $C_6 \in \mathbb{R}^+$ such that $C_6^{-1}e^{nP(\phi)} \leq \Lambda(X, 2\gamma, n) \leq \Lambda(X, 2\gamma, 2\gamma, n) \leq C_6 e^{nP(\phi)}$ for every $n \in \mathbb{N}$.*

Lemma 6.11. *Let $\varepsilon, \delta, \gamma, M, N_1$ be as above. Suppose that $\mathcal{G}_n^M \neq \emptyset$ for all $0 < n < N_1$. Then there exists $C_6' \in \mathbb{R}^+$ so that for all $n \in \mathbb{N}$,*

$$\Lambda(\mathcal{G}^M, 2\gamma, n) \geq C_6' e^{nP(\phi)}.$$

6.5. **An equilibrium state with a Gibbs property.** From now on we fix $\varepsilon > 40\delta > 0$ so the hypotheses of Theorem 5.6 are satisfied:

- for every $M \in \mathbb{N}$ there are $N(M), \tau_M \in \mathbb{N}$ such that $\mathcal{G}^M$ has specification at scale δ for $n \geq N(M)$ with transition time τ_M;
- ϕ is Bowen on $\mathcal{G}$ at scale ε (with constant K);
- $P_{\exp}^{\perp}(\phi, \varepsilon) < P(\phi)$;
- $P(\mathcal{D}^c \cup \mathcal{P} \cup \mathcal{S}, \phi, \delta, \varepsilon) < P(\phi)$.

We fix a scale $\rho \in (5\delta, \varepsilon/8]$ and we let $\rho' := \rho - \delta$. For concreteness, we could keep in mind the values $\varepsilon = 48\delta, \rho = 6\delta, \rho' = 5\delta$. Note that the scales can be sharpened a little as described in Remark 4.13.

For each $n \in \mathbb{N}$, let $E_n \subset X$ be a maximizing (n, ρ')-separated set for $\Lambda(X, n, \rho')$. Then consider the measures

$$\nu_n := \frac{\sum_{x \in E_n} e^{\Phi_0(x,n)} \delta_x}{\sum_{x \in E_n} e^{\Phi_0(x,n)}},$$

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

By compactness there is a sequence $n_k \to \infty$ such that μ_{n_k} converges in the weak* topology. Let $\mu = \lim_k \mu_{n_k}$.

Lemma 6.12. *μ is an equilibrium state for (X, f, ϕ).*

This is immediate from the second part of the proof of [Wal82, Theorem 8.6], and the fact that $P(\phi) = P(\phi, \rho')$ by Lemma 5.9. As in the flow case, this argument yields existence of an equilibrium state for ϕ whenever $P_{\exp}^{\perp}(\phi) < P(\phi)$.

The following two lemmas require that $8\rho \leq \varepsilon$ and ϕ has the Bowen property at scale 3ρ, which is true by assumption. The proofs are a straight adaption of those of Lemma 4.16 and Lemma 4.17. The proof of Lemma 4.17 simplifies because all transition times are integer valued, so the step that involves replacing $\hat{q}$ with q' can be removed.

Lemma 6.13. *For sufficiently large M there is $Q_M > 0$ such that for every $(x, n) \in \mathcal{G}^M$ with $n \geq N(M)$ we have*

$$(6.6) \qquad \mu(B_n(x, \rho)) \geq Q_M e^{-nP(\phi) + \Phi_0(x,n)}.$$

Lemma 6.14. *For sufficiently large M there is $Q_M' > 0$ such that for every $(x_1, n_1), (x_2, n_2) \in \mathcal{G}^M$ with $n_1, n_2 \geq N(M)$ and every $q \geq 2\tau_M$ there is an integer $q' \in [q - 2\tau_M, q]$ such that we have*

$$\mu(B_{n_1}(x_1, \rho) \cap f_{-(t_1 + q')} B_{n_2}(x_2, \rho)) \geq Q_M' e^{-(n_1 + n_2)P(\phi) + \Phi_0(x_1, n_1) + \Phi_0(x_2, n_2)}.$$

6.6. **Adapted partitions and positive measure sets.** The proof of the following lemma is identical to that of Lemma 4.18. Note that ε, δ are as in the previous section.

Lemma 6.15. *Let ν be an equilibrium state for ϕ, and let $\gamma \in (2\delta, \varepsilon/8]$. For every $\alpha \in (0,1)$, there is $C_\alpha > 0$ such that if $\{E_n\}_{n\in\mathbb{N}}$ are maximizing $(n, 2\gamma)$-separated sets for $\Lambda(X, 2\gamma, n)$, and $\mathcal{A}_n$ are adapted partitions for E_n, and $E'_n \subset E_n$ satisfies $\nu\left(\bigcup_{x\in E'_n} w_x\right) \geq \alpha$ for all n, then letting $\mathcal{C} = \{(x, n) : x \in E'_n\}$, we have $\Lambda(\mathcal{C}, 2\gamma, 2\gamma, n) \geq C_\alpha e^{nP(\phi)}$.*

6.7. **No mutually singular equilibrium measures.** Let μ be the equilibrium state we have constructed, and suppose that there exists another equilibrium state $\nu \perp \mu$. Let $P \subset X$ be an f-invariant set such that $\mu(P) = 0$ and $\nu(P) = 1$, and let $\mathcal{A}_n$ be adapted partitions for maximizing $(n, 2\rho)$-separated sets E_n. Applying Lemma 5.10, there exists $U_n \subset \mathcal{A}_n$ such that $\frac{1}{2}(\mu + \nu)(U_n \triangle P) \to 0$. In particular, we have $\nu(U_n) \to 1$ and $\mu(U_n) \to 0$, and we can assume without loss of generality that $\inf_n \nu(U_n) > 0$, and so by Lemma 6.15, for $\mathcal{C} = \{(x, n) : x \in E_n \cap U_n\}$, we have

$$\Lambda(\mathcal{C}, 2\rho, 2\rho, n) \geq Ce^{nP(\phi)},$$

and so by Lemma 6.7, there exists M so that

$$\Lambda(\mathcal{C} \cap \mathcal{G}^M, 2\rho, 2\rho, n) \geq \frac{C}{2}e^{nP(\phi)}$$

for sufficiently large n. In other words, letting $E_n^M = \{x \in E_n \mid (x, n) \in \mathcal{G}^M\}$, and using the Bowen property for $\mathcal{G}$, we have

$$\sum_{x\in E_n^M \cap U_n} e^{\Phi_0(x,n)} \geq e^{-K} \sum_{x\in E_n^M \cap U_n} e^{\Phi_{2\rho}(x,n)} \geq \frac{C}{2}e^{-K}e^{nP(\phi)}.$$

Finally, we use the Gibbs property in Lemma 6.13 to observe that since $U_n \supset B_n(x, \rho)$ for every $x \in E_n \cap U_n$, we have

$$\mu(U_n) \geq \sum_{x\in E_n^M \cap U_n} Q_M e^{-nP(\phi)+\Phi_0(x,n)} \geq Q_M e^{-K}C/2 > 0,$$

contradicting the fact that $\mu(U_n) \to 0$. This contradiction implies that any equilibrium state ν is absolutely continuous with respect to μ.

6.8. **Ergodicity.** The following result is the final ingredient needed to complete the proof of Theorem 5.6. The proof is a straight adaptation of Proposition 4.19.

Proposition 6.16. *The equilibrium state μ constructed above is ergodic.*

In conclusion, §4.7 showed that any equilibrium state ν is absolutely continuous with respect to μ, and since μ is ergodic, this in turn implies that $\nu = \mu$, which completes the proof of Theorem 5.6.

6.9. **Upper Gibbs bound.** The upper Gibbs bound is proved exactly as in Proposition 4.21.

Proposition 6.17. *Let ε, δ be as in Theorem 5.6 and let $\gamma \in (4\delta, \varepsilon/4]$. Then there is $Q > 0$ such that for every $(x, n) \in X \times \mathbb{N}$, the unique equilibrium state μ satisfies $\mu(B_n(x, \gamma)) \leq Q e^{-nP(\phi) + \Phi_\gamma(x,n)}$.*

7. Proof of Theorem 5.7

We can obtain a large deviations result from the upper Gibbs bound in Proposition 6.17 by applying results of Pfister and Sullivan [PS05].

Recall from [PS05, Definition 3.4] that a function $e_\mu \colon X \to \mathbb{R}^+$ is said to be an *upper-energy function* for μ if $-e_\mu$ is upper semicontinuous, bounded, and

$$(7.1) \qquad \lim_{\varepsilon \to 0} \overline{\lim_{n \to \infty}} \sup_{x \in X} \frac{1}{n} \left(\log \mu(B_n(x, \varepsilon)) + S_n e_\mu(x) \right) \leq 0,$$

where $S_n e_\mu(x) = \sum_{i=0}^{n-1} e_\mu(f^i x)$. If e_μ is an upper-energy function for μ, then [PS05, Theorem 3.2] shows that given any closed convex subset A of the space of Borel probability measures on X, we have

$$(7.2) \qquad \overline{\lim_{n \to \infty}} \frac{1}{n} \log \mu\{x \mid \mathcal{E}_n(x) \in A\} \leq \sup_{\nu \in A \cap \mathcal{M}_f(X)} \left(h_\nu(f) - \int e_\mu \, d\nu \right).$$

Comparing (7.2) to (5.9), we see that to prove the remaining claim of Theorem 5.7, it suffices to show that $e_\mu(x) := P(\phi) - \phi(x)$ is an upper energy function for μ.

To this end, observe that $-e_\mu$ is continuous and hence bounded, so it suffices to establish (7.1). For all sufficiently small $\gamma > 0$, Proposition 6.17 gives

$$(7.3) \qquad \begin{aligned} \log \mu(B_n(x, \gamma)) &\leq \log Q - nP(\phi) + \Phi_\gamma(x, n) \\ &\leq \log Q - nP(\phi) + S_n \phi(x) + n \operatorname{Var}(\phi, \gamma), \end{aligned}$$

and so

$$\tfrac{1}{n} (\log \mu(B_n(x, \gamma)) + S_n e_\mu(x)) \leq \operatorname{Var}(\phi, \gamma) + \tfrac{1}{n} \log Q$$

for every $x \in X$ and $n \in \mathbb{N}$. Sending $n \to \infty$ gives

$$\overline{\lim_{n \to \infty}} \sup_{x \in X} \frac{1}{n} (\log \mu(B_n(x, \gamma)) + S_n e_\mu(x)) \leq \operatorname{Var}(\phi, \gamma),$$

and sending $\gamma \to 0$ gives (7.1). By [PS05, Theorem 3.2], this completes the proof of Theorem 5.7.

The above argument is even simpler when μ has the standard upper Gibbs property, which yields $\log \mu(B_n(x, \gamma)) \leq \log Q - nP(\phi) + S_n \phi(x)$. By Proposition 6.17, any unique equilibrium state μ provided by Theorem 5.6 satisfies the weak upper Gibbs property

$$\mu(B_n(x, \gamma)) \leq Q e^{-nP(\phi) + \Phi_\gamma(x,n)}.$$

By Remark 2.1, if ϕ has the Bowen property (in particular, if $\phi = 0$), then this reduces to the standard Gibbs property, and thus μ satisfies the upper bound of the large deviations principle.

Similarly, if (X, f) is expansive at scale γ, then (7.3) gives

$$\tfrac{1}{n}(\log \mu(B_n(x, \gamma)) + S_n e_\mu(x)) \leq \tfrac{1}{n}|\Phi_\gamma(x, n) - \Phi_0(x, n)| + \tfrac{1}{n}\log Q \to 0$$

for the unique equilibrium state μ of Theorem 5.6. and we see once again that μ satisfies the upper bound of the large deviations principle.

8. Proof of Theorem 3.2

We use the notation of §3.1. We fix a homeomorphism f and a metric d, and suppress this from our notation for this section. We begin with the following result, which is an adaptation of [Bow72, Proposition 2.2]. For a set $Z \subset X$, we let

$$\Lambda_n^{\mathrm{span}}(Z, \delta) = \inf \left\{ \sum_{x \in E} e^{\Phi_0(x,n)} \mid E \text{ is } (t, \delta)\text{-spanning for } Z \right\}.$$

Proposition 8.1. *Fix $h > h^*(\mu, \varepsilon)$. For every $\delta, \eta > 0$ there is $c \in \mathbb{R}$ and $Z \subset X$ such that $\mu(Z) > 1 - \eta$ and*

$$(8.1) \qquad \Lambda_n^{span}(B_n(x, \varepsilon), \delta) \leq ce^{hn}$$

for every $x \in Z$ and $n \in \mathbb{N}$.

Proof. Let $Y = \{y \in X \mid h(\Gamma_\varepsilon(x)) \leq h^*(\mu, \varepsilon)\}$, and note that $\mu(Y) = 1$. Fix $h' \in (h^*(\mu, \varepsilon), h)$. Following Bowen's proof, for each $y \in Y$ choose $m(y) \in \mathbb{N}$ such that there is a set $E(y)$ that $(m, \delta/4)$-spans $\Gamma_\varepsilon(y)$ and has $\#E(y) \leq e^{h'm(y)}$. Consider the open set

$$U(y) = \bigcup_{z \in \Gamma_\varepsilon(y)} B_{m(y)}(z, \delta/4).$$

Again following Bowen's proof, note that there is $N = N(y) \in \mathbb{N}$ and an open neighbourhood $V(y)$ of y such that every $u \in V(y)$ has

$$B_{[-N,N]}(u, \varepsilon) := \{y : d(f^i y, f^i u) < \varepsilon \text{ for } i = -N, \ldots, N\} \subset U(y).$$

Let $K \subset X$ be a $\delta/2$-spanning set, and fix $\alpha > 0$ such that $h' + \alpha \log \#Q < h$. Let $W \subset Y$ be a compact set such that $\mu(Y') > 1 - \alpha/2$, and let $y_1, \ldots, y_s \in Y$ be such that $W \subset \bigcup_{i=1}^s V(y_i)$. Let

$$N := 1 + \max\{N(y_1), \ldots, N(y_s), m(y_1), \ldots, m(y_s)\}.$$

Given $n \in \mathbb{N}$ and $x \in X$, let $a_n(x) = \tfrac{1}{n}\#\{0 \leq k < n \mid f^k(x) \notin W\}$. For $\ell \in \mathbb{N}$, let $Z_\ell = \{x \in X \mid a_n(x) \leq \alpha \text{ for all } n \geq \ell\}$. By ergodicity of μ and the Birkhoff ergodic theorem, there is ℓ such that $\mu(Z_\ell) > 1 - \eta$, and we take $Z = Z_\ell$.

Given $x \in Z$ and $n \geq \max\{2N, \ell\}$, we choose integers $0 = t_0 < s_1 \leq t_1 < s_2 \leq t_2 < \cdots s_r \leq t_r = n$. Put $s_1 = N$ and then construct t_i, s_i iteratively as follows.

(1) t_i is the smallest value of $t \geq s_i$ such that $f^t(x) \in y_{k_i}$ for some k_i.

(2) $s_{i+1} = t_i + m(y_{k_i})$.

Once some $s_r \geq n - N$ we put $t_r = n$. Observe that for each $1 \leq i < r$, the set $E(y_{k_i})$ is an $(m, \delta/2)$-spanning set for $U(y_{k_i}) \supset f^{t_i}(B_n(x, \varepsilon))$. And for each $s_i \leq t < t_i$, the point $f^t(x)$ is within $\delta/2$ of some element of K. Similarly for every $0 \leq t < N$ and $n - N \leq t < n$. It follows from Bowen's Lemma 2.1 that

$$\Lambda_n^{\mathrm{span}}(B_n(x, \varepsilon), \delta) \leq (\#K)^{2N + \sum_i (t_i - s_i)} e^{h' \sum_i (s_i - t_i)}.$$

By definition of t_i and Z, we have $\sum_i (t_i - s_i) \leq n a_n(x) \leq \alpha n$, and so

$$\Lambda_n^{\mathrm{span}}(B_n(x, \varepsilon), \delta) \leq (\#K)^{2N + \alpha n} e^{h' n} = (\#K)^{2N} e^{n(h' + \alpha \log(\#K))} \leq (\#K)^{2N} e^{hn},$$

where the last inequality follows from the choice of α. Taking $c = (\#K)^{\max\{2N, \ell\}}$ completes the proof of Proposition 8.1. $\qquad \square$

With $\mathcal{A}$ as above, let $\mathcal{B}$ be another partition (eventually we will send the diameter of $\mathcal{B}$ to 0), and consider the common refinement $\mathcal{A} \vee \mathcal{B}$. Consider the conditional entropy

$$H_\mu(\mathcal{B} \mid \mathcal{A}) = \sum_{A \in \mathcal{A}} \mu(A) \sum_{B \in \mathcal{B}, B \subset A} -\mu(B) \log \mu(B),$$

$$h_\mu(f, \mathcal{B} \mid \mathcal{A}) = \lim_{n \to \infty} \tfrac{1}{n} H_\mu(\mathcal{B} \mid \mathcal{A}).$$

A standard computation shows that $h_\mu(f, \mathcal{A} \vee \mathcal{B}) = h_\mu(f, \mathcal{A}) + h_\mu(f, \mathcal{B} \mid \mathcal{A})$. Moreover, if $\operatorname{diam} \mathcal{B} \to 0$ then $h_\mu(f, \mathcal{A} \vee \mathcal{B}) \to h_\mu(f)$; for a proof, see [Bow72, Lemma 3.2]. Thus Theorem 3.2 will follow immediately from the following proposition.

Proposition 8.2. *For every $\alpha > 0$, there is a partition $\mathcal{B}$ such that $\operatorname{diam} \mathcal{B} < \alpha$ and $h_\mu(f, \mathcal{B} \mid \mathcal{A}) \leq h^*(\mu, \varepsilon)$.*

Before proving Proposition 8.2 we prove the following general result, which is similar in spirit to the Katok entropy formula [Kat80], but simpler to prove because it uses partitions instead of Bowen balls. In what follows, whenever $\mathcal{A}' \subset \mathcal{A}$ for some partition $\mathcal{A}$, we continue to use our notation convention of writing $\mu(\mathcal{A}')$ to represent $\mu \left(\bigcup_{A \in \mathcal{A}'} A \right)$.

Lemma 8.3. *Let (X, f, μ) be an ergodic measure-preserving transformation and let $\mathcal{A}, \mathcal{B}$ be finite partitions of X. Suppose that for some $c, \gamma > 0$ and $h \geq 0$ there are sequences $\mathcal{A}'_n \subset \mathcal{A}^n$ and $\mathcal{C}'_n \subset (\mathcal{A} \vee \mathcal{B})^n$ such that for sufficiently large n we have $\mu(\mathcal{A}'_n) > \gamma$, and that writing $\mathcal{C}'_n(A) = \{C \in \mathcal{C}'_n \mid C \in A\}$, for every $A \in \mathcal{A}'_n$ we have*

(1) $\mu(\mathcal{C}'_n(A)) > \gamma \mu(A)$;

(2) $\#\mathcal{C}'_n(A) \leq c e^{nh}$.

Then $h_\mu(f, \mathcal{B} \mid \mathcal{A}) \leq h$.

Proof. Since $h_\mu(f, \mathcal{B} \mid \mathcal{A}) = h_\mu(f, \mathcal{A} \vee \mathcal{B}) - h_\mu(f, \mathcal{A})$, the Shannon–McMillan–Breiman theorem gives the following for μ-a.e. $x \in X$:

$$h_\mu(f, \mathcal{B} \mid \mathcal{A}) = -\lim_{n\to\infty} \frac{1}{n} \log \left(\frac{\mu((\mathcal{A} \vee \mathcal{B})^n(x))}{\mu(\mathcal{A}^n(x))} \right).$$

Fix $0 < \gamma' \leq \gamma^2/2$. For every $h' < h_\mu(f, \mathcal{B} \mid \mathcal{A})$, there is $E \subset X$ and $N \in \mathbb{N}$ such that $\mu(E) \geq 1 - \gamma'$ and for every $n \geq N$ and $x \in E$, we have

$$(8.2) \qquad \mu\left((\mathcal{A} \vee \mathcal{B})^n(x)\right) \leq e^{-nh'} \mu(\mathcal{A}^n(x)).$$

Using $\mu(X \setminus E) \leq \gamma'$ and $\mu(\mathcal{A}'_n) > \gamma$, we have

$$\frac{\sum_{A\in\mathcal{A}'_n} \mu(A \cap E)}{\sum_{A\in\mathcal{A}'_n} \mu(A)} = \frac{\sum_{A\in\mathcal{A}'_n} \mu(A) - \mu(A \cap (X \setminus E))}{\sum_{A\in\mathcal{A}'_n} \mu(A)} \geq 1 - \frac{\gamma'}{\gamma} \geq 1 - \frac{\gamma}{2}.$$

It follows that there is $A_n \in \mathcal{A}'_n$ with $\mu(A_n \cap E)/\mu(A_n) \geq 1 - \gamma/2$. For this choice of A_n we have

$$\sum_{C\in\mathcal{C}'_n(A_n)} \mu(C \cap E) = \mu(A_n \cap E) - \mu\left(A_n \setminus \bigcup_{C\in\mathcal{C}'_n(A_n)} C \right)$$

$$\geq \left(1 - \frac{\gamma}{2}\right) \mu(A_n) - (1 - \gamma)\mu(A_n) = \frac{\gamma}{2}\mu(A_n).$$

Let $\mathcal{C}''_n = \{C \in \mathcal{C}'_n(A_n) \mid \mu(C \cap E) > 0\}$. Then for every $n \geq N$ and $C \in \mathcal{C}''_n$, it follows from (8.2) that

$$(8.3) \qquad \mu(C) \leq e^{-nh'} \mu(A_n).$$

Together with the estimate on $\#\mathcal{C}'_n(A_n)$ from the hypothesis of the proposition, this gives

$$\frac{\gamma}{2}\mu(A_n) \leq \sum_{C\in\mathcal{C}''_n} \mu(C) \leq ce^{nh}e^{-nh'}\mu(A_n),$$

and so we get $e^{n(h-h')} \geq \gamma/(2c)$ for all large n, whence $h \geq h'$. Since this is true for any $h' < h_\mu(f, \mathcal{B} \mid \mathcal{A})$, we conclude that $h \geq h_\mu(f, \mathcal{B} \mid \mathcal{A})$, as desired. This completes the proof of Lemma 8.3. $\square$

Proof of Proposition 8.2. Let $\mathcal{B}$ be a partition of diameter $< \alpha$ whose boundary $\partial\mathcal{B}$ carries zero μ-measure.

Fix $h > h^*(\mu, \varepsilon)$ and $0 < \beta' < \beta < \frac{1}{2}$. By Stirling's formula there is a constant K such that

$$(8.4) \qquad \sum_{0\leq j\leq \beta n} \binom{n}{j} \leq Ke^{(-\beta\log\beta)n}$$

Our goal is to use this together with Proposition 8.1 to produce $\mathcal{A}'_n, \mathcal{C}'_n$ satisfying the conditions of Lemma 8.3 with

$$(8.5) \qquad \#\mathcal{C}'_n(A) \leq cKe^{(h-\beta\log\beta)n},$$

and then use the fact that h can be taken arbitrarily close to $h^*(\mu, \varepsilon)$, and β can be taken arbitrarily close to 0, to establish Theorem 3.2.

Our next step is to reduce to the collection of orbits that do not spend too much time near the boundary of $\mathcal{B}$. We start by observing that because $\mu(\partial\mathcal{B}) = 0$, there is $\delta > 0$ such that $\mu(B(\partial\mathcal{B}, 2\delta)) < \beta'$, where $B(\partial\mathcal{B}, \delta))$ denotes the δ-neighbourhood of the boundary of $\mathcal{B}$. For each n, we consider the set
$$\mathcal{B}'_n = \{x \mid \#\{0 \le k < n \mid d(f^k x, \partial\mathcal{B}) < 2\delta\} \le \beta n\}.$$
By the Birkhoff ergodic theorem we have $\mu(\mathcal{B}_n) \to 1$ as $n \to \infty$.

Fix $0 < \gamma < \gamma' < 1$. By Proposition 8.1, there is $Z \subset X$ such that $\mu(Z) > \gamma'$ and $c \in \mathbb{R}$ such that (8.1) holds for every $x \in Z$ and $n \in \mathbb{N}$. Let $\mathcal{B}''_n = \mathcal{B}_n \cap Z$ and note that for all sufficiently large n we have $\mu(\mathcal{B}''_n) > \gamma'$.

Now let $\mathcal{C}'_n = \{C \in (\mathcal{A}\vee\mathcal{B})^n \mid \mu(C\cap\mathcal{B}''_n) > 0\}$. As before, for each $A \in \mathcal{A}^n$ write $\mathcal{C}'_n(A) = \{C \in \mathcal{C}'_n \mid C \subset A\}$. Let
$$\mathcal{A}'_n = \{A \in \mathcal{A}^n \mid \mu(\mathcal{C}'_n(A)) > \gamma\mu(A)\}.$$
For n large we have
$$\gamma' < \mu(\mathcal{B}''_n) \le \mu(\mathcal{C}'_n) \le \gamma(1 - \mu(\mathcal{A}'_n)) + \mu(\mathcal{A}'_n),$$
and so $\mu(\mathcal{A}'_n) \ge \frac{\gamma'-\gamma}{1-\gamma}$. Thus the first two hypotheses of Lemma 8.3 are satisfied, and it remains only to get a bound on $\#\mathcal{C}'_n(A)$ for $A \in \mathcal{A}'_n$.

To this end, fix $A \in \mathcal{A}'_n$ and $x \in \mathcal{B}''_n \cap A$. Note that $B_n(x, \varepsilon) \supset A = \mathcal{A}^n(x)$ since $\operatorname{diam}\mathcal{A} < \varepsilon$, and since $x \in \mathcal{B}''_n \subset Z$, by (8.1) we see that the set A admits an (n, δ)-spanning set S with $\#S \le ce^{nh}$.

Now we are in a position to estimate $\#\mathcal{C}'_n(A)$. Let $R \subset \mathcal{B}''_n$ be a set containing exactly one element of $C\cap\mathcal{B}''_n$ for every $C \in \mathcal{C}'_n(A)$. In particular, $\#\mathcal{C}'_n(A) = \#R$, and S is (n, δ)-spanning for R. Given $x, y \in R$, let $d_n^H(x, y)$ denote the number of times $k \in [0, n)$ for which $f^k x, f^k y$ are in different partition elements of $\mathcal{B}$. In other words, $d_n^H(x, y)$ is the Hamming distance between the $\mathcal{B}$-codings of x, y.

Lemma 8.4. *If $d_n^H(x, y) > \beta n$ then $d_n(x, y) > 2\delta$, where d_n is the usual nth Bowen metric.*

Proof. By the definition of $\mathcal{B}_n$, we see that there are at most βn values of $k \in [0, n)$ for which $f^k x, f^k y \in B(\partial\mathcal{B}, 2\delta)$. Thus there is some k for which $f^k x, f^k y$ lie in different elements of $\mathcal{B}$ and one of them is at least 2δ from the boundary. $\square$

Given $x \in R$, the number of $y \in R$ such that $d_n^H(x, y) \le \beta n$ is at most $\sum_{0 \le j \le \beta n} \binom{n}{j}$, which by (8.4) is at most $Ke^{(-\beta\log\beta)n}$. Thus Lemma 8.4 implies given $x \in S$, the Bowen ball $B_n(x, \delta)$ intersects at most $Ke^{(-\beta\log\beta)n}$ elements of R. In particular, we have
$$\#R \le Ke^{(-\beta\log\beta)n}\#S \le cKe^{(h-\beta\log\beta)n}.$$
This proves (8.5), and now we can apply Lemma 8.3 to $\mathcal{A}'_n, \mathcal{C}'_n$ to get $h_\mu(f, \mathcal{B} \mid \mathcal{A}) \le h - \beta\log\beta$. Since h can be taken arbitrarily close to $h^*(\mu, \varepsilon)$ and β can be taken arbitrarily close to 0, we obtain $h_\mu(f, \mathcal{B} \mid \mathcal{A}) \le h^*(\mu, \varepsilon)$, completing the proof of Proposition 8.2. $\square$

Theorem 3.2 follows from Proposition 8.2 and the remarks preceding it.

References

[Abr59] L. M. Abramov, *On the entropy of a flow*, Dokl. Akad. Nauk SSSR **128** (1959), 873–875. MR 0113985 (22 #4816)

[BCFT16] Keith Burns, Vaughn Climenhaga, Todd Fisher, and Daniel J. Thompson, *Unique equilibrium states for geodesic flows in nonpositive curvature*, Work in progress, 2016.

[BF13] Jérôme Buzzi and Todd Fisher, *Entropic stability beyond partial hyperbolicity*, J. Mod. Dyn. **7** (2013), no. 4, 527–552. MR 3177771

[Bow72] Rufus Bowen, *Entropy-expansive maps*, Trans. Amer. Math. Soc. **164** (1972), 323–331. MR 0285689 (44 #2907)

[Bow74] ———, *Maximizing entropy for a hyperbolic flow*, Math. Systems Theory **7** (1974), no. 4, 300–303. MR 0385928 (52 #6787)

[Bow08] ———, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, revised ed., Lecture Notes in Mathematics, vol. 470, Springer-Verlag, Berlin, 2008, With a preface by David Ruelle, Edited by Jean-René Chazottes. MR 2423393

[Bow75] ———, *Some systems with unique equilibrium states*, Math. Systems Theory **8** (1974/75), no. 3, 193–202. MR 0399413 (53 #3257)

[BR75] Rufus Bowen and David Ruelle, *The ergodic theory of Axiom A flows*, Invent. Math. **29** (1975), no. 3, 181–202. MR 0380889 (52 #1786)

[BSS02] Luis Barreira, Benoît Saussol, and Jörg Schmeling, *Higher-dimensional multifractal analysis*, J. Math. Pures Appl. (9) **81** (2002), no. 1, 67–91. MR 1994883 (2004g:37038)

[BW72] Rufus Bowen and Peter Walters, *Expansive one-parameter flows*, J. Differential Equations **12** (1972), 180–193. MR 0341451 (49 #6202)

[CFT15] Vaughn Climenhaga, Todd Fisher, and Daniel J. Thompson, *Unique equilibrium states for the robustly transitive diffeomorphisms of Mañé and Bonatti–Viana*, Preprint, arXiv:1505.06371, 2015.

[Cli16] Vaughn Climenhaga, *Specification and towers in shift spaces*, Preprint, arXiv:1502.00931v2, 2016.

[CP16] Vaughn Climenhaga and Yakov Pesin, *Building thermodynamics for non-uniformly hyperbolic maps*, Preprint, 2016.

[CRL11] Henri Comman and Juan Rivera-Letelier, *Large deviation principles for non-uniformly hyperbolic rational maps*, Ergodic Theory Dynam. Systems **31** (2011), no. 2, 321–349. MR 2776378 (2012c:37059)

[CT12] Vaughn Climenhaga and Daniel J. Thompson, *Intrinsic ergodicity beyond specification: β-shifts, S-gap shifts, and their factors*, Israel J. Math. **192** (2012), 785–817. MR 3009742

[CT13] ———, *Equilibrium states beyond specification and the Bowen property*, J. Lond. Math. Soc. (2) **87** (2013), no. 2, 401–427. MR 3046278

[CT14] ———, *Intrinsic ergodicity via obstruction entropies*, Ergodic Theory Dynam. Systems **34** (2014), no. 6, 1816–1831. MR 3272773

[DKU90] Manfred Denker, Gerhard Keller, and Mariusz Urbański, *On the uniqueness of equilibrium states for piecewise monotone mappings*, Studia Math. **97** (1990), no. 1, 27–36. MR 1074766 (92c:58061)

[Fra77] Ernesto Franco, *Flows with unique equilibrium states*, Amer. J. Math. **99** (1977), no. 3, 486–514. MR 0442193 (56 #579)

[HK82] Franz Hofbauer and Gerhard Keller, *Equilibrium states for piecewise monotonic transformations*, Ergodic Theory Dynam. Systems **2** (1982), no. 1, 23–43. MR 684242 (85f:58069)

[IT10]   Godofredo Iommi and Mike Todd, *Natural equilibrium states for multi-modal maps*, Comm. Math. Phys. **300** (2010), no. 1, 65–94. MR 2725183 (2011m:37048)

[Kat80]   A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Inst. Hautes Études Sci. Publ. Math. (1980), no. 51, 137–173. MR 573822 (81i:28022)

[Kif90]   Yuri Kifer, *Large deviations in dynamical systems and stochastic processes*, Trans. Amer. Math. Soc. **321** (1990), no. 2, 505–524. MR 1025756 (91e:60091)

[Pav16]   Ronnie Pavlov, *On intrinsic ergodicity and weakenings of the specification property*, Adv. Math. **295** (2016), 250–270. MR 3488036

[PP90]   William Parry and Mark Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque (1990), no. 187-188, 268.

[PS05]   C.-E. Pfister and W. G. Sullivan, *Large deviations estimates for dynamical systems without the specification property. Applications to the β-shifts*, Nonlinearity **18** (2005), no. 1, 237–261. MR 2109476 (2005h:37015)

[PSZ16]   Y. Pesin, S. Senti, and K. Zhang, *Thermodynamics of towers of hyperbolic type*, Trans. Amer. Math. Soc. (2016), electronically published on March 1, 2016, DOI: http://dx.doi.org/10.1090/tran/6599 (to appear in print).

[Rue76]   David Ruelle, *A measure associated with axiom-A attractors*, Amer. J. Math. **98** (1976), no. 3, 619–654. MR 0415683 (54 #3763)

[Sar99]   Omri M. Sarig, *Thermodynamic formalism for countable Markov shifts*, Ergodic Theory Dynam. Systems **19** (1999), no. 6, 1565–1593.

[Wal82]   Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79, Springer-Verlag, New York-Berlin, 1982. MR 648108 (84e:28017)

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HOUSTON, HOUSTON, TEXAS 77204

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, 100 MATH TOWER, 231 WEST 18TH AVENUE, COLUMBUS, OHIO 43210

E-mail address: climenha@math.uh.edu

E-mail address: thompson@math.osu.edu